Pontifícia Universidade Católica
do Rio de Janeiro

Gilberto Oliveira Boaretto

Essays concerning inflation forecasting: disaggregation, combination of forecasts, and unstructured data

Tese de Doutorado

Thesis presented to the Programa de Pós-graduação em Economia of PUC-Rio in partial fulfillment of the requirements for the degree of Doutor em Economia.

Advisor: Prof. Marcelo Cunha Medeiros

Rio de Janeiro
May 2023

Gilberto Oliveira Boaretto

Essays concerning inflation forecasting: disaggregation, combination of forecasts, and unstructured data

Thesis presented to the Programa de Pós-graduação em Economia of PUC-Rio in partial fulfillment of the requirements for the degree of Doutor em Economia. Approved by the Examination Committee.

Prof. Marcelo Cunha Medeiros
Advisor
Department of Economics – PUC-Rio

Prof. Márcio Gomes Pinto Garcia
Department of Economics – PUC-Rio

Prof. João Victor Issler
Brazilian School of Economics and Finance – FGV EPGE

Prof. Marcelo Fernandes
São Paulo School of Economics – FGV EESP

Dr. Gabriel Filipe Rodrigues Vasconcelos
Quantitative Research – BOCOM BBM

Rio de Janeiro, May 11th, 2023

Gilberto Oliveira Boaretto

BA in Economics from the Federal University of Uberlândia (UFU) in 2015.

MSc in Economics from the University of São Paulo (USP) at Ribeirão Preto in 2018.

Bibliographic data

Boaretto, Gilberto Oliveira

Essays concerning inflation forecasting: disaggregation, combination of forecasts, and unstructured data / Gilberto Oliveira Boaretto; advisor: Marcelo Cunha Medeiros. – Rio de janeiro: PUC-Rio, Departamento de Economia, 2023.

v., 110 f: il. color. ; 30 cm

Tese (doutorado) - Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Economia.

Inclui bibliografia

1. Economia – Teses. 2. Econometria – Teses. 3. Macroeconomia – Teses. 4. Previsão de Inflação. 5. Aprendizado de Máquina. 6. Análise Desagregada. 7. Combinação de Previsões. 8. Notícias. I. Medeiros, Marcelo Cunha. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Economia. III. Título.

CDD: 330

To my family and friends.

Acknowledgments

To my mother Lucia, for her unwavering support throughout another stage of my life. You have been instrumental in helping me reach this point.

To Marcos, my father Ventomar, my grandparents, and my godparents Elzeli and Geraldo, for their continuous support throughout my academic journey.

To my advisor, Marcelo C. Medeiros, with whom I had the privilege of gaining extensive knowledge during my PhD. I regard him as an exemplary professor, researcher, and highly productive individual who is deeply passionate about his work.

To professors Marcelo Fernandes, João Victor Issler, Gabriel Vasconcelos, and Márcio Garcia for their participation in the examination committee and their valuable suggestions that greatly contributed to the final version of this dissertation.

To all the professors of the Department of Economics at PUC-Rio, from whom I have gained significant knowledge in teaching and research. I am grateful for their guidance and advice, which directly and indirectly contributed to this dissertation.

To the administrative staff of the Department of Economics at PUC-Rio, for their support and assistance, I am grateful.

To professors and my former advisors Cleomar Gomes and Márcio Laurini, who continued to provide guidance and support.

To all old and new friends. While I will not mention their names to avoid any injustice, they all know their respective direct and indirect contributions to the completion of this dissertation.

To PUC-Rio, CAPES, CNPq, and FAPERJ for the financial support provided throughout my PhD.

To the Brazilian State and all its sponsors.

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) – Finance Code 001.

Abstract

Boaretto, Gilberto Oliveira; Medeiros, Marcelo Cunha (Advisor). **Essays concerning inflation forecasting: disaggregation, combination of forecasts, and unstructured data**. Rio de Janeiro, 2023. 110p. Tese de doutorado – Departamento de Economia, Pontifícia Universidade Católica do Rio de Janeiro.

This dissertation consists of three essays concerning inflation forecasting, taking the Brazilian case as an application. In the first essay, we examine the effectiveness of several forecasting methods for predicting inflation, focusing on aggregating disaggregated forecasts. We consider different disaggregation levels for inflation and employ a range of traditional time series techniques, as well as linear and nonlinear machine learning (ML) models that deal with a larger number of predictors. For many forecast horizons, aggregation of disaggregated forecasts performs just as well as survey-based expectations and models generating forecasts directly from the aggregate. Overall, ML methods outperform traditional time series models in predictive accuracy, with outstanding performance in forecasting disaggregates. In our second essay, we investigate the potential benefits of combining individual inflation forecasts by proposing a time-varying bias correction for the average forecast. Our analysis includes estimations using both rolling windows and state-space models that use the recursiveness of the Kalman filter. We achieve good forecast performance for models based on small rolling windows for shorter and intermediate forecast horizons, while a state-space model performs slightly worse than procedures based on rolling windows. In the third essay, we use supervised learning to generate forward-looking indexes based on tweets and news articles for accumulated inflation and investigate whether these indexes can improve inflation forecasting performance. Our results indicate that news-based indexes provide significant predictive gains, particularly for 3- and 12-month-ahead horizons. These findings suggest that incorporating more information sources than just expectations based on experts' opinions can lead to more accurate forecasts.

Keywords

Inflation Forecasting   Machine Learning   Disaggregated Analysis
Combination of Forecasts   News

Resumo

Boaretto, Gilberto Oliveira; Medeiros, Marcelo Cunha. **Ensaios sobre previsão de inflação: desagregação, combinação de previsões e dados não estruturados**. Rio de Janeiro, 2023. 110p. Tese de Doutorado – Departamento de Economia, Pontifícia Universidade Católica do Rio de Janeiro.

Esta tese consiste em três ensaios sobre previsão de inflação, com foco na inflação brasileira. No primeiro ensaio, examinamos a eficácia de vários métodos de previsão para prever a inflação, com foco na agregação de previsões desagregadas. Consideramos diferentes níveis de desagregação para a inflação e empregamos uma variedade de técnicas tradicionais de séries temporais, bem como modelos lineares e não lineares de aprendizado de máquina que lidam com um número grande de preditores. Para muitos horizontes de previsão, a agregação de previsões desagregadas performa tão bem quanto expectativas baseadas em coleta e modelos que geram previsões a partir do agregado. No geral, os métodos de aprendizado de máquina superam os modelos de séries temporais tradicionais em precisão preditiva, com excelente desempenho para os desagregados da inflação. Em nosso segundo ensaio, investigamos os potenciais benefícios de combinar previsões de inflação individuais ao propor uma correção para viés variável no tempo da média de previsões. Nossa análise inclui estimações empregando janelas rolantes e modelos em espaço de estados que usam a recursividade do filtro de Kalman. Obtivemos um bom desempenho de previsão para modelos baseados em janelas rolantes pequenas em horizontes de previsão curtos e intermediários, enquanto um modelo em espaço de estados obtem um desempenho um pouco pior do que os procedimentos baseados em janelas rolantes. No terceiro ensaio, usamos aprendizado supervisionado para gerar índices prospectivos baseados em tweets e notícias para inflação acumulada e investigamos se esses índices podem melhorar o desempenho da previsão de inflação. Nossos resultados indicam que os índices baseados em notícias fornecem ganhos preditivos significativos, principalmente para os horizontes de 3 e 12 meses à frente. Esses achados sugerem que a incorporação de mais fontes de informação do que apenas expectativas baseadas em opiniões de especialistas pode levar a previsões mais precisas.

Palavras-chave

Previsão de Inflação   Aprendizado de Máquina   Análise Desagregada
Combinação de Previsões   Notícias

Table of contents

List of figures

List of tables

"This is nothing but a convenient way of expressing opinions about real phenomena. But the probability concept has the advantage that it is 'analytic', we can derive new statements from it by the rules of logic."

Trygve Magnus Haavelmo, *The Probability Approach in Econometrics*.

Introduction

Forecasting inflation is crucial for a wide range of economic activities, including investment-consumption decisions, risk management, investment decisions, price setting by firms, wage negotiations, and the conduction of monetary and fiscal policies. As a result, economists and econometricians are continually in pursuit of forecasting procedures that can enhance the accuracy of forecasts. The advent of vast amounts of information (data) and the emergence of new statistical and econometric methods provide many ways to tackle this challenge. In addition to traditional structured data, we now have access to unstructured data and a vast amount of online data. Although there is no magical model or technique, multiple options are available and it is crucial to explore them through unstable and challenging contexts for inflation forecasting. Thus, forecasting inflation represents a longstanding inquiry that now presents new research opportunities and practical applications in light of these possibilities.

This dissertation includes three essays on inflation forecasting in Brazil. In the first essay, we examine the effectiveness of several forecasting methods in predicting inflation, focusing on aggregating disaggregated forecasts. For many forecast horizons, this approach performs on par with survey-based expectations and models that generate forecasts directly from the aggregate, with machine learning methods showing superior performance. In our second essay, we investigate the potential benefits of combining individual inflation forecasts by proposing a time-varying bias correction for the average forecast. We achieve good forecast performance for models based on short rolling windows for small and intermediate forecast horizons. Finally, in the third essay, we employ supervised learning to generate forward-looking indexes based on tweets and news articles for accumulated inflation. Our results indicate that indexes derived from news sources offer significant predictive gains, even in an environment with many predictors.

1
Forecasting inflation using disaggregates and machine learning

Abstract. This essay examines the effectiveness of several forecasting methods for predicting inflation, focusing on aggregating disaggregated forecasts – also known in the literature as the bottom-up approach. Taking the Brazilian case as an application, we consider different disaggregation levels for inflation and employ a range of traditional time series techniques as well as linear and nonlinear machine learning (ML) models to deal with a larger number of predictors. For many forecast horizons, the aggregation of disaggregated forecasts performs just as well survey-based expectations and models that generate forecasts using the aggregate directly. Overall, ML methods outperform traditional time series models in predictive accuracy, with outstanding performance in forecasting disaggregates. Our results reinforce the benefits of using models in a data-rich environment for inflation forecasting, including aggregating disaggregated forecasts from ML techniques, mainly during volatile periods. Starting from the COVID-19 pandemic, the random forest model based on both aggregate and disaggregated inflation achieves remarkable predictive performance at intermediate and longer horizons.

Keywords: inflation forecasting; disaggregated inflation; bottom-up approach; data-rich environment; machine learning.

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

1.1 Introduction

Economists and econometricians aim to provide as accurate inflation forecasts as possible by utilizing the most efficient approaches available. An important question is whether considering disaggregated inflation in different markets or economic classifications can enhance the forecasting performance for aggregate inflation. At first, this approach could capture trend dynamics, seasonality, and short-term changes more effectively (Espasa *et al.*, 2002). In other words, using subcomponents would allow the econometric models to capture the heterogeneity underlying the aggregate variable better (Bermingham & D'Agostino, 2014). Given the unknown data-generating process, whether direct or indirect forecasting through aggregating disaggregated forecasts can improve or not forecast accuracy is strictly an empirical question (Lütkepohl, 1984; Hendry & Hubrich, 2011; Faust & Wright, 2013). Nevertheless, a critical challenge that emerges is the increase in estimation uncertainty. To mitigate this problem, this essay implements a disaggregated analysis using machine learning (ML) methods that can deal with the bias-variance trade-off. Studies such as Inoue & Kilian (2008), Garcia *et al.* (2017), and Medeiros *et al.* (2021) point out the benefits of these techniques for inflation forecasting. Our essay employs these techniques in the context of disaggregated analysis, something scarcely explored in the literature.

The broad literature on inflation forecasting documents that the predictive performance of survey-based forecasts is challenging to beat, especially in the short-term horizons – current and immediate next months (Thomas, 1999; Ang *et al.*, 2007; Croushore, 2010). Faust & 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 & Wright, 2013). A potential explanation for this phenomenon is that forecasters are likely to have a richer information set than the econometrician employing a standard set of macroeconomic variables as predictors for inflation (Del Negro & Eusepi, 2011). Thus, including revealed expectations among the predictors is a way to exploit an information set that is not available. Baştürk *et al.* (2014), Altug & Çakmaklı (2016), Garcia *et al.* (2017), Fulton & Hubrich (2021), and Bańbura *et al.* (2021) find evidence favorable to the incorporation of survey-based forecasts into forecasting econometric models.

This essay examines the effectiveness of various forecasting methods for predicting aggregate inflation, focusing on aggregating disaggregated forecasts – also known in the literature as the bottom-up approach. Using the Brazilian case as an example, we compare the predictive performance of the bottom-up approach with traditional approaches in the literature, including survey-based forecasts and direct forecasting based exclusively on the aggregate. We explore different levels of disaggregation to assess how forecasts based on disaggregate price levels fare relative to those that rely solely on aggregate. Granularity is a potential advantage of considering disaggregates. Besides the specific effects of the traditional macro-variables related to money, economic activity, government, and external sector, we include lagged and crossed effects between disaggregates. When we compute our forecasts, we also consider *available* survey-based expectations as a predictor to add information not captured by other variables. Finally, we employ a range of 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-based 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 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 ensemble of non-linear tree-based models;
7. Model combination via average of forecasts for each disaggregation.

The Brazilian case is interesting for several reasons. First, the Broad Consumer Price Index (*Índice de Preços ao Consumidor Amplo* – IPCA), which serves as the official Brazilian price index, is available monthly and boasts a rich structure to be explored. The index contains several disaggregation levels and all time-varying weights of goods and services in the representative consumption basket are readily available. Second, the Central Bank of Brazil conducts the Focus survey, an extensive daily survey of expectations for multiple forecast horizons for some variables, including inflation. This survey reflects experts' opinions, mainly financial market professionals, and may contain private information that is not available to the econometrician. Beyond its utility as a predictor for generating model-based forecasts, Focus' inflation expectations can be used as a benchmark to assess whether improving survey-based forecasts for a given horizon is possible. Third, due to Brazil's inflationary history, in addition to the official price index, the country has several price indexes that may be used as predictors for inflation. Hence, it is pertinent to examine whether this information is valuable for forecasting Brazilian inflation over future horizons.

Findings. Among the main results of this essay are:

(i) It is challenging to outperform the survey-based forecast (Focus) before the COVID-19 pandemic; however, it is achievable post-pandemic (including at short horizons);

(ii) Taking disaggregated inflation into account tends to generate forecasts as good as survey-based expectations and forecasts based on aggregate inflation directly;

(iii) Overall, ML methods tend to outperform traditional time series models in predictive accuracy, with outstanding performance in predicting disaggregates;

(iv) There exists high variability in the type of predictors selected by the adaLASSO and FarmPredict;

(v) The *available* survey-based inflation expectations and price variables are relevant predictors;

(vi) Starting from the pandemic, the RF using aggregate inflation or some disaggregation achieves remarkable predictive performance at intermediate and longer horizons.

Contributions for the literature. We can summarize the main contributions of this essay in two fields. First, this essay advances the literature on inflation forecasting via aggregation of disaggregated forecasts by considering many predictors for each disaggregate, as well as several statistical and econometric methods underexplored in this literature. Many papers employ traditional time series models and a limited number of predictors. In this context, some papers find evidence favoring the bottom-up approach for the Euro Area (Espasa *et al.*, 2002; Espasa & Albacete, 2007) and various countries (Bruneau *et al.*, 2007; Moser *et al.*, 2007; Capistrán *et al.*, 2010; Aron & Muellbauer, 2012; Carlo & Marçal, 2016; Fulton & Hubrich, 2021). On the other hand, some papers find that aggregating forecasts by components does not necessarily improve aggregate inflation forecasting (Benalal *et al.*, 2004; Hubrich, 2005; Hendry & Hubrich, 2011). In turn, Duarte & Rua (2007), Ibarra (2012), and Bermingham & D'Agostino (2014) highlight the benefits of aggregating a large number of disaggregates. By employing a large number of predictors, Florido (2021, Chapter 1) points out the benefits of the disaggregated analysis in inflation nowcasting, while Araujo & Gaglianone (2023) do not find good results by using a disaggregation in multi-period forecasting. We show that the bottom-up approach can generate multi-period forecasts as accurately as survey-based expectations and direct forecasts.

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. The employ of ML methods to directly forecast inflation started with factor and principal component models (Stock & Watson, 1999, 2002; Forni *et al.*, 2003; Bai & Ng, 2008; Ibarra, 2012), and neural network models (Moshiri & Cameron, 2000; Nakamura, 2005; Choudhary & 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 regressions (CSR), but keeping focus on forecast inflation directly from the aggregate (Inoue & Kilian, 2008; Medeiros *et al.*, 2016; Garcia *et al.*, 2017; Zeng, 2017; Baybuza, 2018; Medeiros *et al.*, 2021; Araujo & Gaglianone, 2023). Florido (2021, Chapter 1) considers ML techniques, disaggregated inflation, and a broad set of predictors in inflation nowcasting, finding good results. Araujo & Gaglianone (2023) consider the inflation disaggregated into administered prices, services, industrial goods, and food at home to generate multi-horizon inflation forecasts employing several ML models. However, their results are not favorable to the bottom-up approach. In contrast, our combination between disaggregated analysis, ML, and many predictors yields promising results and opens up new possibilities for further exploration.

We also point out five other minor contributions. First, we corroborate the findings of Baştürk *et al.* (2014), Altug & Çakmaklı (2016), Garcia *et al.* (2017), Fulton & Hubrich (2021), and Bańbura *et al.* (2021) regarding the benefits of incorporating a survey-based expectation as a predictor when econometrician computes their forecasts. The presence of this variable is relevant to improve predictive accuracy even for some disaggregation levels. Second, when estimating a factor-augmented autoregression model using a method that allows predictor selection, we find that the factor that summarizes most of the predictors' variability is not relevant for predicting inflation. Hence, using an estimation method such as the adaLASSO or the approaches of Bai & Ng (2008, 2009) instead of least squares may be beneficial. Third, our essay is one of the first to employ the FarmPredict, a model proposed by Fan *et al.* (2021) that combines factor and sparse linear regressions. We adapt it to allow the simultaneous estimation via adaLASSO of a final model containing lags, common factors, and idiosyncratic components. Fourth, as in Duarte & Rua (2007), Ibarra (2012), and Bermingham & D'Agostino (2014), we also indicate the potential benefits of considering a high level of disaggregation. However, there are caveats about how to improve the bottom-up approach by considering different models predicting different disaggregates. Finally, our analysis underscores the importance of examining sub-periods and emphasizing the benefits of model-based forecasts in volatile periods, as also pointed out by Altug & Çakmaklı (2016) and Medeiros *et al.* (2021).

Outline. This chapter has five more sections in addition to this Introduction. Section 1.2 presents the forecasting methodology, and Section 1.3 describes the models, estimation, metrics, and test to compute and assess the results. Section 1.4 displays the data and setup. Section 1.5 presents the results and provides an economic discussion about them. Finally, Section 1.6 concludes. Appendixes from 1.A to 1.D offer supplementary information and complementary results.

1.2 Forecasting methodology

1.2.1 "Traditional" inflation forecasting

Let π_t be the (aggregate) inflation at period t. We compute the inflation from the percentage change in a price index based on a typical consumption basket. For forecasting purposes, assume there are J predictors for inflation. Let z_t be a J-dimensional vector of these explanatory variables *observed* at t, that is, the information set available to the econometrician to perform the forecasting. Notice that z_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. As the estimation is based on moving windows, the mapping is dependent on time, which we indicate 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. In a rich-data environment, we can consider dimensionality reduction or shrinkage with or without selecting predictors. Whatever the choices, we must be careful to avoid overfitting. Finally, an h-period-ahead forecast is given by

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

where hats indicate estimation.

1.2.2
Aggregation of disaggregated forecasts

Besides the general price index and their percentage change, the aggregate inflation π_t, now consider the availability of N^d disaggregated price indexes (subcomponents of the original price index) indexed by $i = 1, \ldots, N^d$. The letter d indicates the disaggregation level. Let π^d_{it} be the percentage change of the disaggregate i in disaggregation level d at period t. Let ω^d_{it} be the weight of the disaggregate i at disaggregation level d 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 indexes is given by

$$\pi_t = \sum_{i=1}^{N^d} \omega^d_{it}\, \pi^d_{it}, \tag{1.1}$$

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

Let $\boldsymbol{z}^d_t$ be a J^d-dimensional vector of *all* explanatory variables for price changes *observed* at t by the econometrician at disaggregation level d. Note that it is expected that $J^d > J$ 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 disaggregates as predictors for a specific disaggregate. The question arises as to whether capturing and exploring the crossed dependence between disaggregated prices could enhance inflation forecasting. Let $\mathcal{M}^d_{i,t,h}$ be a time-varying mapping between predictors and h-period-ahead price variation of each disaggregate $i = 1, \ldots, N^d$ at disaggregation level d. Following (1.1), an h-period-ahead forecast for aggregate inflation is given by

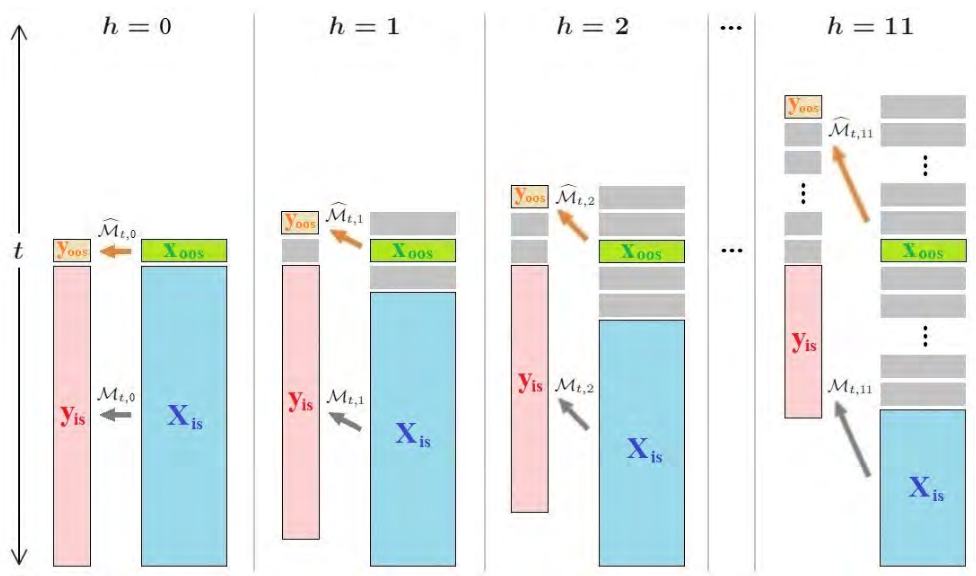
$$\widehat{\pi}_{t+h\,|\,t} = \sum_{i=1}^{N^d} \widetilde{\omega}_{it}^d\, \widehat{\pi}_{i,t+h} = \sum_{i=1}^{N^d} \widetilde{\omega}_{it}^d\, \widehat{\mathcal{M}}_{i,t,h}^d(z_t^d),$$

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

1.2.3 Direct forecasting approach and expanding window scheme

We employ a direct forecast approach considering expanding windows for (monthly) horizons $h \in \{0, 1, \ldots, 11\}$. We take the time-adjusted predictors to fit the mapping between them and inflation in this approach. For example, suppose we want to generate a forecast for the current period ($h = 0$), which is called *nowcasting*. In that case, we consider the most recently *available* information to estimate the desired mapping. Conversely, when computing a one-month-ahead forecast ($h = 1$), we use the information available up until the preceding period in which the forecast is estimated. We continue this way until we calculate the forecast for $h = 11$, utilizing information *available* ten periods prior. Figure 1.1 illustrates the exercise. Following the computation of forecasts based on a given period, we advance the time window by one period and repeat the estimation procedure for each forecast horizon, subsequently calculating new forecasts.

Figure 1.1: Direct forecasting approach with expanding window scheme



Notes: y indicates the target variable. $\boldsymbol{X}$ represents predictor variables. Subscripts "is" and "oos" denote in-sample and out-of-sample, respectively.

1.3 Models and forecast evaluation

1.3.1 Models

To enhance clarity in presenting the following forecast methods, we omit the superscript d that indicates the level of disaggregation, whenever applicable.

1.3.1.1 Benchmarks

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

Historical mean. Also for the aggregated case, a prediction for h periods ahead is given by historical average inflation computed at t, that is,

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

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

Autoregressive model – AR(p). For both aggregated and disaggregated cases, in the direct forecast 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, a h-period-ahead inflation forecast is given by

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

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

$$\pi_t = \mu + \sum_{l=1}^{p} \phi_l \, \pi_{t-h-l+1} + \eta \, \pi^e_{t\,|\,t-h} + \sum_{m=1}^{11} \delta_m \, d_{mt} + \varepsilon_t \tag{1.2}$$

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 a h-period-ahead forecast is given by

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

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

$$\pi_t = \mu + \sum_{l=1}^{p} \phi_l \, \pi_{t-h-l+1} + \eta \, \pi^e_{t\,|\,t-h} + \psi_1 \, g_{t-h} + \psi_2 \, \Delta s_{t-h} + \varepsilon_t$$

where g_{t-h} is some economic activity measure *observed* at $t-h$, and Δs_{t-h} is an exchange rate measure *observed* at $t-h$. We compute the forecast by

$$\widehat{\pi}^{\text{HNKPC}}_{t+h\,|\,t} = \widehat{\mu} + \sum_{l=1}^{p} \widehat{\phi}_l \, \pi_{t-l+1} + \widehat{\eta} \, \pi^e_{t+h\,|\,t} + \widehat{\psi}_1 \, g_t + \widehat{\psi}_2 \, \Delta s_t$$

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

1.3.1.2
Shrinkage-based models

Ridge (with incomplete information). For disaggregated cases, we consider the augmented AR model (1.2) with the addition of other lagged disaggregates:

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

where N^d is the number of subcomponents in the disaggregation level indicated by d. We consider four disaggregation levels in this essay: aggregate inflation, economic categories defined by the BCB, and groups and subgroups from IPCA (IBGE).

We estimate the coefficients employing the 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 predictors, and $\boldsymbol{\beta}_i$ is a vector of coefficients. Chosen λ via information criteria (e.g., BIC), a prediction for h periods ahead is given by

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

adaLASSO (with full information). For all cases, consider the model with full information given by

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

where $\boldsymbol{x}_{t-h} \in \mathbb{R}^{J \cdot p}$ is an expanded vector of potential predictors for π_{it}. We estimate this model employing the adaptive LASSO (adaLASSO). Introduced by Zou (2006), this method selects predictors and their optimization problem is given by

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

where ξ is a regularization parameter, $\boldsymbol{z}_{t-h} \in \mathbb{R}^V$, $V = N^d \cdot p + 12 + J \cdot p$, is a vector of *all* predictors, and $\boldsymbol{\zeta} = (\zeta_1, \ldots, \zeta_V)$ is a vector of weights obtained previously employing LASSO – a estimator that assumes $\zeta_{ij} = 1$, for all j. More precisely, we compute the weights via

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

where we add $T^{-1/2}$ to allow a variable that is not selected in the first stage to have a chance of being selected in the second stage.

As before, a h-period-ahead forecast is $\widehat{\pi}^{\text{adaLASSO}}_{i,t+h\,|\,t} = \widehat{\mu}_i + \widehat{\boldsymbol{\beta}}_{\text{adaLASSO},i} \, \boldsymbol{z}_t$.

1.3.1.3
Factor models

(Augmented) factor model. Consider that all regressors are normalized for both aggregate and disaggregate cases. Thus, for $i = 1, \ldots, N^d$, a factor-augmented autoregression model is described by

$$\boldsymbol{x}_t = \sum_{k=1}^{K} \boldsymbol{\lambda}_k f_{kt} + \boldsymbol{u}_t \tag{1.3}$$

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

from which we compute common factors $\widehat{\boldsymbol{f}}_t = \left(\widehat{f}_{1t}, \ldots, \widehat{f}_{Kt}\right)$ and factor loadings $\boldsymbol{\lambda}_k = \left(\lambda_{1k}, \ldots, \lambda_{Jk}\right)$ by combining principal component analysis (PCA) and OLS. Finally, we compute $\left(\widehat{\mu}_i,\, \widehat{\boldsymbol{\phi}},\, \widehat{\eta}_i,\, \widehat{\boldsymbol{\beta}}\right)$ via adaLASSO.

For identification purposes, we assume that

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

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

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

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

Target factor model. Proposed by Bai & Ng (2008), in this "hard thresholding" version, this approach controls for the participation of normalized explanatory variables in the factor construction. In a previous stage, for each predictor indexed by $j = 1, \ldots, J$, and disaggregate indexed by $i = 1, \ldots, N^d$, we estimate

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

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 factor-augmented autoregressive model: we perform

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

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.

1.3.1.4
FarmPredict

Some idiosyncratic errors of the factor model, that is, some $\boldsymbol{u}_t$ entries in Equation (1.3), can impact the price variation, which the common factor structure does not capture. 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:

$$\begin{aligned}\pi_{it} = \mu_i + \sum_{i'=1}^{N^d} \sum_{l=1}^{p} \phi_{i'l}\, \pi_{i',t-h-l+1} + \gamma_i\, \pi^e_{t\,|\,t-h} + \sum_{m=1}^{11} \delta_{im}\, d_{mt} \\ + \sum_{k=1}^{K} \sum_{l=1}^{p} \beta_{ikl}\, \widehat{f}_{k,t-h-l+1} + \sum_{j=1}^{J} \sum_{l=1}^{p} \theta_{ijl}\, \widehat{u}_{j,t-h-l+1} + \varepsilon_{it}. \end{aligned} \tag{1.4}$$

This model is a specific form of a general model called FarmPredict proposed by Fan *et al.* (2021). Here, we estimate the "final equation" (1.4) with all regressors *simultaneously* employing the adaLASSO. Next, we compute the forecast.

1.3.1.5
Complete subset regression (CSR)

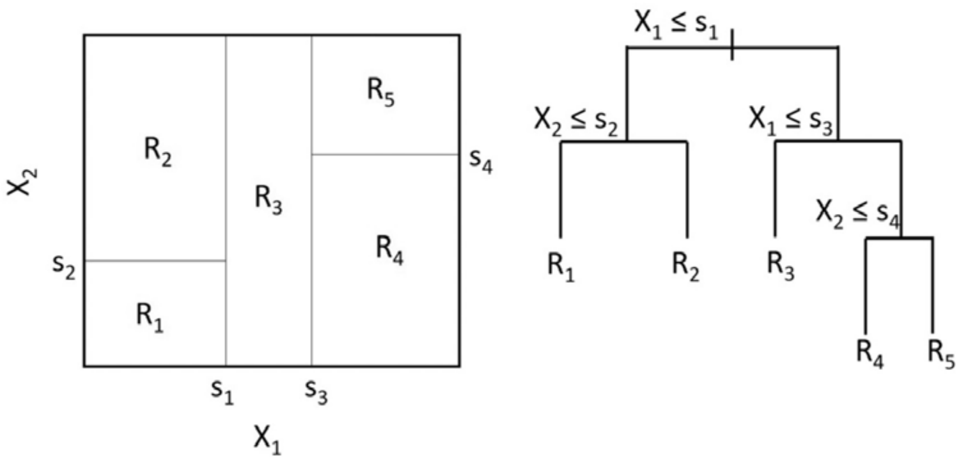
Introduced by Elliott *et al.* (2013, 2015), this ensemble method combines estimates from all (or several) possible linear regression models, keeping the number of predictors fixed. Let p be the total available predictors and $k \leqslant p$ be the number of "selected" predictors (complete subsets). 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 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-)select $\widetilde{p} \leqslant p$ predictors based on a ranking of t-statistics in absolute value, as in

Garcia *et al.* (2017) and Medeiros *et al.* (2021). 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 $\widetilde{p}$ pre-selected predictors.

1.3.1.6
Random forest (RF)

Breiman (2001) introduces the random forest (RF), 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 1.2.

Figure 1.2: A regression tree with two explanatory variables (X_1, X_2)



Notes: $s_i,\ i = 1, \ldots, 4$, indicate splits, while $R_k,\ k = 1, \ldots, 5$, denote regions. Example extracted from Medeiros *et al.* (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 variables to incorporate in the estimation of 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,

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

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

1.3.2
Model combinations via average of forecasts

Methods may perform differently for distinct disaggregates or even over time for the same disaggregate. To mitigate instabilities associated with some method for some disaggregate or at some point in time, for each disaggregation, we will compute a combined forecast given by the average of forecasts generated by all methods applied to this disaggregation and Focus expectations *available* when the econometrician computes their forecasts, that is,

$$\widehat{\pi}^{\mathrm{Comb},\,d}_{t+h\,|\,t} = \frac{1}{M^d+1}\left(\sum_{m=1}^{M^d}\sum_{i=1}^{N^d} \widetilde{\omega}^d_{it}\, \widehat{\pi}^{m,\,d}_{i,t+h\,|\,t} + \pi^e_{t+h\,|\,t}\right),$$

where d indicates one of four possible disaggregations levels addressed in this essay (aggregate inflation, BCB categories, IBGE groups, and IBGE subgroups), m indicates a method, M^d is the number of methods employed to forecast the inflation for the disaggregation d, N^d is the number of disaggregates in the disaggregation d, and $\widetilde{\omega}^d_{it}$ is the weight of disaggregate i of the disaggregation level d in the aggregate index *observed* at t by the econometrician. The idea is to investigate whether this simple combination leads to improvements in forecast performance.

1.3.3
Evaluation: metrics and test

Metrics. We use out-of-sample root mean squared error (RMSE) as the main metric to evaluate the forecast performance. For each horizon h, this metric is described by

$$\mathrm{RMSE}^{m,\,d}_h = \left[\frac{1}{T}\sum_{t=1}^{T}\left(\pi_{t+h} - \widehat{\pi}^{m,\,d}_{t+h\,|\,t}\right)^2\right]^{1/2}$$

where $\widehat{\pi}^{m,d}_{t+h\,|\,t}$ indicates a forecast generated by the model m considering the disaggregation level d. The smaller the $\mathrm{RMSE}_{h,\,m}$, the better the model's predictive performance. For the Diebold-Mariano test, we consider the mean squared error (MSE) defined by

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

Test. To assess the results, we consider the widely employed test developed by Diebold & Mariano (1995). Let $\widehat{v}_{t+h\,|\,t}^{m} = \pi_{t+h} - \widehat{\pi}_{t+h\,|\,t}^{m}$ be a forecast error of the model m. Here, we omit the disaggregation level d. 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

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

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

1.4 Data and setup

Data. We analyze the period from January 2004 to June 2022, totalizing 18,5 years of monthly data. For aggregate inflation, we employ the IPCA, the official Brazilian price index computed by the 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. Subgroups are subdivisions and, in some cases, the group itself. For definition of groups and subgroups, and their respective average weights in the IPCA, see Table 1.A.1 in Appendix 1.A. In addition, we use a 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 lags of other disaggregates in 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 private 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 eighty-nine other predictor variables (and their lags) divided into ten categories: prices and money (17), commodities prices (4), economic activity (19), employment (5), electricity (4), confidence (3), finance (12), credit (4), government (12), and exchange and international transactions (9). In Appendix 1.B, Table 1.B.1 presents a description of these variables, the delay for each to become available and transformations implemented to guarantee the stationarity.

Setup. The reference day to compute our forecasts is the last business day of each month. For the results shown in the following section, we consider three lags for all predictive variables, including variables mentioned above, factors in factor models, idiosyncratic components in FarmPredict, and lags of aggregate and 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 1.2.3, the main results are generated based on expanding windows. In this setup, we generate 114 forecasts for each horizon. 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 & 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 allow the trees to grow until five observations by leaf. We set the proportion of selected variables 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 *et al.* (2017) and Medeiros *et al.* (2021). Finally, to estimate the empirical Phillips curve, we use the Central Bank of Brazil's economic activity index (IBC-Br) and BIS' real effective exchange rate (REER) as a proxy for economic activity and exchange rate, respectively.

1.5 Results

1.5.1 Entire period: forecasts from January 2014 to June 2022

Table 1.1 exhibits the results of forecast performance in terms of root mean squared errors (RMSE) for different models and horizons ranging from nowcasting ($h = 0$) to eleven months ahead ($h = 11$), as well as for 12-month accumulated inflation. We normalize every RMSE to relative terms by computing their ratio to the RMSE of the Focus consensus – the median expectation of the available Focus survey. Thus, a value lower than one indicates that a model numerically outperforms the Focus consensus, while a value greater than one suggests underperformance compared to the same benchmark. At this first moment, the results consider the entire period for which we compute predictions, from January 2014 to June 2022. Each panel of Table 1.1 considers a group of competitors. In panel A, we have the available and *ex-post* Focus, the latter released by the Central Bank in the following week reflecting the experts' opinions on the same day we compute our forecasts. We note virtually no difference between the available and ex-post Focus for longer horizons. However, in the short term ($h \leqslant 3$), there is evidence that ex-post Focus statistically outperforms available Focus. Despite being only a few days apart, the informational gain is considerable for shorter horizons, which does not occur for more distant periods since it is unlikely that very relevant information about them will emerge within a few days.

Panels B to E of Table 1.1 show the results for each model considering different levels of disaggregation: aggregate inflation, disaggregations from the Central Bank of Brazil (BCB), and disaggregations into IBGE groups and subgroups, respectively. Perhaps not surprisingly, it is hard to outperform the Focus survey in nowcasting. Exceptions are due to models that forecast aggregate inflation directly. However, such models perform better only than the available Focus. Considering the whole period, no alternative beats the ex-post Focus, which delivers almost 7% RMSE reduction compared to the available survey. However, it is appreciable that some models are competitive with the ex-post Focus. Specialists who report their expectations to the BCB often have access to information unavailable to econometricians, such as private data. Since models do not have this additional information and other advantages, such as including personal judgments, as pointed out by Faust & Wright (2013), their ability to outperform available expectations and get closer to *ex-post* survey-based expectations is a great result. We note that models forecasting disaggregates do not deliver good per-

Table 1.1: Out-of-sample RMSE with respect to the available Focus: Jan/2014 to Jun/2022

Estimator/Model	$h=0$	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$	$h=6$	$h=7$	$h=8$	$h=9$	$h=10$	$h=11$	$\Sigma\,12m$
A. Survey													
Focus (available)	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Focus (*ex-post*)	**0.933*****	0.972***	0.993***	1.001	1.000	1.000	1.000	0.999	0.999	1.000	1.001	1.001	0.999
B. Aggregate inflation													
RW	2.782	1.471	1.212	1.251	1.279	1.271	1.234	1.180	1.087	1.063	1.102	1.138	1.485
Historical Mean	3.117	1.342	1.084	1.053	1.034	1.018	0.998	0.990	0.987	0.983	0.982	0.984	1.041
AR	2.534	1.260	1.053	1.077	1.090	1.064	1.008	0.961	0.939*	0.941*	*0.964*	0.988	0.987
HNKPC	*0.946***	*0.957***	0.972*	*0.981*	1.036	1.030	0.980	0.968	0.963*	0.974	0.998	1.021	0.984
Augmented AR	0.970	0.981	0.989	1.012	1.023	0.999	0.963	*0.943**	*0.936***	0.961	0.993	1.016	0.954*
adaLASSO	*0.944***	*0.948****	0.978*	0.993	*0.977***	0.989	0.976	0.976	0.989	1.004	1.043	0.992	0.961*
Factor	*0.943***	*0.931****	*0.966***	**0.964****	*0.974*	0.980	0.989	0.986	1.016	1.016	1.052	1.045	0.989
FarmPredict	*0.944***	*0.942****	*0.961***	0.987	0.999	1.001	1.004	1.020	1.020	1.032	1.027	1.051	1.001
Target Factor	1.280	1.129	1.045	1.155	1.122	1.069	**0.937***	0.989	0.990	0.980	1.019	1.102	0.970
CSR	0.961	**0.927****	*0.962**	0.983	1.022	0.996	0.969	*0.940***	0.940**	0.975	1.042	1.071	0.935**
Random Forest	1.665	1.110	1.025	1.012	0.996	0.987	0.959	0.949	**0.920****	*0.918***	*0.908***	**0.909****	*0.918**
C. Disaggregation: tradable, nontradable and monitored prices (BCB)													
AR	2.579	1.241	1.039	1.057	1.089	1.085	1.015	1.037	0.990	0.989	1.026	1.036	1.050
Augmented AR	0.978	0.978	1.012	1.061	1.094	1.072	1.010	1.010	0.982	1.003	1.052	1.073	1.048
Ridge	1.018	1.015	1.007	1.019	1.045	1.083	0.999	0.978	0.969	0.960*	1.013	0.981	1.032
adaLASSO	1.099	0.998	*0.968*	*0.969*	*0.977*	*0.948***	0.962*	0.957*	0.979	0.964	0.972	*0.957**	*0.903***
Factor	1.023	0.988	0.991	0.990	*0.977*	**0.934****	*0.940***	0.977	1.000	1.007	1.009	1.009	0.943*
FarmPredict	1.129	1.024	1.003	*0.981*	**0.969**	*0.944***	0.956*	0.984	1.003	1.009	1.008	0.991	0.963
Target Factor	1.331	1.052	0.976	1.090	1.131	1.079	*0.943*	0.962	0.964	0.991	1.016	1.077	0.936
CSR	2.176	1.178	1.016	1.015	1.018	0.983	0.957	0.977	0.959	0.967	1.022	1.052	0.952
Random Forest	1.934	1.177	1.020	1.051	1.014	1.001	0.976	0.963	*0.930**	**0.912****	**0.902****	*0.915***	*0.923*
D. Disaggregation: groups (IBGE)													
AR	2.731	1.296	1.048	1.063	1.054	1.032	1.067	1.087	1.118	1.145	1.105	1.095	1.067
Augmented AR	1.004	1.046	1.042	1.056	1.088	1.046	1.067	1.065	1.069	1.105	1.134	1.180	1.028
Ridge	1.656	1.288	1.044	1.032	1.016	1.010	0.986	0.989	0.972	0.983	0.989	0.999	1.005
adaLASSO	1.223	1.090	1.026	1.013	1.003	0.994	1.019	0.989	0.972	1.004	0.988	1.000	0.954
Factor	1.219	1.075	1.034	1.031	1.035	1.021	1.024	0.967	0.974	1.021	1.000	0.990	1.005
FarmPredict	1.334	1.089	1.025	1.049	1.040	1.016	1.023	0.991	0.971	1.029	0.997	0.998	1.020
Target Factor	1.268	1.082	1.068	1.142	1.047	0.997	0.976	0.961	1.003	0.970	1.026	1.092	0.977
CSR	2.146	1.144	1.048	1.047	1.002	0.980	*0.940**	**0.926***	0.975	0.981	0.981	0.985	**0.900***
Random Forest	2.096	1.217	1.026	1.047	1.009	0.998	0.971	0.953	0.944	*0.931**	*0.920***	*0.921**	0.955
E. Disaggregation: subgroups (IBGE)													
AR	3.111	1.502	1.186	1.212	1.244	1.184	1.212	1.253	1.264	1.316	1.275	1.362	1.318
Augmented AR	1.197	1.260	1.290	1.288	1.290	1.191	1.164	1.212	1.239	1.310	1.314	1.440	1.208
Ridge	3.100	1.308	1.058	1.045	1.023	1.026	1.008	1.002	1.000	0.993	0.995	1.016	1.068
adaLASSO	1.380	1.122	1.096	1.056	1.030	1.010	0.990	0.990	0.983	1.011	1.049	1.039	1.005
Factor	1.364	1.147	1.068	1.050	1.039	1.029	1.009	0.985	0.996	1.027	1.040	1.005	1.034
FarmPredict	1.384	1.149	1.047	1.040	1.028	1.008	1.010	0.980	0.982	1.033	1.060	1.042	1.036
Target Factor	1.230	1.059	1.145	1.126	1.117	1.059	0.976	*0.937*	1.028	0.979	1.051	1.116	1.021
CSR	2.222	1.210	1.076	1.067	1.027	1.028	0.977	0.955	*0.939*	0.987	1.034	1.030	0.974
Random Forest	2.106	1.232	1.032	1.063	1.021	1.022	0.990	0.962	0.949	*0.925**	*0.917***	*0.921**	0.974
F. Model combinations for disaggregates													
Aggreg. Comb.	1.201	0.989	**0.952****	*0.972**	0.982	*0.969*	*0.942***	*0.934***	*0.935***	*0.939***	0.971	0.986	*0.935***
BCB Comb.	1.243	1.014	0.969	0.993	0.997	0.977	0.945*	0.954*	0.950*	0.952*	0.975	0.986	0.950
Groups Comb.	1.401	1.068	0.976	1.001	0.988	*0.969*	0.970	0.954	0.959	0.982	0.978	0.995	0.961
Subgroups Comb.	1.583	1.129	1.035	1.045	1.031	1.007	0.981	0.966	0.985	1.004	1.026	1.046	1.029

Notes: ***, **, and * indicate that for a specific forecast horizon, a model m performed statistically better than the median of the available Focus at 1, 5, and 10% significance levels in a one-tailed Diebold-Mariano test with $\mathbb{H}_0 : \text{MSE}(\widehat{\pi}^{m}_{t+h\,|\,t}) = \text{MSE}(\pi^{\text{Focus}}_{t+h\,|\,t})$ *versus* $\mathbb{H}_1 : \text{MSE}(\widehat{\pi}^{m}_{t+h\,|\,t}) < \text{MSE}(\pi^{\text{Focus}}_{t+h\,|\,t})$. The value highlighted in bold blue indicates the best model for each horizon in terms of RMSE ratio with respect to *ex-post* Focus, and the values in blue italics indicate the second and third best models.

formance for nowcasting. Lastly, the combinations of models in each level of disaggregation, whose results are shown in Panel F, also do not generate forecasts better.

For other horizons, the contribution of the models becomes more effective. Despite the challenge of surpassing survey-based expectations for short-term horizons such as $h = 1$ and $h = 2$, several models for aggregate inflation (Panel B) achieve good results for these horizons. Like occurred for $h = 0$, the hybrid Phillips curve, adaLASSO, factor model, FarmPredict, and, additionally, the complete subset regression (CSR), deliver the best forecast performances for one and two months ahead. All are statistically superior to the available Focus according to the Diebold-Mariano (DM) test considering at least the more slack significance level (i.e., 10%). Furthermore, these models also numerically outperform the ex-post Focus. On the other hand, the adaLASSO using BCB disaggregation (Panel C) is the only model employing any disaggregation among the best models. However, this model is not statistically superior to the available Focus by the DM test. Finally, regarding these shorter horizons, it is worth highlighting the performance of the average forecast of the models for aggregate inflation, which achieves the highest accuracy for $h = 2$ by presenting a statistically significant reduction of almost 5% in RMSE.

Models considering some disaggregation for the inflation yield better results starting from the 4-month horizon. The adaLASSO, factor model, and FarmPredict, all using the BCB disaggregation, perform well for forecast horizons ranging from fourth to seventh months. These models are statistically superior to available or ex-post Focus at various periods. Regarding the use of disaggregated inflation data in groups from the IBGE, it is worth mentioning the good performance of the CSR, which achieves the best result among all the options for $h = 7$. Another highlight is the combination of forecasts generated by models that directly forecast the aggregate inflation, which achieves a statistically significant reduction of 6% in RMSE from 6 to 9 months ahead. For $h \geqslant 8$, there is broad dominance of the random forest (RF), whether using aggregate inflation or some disaggregation. Frequently, for these more distant horizons, the RF registers a statistically significant reduction in RMSE ranging from 7% to 10% in comparison to the survey-based expectations. This result highlights that the RF, employing IBGE group disaggregation, achieves the best performance among all competitors for inflation accumulated over 12 months (see last column of Table 1.1), closely followed by the adaLASSO using BCB disaggregation, which achieves a similar RMSE reduction.

Remarks. Considering the forecast performance of various models from January 2014 to June 2022, we observe that different approaches are more effective at different times. In the short term, machine learning models that deal directly with aggregate inflation perform better, whereas for intermediate horizons of 4 to 7 months, considering the BCB disaggregation lead to significant benefits. For the period between 6 and 9 months ahead, the average of forecasts obtained from models that used only aggregate inflation also perform well. Finally, for longer horizons of 8 months or more, regardless of the approach, the RF delivers the best forecast performances. While Garcia *et al.* (2017) points to the superiority of the CSR in several horizons, we only verify the prevalence of the CSR for $h = 1$ and other isolated good performances. The RF's performance in predicting inflation had already been pointed out by Medeiros *et al.* (2021) when analyzing the case of the United States and highlighting the benefits of this method for dealing with non-linearities. The advent of the COVID-19 pandemic in Brazil changes the price dynamics considerably from 2020 onwards. Because of that, in what follows, we divide our analysis into two sub-periods: (i) before the pandemic, from January 2014 to February 2020, and (ii) after the pandemic, from March 2020 onwards.

1.5.2
Forecasts before and after of COVID-19 pandemic

Table 1.2 shows that the sub-period between January 2014 and February 2020 is quite challenging for model-based forecasts. Almost no RMSE ratios are below 1, with exceptions mainly in the short term. Nevertheless, no model is able to beat the ex-post Focus or be statistically superior to the available Focus for nowcasting. For $h = 1$, only the factor model and FarmPredict are statistically superior to the available Focus at the 10% significance level, with these two tying with the predictive performance of the ex-post Focus. For the 3-month forecast, the hybrid Phillips curve for aggregate inflation is subtly superior to the Focus in numerical terms but without statistical significance. For the other horizons, no model performs better than the survey-based expectations. There are some potential explanations for this poor performance of the models. First, since we are analyzing the first sub-period, a small sample may have affected the estimates, contributing to the models' poor forecast performance. Second, the instabilities in the Brazilian economy in 2014 and 2015 that resulted in a sharp increase in inflation in 2015, as well as the rapid disinflation that occurred from the second half of 2016, are challenging events to anticipate, especially without extensive historical data. Conversely, from 2017 until the begging of the COVID-19 pandemic, Brazilian inflation remained reasonably controlled and close to the inflation target, leaving limited opportunities for models to enhance survey-based expectations. These dynamics of the Brazilian inflation can be observed from Figure 1.3 later on.

Table 1.2: Out-of-sample RMSE with respect to the available Focus: Jan/2014 to Feb/2020

Estimator/Model	$h=0$	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$	$h=6$	$h=7$	$h=8$	$h=9$	$h=10$	$h=11$	$\Sigma\,12m$
A. Survey													
Focus (available)	1.000	*1.000*	*1.000*	*1.000*	*1.000*	**1.000**	*1.000*	*1.000*	*1.000*	*1.000*	*1.000*	**1.000**	*1.000*
Focus (*ex-post*)	**0.932*****	**0.969*****	**0.995****	*1.001*	**0.999**	*1.001*	**0.999**	**1.000**	**1.000**	**1.000**	**0.999**	*1.002*	**0.996*****
B. Aggregate inflation													
RW	3.088	1.752	1.428	1.474	1.574	1.547	1.516	1.427	1.260	1.240	1.307	1.382	2.114
Historical Mean	3.309	1.492	1.246	1.208	1.184	1.173	1.155	1.146	1.133	1.125	1.127	1.134	1.422
AR	2.781	1.466	1.225	1.240	1.266	1.244	1.172	1.122	1.093	1.098	1.121	1.155	1.339
HNKPC	*0.966*	1.008	1.022	**0.994**	1.083	1.087	*1.061*	*1.043*	*1.046*	*1.050*	*1.072*	1.079	*1.089*
Augmented AR	1.015	1.067	1.081	1.090	1.114	1.101	1.082	1.076	1.068	1.086	1.117	1.126	1.161
adaLASSO	*0.990*	*0.979*	*1.009*	1.052	*1.031*	1.074	1.086	1.084	1.134	1.136	1.176	1.097	1.136
Factor	*0.990*	*0.970**	*1.001*	*1.016*	*1.056*	*1.055*	*1.056*	*1.043*	1.106	1.121	1.113	1.115	*1.109*
FarmPredict	*0.990*	*0.970**	*1.003*	*1.024*	*1.065*	*1.040*	1.062	1.094	1.113	1.129	1.105	1.163	*1.102*
Target Factor	1.499	1.225	1.122	1.338	1.334	1.215	1.110	1.223	1.212	1.231	1.171	1.118	1.333
CSR	*0.968*	1.011	1.035	1.081	1.139	1.104	1.086	1.063	*1.056*	*1.047*	*1.043*	*1.054*	1.113
Random Forest	1.918	1.276	1.156	1.151	1.137	1.144	1.130	1.130	1.101	1.113	1.110	1.127	1.319
C. Disaggregation: tradable, nontradable and monitored prices (BCB)													
AR	2.943	1.535	1.268	1.275	1.304	1.318	1.241	1.261	1.150	1.086	1.079	1.077	1.519
Augmented AR	1.049	1.120	1.131	1.150	1.182	1.189	1.164	1.186	1.137	1.163	1.189	1.153	1.335
Ridge	1.100	1.190	1.134	1.129	1.167	1.242	1.125	1.111	1.088	1.069	1.144	*1.060*	1.363
adaLASSO	1.225	1.098	1.060	1.092	1.112	*1.049*	1.064	1.069	1.110	1.095	1.106	*1.062*	1.168
Factor	1.094	1.098	1.120	1.142	1.142	1.082	1.075	1.074	1.089	1.084	1.098	1.101	1.178
FarmPredict	1.226	1.118	1.095	1.116	1.110	1.072	1.072	1.050	1.092	1.107	1.101	1.085	1.180
Target Factor	1.569	1.183	1.078	1.221	1.372	1.260	1.127	1.215	1.123	1.239	1.175	1.118	1.309
CSR	2.364	1.401	1.221	1.188	1.196	1.162	1.168	1.184	1.133	1.123	1.153	1.105	1.362
Random Forest	2.223	1.322	1.161	1.167	1.163	1.182	1.164	1.161	1.131	1.107	1.112	1.134	1.380
D. Disaggregation: groups (IBGE)													
AR	3.236	1.659	1.301	1.319	1.272	1.199	1.176	1.221	1.236	1.268	1.237	1.314	1.465
Augmented AR	1.084	1.305	1.270	1.263	1.276	1.238	1.199	1.195	1.177	1.261	1.269	1.346	1.380
Ridge	2.043	1.443	1.205	1.171	1.149	1.136	1.111	1.107	1.084	1.098	1.106	1.116	1.329
adaLASSO	1.361	1.273	1.159	1.146	1.128	1.110	1.154	1.155	1.163	1.186	1.148	1.146	1.302
Factor	1.411	1.270	1.164	1.163	1.157	1.157	1.127	1.105	1.105	1.114	1.109	1.119	1.289
FarmPredict	1.510	1.226	1.148	1.174	1.181	1.138	1.108	1.113	1.089	1.139	1.128	1.149	1.322
Target Factor	1.553	1.323	1.227	1.315	1.215	1.172	1.166	1.181	1.244	1.167	1.210	1.185	1.358
CSR	2.410	1.434	1.318	1.285	1.231	1.159	1.083	1.074	1.171	1.148	1.138	1.111	1.248
Random Forest	2.282	1.342	1.174	1.216	1.185	1.170	1.154	1.170	1.167	1.146	1.120	1.127	1.409
E. Disaggregation: subgroups (IBGE)													
AR	3.467	1.745	1.461	1.503	1.503	1.464	1.448	1.538	1.510	1.501	1.453	1.665	1.788
Augmented AR	1.345	1.529	1.547	1.547	1.545	1.496	1.418	1.462	1.498	1.494	1.476	1.745	1.634
Ridge	3.208	1.448	1.214	1.185	1.151	1.142	1.124	1.107	1.108	1.111	1.110	1.118	1.390
adaLASSO	1.533	1.261	1.264	1.174	1.141	1.117	1.105	1.166	1.149	1.150	1.200	1.166	1.308
Factor	1.521	1.266	1.180	1.149	1.156	1.133	1.108	1.106	1.104	1.110	1.137	1.114	1.280
FarmPredict	1.541	1.273	1.136	1.148	1.146	1.121	1.117	1.098	1.081	1.143	1.176	1.173	1.302
Target Factor	1.450	1.296	1.337	1.278	1.280	1.184	1.141	1.153	1.255	1.171	1.218	1.244	1.407
CSR	2.457	1.394	1.319	1.244	1.201	1.155	1.096	1.107	1.104	1.138	1.177	1.119	1.253
Random Forest	2.325	1.360	1.180	1.224	1.194	1.199	1.182	1.168	1.150	1.117	1.106	1.117	1.407
F. Model combinations for disaggregates													
Aggreg. Comb.	1.273	1.077	1.024	1.042	1.076	1.059	*1.048*	*1.046*	*1.050*	*1.058*	*1.075*	1.083	1.142
BCB Comb.	1.393	1.168	1.097	1.119	1.138	1.122	1.096	1.107	1.084	1.086	1.093	1.066	1.275
Groups Comb.	1.610	1.271	1.142	1.158	1.133	1.100	1.083	1.090	1.104	1.117	1.107	1.126	1.290
Subgroups Comb.	1.739	1.290	1.187	1.180	1.163	1.138	1.105	1.118	1.133	1.130	1.143	1.172	1.341

Notes: see Table 1.1.

Since the COVID-19 pandemic, most models perform statistically better than the Focus survey, as seen in Table 1.3. Models that look directly at aggregate inflation or use some inflation disaggregation tend to perform very well at all horizons, including nowcasting. Specifically, the adaLASSO, factor model, and FarmPredict working directly with the aggregate inflation achieve RMSE 10% lower than the RMSE of the available Focus and smaller RMSE than the ex-post Focus, something challenging to imagine before the pandemic. Meanwhile, augmented AR and ridge, both employing BCB disaggregation, also perform well. One month ahead, even some models employing the highest level of disaggregation (i.e., subgroups) deliver good results. From the results in this second sub-period, we understand the good performance of the RF for the entire period. Considering aggregate inflation, the method already registers the best performances from $h = 3$, with similar results when using any disaggregation. Frequently, the RF can reduce the RMSE of the available and ex-post Focus by up to 30% for longer horizons, regardless of whether considering aggregate inflation or some disaggregation. For 12-month accumulated inflation, the RF obtains the best performance when using the BCB disaggregation: a 35% reduction compared to the RMSE of the main benchmarks. Finally, regarding model combinations, each is statistically superior to Focus for all $h \geqslant 1$, often at a 1% significance level. However, the combinations do not beat some individual models with good predictive performance in the sub-period.

Table 1.3: Out-of-sample RMSE with respect to the available Focus: Mar/2020 to Jun/2022

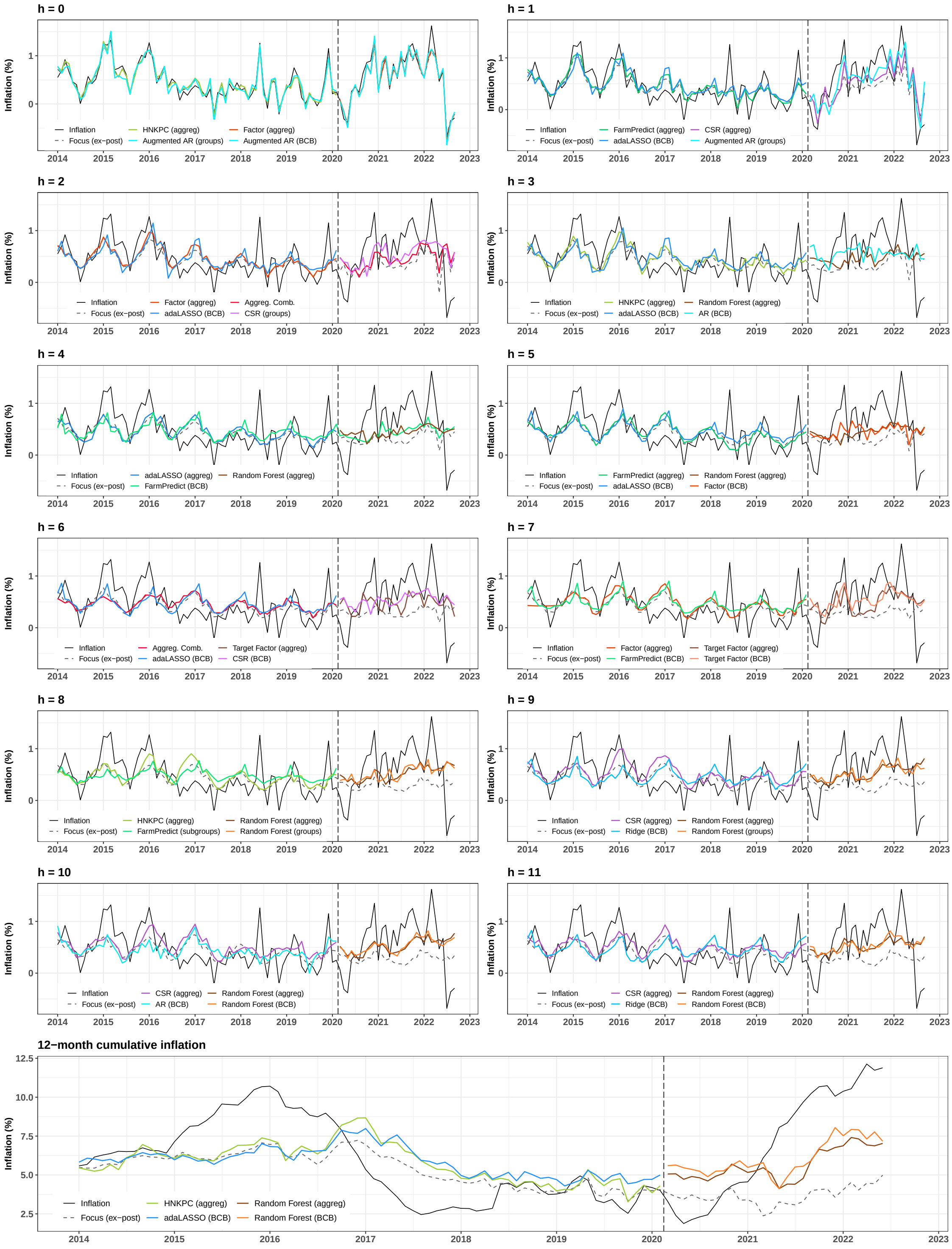
Estimator/Model	$h=0$	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$	$h=6$	$h=7$	$h=8$	$h=9$	$h=10$	$h=11$	$\Sigma\,12m$
A. Survey													
Focus (available)	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Focus (*ex-post*)	0.934**	0.975***	0.991***	1.001	1.000	0.999	1.001	0.999	0.999	1.000	1.002	1.001	1.000
B. Aggregate inflation													
RW	2.443	1.189	1.023	1.056	1.003	1.024	0.983	0.965	0.939	0.908	0.920	0.920	1.128
Historical Mean	2.915	1.203	0.948	0.921**	0.906**	0.888***	0.868***	0.862***	0.863***	0.862***	0.860***	0.857***	0.835***
AR	2.264	1.060	0.904*	0.938*	0.937*	0.910**	0.871***	0.827***	0.807***	0.803***	*0.828****	*0.846****	0.799***
HNKPC	0.927*	0.913***	0.933***	0.971	0.999	0.986	0.917***	0.911**	0.896***	0.913***	0.940***	0.976*	0.938***
Augmented AR	0.924*	0.904**	0.915**	0.951	0.949*	0.916**	0.868***	0.834***	0.825***	0.856***	0.889***	0.927**	0.854***
adaLASSO	*0.896***	0.922**	0.954***	0.948***	0.935***	0.921***	0.889***	0.890***	0.868***	0.893***	0.933***	0.908***	0.879***
Factor	**0.894****	0.896***	0.939***	0.923***	0.907***	0.921***	0.938***	0.943***	0.944***	0.931***	1.005	0.990	0.935***
FarmPredict	*0.896***	0.918***	0.929***	0.958**	0.947**	0.971	0.959**	0.964**	0.947**	0.953**	0.965*	0.962*	0.956**
Target Factor	1.018	1.043	0.986	0.998	0.934	0.948	*0.789****	*0.778****	0.788***	*0.740***	0.889*	1.089	0.772***
CSR	0.954	0.851***	0.904***	0.903**	0.925**	0.909**	0.875***	0.840***	0.845***	0.917**	1.041	1.083	0.851***
Random Forest	1.370	0.951	0.916*	0.896***	0.877***	0.855***	0.815***	0.794***	*0.761****	*0.740****	*0.723****	**0.710*****	*0.689****
C. Disaggregation: tradable, nontradable and monitored prices (BCB)													
AR	2.161	0.931	*0.831***	*0.861***	0.895*	0.876**	0.815***	0.841***	0.853***	0.910*	0.985	1.006	0.779***
Augmented AR	*0.903**	*0.841***	0.914**	0.991	1.024	0.977	0.883***	0.863***	0.849***	0.865***	0.938	1.011	0.904***
Ridge	0.930*	0.842**	0.903**	0.929**	0.944**	0.950*	0.898***	0.870***	0.871***	0.870***	0.904***	0.919***	0.859***
adaLASSO	0.959	0.907*	0.895***	0.866***	0.864***	0.867***	0.881***	0.868***	0.869***	0.854***	0.859***	0.873***	0.769***
Factor	0.948	0.886***	0.884***	*0.861****	*0.832****	**0.808*****	0.829***	0.901***	0.928***	0.946***	0.937***	0.938***	0.827***
FarmPredict	1.024	0.940**	0.928***	0.867***	*0.849****	*0.838****	0.863***	0.934***	0.932***	0.929***	0.933***	0.916***	0.858***
Target Factor	1.046	0.931	0.893**	0.982	0.910	0.924	*0.784****	**0.727*****	0.827***	0.754***	0.881*	1.046	0.729***
CSR	1.973	0.954	*0.834****	*0.864***	0.860**	*0.827****	**0.769*****	0.796***	0.808***	0.831***	0.912**	1.011	*0.719****
Random Forest	1.600	1.042	0.902*	0.955	0.886***	*0.845****	*0.814****	0.790***	*0.748****	*0.733****	**0.707*****	*0.716****	**0.648*****
D. Disaggregation: groups (IBGE)													
AR	2.118	0.888	*0.814**	**0.825****	*0.856***	0.890**	0.981	0.979	1.022	1.043	0.995	0.903**	0.850**
Augmented AR	0.918	**0.767****	0.836**	0.871*	0.924*	0.878**	0.960	0.961	0.983	0.973	1.024	1.043	0.841***
Ridge	1.153	1.143	0.907**	0.916**	0.904**	0.907***	0.884***	0.894***	0.881***	0.887***	0.893***	0.905***	0.835***
adaLASSO	1.070	0.910*	0.915**	0.901**	0.898***	0.900***	0.909**	0.849***	0.803***	0.844***	0.851***	0.879***	0.767***
Factor	0.993	0.880**	0.926**	0.922**	0.934**	0.909***	0.942**	0.854***	0.866***	0.946**	0.911***	0.885***	0.861***
FarmPredict	1.134	0.960	0.923**	0.946*	0.922**	0.916**	0.957*	0.893***	0.874***	0.938**	0.888***	0.871***	0.865***
Target Factor	*0.904*	*0.831***	0.934	0.995	0.901*	0.846**	*0.812****	*0.765****	*0.779****	0.792***	0.865**	1.019	0.768***
CSR	1.849	*0.831**	**0.792*****	*0.828****	**0.789*****	*0.825****	0.823***	0.802***	0.799***	0.834***	0.846***	0.882***	*0.710****
Random Forest	1.895	1.102	0.902*	0.902**	*0.855***	0.851***	0.815***	*0.760****	**0.738*****	**0.731*****	*0.737****	*0.735****	*0.685****
E. Disaggregation: subgroups (IBGE)													
AR	2.715	1.264	0.931	0.939	1.008	0.927	1.009	0.999	1.047	1.158	1.126	1.089	1.067
Augmented AR	1.030	0.982	1.060	1.056	1.061	0.905	0.939	0.995	1.006	1.153	1.180	1.168	0.980
Ridge	2.990	1.179	0.927*	0.927**	0.915**	0.931**	0.916***	0.920**	0.912**	0.896***	0.901***	0.934**	0.903***
adaLASSO	1.210	0.992	0.954	0.960	0.938**	0.925**	0.898***	0.841***	0.841***	0.893***	0.922***	0.936***	0.849***
Factor	1.189	1.037	0.977	0.971	0.943**	0.946**	0.932***	0.887***	0.909***	0.961*	0.962**	0.918***	0.914***
FarmPredict	1.209	1.034	0.975	0.951	0.930**	0.916**	0.925***	0.886***	0.903***	0.942**	0.965**	0.935***	0.904***
Target Factor	0.964	*0.810**	0.979	0.998	0.977	0.959	0.838**	*0.744****	0.823***	0.807**	0.908	1.012	0.810***
CSR	1.963	1.033	0.853***	0.914**	0.875***	0.923***	0.882***	0.829***	0.795***	0.857***	0.913***	0.961	0.832***
Random Forest	1.865	1.114	0.908*	0.926*	0.869**	0.870**	0.825***	0.781***	*0.768****	*0.750****	*0.746****	*0.746****	0.725***
F. Model combinations for disaggregates													
Aggreg. Comb.	1.127	0.909***	0.895***	0.918***	0.906***	0.898***	0.858***	0.845***	0.840***	0.840***	0.886***	0.910***	0.835***
BCB Comb.	1.075	0.866***	0.862***	0.888***	0.877***	0.856***	0.820***	0.827***	0.838***	0.839***	0.878***	0.923***	0.777***
Groups Comb.	1.159	0.864**	*0.833****	0.866***	0.864***	0.861***	0.880***	0.842***	0.836***	0.868***	0.869***	0.888***	0.786***
Subgroups Comb.	1.411	0.975	0.906**	0.932**	0.921***	0.899***	0.882***	0.840***	0.860***	0.898**	0.929**	0.944**	0.869***

Notes: see Table 1.1.

Figure 1.3 presents the temporal evolution of actual inflation, Focus survey expectations, and the best aggregate- and disaggregates-based models for each horizon. Looking at the projections for $h = 0$, we partially understand why it is not easy to outperform the survey in the very-short term. The Focus consensus is very close to the actual values. Furthermore, as we will see in Subsection 1.5.3, available inflation expectations are the primary predictor for model-based nowcasting. The survey contains much relevant information unavailable to the econometrician, so we already expect this result. When analyzing the other horizons ($h \geqslant 1$), we note that it is challenging for the survey and models to predict peaks and valleys of inflation. Already at $h = 1$, we observe outstanding forecasting errors. One consequence of COVID-19, which start is highlighted by vertical dashed lines on each plot, is that the Focus survey initially overestimated inflation and afterward systematically underestimated one. Despite the challenges of generating accurate forecasts in such an uncertain period, several models perform better than expert forecasts for all horizons. A punctual example is the adaLASSO that, using both aggregates and BCB disaggregations, as well as other models, achieves a great result in forecasting the peak observed in December 2020 at $h = 1$, a point at which the available inflation expectation is far from the actual value. Thus, other variables besides the available inflation expectations are fundamental for the performance of model-based forecasts.

Remarks. The good performance of the RF is mainly due to its ability to capture the higher level of future inflation from the second half of 2020. The model generates forecasts closer to the actual inflation than the Focus expectations, which systematically underestimate inflation in that period. Since the pandemic, models for disaggregated inflation tend to provide more accurate forecasts than models for aggregate inflation, except for nowcasting. For each $h \geqslant 2$, we note that adaLASSO, factor model, FarmPredict, and CSR using any disaggregation of inflation deliver forecasts with lower RMSE than the respective models using aggregate inflation, with a few exceptions for the CSR using groups and subgroups that do not outperform the CSR using aggregate inflation. In turn, the RF performs well regardless of the target variable. These findings underscore the use of models in inflation forecasting, including the junction between disaggregated analysis and machine learning techniques, particularly during periods of higher economic instability, such as a pandemic. Previous studies such as Altug & Çakmaklı (2016) and Medeiros *et al.* (2021) have also shown that models perform well during more volatile periods. Next, we will analyze the forecasts for disaggregates and identify the predictors selected by adaLASSO and FarmPredict, methods that allow variable selection.

Figure 1.3: Forecasts by each horizon and 12-month cumulative period



Notes: Black solid lines indicate the actual inflation. Gray dashed lines indicate the median of the ex-post inflation expectations from the Focus survey on the last business day of each month. Solid-colored lines indicate forecasts generated by different models or a combination of models. The vertical dashed lines separate the period before and after the COVID-19 pandemic.

1.5.3
Forecast of disaggregates and variable selection

1.5.3.1
Disaggregation into BCB categories

Predictive performance. Now we consider the predictive performance of the models using different disaggregations, starting with the disaggregation from the BCB. Since we lack long time series for survey-based expectations for disaggregated inflation, we use the AR model as a benchmark. Of all the BCB disaggregations, monitored prices are the most challenging to forecast since they are subject to many unexpected changes resulting from government decisions that often do not freely follow supply and demand movements. According to the results displayed in Table 1.4, other models manage to beat the AR only in short or more distant horizons for this disaggregate (Panel A). In nowcasting, other methods perform significantly better than the AR, with augmented AR and Ridge standing out by obtaining more than a 30% reduction in RMSE. For $h = 1$, augmented AR achieves a 10% reduction in RMSE, the only statistically significant at any level. Some models present minor RMSE for intermediate horizons, but the results are not statistically significant according to the DM test. For ten and eleven months ahead, Ridge delivers reductions of 4% and 6% in RMSE, respectively, compared to the AR, and, specifically for $h = 11$, adaLASSO, factor model, and FarmPredict also statistically outperform the AR, with RMSE reductions ranging from 3% and 4%. Putting all horizons together, Ridge, adaLASSO, FarmPredict, and factor model generate more accurate forecasts than the AR by delivering RMSEs 2% to 4% lower than the AR model, all statistically significant at the 1% level. The good result of the augmented AR model is restricted to the short term.

Looking at the other BCB disaggregates, namely non-tradable and tradable items, we notice that machine learning models deliver better results than the traditional AR model (Panels B and C of Table 1.4). For non-tradables, once again augmented AR and Ridge stand out in nowcasting with RMSE reductions of 18% and 16%, while adaLASSO and FarmPredict achieve the best performances between one and five months ahead, with RMSE reductions oscillating between 20% and 27%. For all $h \geqslant 5$, the random forest dominates by delivering the lowest RMSE. Aggregating all horizons, all ML models perform statistically better than AR for non-tradable items. In turn, the results for tradables are similar, with the ML methods yielding subtly smaller improvements. Also for tradables, there is a predominance of the RF: this method obtains the best or second-best performance for all $h \geqslant 1$, with RMSE reductions ranging from

Table 1.4: Out-of-sample RMSE for BCB disaggregate (in terms of RMSE of the AR model): Jan/2014 to Jun/2022, by disaggregate and horizon

Estimator/Model	$h=0$	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$	$h=6$	$h=7$	$h=8$	$h=9$	$h=10$	$h=11$	all h
							A. Monitored Prices						
AR	1.000	1.000	1.000	1.000	1.000	1.000	*1.000*	1.000	1.000	1.000	1.000	1.000	1.000
Augmented AR	**0.694*****	**0.903*****	1.015	1.022	1.025	1.023	1.026	1.035	1.038	1.037	1.021	1.033	0.996
Ridge	*0.709****	*0.954*	*0.987*	**0.984**	**0.984**	*0.988*	1.001	*0.985*	**0.988**	**0.986**	**0.957****	**0.943****	**0.960*****
adaLASSO	0.756***	0.960	**0.986**	*0.988*	*0.987*	**0.987**	1.004	0.989	*0.996*	1.004	0.971	0.962*	*0.969****
Factor	0.734***	0.970	1.008	1.003	0.999	0.999	1.016	1.009	1.006	*0.996*	0.976	0.961*	0.977***
FarmPredict	0.774***	0.980	1.002	0.998	0.997	1.009	**0.998**	**0.982**	0.996	0.998	0.974	0.965*	0.975***
Target Factor	0.750***	1.033	1.045	1.082	1.063	1.079	1.086	1.149	1.143	1.097	1.073	1.080	1.063
CSR	0.974	1.002	1.022	1.034	1.061	1.053	1.034	1.046	1.066	1.057	1.021	1.027	1.034
Random Forest	0.872***	1.020	1.039	1.060	1.058	1.076	1.098	1.068	1.061	1.033	*0.967*	*0.956*	1.027
							B. Non-Tradables						
AR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Augmented AR	*0.843***	0.831***	0.844***	0.820***	0.810***	0.787***	0.829***	0.853**	0.907**	0.963	0.985	0.975	0.869***
Ridge	**0.824*****	0.817***	0.830***	0.808***	0.796***	0.780***	0.823***	*0.847****	*0.900***	*0.944*	0.957	0.950	0.854***
adaLASSO	0.896*	**0.784*****	**0.778*****	*0.777****	*0.748****	*0.733****	*0.812****	0.850**	0.944	0.977	0.955	0.962	*0.847****
Factor	0.889**	0.809***	0.806***	0.804***	0.765***	0.774***	0.823***	0.866**	0.951	1.006	0.968	0.950	0.864***
FarmPredict	0.895*	*0.798****	*0.782****	**0.768*****	**0.744*****	0.739***	0.842***	0.855**	0.955	1.000	0.968	*0.938*	0.853***
Target Factor	1.040	0.846**	0.896	1.022	0.935	0.799***	0.885*	1.030	1.026	1.015	*0.944*	1.045	0.954**
CSR	0.892***	0.864***	0.887***	0.887***	0.867***	0.816***	0.856***	0.875***	0.931*	0.956	0.968	1.013	0.899***
Random Forest	0.868***	0.855***	0.816***	0.803***	0.778***	**0.729*****	**0.791*****	**0.805*****	**0.854*****	**0.886****	**0.893****	**0.918***	**0.829*****
							C. Tradables						
AR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Augmented AR	0.832**	0.911**	0.995	1.070	1.082	1.051	0.998	0.948**	0.951*	0.989	1.044	1.055	1.003
Ridge	0.810***	0.920**	0.954**	0.982	0.960**	1.042	0.899***	0.874***	0.917**	0.935**	1.030	0.889***	0.940***
adaLASSO	**0.763*****	0.941*	0.943**	*0.940****	*0.909****	*0.844****	0.848***	0.861***	0.918**	0.886***	*0.912***	*0.834****	*0.886****
Factor	0.814***	0.928**	0.972	0.956*	0.927**	0.860***	0.845***	0.882***	0.932**	0.966	0.937*	0.906***	0.913***
FarmPredict	0.811***	0.967	0.980	0.967	0.929**	0.872***	0.858***	0.897***	0.940*	0.956*	0.937*	0.899***	0.920***
Target Factor	*0.792***	**0.886****	*0.907***	0.979	1.000	0.935	*0.821****	**0.769*****	**0.831*****	0.896**	0.990	0.926*	0.900***
CSR	0.884**	1.045	1.079	1.003	0.917*	0.886***	0.909***	0.884***	0.866***	*0.873****	0.925**	0.932***	0.933***
Random Forest	0.820***	*0.903****	**0.906*****	**0.902*****	**0.853*****	**0.824*****	**0.819*****	*0.810****	*0.836****	**0.848*****	**0.841*****	**0.810*****	**0.847*****

Notes: ***, **, and * indicate that for a specific disaggregate and forecast horizon, a model m performed statistically better than an AR model at 1, 5, and 10% significance levels in a one-tailed Diebold-Mariano test with $\mathbb{H}_0 : \text{MSE}(\widehat{\pi}^{m}_{i,t+h\,|\,t}) = \text{MSE}(\pi^{\text{AR}}_{i,t+h\,|\,t})$ *versus* $\mathbb{H}_1 : \text{MSE}(\widehat{\pi}^{m}_{i,t+h\,|\,t}) < \text{MSE}(\pi^{\text{AR}}_{i,t+h\,|\,t})$. The value highlighted in bold blue indicates the best model for each horizon in terms of RMSE ratio with respect to the AR model, and the values in blue italics indicate the second and third best models. The average weights of each disaggregate in IPCA are: monitored prices (25%); non-tradables (41.5%); and tradables (33.5%).

10% to 19%. Other methods that stand out are adaLASSO, which obtains the best result in nowcasting (almost 24% reduction in RMSE), and target factor, which registers the best performance one, seven, and eight months ahead. By gathering the forecasts for all periods, the RF obtains an average reduction of 15% in RMSE compared to the AR model, with adaLASSO coming close behind, with a reduction of 11% in RMSE. All models, except the augmented AR, outperform the AR at the 1% significance level. These findings suggest that models that include more predictors, impose restrictions on parameters, or assume other functional forms can be more advantageous in inflation forecasting than traditional time-series models such as the AR model. Lastly, we note that the improvements due to ML methods in a data-rich environment are not just observed for the pandemic period, as seen in Tables 1.C.1 and 1.C.2 in Appendix 1.C.

Variable selection. To explore potential economic intuitions for the results from aggregate inflation and BCB disaggregation, we compare what is behind the two approaches in terms of variable selections by adaLASSO and FarmPredict, as shown in Figures 1.4 and 1.5, respectively. In both Figures, panel A brings the predictors that each method selected to forecast the aggregate inflation directly. Meanwhile, 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 at least 20% and 11.5% of the time in at least one forecast horizon, respectively. Variables definitions are shown in Table 1.B.1 in Appendix 1.B. The prefix "u_" in some variables in Figure 1.5 indicates that the variable had the common factors "discounted" and, therefore, only its idiosyncratic component is left. These variables are indicated by $\widehat{u}_j$ in Equation (1.4).

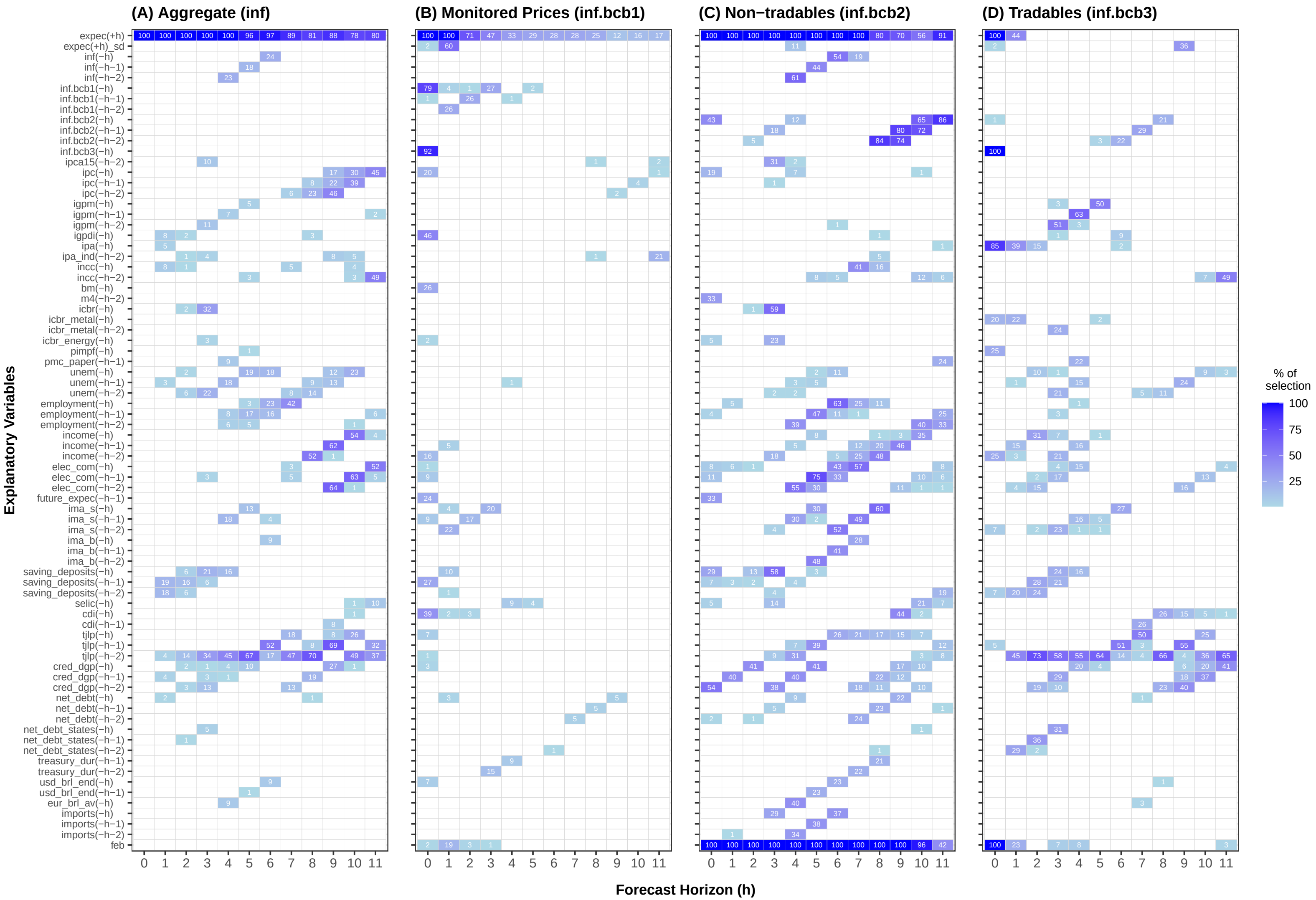
We can summarize the results of the variable selections in the following topics:

1. *High variability in the type of selected predictors*. The adaLASSO selects all ten classes of predictors (see Table 1.B.1 in Appendix 1.B). Even controlling for common factors (FarmPredict), each class of predictors still appears. 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 when we control for common factors via FarmPredict and almost never for adaLASSO. There exists some correlation between activity variables and monetary base, M1 and M2 money supplies, variables occasionally select, for example. Beyond that, the information linked to economic activity variables may be contained in other relevant variables (e.g., inflation expectations), including the possibility of a non-linear relationship between these variables. Thus, we must be careful not to conclude that economic activity variables are irrelevant for forecasting inflation.

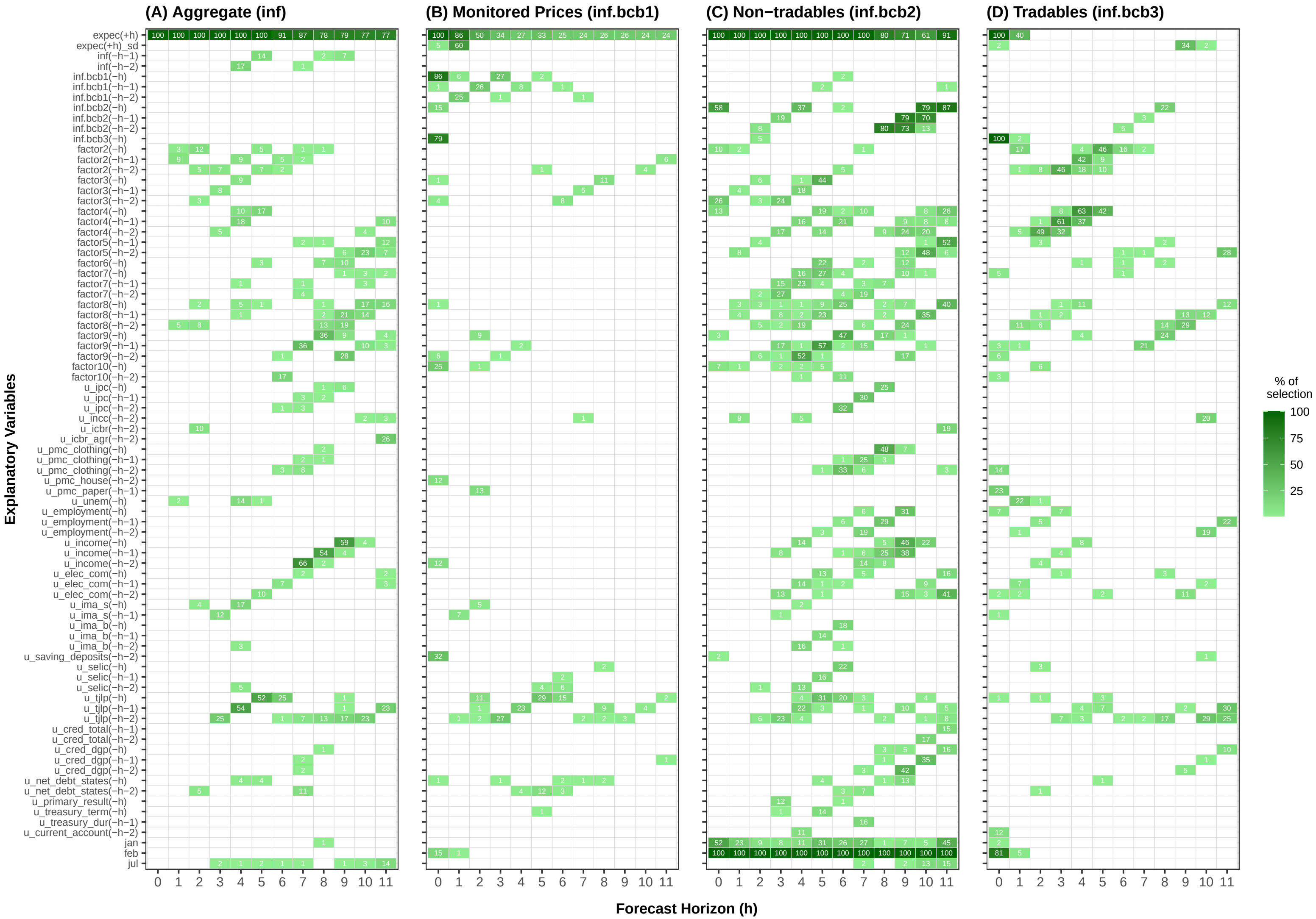
3. *Available inflation expectations (survey) are frequently picked*. adaLASSO and FarmPredict for aggregate and non-tradable items frequently select the inflation expectation (`expec`) at all horizons. For administrated and tradable items, the selection of the expectation decreases as the horizon increases, appearing only in the very short term for tradables. We note that the expectation is about aggregate inflation, so it is reasonable that it is not relevant to explain some specific disaggregate. Furthermore, due to their greater share, non-tradables present a more remarkable similarity with aggregate inflation than the other breakdowns.

Figure 1.4: adaLASSO selection: aggregate inflation and BCB disaggregates (% of extending windows)



Notes: We cut out variables that do not exceed 20% of selection at least one forecast horizon. The definitions of the variables are in Table 1.B.1.

Figure 1.5: FarmPredict selection: aggregate inflation and BCB disaggregates (% of extending windows)



Notes: We cut out variables that do not exceed 11.5% of selection at least one forecast horizon. The definitions of the variables are in Table 1.B.1.

4. *Prices (including commodities) are often chosen*. According to the variable selections, Brazilian inflation indexes and commodities price variations are relevant to forecast official Brazilian inflation. Various indexes and price variations carry relevant information to forecast Brazilian inflation. In addition, due to past inflationary history, Brazil has several monthly indexes calculated by different organizations that cover different periods (e.g., days 1 to 30, 11 to 10, and 21 to 30). Thus, 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 (i.e., `factor1`) is rarely select to forecast aggregate inflation or any disaggregate. Instead, the following factors up to the tenth are chosen for various horizons, mainly to predict non-tradables, lesser extent for aggregate and tradables, and very little for administrated ones.

6. *Non-tradables record the richer structure of predictors; monitored items, the poorer.* The good predictability of non-tradables is potentially related to the larger number of predictors. Note that non-tradables have more predictors than aggregate inflation itself. This finding underscores the importance of looking at disaggregates. For example, the February dummy's relevance for forecasting inflation only appears when we consider the forecast of non-tradable inflation. This dummy is crucial because it captures the variation in education prices, a sector whose contracts are usually updated in January. In turn, the price variation of monitored items has few predictors, contributing to this disaggregate being the most challenging for forecasting.

1.5.3.2
Disaggregation into IBGE groups

Now we address the predictive performance for each IBGE group. The identification of each group, as well as their respective participation in the IPCA, are available in Table 1.A.1 in Appendix 1.A. From Table 1.5, we note that ML methods perform statistically better than the AR model and numerically better than the augmented AR for all components by stacking the horizons. An exception is the target factor, which does not perform well for some groups. Furthermore, from Table 1.C.3 in Appendix 1.C, which presents detailed results of each disaggregate by forecast horizon, we notice that there are infrequent horizons that do not have an ML model performing statistically better than the benchmark.

Table 1.5: Out-of-sample RMSE for IBGE groups (in terms of RMSE of the AR model): Jan/2014 to Jun/2022, by disaggregate, joining all horizons

Estimator/Model	inf.g1	inf.g2	inf.g3	inf.g4	inf.g5	inf.g6	inf.g7	inf.g8	inf.g9
AR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Augmented AR	1.007	0.977**	1.062	0.874***	1.001	**0.881*****	1.006	0.451***	1.066
Ridge	*0.886****	**0.908*****	0.956***	**0.836*****	0.966**	*0.889****	0.917***	**0.446*****	0.805***
adaLASSO	0.900***	*0.913****	0.905***	*0.844****	0.961***	0.995	0.886***	0.457***	**0.790*****
Factor	0.911***	0.915***	0.900***	0.863***	*0.959****	0.994	**0.876*****	*0.450****	*0.795****
FarmPredict	0.909***	0.919***	0.912***	0.920***	0.966***	1.015	*0.882****	0.457***	0.798***
Target Factor	0.932***	0.992	0.905***	0.884***	1.026	0.909***	1.008	0.487***	1.164
CSR	0.954**	0.949***	*0.849****	0.898***	**0.951*****	0.974*	0.927***	0.698***	0.863***
Random Forest	**0.844*****	0.961***	**0.845*****	0.866***	0.968***	0.945***	0.931***	0.573***	0.838***

Notes: see Table 1.4. For a definition of the groups and their respective weights in the IPCA, see Table 1.A.1 in Appendix 1.A.

Different models perform better for different disaggregates. Education (`inf.g8`) is the group most benefited from using other techniques. However, given the good performance of the augmented AR, we infer that predictive improvement is mainly due to the inclusion of the February dummy. Except for augmented AR and target factor, the models also achieve predictive improvement for communication (`inf.g9`), including all horizons individually. Transportation is the group for which the models beat the AR model by stacking the horizons with the smallest margin (`inf.g5`). In addition, the ML methods are not statistically superior to AR in half of the forecast horizons for this disaggregate. The transportation group comprises public transport fares and expenses with own vehicle and fuel, mostly items whose prices are administered by the government, which are difficult to forecast. However, except once again for augmented AR and target factor, the models deliver a statistically significant average reduction of at least 3% in RMSE compared to the AR model, with Ridge's predictive gain around 10% between 6- and 9-month-ahead. Lastly, it is worth highlighting the good performance of the RF to forecast the price variation of foods and beverages (`inf.g1`) at all horizons.

1.5.3.3
Disaggregation into IBGE subgroups

Lastly, we examine the predictive performance for each IBGE subgroup. The results stacking all horizons are shown in Table 1.6, and results for each horizon are in Table 1.C.4 in Appendix 1.C. Descriptions of subgroups are in Table 1.A.1 in Appendix 1.A. Similar to what happens with the disaggregation into groups, ML models achieve more accurate forecasts in comparison to AR models for each subgroup, but with no single model emerging as a dominant predictor for different disaggregates. The reduction of RMSE reaches 71% in the case of courses, reading, and stationery (`inf.sg18`), 40% for communication (`inf.sg19`), and 35% for household operations (`inf.sg7`) and personal services (`inf.sg16`). RF stout out for delivering the best predictive performances considering all forecast horizons for food at home (`inf.sg1`), appliances (`inf.sg6`), household operations (`inf.sg7`), and fabrics (`inf.sg10`). The Ridge performs well at all horizons for domestic fuels and energy (`inf.sg4`) and jewelry (`inf.sg10`), and the adaLASSO performs well for communication (`inf.sg19`) at all horizons as well. Furthermore, we show outstanding performances for specific horizons: adaLASSO performs well in the long-term for food away from home (`inf.sg2`), target factor in short- and intermediate-term for pharmaceutical and optical products (`inf.sg13`), while the RF performs well at short term for that one and at more distant horizons for the latter.

Table 1.6: Out-of-sample RMSE for IBGE subgroups (in terms of RMSE of the AR model): Jan/2014 to Jun/2022, by disaggregate, joining all horizons

Estimator/Model	inf.sg1	inf.sg2	inf.sg3	inf.sg4	inf.sg5	inf.sg6	inf.sg7	inf.sg8	inf.sg9	inf.sg10
AR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Augmented AR	1.062	1.053	1.080	1.000	1.063	1.085	1.050	0.902***	1.040	1.106
Ridge	*0.742****	0.798***	0.804***	**0.794*****	0.854***	0.827***	0.681***	0.886***	0.832***	**0.693*****
adaLASSO	0.766***	**0.759*****	*0.781****	*0.803****	*0.823****	0.826***	*0.677****	*0.737****	*0.768****	0.721***
Factor	0.751***	0.803***	0.792***	0.805***	0.838***	*0.810****	0.684***	0.759***	0.780***	0.725***
FarmPredict	0.757***	0.795***	0.793***	0.810***	0.839***	0.815***	0.687***	0.809***	0.784***	0.728***
Target Factor	0.786***	0.811***	0.857***	0.954	0.858***	0.836***	0.812***	**0.731*****	0.802***	0.808***
CSR	0.779***	0.796***	0.785***	0.850***	**0.801*****	0.819***	0.683***	0.810***	0.791***	0.730***
Random Forest	**0.706*****	*0.772****	**0.770*****	0.858***	0.849***	**0.757*****	**0.650*****	0.759***	**0.734*****	*0.694****

Estimator/Model	inf.sg11	inf.sg12	inf.sg13	inf.sg14	inf.sg15	inf.sg16	inf.sg17	inf.sg18	inf.sg19
AR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Augmented AR	1.147	1.037	0.889***	1.010	1.023	1.117	1.061	0.436***	1.022
Ridge	0.736***	0.874***	0.941**	0.987	0.899***	0.786***	0.742***	0.415***	0.613***
adaLASSO	*0.706****	0.861***	0.827***	0.979*	0.898***	0.713***	*0.737****	0.389***	**0.599*****
Factor	0.736***	0.863***	0.892***	0.970**	0.899***	0.719***	**0.734*****	**0.383*****	0.605***
FarmPredict	0.739***	*0.861****	0.877***	0.978*	**0.896*****	0.730***	0.738***	*0.389****	*0.604****
Target Factor	0.774***	0.923***	**0.729*****	**0.958****	0.924***	0.807***	0.831***	0.400***	0.829***
CSR	0.722***	0.867***	0.848***	*0.967***	*0.897****	*0.662****	0.776***	0.604***	0.653***
Random Forest	**0.673*****	**0.860*****	*0.798****	0.969**	0.913***	**0.643*****	0.761***	0.480***	0.639***

Notes: see Table 1.4. For a definition of the subgroups and their respective weights in the IPCA, see Table 1.A.1 in Appendix 1.A.

Some subgroups are equivalent to groups. It is the case of transportation (`inf.sg12` and `inf.g5`), courses, reading, and stationary (`inf.sg18`) that is equivalent to education (`inf.g8`), and communication (`inf.sg19` and `inf.g9`). Interestingly, the predictive accuracy obtained from disaggregation into subgroups is superior to that based on groups, both stacking the horizons and considering them individually. In particular, the improvements are remarkable for transportation and communication. The difference between both approaches is that, while one includes lags from other subgroups and not groups, the other includes lags from different groups and not subgroups. Note that, stacking the forecast horizons, while CSR reduces the RMSE of the AR model for *group* transportation (`inf.g5`) by 5%, the RF obtains an average reduction of 14% for *subgroup* transportation (`inf.sg12`). For transportation, while the ML methods using disaggregation into groups do not statistically outperform the AR model for horizons 1 to 4, 10, and 11 months ahead, some of these methods employing disaggregation into subgroups generate statistically significant improvements at the most demanding level (i.e., 1%) for all these horizons. For communication, disaggregation into subgroups can improve the already good performance of methods when employing disaggregation into groups.

1.5.3.4 Further remarks

We find that ML models considering several predictors beat the AR model by a wide and statistically significant margin for all disaggregates considered. Comparing the use of group and subgroup disaggregations, the predictive benefits associated with a higher level of disaggregation appear to align with the results of Duarte & Rua (2007), for Portugal, Ibarra (2012), for Mexico, and Bermingham & D'Agostino (2014), for the US and Euro Area. However, upon revisiting the results of the aggregation of disaggregated forecasts (Tables 1.1 to 1.3), we note that the use of disaggregation into subgroups does not generate more accurate forecasts than the direct approach or aggregation from less profound disaggregations. A potential explanation for this result is that we do not consider the aggregation of disaggregated forecasts generated by different models in this essay. As we have seen, a single model does not dominate the forecasts of all disaggregates and often does not even exhibit dominance over time for the same disaggregate (see Figures from 1.D.1 to 1.D.4 in Appendix 1.D). Additionally, there is a possibility of inaccurate disaggregated forecasts occurring at some point in time, which can lead to a deterioration of the predictive accuracy of the aggregation of disaggregated forecasts. Importantly, this deterioration cannot be attributed to the use of the most recent available weights for each item in the consumption basket. When we conduct the forecasting exercises again using the actual weights, the results demonstrate no significant changes.

1.6 Conclusion

In this essay, we investigate the use of inflation disaggregations to forecast the aggregate via aggregation of disaggregated forecasts – what became known as the bottom-up approach. We innovate by considering multi-horizon forecasts of several inflation disaggregates in a data-rich environment (i.e., considering many predictors), which is only possible by employing machine learning methods. Analyzing Brazilian inflation and exploring different levels of disaggregation for inflation, we conduct forecasting exercises from both direct and bottom-up forecast approaches. We highlight the relevance of considering the combination of disaggregated analysis and machine learning methods in the econometrician's toolbox.

For many forecast horizons, the aggregation of disaggregated forecasts performs as well as survey-based expectations and models that generate forecasts directly from the aggregate. Our results reinforce the benefits of using models in a data-rich environment for inflation forecasting, including aggregating disaggregated forecasts generated from machine learning techniques, mainly during volatile periods. During the COVID-19 pandemic, model-based forecasts, including those based on disaggregated data, tend to provide more accurate forecasts than survey-based expectations. For example, the random forest model based on both aggregate and disaggregated inflation delivers great results for intermediate and longer horizons. The selection of predictors obtained by the adaLASSO and FarmPredict indicates the importance of considering a broad and diversified set of variables when forecasting inflation. Regarding the prediction of individual disaggregates, we find that ML models considering several predictors beat the AR model by a wide and statistically significant margin for all disaggregates and the vast majority of horizons.

This essay can be extended in many ways in future research. Firstly, it is possible to replicate the procedures analyzed here for other developed and emerging countries. Secondly, an important possibility is the formulation and implementation of a methodology that combines different models predicting different disaggregates. As we have seen, there is no dominant technique. Thus, to fully exploit the potential of disaggregated forecasting, it is necessary to combine different forecasts to obtain the aggregate forecast. Beyond combining different models in the dimension of disaggregates, one can also combine different models over horizons to improve the forecast of time-accumulated inflation (e.g., 12-month cumulative inflation). Thirdly, combining different levels of disaggregation could be valuable since increasing the level in some branches may be advantageous while for others, it does not. For example, in the Brazilian price index employed in this essay, some groups may be worth keeping in the final combination, while others may benefit from further breakdown into subgroups or deeper disaggregation. Finally, a fourth possibility not explored in this essay is to use breakeven inflation as a predictor and benchmark in inflation forecasting.

1.A
Groups and subgroups of the IPCA

Table 1.A.1: Description of groups e subgroups of the IPCA

Groups / subgroups	**Weight**
Foods and beverages	23.7
Food at home	*15.8*
Food away from home	*7.9*
Housing	15.5
Utilities and maintenance	*10.3*
Domestic fuels and energy	*5.2*
Household goods	4.1
Furniture and fixtures	*2.0*
Appliances	*1.7*
Household operations	*0.3*
Apparel	5.5
Clothes	*3.6*
Footwear and accessories	*1.6*
Jewelry	*0.3*
Fabrics	*0.1*
Transportation	19.0
Transportation	*19.0*
Medical and personal care	12.1
Pharmaceutical and optical products	*3.6*
Medical services	*5.5*
Personal care	*3.0*
Personal expenses	10.6
Personal services	*6.6*
Recreation and tobbaco	*4.0*
Education	5.1
Courses, reading, and stationery	*5.1*
Communication	4.3
Communication	*4.3*

Notes: The column "Weight" shows the average weight of each group and subgroup in the IPCA from January 2014 to June 2022. The text indicates each group and subgroup by `inf.g#` and `inf.sg#`, respectively, where the # indicates the order in which each appears in this table.

1.B
Predictive variables

Table 1.B.1: Description of predictive variables

#	Abbreviation	Description	Unit	Source	Lag	Transformation
		A. Prices and Money				
1	inf, inf.bcb#, inf.g#, inf.sg#	IPCA and disaggregations	index	IBGE	1	% change
2	expec	Focus-based inflation expectation (available)	% per month	BCB	0	–
3	inpc	INPC	index	IBGE	1	% change
4	ipca15	IPCA-15	index	IBGE	0	% change
5	ipc	IPC-Br	index	FGV	1	% change
6	igpm	IGP-M	index	FGV	1	% change
7	igpdi	IGP-DI	index	FGV	1	% change
8	igp10	IGP-10	index	FGV	1	% change
9	ipc_fipe	IPC-Fipe	index	FGV	1	% change
10	ipa	IPA	index	FGV	1	% change
11	ipa_ind	IPA	index	FGV	1	% change
12	ipa_agr	IPA	index	FGV	1	% change
13	incc	INCC	index	FGV	1	% change
14	bm_broad	Broad Monetary Base – end-of-period balance	index	BCB	2	% change
15	bm	Monetary Base – working day balance average	index	BCB	2	% change
16	m1	Money supply M1 – working day balance average	index	BCB	2	% change
17	m2	Money supply M2 – end-of-period balance	index	BCB	2	% change
18	m3	Money supply M3 – end-of-period balance	index	BCB	2	% change
19	m4	Money supply M4 – end-of-period balance	index	BCB	2	% change
		B. Commodities Prices				
20	icbr	Brazilian Commodity index (all)	index	BCB	1	% change
21	icbr_agr	Brazilian Commodity index – agriculture	index	BCB	1	% change
22	icbr_metal	Brazilian Commodity index – metal	index	BCB	1	% change
23	icbr_energy	Brazilian Commodity index – energy	index	BCB	1	% change
		C. Economic Activity				
24	ibcbr	Brazilian IBC-Br Economic Activity index	index	BCB	3	% change
25	month_gdp	GDP monthly – current prices	R$ million	BCB	1	% change
26	tcu	Total capacity utilization – manufacturing industry	%	FGV	1	first difference
27	pimpf	Industry Production – general	index	IBGE	2	% change
28	pmc_total	Retail sales volume – total	index	IBGE	2	% change
29	pmc_fuel	Retail sales volume – fuels and oils	index	IBGE	2	% change
30	pmc_supermarket	Retail sales volume – supermarkets and food products	index	IBGE	2	% change
31	pmc_clothing	Retail sales volume – fabrics, clothing and shoes	index	IBGE	2	% change
32	pmc_house	Retail sales volume – furniture and appliances	index	IBGE	2	% change
33	pmc_drugstore	Retail sales volume – pharmaceutical and cosmetic articles	index	IBGE	2	% change
34	pmc_paper	Retail sales volume – books, newspapers and stationery	index	IBGE	2	% change
35	pmc_office	Retail sales volume – office and eletronical equipments	index	IBGE	2	% change
36	pmc_others	Retail sales volume – others	index	IBGE	2	% change
37	pmc_building	Retail sales volume – building material	index	IBGE	2	% change
38	pmc_auto	Retail sales volume – automotive and parts	index	IBGE	2	% change
39	steel	Steel production	index	BCB	1	–
40	prod_vehicles	Vehicle production – total	units	Anfavea	1	% change
41	prod_agr_mach	Production of agricultural machinery – total	units	Anfavea	1	% change
42	vehicle_sales	Vehicle sales by dealerships – total	units	Fenabrave	1	% change
		D. Employment				
43	unem	Unemployment (combination of PME and PNADC)	%	IBGE	2	first difference
44	employment	Registered employess by economic activity - Total	units	IBGE	1	% change
45	aggreg_wage	Overall Earnings (broad wage income)	R$ (million)	BCB	1	% change
46	min_wage	Federal Minimum Wage	R$	MTb	0	% change
47	income	Households gross disposable national income	R$ (million)	BCB	2	% change
		E. Electricity				
48	elec	Electricity consumption – total	GWh	Eletrobrás	2	% change
49	elec_res	Electricity consumption – residential	GWh	Eletrobrás	2	% change
50	elec_com	Electricity consumption – commercial	GWh	Eletrobrás	2	% change
51	elec_ind	Electricity consumption – industry	GWh	Eletrobrás	2	% change
		F. Confidence				
52	cons_confidence	Consumer Confidence index	index	Fecomercio	1	% change
53	future_expec	Future expectations index	index	Fecomercio	1	% change
54	conditions	Current economic conditions index	index	Fecomercio	1	% change

(continued on next page)

Table 1.B.1 Description of predictive variables (cont.)

#	Abbreviation	Description	Unit	Source	Lag	Transformation
		G. Finance				
55	ibovespa	Ibovespa index	% per month	B3	0	-
56	irf_m	Anbima Market Index of the prefixed federal bonds	index	Anbima	1	% change
57	ima_s	Anbima Market Index of the federal bonds tied to the Selic	index	Anbima	1	% change
58	ima_b	Anbima Market Index of the federal bonds tied to the IPCA	index	Anbima	1	% change
59	ima	General Anbima Market index	index	Anbima	1	% change
60	saving_deposits	Saving deposits – end-of-period balance	R$ (mil)	BCB	2	% change
61	selic	Selic Basic Interest rate	% per month	BCB	0	–
62	cdi	Cetip DI Interbank Deposits rate	% per month	Cetip	0	–
63	tjlp	TJLP Long-term Interest rate	% per year	BCB	1	–
64	ntnb	3-Year Treasury (real) Rate indexed to the IPCA (NTN-B)	% per year	Anbima	0	–
65	embi	Emerging Markets Bond Index Plus – Brazil	b.p. acc. month	JP Morgan	0	first difference
66	vix	CBOE Volatility Index (VIX)	index	CBOE	0	
		H. Credit				
67	cred_total	Credit outstanding – total	R$ (million)	BCB	2	% change
68	cred_dgp	Credit outstanding as a percentage of GDP	% of GDP	BCB	2	first difference
69	indebt_house1	Household debt to income – total	% of 12m income	BCB	2	first difference
70	indebt_house2	Household debt without mortgage loans	% of 12m income	BCB	2	first difference
		I. Government				
71	net_debt_gdp	Net public debt (% GDP) -Consolidated public sector	% of GDP	BCB	2	first difference
72	net_debt	Net public debt – Total – Consolidated public sector	R$ (million)	BCB	2	% change
73	net_debt_fedgov_bcb	Net public debt – Federal Government and Central Bank	R$ (million)	BCB	2	% change
74	net_debt_states	Net public debt – State governments	R$ (million)	BCB	2	% change
75	net_debt_cities	Net public debt – Municipal governments	R$ (million)	BCB	2	% change
76	primary_result	Primary result – Consolidated public sector	R$ (million)	BCB	2	% change
77	debt_fedgov_old	Gross general government debt – Method used until 2007	R$ (million)	BCB	2	% change
78	debt_fedgov_new	Gross general government debt – Method used since 2008	R$ (million)	BCB	2	% change
79	treasury_emit	National Treasuary domestic securities – Total issued	R$ (million)	BCB	2	% change
80	treasury_mkt	National Treasuary domestic securities – Total on market	R$ (million)	BCB	2	% change
81	treasury_term	National Treasury securities debt – medium term	months	BCB	2	first difference
82	treasury_dur	National Treasury securities debt – medium duration	months	BCB	2	first difference
		J. Exchange and International Transactions				
83	reer	Real Effective Exchange Rate	R$/other	BIS	1	% change
84	usd_brl_end	USD-BRL rate – end-of-period	R$/US$	BCB	0	% change
85	usd_brl_av	USD-BRL rate – monthly average	R$/US$	BCB	1	% change
86	eur_brl_end	EUR-BRL rate – end-of-period	R$/€	Bloomberg	0	% change
87	eur_brl_av	EUR-BRL rate – monthly average	R$/€	Bloomberg	1	% change
88	current_account	Current account – net	US$ (million)	BCB	2	% change
89	trade_balance	Balance on goods and services – net (Brazilian trade balance)	US$ (million)	BCB	2	% change
90	exports	Exports	US$ (million)	BCB	2	% change
91	imports	Imports	US$ (million)	BCB	2	% change

Notes: "Lag" column indicates the delay for each variable to become available and "Transformations" column indicates transformation implemented to guarantee the stationarity of the series.

1.C
Forecast performance for disaggregates by horizon

1.C.1
BCB disaggregates before and after the COVID-19 pandemic

Table 1.C.1: Out-of-sample RMSE for BCB disaggregate (in terms of RMSE of the AR model): Jan/2014 to Feb/2020, by disaggregate and horizon

Estimator/Model	$h=0$	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$	$h=6$	$h=7$	$h=8$	$h=9$	$h=10$	$h=11$	all h
A. Monitored Prices													
AR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	*1.000*	1.000	1.000	1.000	1.000
Augmented AR	**0.740*****	**0.860*****	**0.971**	0.985	1.005	1.010	1.030	1.040	1.040	1.015	1.012	1.024	0.983*
Ridge	*0.744****	0.952	0.990	**0.979**	**0.987**	*0.975*	*0.994*	*0.991*	**0.998**	**0.977**	**0.942****	**0.929***	**0.958*****
adaLASSO	0.795***	0.929*	*0.982*	*0.985*	*0.991*	**0.972**	0.999	0.998	1.012	1.010	0.967	0.963	*0.970****
Factor	0.777***	0.937*	1.018	1.011	1.013	0.995	1.019	1.028	1.036	*0.994*	0.976	*0.961*	0.983*
FarmPredict	0.815***	0.957	1.008	1.004	1.009	1.011	**0.988**	**0.983**	1.013	0.998	0.972	0.967	0.979**
Target Factor	0.796***	*0.921*	1.004	1.011	1.069	1.076	1.056	1.202	1.151	1.091	1.052	1.012	1.043
CSR	0.950*	0.961*	0.987	0.996	1.013	1.030	1.028	1.041	1.030	1.008	0.997	1.012	1.005
Random Forest	0.913*	0.991	1.032	1.026	1.060	1.076	1.085	1.050	1.047	1.014	*0.967**	0.977	1.021
B. Non-Tradables													
AR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	*1.000*	*1.000*	**1.000**	1.000
Augmented AR	*0.784****	0.801**	0.830***	0.795***	0.781***	0.735***	0.819***	0.831**	0.914*	1.063	1.112	1.056	0.867***
Ridge	**0.763*****	*0.791****	*0.821****	*0.792****	0.775***	0.735***	0.808***	0.820***	*0.903***	1.024	1.059	*1.015*	*0.850****
adaLASSO	0.850**	**0.790****	**0.783*****	**0.773*****	**0.746*****	**0.709*****	*0.786****	*0.778****	0.922*	1.039	1.089	1.087	0.851***
Factor	0.875*	0.837**	0.843**	0.828***	0.791***	0.758***	0.814**	0.813**	0.911**	1.043	1.053	1.026	0.872***
FarmPredict	0.878*	0.827**	0.835***	0.799***	*0.770****	0.741***	0.814***	0.784***	0.923*	1.051	1.070	1.038	0.866***
Target Factor	1.054	0.805**	0.895*	1.027	0.965	0.750***	0.832**	0.907	1.004	1.085	1.058	1.105	0.949**
CSR	0.882***	0.863***	0.887***	0.861***	0.834***	0.799***	0.843***	0.864**	0.979	1.076	1.074	1.035	0.906***
Random Forest	0.820***	0.803***	0.828***	0.797***	0.775***	*0.721****	**0.765*****	**0.777*****	**0.841*****	**0.942**	**0.997**	1.038	**0.832*****
C. Tradables													
AR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Augmented AR	0.786**	*0.869****	0.966	1.047	1.074	1.049	1.000	0.972	0.979	1.011	1.061	1.039	0.996
Ridge	0.763**	0.884**	**0.900****	0.951	*0.948**	1.060	0.857**	0.846**	0.869**	**0.855*****	1.047	**0.842*****	0.907***
adaLASSO	*0.727****	0.911**	0.930*	*0.948**	0.956*	**0.856****	**0.831*****	0.875**	0.938	*0.871***	0.932	*0.853****	0.889***
Factor	0.741***	**0.866*****	0.952	1.002	0.993	0.909	0.858**	*0.844***	**0.861****	0.892**	*0.885***	0.888**	0.894***
FarmPredict	0.747***	0.894**	0.914**	0.997	0.985	0.910	*0.853***	**0.843****	*0.868***	0.887***	**0.863****	0.858***	*0.887****
Target Factor	**0.722****	0.936	0.932	1.004	1.119	1.010	0.903*	0.863*	0.892*	1.001	1.000	0.964	0.953**
CSR	0.844***	1.046	1.176	1.065	0.979	0.928	0.965	0.949	0.919*	0.927*	0.971	0.973	0.981
Random Forest	0.768***	0.880**	*0.913***	**0.923****	**0.888****	*0.880**	0.868**	0.879*	0.905*	0.894**	0.896**	0.883**	**0.884*****

Notes: ***, **, and * indicate that for a specific disaggregate and forecast horizon, a model m performed statistically better than an AR model at 1, 5, and 10% significance levels in a one-tailed Diebold-Mariano test with $\mathbb{H}_0 : \text{MSE}(\widehat{\pi}^{m}_{i,t+h\,|\,t}) = \text{MSE}(\pi^{\text{AR}}_{i,t+h\,|\,t})$ *versus* $\mathbb{H}_1 : \text{MSE}(\widehat{\pi}^{m}_{i,t+h\,|\,t}) < \text{MSE}(\pi^{\text{AR}}_{i,t+h\,|\,t})$. The value highlighted in bold blue indicates the best model for each horizon in terms of RMSE ratio with respect to the AR model, and the values in blue italics indicate the second and third best models.

Table 1.C.2: Out-of-sample RMSE for BCB disaggregate (in terms of RMSE of the AR model): Mar/2020 to Jun/2022, by disaggregate and horizon

Estimator/Model	$h=0$	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$	$h=6$	$h=7$	$h=8$	$h=9$	$h=10$	$h=11$	all h
A. Monitored Prices													
AR	1.000	1.000	1.000	1.000	1.000	**1.000**	**1.000**	1.000	1.000	1.000	1.000	1.000	1.000
Augmented AR	**0.634*****	**0.955**	1.066	1.066	1.049	1.040	1.022	1.029	1.036	1.064	1.034	1.046	1.012
Ridge	*0.665****	*0.955*	**0.984**	**0.991**	**0.981**	*1.006*	*1.011*	**0.978**	0.975	**0.997**	*0.977*	*0.962*	**0.962*****
adaLASSO	0.706***	0.999	*0.991*	**0.991**	**0.981**	*1.006*	*1.011*	*0.978*	*0.975*	0.997	*0.977*	*0.962*	*0.968***
Factor	0.679***	1.010	0.997	**0.991**	**0.981**	*1.006*	1.012	0.985	**0.971**	**0.997**	*0.977*	*0.962*	0.969**
FarmPredict	0.722***	1.007	0.996	**0.991**	**0.981**	*1.006*	*1.011*	0.980	0.975	**0.997**	*0.977*	*0.962*	0.971**
Target Factor	0.692***	1.160	1.092	1.165	1.056	1.083	1.121	1.083	1.134	1.103	1.101	1.163	1.088
CSR	1.001	1.053	1.063	1.080	1.117	1.081	1.042	1.052	1.106	1.116	1.052	1.047	1.068
Random Forest	0.821**	1.056	1.048	1.101	1.055	1.076	1.115	1.089	1.077	1.058	**0.967**	**0.928**	1.036
B. Non-Tradables													
AR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Augmented AR	0.973	0.889	0.863	0.860	0.855	0.878	0.849	0.907	0.893	0.818**	0.844**	0.891*	0.872***
Ridge	0.957	0.869	0.841	0.833*	0.828	0.858	0.853	0.910	0.894	0.830**	0.848**	0.884**	0.862***
adaLASSO	1.001	0.773**	0.771*	0.784**	0.752**	0.775*	0.862*	1.005	0.989	0.891*	*0.805****	*0.825****	0.841***
Factor	*0.921*	*0.750***	*0.754**	*0.763***	*0.720***	0.803*	*0.841**	0.987	1.031	0.957	0.878**	0.872**	0.851***
FarmPredict	0.938	**0.736****	**0.707****	**0.715****	**0.699****	**0.737****	0.896	1.009	1.018	0.932	0.857***	0.833**	*0.832****
Target Factor	1.003	0.925	0.898	1.014	0.884	0.883	0.983	1.286	1.070	0.919	0.818*	0.985	0.963
CSR	**0.916**	0.866**	0.886*	0.928	0.920	0.848*	0.882*	*0.903*	**0.821****	**0.775*****	0.853***	0.991	0.887***
Random Forest	0.975	0.951	0.799**	0.814**	0.784**	*0.743****	**0.840****	**0.869**	*0.880*	*0.810**	**0.780*****	**0.788*****	**0.824*****
C. Tradables													
AR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Augmented AR	0.932	0.978	1.033	1.100	1.092	1.053	0.994	0.915*	0.913*	0.955	1.019	1.076	1.012
Ridge	*0.912*	0.976	1.024	1.020	0.974	1.021	0.953*	0.909***	0.977	1.046	1.006	0.946*	0.983*
adaLASSO	**0.842***	0.989	0.959	0.930**	0.849***	0.830***	0.871***	0.842***	0.891**	0.908*	0.883**	*0.810****	0.881***
Factor	0.964	1.024	1.000	*0.895****	0.841***	*0.798****	0.825***	0.929**	1.018	1.070	1.006	0.928**	0.938***
FarmPredict	0.944	1.079	1.063	0.928**	0.857***	0.825***	0.864***	0.962*	1.027	1.053	1.031	0.949	0.963***
Target Factor	0.938	**0.796****	**0.872***	0.946	0.838**	0.838***	**0.694*****	**0.627*****	*0.741***	**0.706*****	0.975	0.875*	*0.820****
CSR	0.972	1.044	0.926*	0.918*	*0.838***	0.834***	0.826***	0.792***	0.789***	0.784***	*0.858****	0.877***	0.863***
Random Forest	0.933	*0.941*	*0.897***	**0.875*****	**0.808*****	**0.753*****	*0.745****	*0.710****	**0.734*****	*0.773****	**0.757*****	**0.703*****	**0.792*****

Notes: see Table 1.C.1.

1.C.2
Groups

Table 1.C.3: Out-of-sample RMSE for IBGE groups (in terms of RMSE of the AR model): Jan/2014 to Jun/2022, by disaggregate and horizon

Estimator/Model	$h=0$	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$	$h=6$	$h=7$	$h=8$	$h=9$	$h=10$	$h=11$	all h
A. Foods and beverages (inf.g1)													
AR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Augmented AR	0.959	0.979	1.036	1.034	1.045	1.030	1.006	0.994	0.938	0.929**	1.025	1.069	1.007
Ridge	0.898*	0.916	0.884*	0.903**	0.899**	0.951	0.964	0.965	0.913**	*0.831****	*0.811****	*0.776****	*0.886****
adaLASSO	*0.805***	*0.839****	*0.877***	*0.877***	*0.896***	0.937*	1.012	1.018	0.991	0.896**	0.884*	0.788***	0.900***
Factor	0.841**	0.889**	0.962	0.992	0.964	0.995	1.014	0.966	0.910**	0.833***	0.820***	0.797***	0.911***
FarmPredict	0.847**	0.888*	0.907	0.962	0.933	0.984	1.024	0.966	0.913**	0.872***	0.856**	0.799***	0.909***
Target Factor	0.925	0.908*	0.945	1.010	0.952	*0.915**	**0.894****	**0.924**	**0.862****	0.966	1.013	0.847***	0.932***
CSR	0.886**	0.955	1.067	1.058	0.983	0.944	0.935*	0.938	1.023	0.886***	0.920	0.856**	0.954**
Random Forest	**0.789*****	**0.831*****	**0.836*****	**0.859*****	**0.848*****	**0.890****	*0.926**	*0.927**	*0.890***	**0.817*****	**0.792*****	**0.761*****	**0.844*****
B. Housing (inf.g2)													
AR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Augmented AR	**0.780*****	**0.886****	0.981	1.000	0.992	0.967	1.006	1.012	1.032	0.999	1.017	1.017	0.977**
Ridge	0.873**	0.914**	**0.952**	**0.948****	**0.928****	*0.883****	**0.895*****	*0.936**	*0.900***	0.902**	*0.861***	0.914*	**0.908*****
adaLASSO	0.862***	0.946*	**0.952**	0.949**	0.948*	0.884***	0.899**	**0.935***	0.907**	0.907**	**0.860****	0.914*	*0.913****
Factor	0.889**	0.967	0.952	**0.948****	*0.948**	0.884***	*0.896***	0.936**	**0.897****	*0.901***	0.861**	*0.913**	0.915***
FarmPredict	0.875**	0.949	0.957	0.979	0.965	**0.879*****	0.901**	0.944*	0.917*	**0.898****	0.862**	**0.913***	0.919***
Target Factor	*0.807****	*0.901***	1.184	1.051	1.044	0.930*	0.981	1.068	0.976	1.033	0.943	0.971	0.992
CSR	0.916***	0.934***	0.986	1.027	0.977	0.943*	0.945**	0.984	0.925**	0.929**	0.907**	0.924*	0.949***
Random Forest	0.898**	0.948	1.001	0.980	0.966	0.944*	0.987	1.002	0.972	0.974	0.903*	0.960	0.961***
C. Household goods (inf.g3)													
AR	*1.000*	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Augmented AR	1.053	1.046	1.025	0.984	0.999	1.055	1.064	1.041	1.078	1.091	1.125	1.109	1.062
Ridge	1.140	1.077	1.015	0.976	1.040	0.955	0.926**	0.878***	0.929**	0.908**	0.919**	0.888***	0.956***
adaLASSO	1.034	1.031	1.030	0.996	0.981	0.904**	0.855***	0.800***	0.856***	0.858***	0.854***	0.849***	0.905***
Factor	1.015	0.913*	*0.940*	**0.905****	0.966	0.869***	0.853***	0.885***	0.937*	0.883**	0.885**	0.844***	0.900***
FarmPredict	**0.995**	*0.912**	0.951	0.928*	0.972	0.883**	0.872***	0.890**	0.942*	0.900**	0.906**	0.870***	0.912***
Target Factor	1.030	1.000	0.954	*0.914**	1.152	0.905**	0.864***	0.849***	0.879***	0.839***	0.829***	0.800***	0.905***
CSR	1.125	**0.893***	**0.928***	0.945	**0.885****	**0.834*****	**0.800*****	**0.757*****	*0.817****	*0.807****	*0.812****	*0.797****	*0.849****
Random Forest	1.053	0.914	0.961	0.919*	*0.916**	*0.839****	*0.809****	*0.764****	**0.800*****	**0.790*****	**0.803*****	**0.777*****	**0.845*****
D. Apparel (inf.g4)													
AR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Augmented AR	*0.718****	**0.710*****	*0.731****	0.896**	0.971	0.876**	0.808***	0.801***	0.912*	0.975	1.082	1.026	0.874***
Ridge	**0.714*****	*0.710****	**0.709*****	0.835***	0.893**	*0.834****	0.797***	0.797***	0.858***	0.941*	0.987	0.966	**0.836*****
adaLASSO	0.752***	0.796***	0.797***	*0.798****	*0.880***	**0.805*****	0.827***	0.789***	**0.806*****	0.974	1.001	0.931*	*0.844****
Factor	0.739***	0.816***	0.793***	**0.775*****	**0.878****	0.848***	0.823***	0.849***	0.976	0.929*	0.984	*0.924***	0.863***
FarmPredict	0.907*	0.912*	0.867**	0.824***	0.930	0.865***	0.905**	0.908**	1.001	0.950	1.000	0.948	0.920***
Target Factor	0.755***	0.801***	0.800***	0.909*	0.959	0.862*	**0.742*****	**0.775*****	0.856***	1.030	1.152	1.014	0.884***
CSR	0.922	0.996	0.938*	0.889**	0.918*	0.861**	*0.789****	0.829***	0.879***	*0.903**	*0.933*	0.944*	0.898***
Random Forest	0.831***	0.846***	0.902**	0.897**	0.969	0.869**	0.810***	*0.784****	*0.823****	**0.902****	**0.921***	**0.906****	0.866***
E. Transportation (inf.g5)													
AR	1.000	1.000	*1.000*	*1.000*	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Augmented AR	0.854***	1.002	1.061	1.017	*0.982*	*0.964*	0.977	1.003	0.962	0.994	1.062	1.121	1.001
Ridge	0.859***	1.139	1.116	1.057	0.994	0.970	**0.907****	**0.868*****	**0.877*****	**0.898****	0.997	1.004	0.966**
adaLASSO	*0.786****	*0.998*	1.108	1.070	1.002	0.983	0.927**	0.889***	0.893***	0.917**	0.996	1.001	0.961***
Factor	**0.778*****	**0.966**	1.112	1.070	0.997	0.992	0.932**	0.889***	*0.891***	0.927**	*0.985*	*0.988*	*0.959****
FarmPredict	0.790***	1.014	1.097	1.065	0.999	0.989	0.933**	0.910***	0.898**	0.932*	0.986	1.004	0.966***
Target Factor	0.870**	1.126	1.213	1.188	1.037	1.023	0.951	0.902**	0.937	0.943	1.041	1.155	1.026
CSR	0.891**	1.001	**0.990**	**0.994**	**0.947**	**0.942****	0.925**	*0.876****	0.932*	0.942*	0.995	1.015	**0.951*****
Random Forest	0.896**	1.035	1.122	1.075	1.012	0.985	*0.921***	0.891***	0.909***	*0.911***	**0.972**	**0.959**	0.968***
F. Medical and personal care (inf.g6)													
AR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Augmented AR	0.991	0.987	**0.861*****	*0.792****	**0.781*****	**0.803*****	**0.833*****	**0.872*****	*0.914***	*0.898***	**0.900*****	0.996	**0.881*****
Ridge	*0.972*	*0.977***	*0.896****	0.822***	*0.798****	*0.811****	0.846***	*0.883****	**0.913*****	0.921***	0.916***	*0.959***	*0.889****
adaLASSO	0.974	1.016	0.897**	0.905*	0.931	1.218	1.030	1.051	0.984	1.013	0.944	0.961	0.995
Factor	**0.970**	0.996	0.977	0.998	0.923*	0.992	0.993	1.008	1.081	1.023	0.981	0.994	0.994
FarmPredict	0.974	1.041	1.045	1.013	0.988	1.023	1.020	1.025	1.043	1.004	0.980	1.023	1.015
Target Factor	0.990	**0.958**	0.948	**0.764*****	0.826**	1.026	*0.834****	0.926	0.919**	**0.877****	0.914*	0.964	0.909***
CSR	1.058	1.050	0.918**	0.857***	0.965	1.062	1.037	1.020	0.957	0.915*	0.919	**0.955**	0.974*
Random Forest	0.985	1.002	0.899*	0.908*	0.940	1.056	0.954	0.907**	0.929*	0.908**	*0.904***	0.969	0.945***

(continued on next page)

Table 1.C.3: Out-of-sample RMSE for IBGE groups (in terms of RMSE of the AR model): Jan/2014 to Jun/2022, by disaggregate and horizon (cont.)

Estimator/Model	$h=0$	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$	$h=6$	$h=7$	$h=8$	$h=9$	$h=10$	$h=11$	all h
							G. Personal expenses (`inf.g7`)						
AR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	*1.000*	1.000
Augmented AR	0.953	0.962	1.006	0.981	0.997	1.023	1.040	1.039	1.058	1.015	0.977	1.034	1.006
Ridge	0.936	0.884**	0.884***	0.806***	0.878***	0.978	0.967	0.875**	0.949	0.937	0.943	1.021	0.917***
adaLASSO	0.871**	**0.840*****	**0.845*****	**0.774*****	0.840***	0.969	0.949	**0.863****	**0.927***	*0.897***	*0.907*	1.005	0.886***
Factor	*0.864***	0.856***	0.858***	0.783***	**0.789*****	**0.895****	**0.905***	*0.863***	0.946	0.899**	0.919	**0.982**	**0.876*****
FarmPredict	0.900*	*0.850****	0.882**	*0.781****	*0.812****	*0.909**	*0.941*	0.876**	*0.931*	**0.871****	**0.874***	1.001	*0.882****
Target Factor	0.998	0.973	1.061	0.914	0.971	1.026	0.984	0.931	1.098	0.936	1.070	1.178	1.008
CSR	**0.839*****	0.946	0.888**	0.822***	0.936	0.957	0.949	0.881**	1.066	0.922	0.936	1.018	0.927***
Random Forest	0.881**	0.851***	*0.852****	0.838***	0.943	1.040	1.024	0.903*	0.955	0.933	0.949	1.048	0.931***
							H. Education (`inf.g8`)						
AR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Augmented AR	0.443***	*0.396****	0.395***	**0.367*****	*0.351****	**0.358*****	0.422***	*0.411****	0.507***	0.885	*0.869*	*0.879*	0.451***
Ridge	**0.429*****	**0.386*****	**0.377*****	0.383***	0.364***	0.380***	0.419***	0.418***	0.495***	**0.823***	**0.853**	**0.835**	**0.446*****
adaLASSO	*0.432****	0.400***	0.388***	0.385***	0.361***	0.384***	0.420***	0.419***	0.488***	*0.874*	0.933	0.897	0.457***
Factor	0.432***	0.408***	0.381***	*0.374****	**0.349*****	*0.370****	**0.409*****	**0.404*****	**0.478*****	0.878*	0.923*	0.881	*0.450****
FarmPredict	*0.432****	0.400***	0.387***	0.385***	0.362***	0.384***	*0.418****	0.419***	*0.486****	0.882	0.930	0.897	0.457***
Target Factor	0.536***	0.425***	*0.380****	0.390***	0.377***	0.442***	0.455***	0.457***	0.504***	0.894	0.915	0.943	0.487***
CSR	0.636***	0.823	0.740***	0.663***	0.630***	0.668***	0.654***	0.528***	0.629***	0.924	1.031	1.020	0.698***
Random Forest	0.569***	0.508***	0.538***	0.542***	0.490***	0.543***	0.520***	0.547***	0.607***	0.927	0.915	0.916	0.573***
							I. Communication (`inf.g9`)						
AR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Augmented AR	1.095	1.040	1.002	1.029	1.067	1.060	1.086	1.041	1.077	1.074	1.183	1.030	1.066
Ridge	0.748***	**0.806*****	*0.769****	0.836**	*0.822**	0.859**	0.864**	0.740**	0.771***	0.773***	0.842**	0.874*	0.805***
adaLASSO	**0.733*****	*0.810****	0.773***	*0.828****	0.823*	**0.848****	**0.855****	**0.732****	*0.767****	**0.712*****	0.819*	**0.824***	**0.790*****
Factor	*0.747****	0.820***	**0.764*****	**0.820*****	0.830**	0.868**	0.868**	*0.733***	0.777***	0.733***	**0.787****	*0.835**	*0.795****
FarmPredict	0.748***	0.832***	0.770***	0.832**	**0.818***	*0.853***	*0.859***	0.735**	**0.764*****	*0.733****	*0.817***	0.858*	0.798***
Target Factor	1.031	1.426	1.096	1.210	1.221	1.149	1.552	0.922	1.104	0.916	1.230	1.103	1.164
CSR	0.776***	0.865**	0.842**	0.915*	0.930	0.894**	0.925	0.781***	0.851**	0.889**	0.832*	0.887*	0.863***
Random Forest	0.768***	0.822***	0.831**	0.880*	0.873	0.900*	0.893*	0.762**	0.796**	0.798***	0.921	0.852*	0.838***

Notes: see Table 1.C.1.

1.C.3 Subgroups

Table 1.C.4: Out-of-sample RMSE for IBGE subgroups (in terms of RMSE of the AR model): Jan/2014 to Jun/2022, by disaggregate and horizon

Estimator/Model	$h=0$	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$	$h=6$	$h=7$	$h=8$	$h=9$	$h=10$	$h=11$	all h
A. Food at home (inf.sg1)													
AR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Augmented AR	0.991	1.105	1.123	1.026	1.053	1.058	1.019	1.064	1.016	0.975	1.152	1.106	1.062
Ridge	0.940	0.832***	*0.790****	*0.765****	*0.766****	0.788***	0.745***	0.730***	0.754***	*0.705****	*0.719****	0.562***	*0.742****
adaLASSO	*0.718****	*0.831****	0.916	0.790***	0.777***	0.797***	0.759***	0.790***	0.815***	0.759***	0.788***	*0.561****	0.766***
Factor	0.799**	0.857***	0.834***	0.797***	0.804***	0.814***	0.777***	0.730***	0.754***	*0.705****	*0.719****	0.574***	0.751***
FarmPredict	0.815**	0.843***	0.819***	0.792***	0.785***	0.804***	0.776***	0.736***	*0.753****	0.758***	0.768***	0.579***	0.757***
Target Factor	0.762**	0.864***	0.948	0.870**	0.829***	*0.785****	**0.703*****	0.718***	0.773***	0.816***	0.831**	0.623***	0.786***
CSR	0.783***	0.878**	0.970	0.885*	0.825***	0.807***	0.730***	**0.707*****	0.753***	0.714***	0.804***	0.610***	0.779***
Random Forest	**0.709*****	**0.766*****	**0.761*****	**0.738*****	**0.731*****	**0.747*****	*0.713****	*0.708****	**0.744*****	**0.702*****	**0.703*****	**0.554*****	**0.706*****
B. Food away from home (inf.sg2)													
AR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Augmented AR	1.058	0.982	1.106	1.206	1.088	1.051	0.987	1.038	0.938	1.030	1.074	1.097	1.053
Ridge	0.883*	0.899*	0.863**	0.903**	0.869**	0.835***	0.768***	0.769***	0.763**	*0.742****	*0.672****	*0.749****	0.798***
adaLASSO	*0.796****	*0.803****	0.821***	0.883**	**0.795*****	*0.783****	**0.741*****	0.766***	**0.747*****	**0.730*****	**0.628*****	**0.720*****	**0.759*****
Factor	0.822**	0.822***	0.827***	0.902**	0.893**	0.825***	0.796***	**0.738*****	*0.756****	0.797***	0.714***	0.823***	0.803***
FarmPredict	0.826**	0.846**	*0.802****	0.890**	0.868**	0.827***	0.768***	*0.760****	0.764**	0.783***	0.712***	0.775***	0.795***
Target Factor	0.881**	0.890*	0.879**	0.901**	0.826***	**0.777*****	0.787***	0.764***	0.781**	0.791***	0.728***	0.820***	0.811***
CSR	0.813***	0.822***	0.808***	**0.839*****	*0.804****	0.796***	*0.754****	0.791***	0.844*	0.750***	0.792**	0.770***	0.796***
Random Forest	**0.784*****	**0.799*****	**0.788*****	*0.860****	0.831***	0.811***	0.763***	0.768***	0.766**	0.746***	0.674***	0.754***	*0.772****
C. Utilities and maintenance (inf.sg3)													
AR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Augmented AR	1.031	1.061	1.088	1.090	1.007	1.064	1.019	1.074	1.058	1.143	1.157	1.140	1.080
Ridge	0.908	0.853**	0.801**	0.857**	0.797***	0.831***	0.780***	0.814***	*0.721****	0.854***	0.805***	0.711***	0.804***
adaLASSO	0.763***	0.788***	0.779***	0.816***	0.798***	0.824***	0.785***	*0.811****	0.727***	**0.824*****	*0.798****	*0.690****	*0.781****
Factor	*0.762****	0.806***	0.774***	0.843**	0.785***	0.831***	**0.774*****	0.819***	0.730***	0.869***	0.825***	0.722***	0.792***
FarmPredict	**0.760*****	0.809***	0.780***	0.839***	0.785***	0.820***	0.783***	0.820***	0.729***	0.868**	0.831**	0.730***	0.793***
Target Factor	0.923	**0.766*****	*0.763****	0.958	0.848***	1.026	0.841***	0.832***	0.752***	0.884**	0.844**	0.863**	0.857***
CSR	0.836	0.849***	0.781***	*0.781****	**0.747*****	**0.784*****	*0.776****	**0.793*****	**0.712*****	*0.831****	0.822**	0.754***	0.785***
Random Forest	0.805**	*0.770****	**0.711*****	**0.771*****	*0.772****	*0.806****	0.783***	0.825***	0.726***	0.847***	**0.792*****	**0.675*****	**0.770*****
D. Domestic fuels and energy (inf.sg4)													
AR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Augmented AR	0.802***	0.912**	1.033	1.050	0.976	0.965*	1.003	1.025	1.074	1.035	1.056	1.055	1.000
Ridge	0.776***	**0.760*****	**0.849*****	*0.838****	**0.832*****	**0.780*****	**0.763*****	**0.837*****	**0.812*****	*0.772****	**0.723*****	**0.815*****	**0.794*****
adaLASSO	*0.756****	0.782***	0.887**	*0.838****	*0.856****	**0.780*****	**0.763*****	0.843***	0.816***	**0.769*****	0.743***	*0.832****	*0.803****
Factor	0.773***	0.786***	0.887**	**0.836*****	0.857***	**0.780*****	0.763***	0.843***	0.817***	0.773***	*0.743****	0.834***	0.805***
FarmPredict	0.771***	0.784***	*0.874***	0.848***	0.889**	0.788***	0.784***	*0.840****	*0.816****	0.775***	0.753***	0.834***	0.810***
Target Factor	**0.720*****	*0.772****	1.545	0.984	0.907**	0.898**	0.901**	0.964	0.985	0.931	0.833**	0.911*	0.954
CSR	0.788***	0.831***	0.905**	0.892**	0.894**	0.860**	0.835***	0.930*	0.849***	0.815***	0.780***	0.847***	0.850***
Random Forest	0.789***	0.816***	0.925*	0.878***	0.895**	0.866***	0.875***	0.920**	0.888**	0.841**	0.769***	0.866**	0.858***
E. Furniture and fixtures (inf.sg5)													
AR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Augmented AR	1.071	1.015	1.049	1.129	1.013	1.003	1.025	1.056	1.039	1.138	1.159	1.081	1.063
Ridge	1.152	0.976	0.888**	0.867**	0.832***	0.734***	0.775***	0.842***	0.795***	0.894**	0.871***	0.816***	0.854***
adaLASSO	0.888*	0.910*	0.888**	*0.853****	**0.797*****	**0.724*****	*0.770****	0.851***	0.784***	0.869***	0.837***	0.799***	*0.823****
Factor	**0.878***	*0.876***	0.832***	0.905**	0.868***	0.737***	0.780***	0.856***	0.812***	0.896**	0.862***	0.823***	0.838***
FarmPredict	*0.882**	0.895**	**0.819*****	0.882**	0.844***	0.734***	0.784***	0.865***	0.821***	0.914**	0.877***	0.825***	0.839***
Target Factor	0.927	0.927	0.844**	0.915*	0.989	0.816***	0.813***	*0.822****	*0.760****	0.858***	0.835***	0.847**	0.858***
CSR	0.899*	**0.848*****	*0.824****	**0.826*****	*0.809****	*0.732****	**0.767*****	**0.815*****	**0.758*****	*0.830****	*0.811****	*0.782****	**0.801*****
Random Forest	0.926	0.905**	0.917*	0.968	0.964	0.820***	0.843**	0.861***	0.767***	**0.791*****	**0.753*****	**0.712*****	0.849***
F. Appliances (inf.sg6)													
AR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Augmented AR	1.052	1.099	1.007	1.004	1.031	1.084	1.111	1.073	1.153	1.106	1.083	1.165	1.085
Ridge	0.954	0.875**	0.850***	0.883**	0.883**	0.824***	0.775***	0.735***	0.780***	0.803***	0.795***	0.839**	0.827***
adaLASSO	0.911	0.865**	0.836***	0.876**	0.880**	0.814***	0.757***	*0.714****	0.789**	0.831***	0.829**	0.868**	0.826***
Factor	0.903*	*0.817****	0.840***	0.850***	0.876**	*0.794****	0.768***	0.732***	0.771***	0.804***	0.798***	*0.830***	*0.810****
FarmPredict	0.903*	0.820***	0.858**	*0.846****	0.879**	0.811***	0.787***	0.728***	0.780***	0.800***	0.797***	0.831**	0.815***
Target Factor	*0.874**	0.870**	0.848***	0.897**	0.932	0.800***	0.861**	0.786***	*0.749****	0.781***	0.824***	0.852**	0.836***
CSR	1.057	0.836**	*0.815****	0.886**	*0.852****	0.798***	*0.750****	0.786***	0.753***	*0.772****	*0.766****	0.833**	0.819***
Random Forest	**0.827****	**0.794*****	**0.781*****	**0.806*****	**0.800*****	**0.755*****	**0.718*****	**0.688*****	**0.712*****	**0.731*****	**0.734*****	**0.795*****	**0.757*****

(continued on next page)

Table 1.C.4: Out-of-sample RMSE for IBGE subgroups (in terms of RMSE of the AR model): Jan/2014 to Jun/2022, by disaggregate and horizon (cont.)

Estimator/Model	$h=0$	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$	$h=6$	$h=7$	$h=8$	$h=9$	$h=10$	$h=11$	all h
G. Household operations (`inf.sg7`)													
AR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Augmented AR	1.079	1.047	0.995	1.062	0.989	1.064	1.066	1.103	1.107	1.040	0.995	1.108	1.050
Ridge	_0.705_***	_0.700_***	0.607***	0.681***	0.649***	0.733***	0.817***	0.730***	0.738***	0.630***	0.594***	0.685***	0.681***
adaLASSO	0.720***	0.708***	_0.606_***	_0.679_***	0.653***	0.749***	0.798***	_0.714_***	_0.716_***	0.632***	0.592***	**0.651*****	_0.677_***
Factor	0.712***	0.711***	0.623***	0.696***	_0.646_***	0.733***	0.820***	0.731***	0.724***	0.632***	0.594***	0.682***	0.684***
FarmPredict	0.709***	0.721***	0.613***	0.688***	0.652***	0.735***	0.817***	0.725***	0.738***	0.640***	0.607***	0.695***	0.687***
Target Factor	0.839**	0.772***	0.698***	0.858*	0.783***	0.837***	0.926	0.822**	0.851**	0.811**	0.688**	0.938	0.812***
CSR	0.726***	0.711***	0.657***	0.712***	0.670***	_0.697_***	_0.766_***	0.741***	0.733***	_0.612_***	_0.579_***	0.671***	0.683***
Random Forest	**0.668*****	**0.684*****	**0.593*****	**0.657*****	**0.610*****	**0.684*****	**0.756*****	**0.695*****	**0.709*****	**0.607*****	**0.556*****	_0.663_***	**0.650*****
H. Clothes (`inf.sg8`)													
AR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Augmented AR	0.824***	0.874*	0.804***	0.838***	0.891*	0.871*	0.883*	0.897*	0.924	1.024	1.014	0.940	0.902***
Ridge	0.804***	0.887**	0.781***	0.861**	0.857*	0.966	0.902*	0.875**	0.880**	0.939	0.938	0.912	0.886***
adaLASSO	0.745***	0.818***	_0.704_***	**0.629*****	**0.684*****	**0.701*****	0.770***	_0.760_***	**0.687*****	0.775**	0.796***	0.781***	_0.737_***
Factor	**0.722*****	_0.793_***	0.720***	_0.631_***	_0.694_***	_0.711_***	_0.754_***	0.845***	0.847**	0.779**	0.806***	_0.772_***	0.759***
FarmPredict	0.873**	0.896*	0.789***	0.688***	0.735***	0.721***	0.819***	0.876**	0.890*	0.781**	0.834***	0.800***	0.809***
Target Factor	_0.723_***	**0.715*****	**0.659*****	0.646***	0.704***	0.754***	**0.701*****	**0.686*****	0.779***	0.769**	0.829**	0.782***	**0.731*****
CSR	0.966	1.039	0.784***	0.745***	0.791**	0.780***	0.818***	0.802***	_0.717_***	_0.755_**	_0.772_***	0.773***	0.810***
Random Forest	0.813***	0.834***	0.776***	0.698***	0.745***	0.731***	0.780***	0.765***	0.725***	**0.746****	**0.762*****	**0.749*****	0.759***
I. Footwear and accessories (`inf.sg9`)													
AR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Augmented AR	1.023	1.063	0.957	0.959	1.011	1.049	0.978	0.932*	1.000	1.145	1.218	1.062	1.040
Ridge	0.938	0.925	0.874*	0.776**	0.884	0.951	0.838**	0.794***	0.800**	0.730*	0.791*	0.798***	0.832***
adaLASSO	0.812***	_0.782_***	**0.780*****	_0.672_***	_0.769_*	_0.860_**	0.876**	0.789***	0.794**	_0.668_*	0.734**	0.738***	_0.768_***
Factor	0.830***	0.860**	0.792***	**0.670*****	0.786*	0.914*	0.862**	_0.787_***	0.817*	0.670*	_0.717_**	_0.731_***	0.780***
FarmPredict	0.884**	0.919	0.837**	0.676***	0.773*	0.885*	_0.805_***	0.792***	0.797**	0.676*	0.727**	0.739***	0.784***
Target Factor	**0.741*****	0.788***	_0.783_***	0.741***	0.818*	0.911*	0.831**	0.810**	0.841*	0.851	0.731**	0.765***	0.802***
CSR	0.889	1.004	0.802***	0.685***	0.789*	0.910*	0.831***	0.824**	**0.719****	0.670*	0.737**	0.736***	0.791***
Random Forest	_0.769_***	**0.779*****	0.792***	0.681***	**0.763***	**0.840*****	**0.752*****	**0.715*****	_0.732_**	**0.652****	**0.698****	**0.707*****	**0.734*****
J. Jewelry (`inf.sg10`)													
AR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Augmented AR	1.084	1.097	1.129	1.074	1.119	1.029	1.106	1.110	0.940	1.166	1.137	1.280	1.106
Ridge	0.714	0.691***	_0.704_***	_0.697_***	_0.681_***	**0.677*****	**0.662*****	**0.717*****	**0.678*****	**0.697*****	**0.666*****	_0.739_***	**0.693*****
adaLASSO	**0.678*****	**0.652*****	0.716	0.732***	0.713***	0.727***	0.693***	0.762***	0.722***	0.739***	0.691***	0.827***	0.721***
Factor	0.722**	0.705***	0.732***	0.751***	0.764***	0.723***	0.709***	_0.745_***	0.683***	**0.697*****	_0.671_***	0.810***	0.725***
FarmPredict	0.723**	0.704***	0.731***	0.740***	0.725***	0.733***	0.726***	0.822***	_0.679_***	0.705***	0.680***	0.774***	0.728***
Target Factor	0.771**	0.777***	0.811	0.798**	0.765***	0.802***	0.867***	0.883***	0.807***	0.788***	0.801**	0.814***	0.808***
CSR	0.705***	0.683**	0.710	0.699*	0.717***	0.712***	0.701***	0.776***	0.761***	0.746***	0.726***	0.822***	0.730***
Random Forest	_0.683_***	_0.664_***	**0.674*****	**0.682*****	**0.679*****	_0.687_***	_0.682_***	0.746***	0.706***	0.722***	0.675***	**0.730*****	_0.694_***
K. Fabrics (`inf.sg11`)													
AR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Augmented AR	1.170	1.134	1.125	1.114	1.155	1.140	1.130	1.099	1.076	1.343	1.136	1.041	1.147
Ridge	0.811	0.833***	0.759***	0.766***	0.724***	0.734***	0.820***	0.807***	0.789***	0.643***	0.595***	0.690***	0.736***
adaLASSO	0.799***	0.833***	0.758	_0.726_***	0.730***	_0.698_***	_0.766_***	0.770***	0.755***	_0.600_***	**0.547*****	_0.639_***	_0.706_***
Factor	_0.773_**	0.831***	_0.747_***	0.759***	0.727***	0.738***	0.822***	0.816***	0.804***	0.648***	0.611***	0.685***	0.736***
FarmPredict	0.790**	_0.818_***	0.747***	0.761***	0.724***	0.743***	0.853***	0.836***	0.794***	0.643***	0.605***	0.693***	0.739***
Target Factor	0.825**	0.900***	0.796	0.750**	_0.723_***	0.762***	0.828***	0.972***	0.811***	0.719***	0.600**	0.740***	0.774***
CSR	0.820***	0.838**	0.757	0.846*	0.736***	0.719***	0.769***	_0.756_***	**0.706*****	**0.584*****	0.592***	0.677***	0.722***
Random Forest	**0.729*****	**0.745*****	**0.694*****	**0.695*****	**0.660*****	**0.678*****	**0.746*****	**0.733*****	_0.727_***	0.604***	_0.550_***	**0.638*****	**0.673*****
L. Transportation (`inf.sg12`)													
AR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Augmented AR	0.858	1.000	1.116	1.063	1.022	1.000	0.977	1.042	1.052	1.039	1.083	1.135	1.037
Ridge	0.977*	0.915*	0.963**	_0.925_**	_0.885_**	_0.880_***	_0.832_***	**0.830*****	_0.839_**	0.787***	_0.877_***	_0.844_***	0.874***
adaLASSO	0.709***	0.830***	0.955***	0.955**	0.907***	**0.880*****	**0.825*****	_0.830_***	0.852***	0.804***	0.900***	0.878***	0.861***
Factor	**0.683****	**0.821*****	0.967***	0.946**	0.890**	0.900***	0.837***	0.839***	0.853***	0.818***	0.908***	0.866***	0.863***
FarmPredict	_0.705_**	_0.827_**	_0.951_***	0.930**	0.887**	0.895***	0.836***	0.842***	0.858**	0.817***	0.899***	0.871***	_0.861_***
Target Factor	0.783**	0.912*	1.071**	1.082**	0.938***	0.966***	0.900***	0.855***	0.877**	**0.774*****	0.930***	0.999***	0.923***
CSR	0.812***	0.890***	**0.903*****	**0.901*****	**0.881*****	0.901***	0.859***	0.843***	**0.830***	0.847***	0.884**	0.859***	0.867***
Random Forest	0.789***	0.861***	0.954***	0.945***	0.897***	0.907***	0.844***	0.852***	0.864**	_0.780_***	**0.840*****	**0.806*****	**0.860*****
M. Pharmaceutical and optical products (`inf.sg13`)													
AR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Augmented AR	1.086	1.047	0.812	_0.797_	_0.800_	_0.814_	0.852	0.868	0.910	0.885	0.907	1.042	0.889
Ridge	1.273	1.217**	0.937**	0.895**	0.903***	0.897***	0.847***	0.882***	0.990***	0.902***	0.853***	1.010***	0.941***
adaLASSO	**0.941*****	0.981***	0.814***	0.822***	0.925***	0.909***	0.874***	0.755***	0.755***	0.796***	0.718***	_0.774_***	0.827***
Factor	0.965***	0.996***	1.020***	0.932**	0.943***	0.918***	0.871***	0.811***	0.947***	0.822***	0.798***	0.780***	0.892***
FarmPredict	0.971***	0.984***	1.019***	0.924***	0.946***	0.914***	0.874***	0.767***	0.865***	0.805**	0.795**	0.780***	0.877***
Target Factor	_0.951_	**0.937*****	**0.740*****	**0.698**	**0.717*****	**0.671**	**0.669*****	**0.681*****	**0.697*****	_0.727_**	0.714**	0.775**	**0.729*****
CSR	1.126	1.061***	0.805***	0.922***	0.943***	0.928***	0.875***	0.764***	0.755***	0.734***	_0.708_**	0.788***	0.848***
Random Forest	1.043**	_0.967_***	_0.743_***	0.809***	0.955***	0.900***	_0.802_***	_0.697_***	_0.737_***	**0.694*****	**0.673*****	**0.756*****	_0.798_***

(continued on next page)

Table 1.C.4: Out-of-sample RMSE for IBGE subgroups (in terms of RMSE of the AR model): Jan/2014 to Jun/2022, by disaggregate and horizon (cont.)

Estimator/Model	$h=0$	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$	$h=6$	$h=7$	$h=8$	$h=9$	$h=10$	$h=11$	all h
N. Medical services (`inf.sg14`)													
AR	1.000	1.000	1.000	1.000	1.000	1.000	*1.000*	1.000	1.000	1.000	1.000	1.000	1.000
Augmented AR	1.009***	1.012**	1.018	0.989	1.013	1.035*	1.002	1.025	1.007	1.028	0.999	0.980	1.010
Ridge	0.983***	0.984***	0.989***	0.987***	0.979***	0.971***	1.012***	0.984***	1.003***	1.006***	0.978***	**0.977*****	0.987***
adaLASSO	0.952***	0.903***	*0.904***	1.009***	1.025***	0.978***	1.039***	0.964***	0.991***	1.032***	0.964***	0.998***	0.979***
Factor	0.945***	*0.901****	0.934**	1.011***	0.982***	0.980***	1.008***	*0.962****	0.982***	1.012***	0.957***	*0.979****	0.970***
FarmPredict	**0.931*****	0.906***	0.928**	1.043***	0.988**	0.995***	1.046***	0.967***	*0.972****	1.035***	*0.948****	0.980***	0.978***
Target Factor	0.949***	1.019***	0.967	**0.932**	0.979**	**0.919****	**0.989****	**0.852**	**0.938**	*0.984*	**0.945****	1.009*	**0.958**
CSR	*0.938****	**0.899*****	**0.902****	0.979**	*0.960***	*0.966***	1.020***	0.973*	0.976***	1.030***	0.987***	0.998***	*0.967****
Random Forest	0.947***	0.944***	0.916*	*0.967****	**0.957****	0.967***	1.015***	0.988**	0.975**	**0.975****	0.974***	1.023**	0.969***
O. Personal care (`inf.sg15`)													
AR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Augmented AR	0.973	0.995	1.034	1.059	1.042	1.019	1.001	0.984	1.071	1.054	1.005	1.021	1.023
Ridge	0.905	0.901	*0.880***	*0.863***	*0.858****	0.953***	0.982***	0.923***	0.901***	0.907**	0.914***	0.833***	0.899***
adaLASSO	0.922*	*0.893**	0.895**	0.881***	0.870***	0.954***	0.975***	0.918***	*0.878****	*0.889****	**0.892*****	0.833***	0.898***
Factor	**0.881***	0.905**	0.882***	0.867**	0.871***	0.950***	0.981***	0.917***	0.924***	0.896**	*0.910****	0.834***	0.899***
FarmPredict	*0.885**	0.900**	**0.869*****	**0.851****	**0.849*****	*0.950****	0.983***	0.925***	0.916***	0.906**	0.918***	*0.832****	**0.896*****
Target Factor	0.950	**0.881**	0.921**	0.883*	0.881	**0.898*****	**0.923*****	**0.852*****	0.935***	0.966***	1.084***	0.903**	0.924***
CSR	0.924*	0.903***	0.888***	0.878***	0.889***	0.952***	*0.966****	*0.898****	**0.863*****	**0.869*****	0.915***	0.840***	*0.897****
Random Forest	0.899	0.925**	0.924*	0.903	0.868	0.951***	0.988**	0.922***	0.913***	0.928***	0.926***	**0.831*****	0.913***
P. Personal services (`inf.sg16`)													
AR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Augmented AR	1.097	1.138	1.027	1.068	1.184	1.138	1.066	0.987	1.225	1.260	1.098	1.151	1.117
Ridge	0.864	0.778**	0.765***	0.786**	0.826**	0.733***	0.781***	0.727***	0.821***	0.972***	0.817***	0.675**	0.786***
adaLASSO	0.698	0.698**	0.707***	0.767**	0.816**	0.712***	0.753***	0.689***	0.663**	**0.769*****	**0.649****	0.663**	0.713***
Factor	0.711*	0.719***	0.734***	0.755***	0.833**	0.725***	0.749***	0.700***	**0.651*****	*0.778****	*0.657****	0.649**	0.719***
FarmPredict	0.742*	0.758***	0.747**	0.793***	0.821**	0.707***	0.757***	0.711***	*0.656****	0.783***	0.666***	0.660**	0.730***
Target Factor	0.792*	0.742**	0.807***	0.878**	0.786	0.786***	0.977**	0.735***	0.781***	0.838***	0.742***	0.804**	0.807***
CSR	*0.655*	*0.571***	*0.609****	*0.636***	**0.642*****	*0.605****	*0.630****	*0.629****	0.723***	0.929***	0.728***	*0.646***	*0.662****
Random Forest	**0.606****	**0.564*****	**0.576*****	**0.619*****	*0.659****	**0.559*****	**0.629*****	**0.605*****	0.716***	0.884***	0.731***	**0.619*****	**0.643*****
Q. Recreation and tobbaco (`inf.sg17`)													
AR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Augmented AR	0.988	1.019	1.021	1.135	1.042	1.032	1.017	1.076	1.144	1.129	1.029	1.056	1.061
Ridge	0.730***	0.786***	0.791***	*0.740****	0.695***	0.773***	0.749***	0.756***	0.703***	0.684***	0.774***	*0.747****	0.742***
adaLASSO	*0.718****	0.791***	0.798***	0.744***	0.676***	*0.759****	*0.740****	**0.752*****	0.699***	*0.678****	0.772***	0.750***	*0.737****
Factor	**0.705*****	**0.781*****	**0.776*****	0.742***	**0.654*****	**0.756*****	**0.737*****	0.756***	0.705***	0.690***	0.792***	**0.744*****	**0.734*****
FarmPredict	0.719***	*0.783****	0.787***	0.740***	*0.666****	0.781***	0.767***	*0.753****	*0.698****	0.680***	*0.772****	0.747***	0.738***
Target Factor	0.781**	0.810***	0.896***	1.108*	0.809***	0.824***	0.839	0.811**	0.754**	0.699**	0.835**	0.773	0.831***
CSR	0.723***	0.854***	0.795***	**0.732*****	0.727***	0.774***	0.767***	0.862***	0.794***	0.729***	0.796***	0.778***	0.776***
Random Forest	0.753***	0.796***	*0.784****	0.776***	0.757***	0.850***	0.815***	0.769***	**0.693*****	**0.652*****	**0.763*****	0.747***	0.761***
R. Courses, reading, and stationery (`inf.sg18`)													
AR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Augmented AR	0.419***	0.411*	0.387***	0.330***	0.336*	0.328*	0.397*	0.443*	0.543	0.830	0.869	0.928	0.436***
Ridge	0.410***	0.364**	0.345***	0.330**	0.349*	0.342	0.389*	0.440**	0.484**	**0.765**	**0.794**	0.844	0.415***
adaLASSO	**0.346*****	**0.350*****	**0.328*****	0.303***	*0.307****	0.311***	0.354***	0.400***	0.453***	0.811**	0.823***	*0.819****	0.389***
Factor	0.353***	0.358***	0.338***	*0.295****	**0.304*****	**0.301*****	**0.345*****	**0.372*****	**0.416****	0.796**	0.810***	0.820***	**0.383*****
FarmPredict	*0.346***	*0.350**	*0.328****	0.303***	0.308***	0.311***	*0.354****	*0.400***	*0.452**	0.803**	0.821***	0.822***	*0.389****
Target Factor	0.384***	0.369***	0.335***	**0.295*****	0.330***	*0.304****	0.395***	0.426***	0.456***	*0.787***	0.845**	**0.761*****	0.400***
CSR	0.466	0.801	0.603***	0.508***	0.550**	0.528***	0.568***	0.502***	0.693***	0.834**	0.917***	0.910***	0.604***
Random Forest	0.454***	0.455***	0.445***	0.425***	0.414***	0.420***	0.428***	0.497***	0.556***	0.848**	*0.802****	0.843***	0.480***
S. Communication (`inf.sg19`)													
AR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Augmented AR	1.141	1.108	1.062	1.064	1.058	1.091	0.972	0.961*	0.919	0.928	0.988	1.030	1.022
Ridge	0.653	0.705	*0.618**	*0.642***	0.622	0.664	0.677**	0.616***	0.553**	0.574*	0.535*	0.573***	0.613***
adaLASSO	**0.638*****	**0.700*****	**0.613*****	**0.636*****	**0.616***	**0.656****	*0.673***	**0.609*****	*0.551***	**0.528***	*0.506***	*0.537****	**0.599*****
Factor	0.650***	*0.702***	0.619***	0.647***	*0.621**	*0.657**	0.674**	*0.609****	0.569*	*0.539**	**0.502****	0.545***	0.605***
FarmPredict	*0.645***	0.720	0.621**	0.643***	0.622*	0.660*	**0.673*****	0.611***	**0.549****	0.544*	0.514**	**0.532*****	*0.604****
Target Factor	0.856***	1.154***	0.874***	0.986***	0.912*	0.864*	0.934**	0.732**	0.667*	0.695	0.708**	0.660***	0.829***
CSR	0.672	0.761	0.646***	0.684***	0.680*	0.679*	0.723***	0.631**	0.599**	0.657*	0.511**	0.663***	0.653***
Random Forest	0.656***	0.711***	0.652***	0.662***	0.668*	0.694**	0.734***	0.649***	0.580**	0.602**	0.576**	0.560***	0.639***

Notes: see Table 1.C.1.

1.D
Frequency of models with least squared forecast error

Figure 1.D.1: Frequency each model attained the least forecast squared error for BCB disaggregation, by horizon and disaggregate (%)

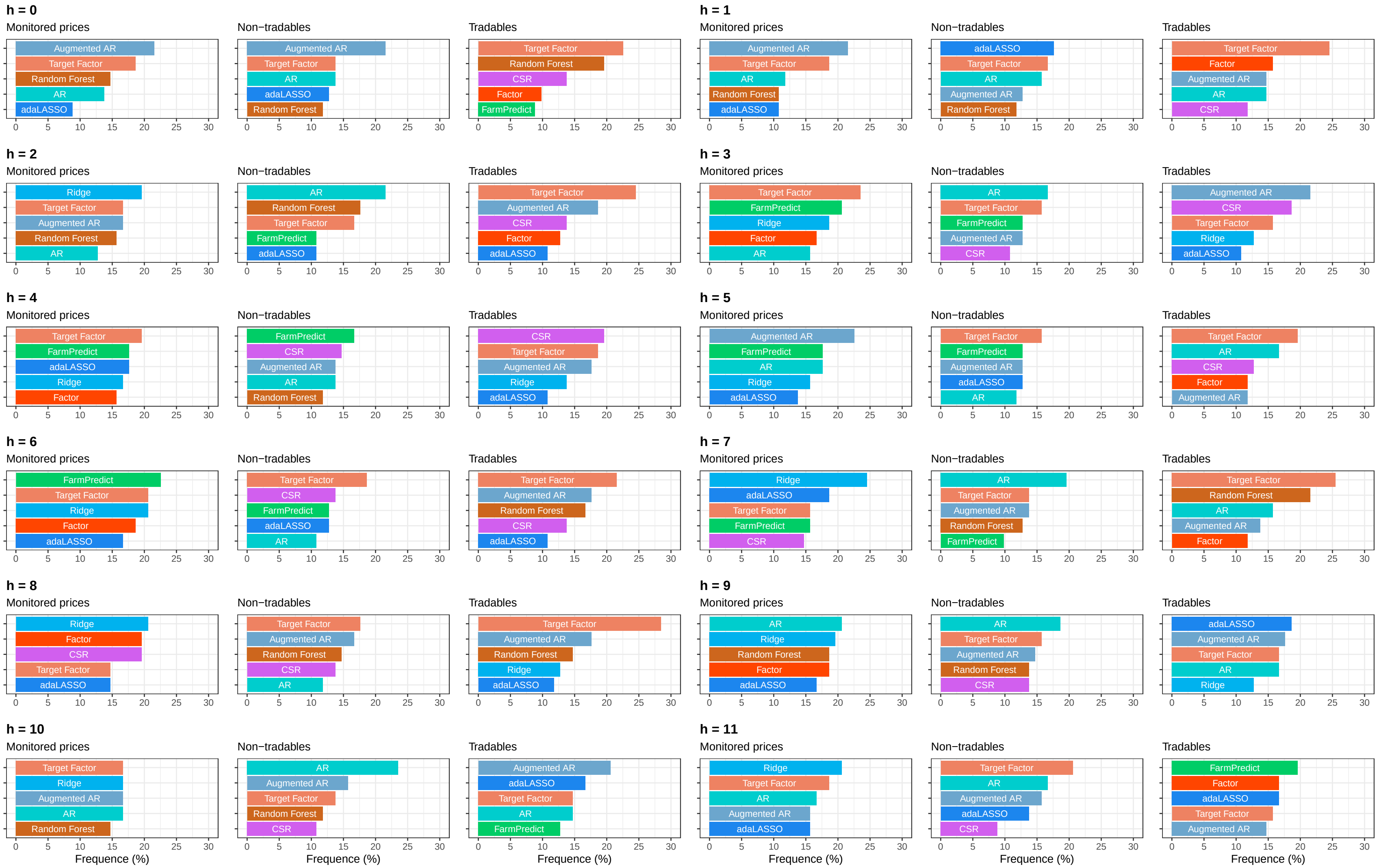


Figure 1.D.2: Frequency each model attained the least forecast squared error for BCB disaggregation, by disaggregate and stacking the horizons (%)

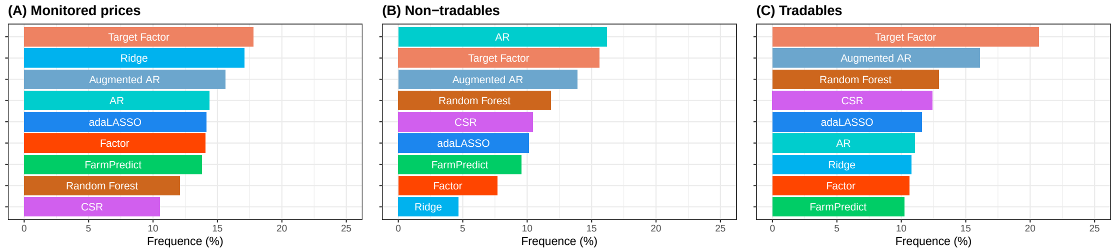


Figure 1.D.3: Frequency each model attained the least forecast squared error for IBGE groups, by disaggregate and stacking the horizons (%)

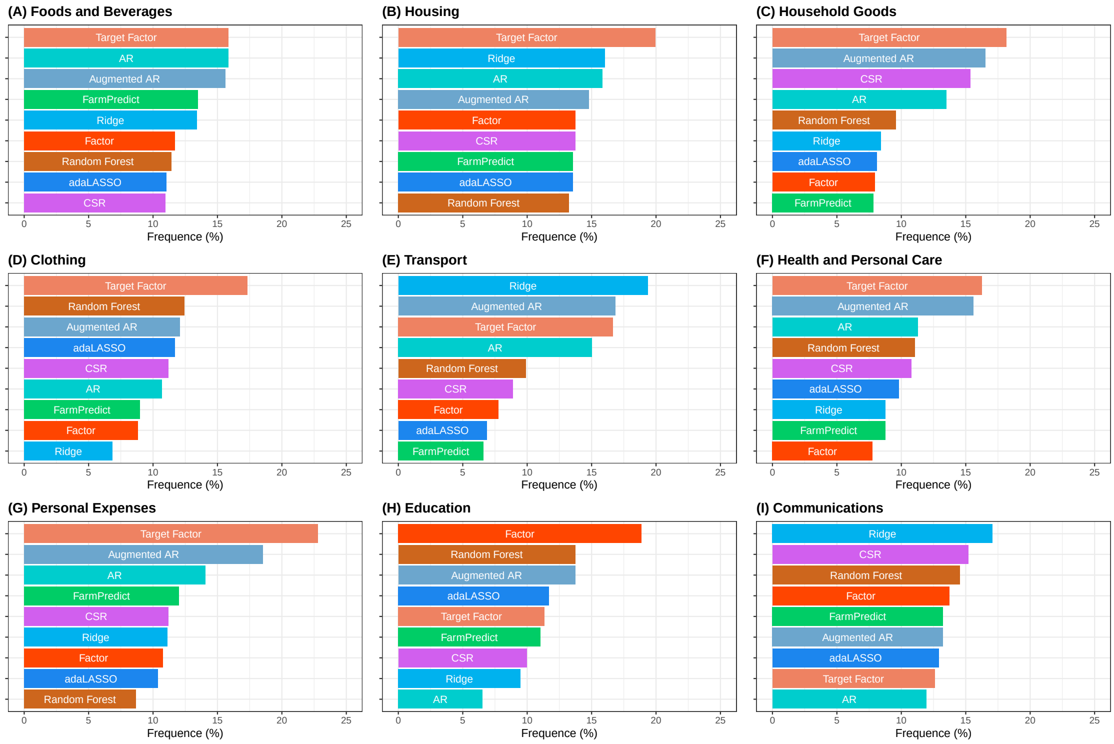
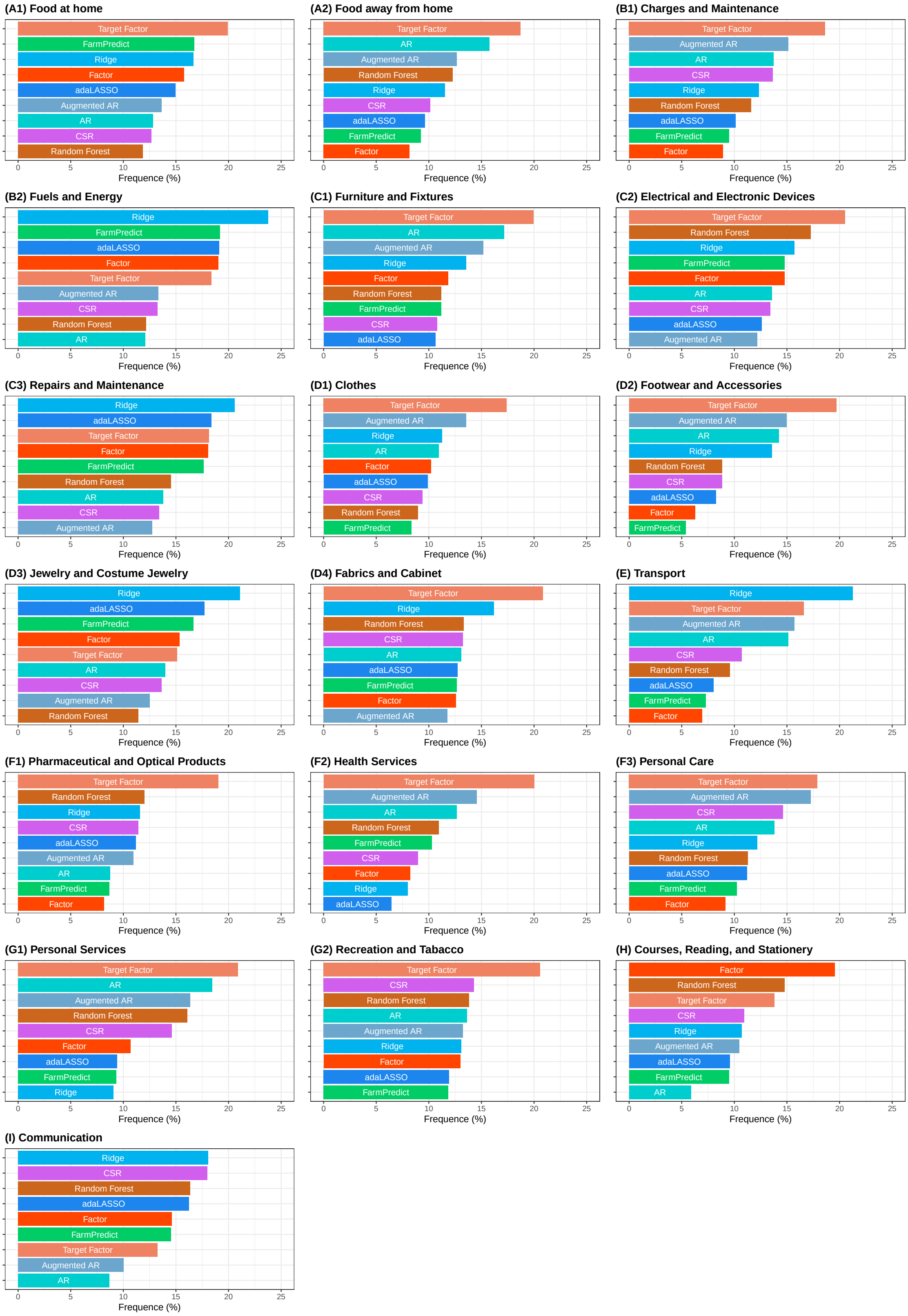


Figure 1.D.4: Frequency each model attained the least forecast squared error for IBGE subgroups, by disaggregate and stacking the horizons (%)



2
Time-varying bias-corrected average forecast

Abstract. We propose a bias correction for the average of a set of individual inflation expectations considering the possibility that intercept and slope biases may vary over time. We proceed in two ways. Firstly, we consider estimations based on rolling windows. Secondly, we employ a state-space model to obtain time-varying intercept and slope biases using the recursiveness of the Kalman filter. The latter approach has the advantage of circumventing the choice of the rolling window size. We also proceed with estimations based on expanding windows, a procedure that is close to what has been done in the literature. We achieve good forecast performance for models based on small rolling windows for shorter and intermediate forecast horizons. In turn, a state-space model that includes corrections for intercept and slope biases varying over time tends to perform slightly worse than procedures based on rolling windows.

Keywords: inflation forecasting; combination of forecasts; bias correction; time-varying parameters; Kalman filter.

JEL Codes: C32, C38, C51, C52, C53, E37.

2.1 Introduction

It is unlikely that some individual forecasting model will consistently outperform all others over time and in different economic conditions since models likely present inaccuracies in their specifications (Elliott & Timmermann, 2016). The aim of combining forecasts is to reduce uncertainty and increase the accuracy of the last forecast. Bates & Granger (1969) show that the combination of *unbiased* forecasts could yield lower mean-squared error than either of the original forecasts. More specifically, they demonstrate that two unbiased forecasts can be optimally combined to obtain a variance no greater than the smaller of the two individual variances. Reid (1969) extends this approach by proposing the combination of more than two unbiased forecasts, and Newbold & Granger (1974) achieve good results by applying this procedure. Granger & Ramanathan (1984) add an intercept in a linear model that combines *potentially biased* individual forecasts. However, there is a challenge when we deal with a large number of individual forecasts and few time observations. The estimation uncertainty could compromise the results, or we might not even be able to estimate the model if we have more individual forecasts than observations.

A solution adopted in the literature to combine multiple potentially biased forecasts avoiding over-parameterization has been to use the average forecast directly. After decomposing individual forecast errors into time-fixed forecast bias, time-varying aggregate uncertainty of the forecasts, and idiosyncratic terms, Palm & Zellner (1992) average equations over cross-sectional dimension and obtains a combination that corrects for additive (intercept) bias. Using panel-data sequential asymptotics, Issler & Lima (2009) show that a bias-correction average forecast (BCAF) obtained via average forecast error is equivalent to the conditional expectation and has an optimal limiting mean-squared error. However, if there is not only additive but also multiplicative (slope) bias, the correction only for intercept bias (BCAF) is no longer optimal. Capistrán & Timmermann (2009) considers both additive and multiplicative bias for bias-adjusting the equal-weighted forecast. Gaglianone & Issler (2023) propose an extended bias-corrected average forecast (EBCAF), which also considers the correction for both types of biases. Additionally, they highlight the implications of the existence of public and private information for the combination of forecasts.

We propose time-varying (extended) bias-corrected average forecast models (TV-BCAF and TV-EBCAF), which account for intercept and slope biases varying over time. Initially, we consider some time variation by means of OLS estimation based on rolling windows. We apply the procedure to different sets of individual forecasts for Brazilian inflation, including the median of available inflation expectations from the Focus survey (the Focus consensus) and forecasts generated by models discussed in Chapter 1. We find that rolling-window-based models perform well, particularly for short windows when forecasting inflation at shorter and intermediate horizons (one to six months ahead). We then propose a state-space model that uses the recursiveness of the Kalman filter to obtain time-varying intercept and slope biases. This approach avoids the need for a discretionary choice of the rolling window size. Overall, models based on the Kalman filter including both types of biases perform slightly worse than some rolling window-based procedures. However, we encourage exploring variations in the specification of the state-space models and finding alternatives for reducing the variance of the estimated time-varying biases.

Outline. This chapter comprises four additional sections following this Introduction. Section 2.2 outlines some extant procedures in the literature for correcting biases and presents their results. Section 2.3 describes models considering time-varying bias correction for both intercept and slope. In Section 2.4, we present and analyze the forecast performances of models that incorporate time-varying bias, while comparing them to traditional combination methods. Section 2.5 brings the final considerations of this chapter. Appendix 2.A provides a brief description of the Kalman filter while Appendix 2.B contains figures displaying the temporal evolution of intercept bias estimated via different approaches.

2.2 Methodologies for combining forecasts

For a given period T and a forecast horizon of h months ahead, consider a set of N inflation forecasts designated by $\{\widehat{\pi}_{i,T+h|T} : i = 1, \ldots, N\}$. Each individual forecast is generated by a model or comes from a survey of experts, for example. For simplicity, we will treat all original forecasts as coming from an unknown model. For a survey of forecasts, this is natural. For projections originating from estimated models, this is a strong simplification. However, it does away with the difficulties inherent in considering a combination of *models* rather than a combination of *forecasts* whose generating process is unknown. Furthermore, the practitioner wants to know whether a combination of forecasts generates empirically satisfactory results.

2.2.1
Average forecast

The simplest way to combine forecasts is to compute the average of the available forecasts. Thus, we compute

$$\widehat{\pi}^{\text{av}}_{T+h\,|\,T} = \frac{1}{N}\sum_{i=1}^{N}\widehat{\pi}_{i,\,T+h\,|\,T}. \tag{2.1}$$

The main advantage of this approach is that it does not require the estimation of weights for available forecasts, which would require training-sample observations. In this setup, the big challenge would be to obtain stable weights over time since we have a small sample of individual forecasts. On the other hand, in the average forecast, the weights assigned to individual forecasts are equal to $\frac{1}{N}$. However, one of the main drawbacks of the average forecast is that assigning the same weight to inaccurate individual forecasts can lead to a sub-optimal combined forecast. That might be true if the set of individual forecasts includes biased forecasts.

2.2.2
A bias-corrected average forecast (BCAF)

We can make a modification to achieve an unbiased combined forecast, which is particularly useful when we potentially consider biased forecasts within the pool of individual forecasts. As in Issler & Lima (2009), consider $\mathbb{E}_{t-h}(\pi_t) = \mathbb{E}(\pi_t\,|\,\mathcal{F}_{t-h})$, the expectation of π_t conditional to a information set available at $t-h$, $\mathcal{F}_{t-h}$, is an optimal device. Then, the econometrician's aim is to approximate this function.

Two-way decomposition or error-component decomposition. Let us consider that an individual forecast $\widehat{\pi}_{i,t\,|\,t-h}$ aims to approximate $\mathbb{E}_{t-h}(\pi_t)$. Thus, we can define the approximation error as

$$\mathbb{E}_{t-h}(\pi_t) - \widehat{\pi}_{i,t\,|\,t-h} = \delta_i^h + \varepsilon_{i,t}^h, \qquad i = 1,\ldots,N, \tag{2.2}$$

where δ_i^h is the individual model time-invariant bias considering the forecast horizon h, and $\varepsilon_{i,t}^h$ is the individual model error term in approximating $\mathbb{E}_{t-h}(\pi_t)$ with $\mathbb{E}(\widetilde{\varepsilon}_{i,t}^h) = 0$ for all i, t, and h. In turn, consider the error for conditional expectation $\mathbb{E}_{t-h}(\pi_t)$ given by

$$\pi_t - \mathbb{E}_{t-h}(\pi_t) = \theta_t^h \tag{2.3}$$

where θ_t^h is an unpredictable time-component with $\mathbb{E}_{t-h}(\theta_t^h) = 0$ for all t and h. Finally, combining (2.2) and (2.3), we obtain the forecast error

$$\pi_t - \widehat{\pi}_{i,t\,|\,t-h} = \delta_i^h + \theta_t^h + \varepsilon_{i,t}^h. \tag{2.4}$$

The term δ_i^h captures a fixed long-term effect on the forecast generated by a model or survey respondent. The term θ_t^h captures time effects arising from the lack of future information between $t-h$ and t, which equally affects all models or respondents. Finally, the term $\varepsilon_{i,t}^h$ captures idiosyncratic errors that affect individuals differently over time (Issler & Lima, 2009).

Issler and Lima's bias-corrected average forecast. Consider the following assumptions:

(i) δ_i^h, θ_t^h, and $\varepsilon_{i,t}^h$ are independent of each other for all i and t;

(ii) δ_i^h is an identically distributed random variable in the cross-sectional dimensional i with mean δ^h and variance σ_δ^2;

(iii) θ_t^h is a stationary and ergodic MA process of order at most $h-1$ with zero mean and finite variance;

(iv) limited degree of cross-sectional dependence of erros, $\varepsilon_{i,t}^h$.

Under these assumptions, Issler & Lima (2009) show that a bias-corrected average forecast (BCAF) given by

$$\widehat{\pi}_{T+h\,|\,T}^{\text{BCAF}} = \widehat{\delta}^h + \widehat{\pi}_{T+h\,|\,T}^{\text{av}}, \tag{2.5}$$

where

$$\widehat{\delta}^h = \frac{1}{N}\sum_{i=1}^{N}\widehat{\delta}_i^h, \qquad \text{with} \qquad \widehat{\delta}_i^h = \frac{1}{T}\sum_{t=1}^{T}\pi_t - \frac{1}{T}\sum_{t=1}^{T}\widehat{\pi}_{i,t-h\,|\,t}, \quad i = 1,\ldots,N.$$

is an optimal forecasting device. Notice that $\widehat{\pi}_{T+h\,|\,T}^{\text{av}}$ is the average forecast defined in (2.1).

For the sake of convenience, we define the forecast error in the usual way, i.e., as $y - \widehat{y}$, where y is the actual variable and $\widehat{y}$ is a forecast for y. Note that Issler & Lima (2009) proceed in an inverse way, as is usual in the literature on error-component decomposition (e.g. Palm & Zellner, 1992). For the model just presented, this difference is not relevant. The relevant difference, which appears in the presentation of the results in Subsection 2.2.3, is that we employ rolling windows of different sizes (including very short ones) to obtain the intercept bias. The procedure suggested by Issler & Lima (2009) assumes that both N and T diverge, which is not compatible with limiting ourselves to using only part of the sample to obtain the bias. However, we also obtain the intercept bias using data in an extended window, that is, considering all available information, which is closer to Issler and Lima's procedure.

Extended error-component decomposition. Now, consider an extended error-component decomposition in which besides intercept bias, we consider the possibility that there is a slope bias as well. Therefore, the model is able to capture both additive and multiplicative biases. In this setup, a decomposition can be written as

$$\mathbb{E}_{t-h}(\pi_t) - \beta^h\, \widehat{\pi}_{i,t\,|\,t-h} = \delta_i^h + \varepsilon_{i,t}^h, \qquad i = 1, \ldots, N \tag{2.6}$$

$$\pi_t - \mathbb{E}_{t-h}(\pi_t) = \theta_t^h. \tag{2.7}$$

By combining the Equations (2.6) and (2.7), we obtain a forecast error given by

$$\pi_t - \beta^h\, \widehat{\pi}_{i,t\,|\,t-h} = \delta_i^h + \theta_t^h + \varepsilon_{i,t}^h, \qquad i = 1, \ldots, N. \tag{2.8}$$

Averaging (2.8) over the cross-sectional dimension i and solving for π_t yields

$$\pi_t = \delta^h + \beta^h\, \widehat{\pi}_{t\,|\,t-h}^{\text{av}} + u_t^h, \tag{2.9}$$

where $u_t^h = \theta_t^h + \varepsilon_t^h$.

Model (2.9) is exactly a model estimated in Capistrán & Timmermann (2009), basically a bias adjustment of the equal-weight forecast. We can also interpret it as an aggregated version of the Mincer & Zarnowitz's (1969) equation used as a first step to test the rationality of a forecast. In this case, rationality is corroborated under $\delta^h = 0$ and $\beta^h = 1$, for all h. Thus, we can proceed with a test for the hypothesis of the rationality of expectations. Several papers point to violating the rationality hypothesis. Then, a procedure that adjusts both intercept and slope bias can be very useful for obtaining an unbiased final forecast. In this line, for each horizon forecast h, an extended bias-corrected average forecast (EBCAF) is given by

$$\widehat{\pi}_{T+h\,|\,T}^{\text{EBCAF}} = \widehat{\delta}^h + \widehat{\beta}^h\, \widehat{\pi}_{T+h\,|\,T}^{\text{av}}, \tag{2.10}$$

where $\widehat{\delta}^h$ and $\widehat{\beta}^h$ are OLS estimates.

The approach to intercept and slope bias-correction presented here differs from that suggested by Gaglianone & Issler (2023). They also assume that forecasts obey a factor model with an affine structure and derive a model in which the slope coefficient appears by multiplying the conditional expectation of π_t rather than the average forecast. Suggested by Issler & Lima (2009), this approach is used in Gaglianone *et al.* (2017) that, as recommended by Gaglianone & Issler (2023), stacking the models along the different forecast horizons to proceed with the estimation using GMM, which can lead to efficiency gains. Another advance of Gaglianone & Issler (2023) in relation to Issler & Lima (2009) is to show that the procedure is consistent with T diverging and N being kept fixed. Finally, due to the possibility of the existence of private information, Gaglianone and Issler show that the average forecast will be correlated with the error term, so that the aggregate Mincer-Zarnowitz regression should be estimated using instrumental variables. However, since in this essay we use the median of the available Focus and forecasts based on public information, this limits the possibility of endogeneity through the channel raised by the authors. Along these lines, the main theoretical problem with our approach remains the fact that we know the generating process of model-based forecasts.

2.2.3
Results of a pseudo-out-of-sample forecast exercise

Setup. We set January 2015 as the starting point for generating combined forecasts to ensure we have at least 25 time periods available when we use a bias-corrected average forecast (BCAF) based on extending windows since individual forecasts began in January 2012[1] (see Chapter 1). We explore BCAF models implemented with rolling windows of different lengths: 3, 6, 7, 8, 9, 10, 12, and 24 months. However, to avoid redundancy, besides the average forecast, and BCAF and EBCAF models based on expanding windows, we report results for BCAF based on rolling windows of 3, 6, 8, 9, and 12 months. We choose the root mean squared error (RMSE) as the metric to evaluate the accuracy of forecasting results. All out-of-sample RMSE are normalized with respect to the median of the available Focus survey's inflation expectations. We also report the RMSE ratio of the median of the *ex-post* Focus survey's inflation expectations.

[1] Initially, there are a total of 25 periods available for $h = 12$ since the 12-month-ahead forecast for Jan/2015 is calculated in Feb/2014. On the other hand, note that there are 36 available periods for $h = 0$.

Forecast performance. Table 2.1 presents results for the period from January 2015 to June 2022. Forecast performance varies depending on the forecast horizon and the rolling window length. For nowcasting ($h = 0$), no forecast combination outperforms the median of ex-post Focus' inflation expectations (Panel A). For longer horizons ($h \geqslant 6$), the average forecast (Panel B) exhibits the best performances. For intermediate horizons ($1 \leqslant h \leqslant 7$), the BCAF models based on rolling windows of 6, 8, 9, and 12 months (Panels D to G) register good predictive performances. Notably, for this subset of horizons and rolling window sizes, the predictive accuracy for shorter horizons improves as the rolling window size increases.

However, choosing the rolling window size can be challenging for the econometrician. In addition, it is important to note that, despite being numerically superior for some horizons, most of the bias-correction-based combinations results are not statistically significant compared to the predictive performance of the median of available Focus' expectations, according to Diebold-Mariano tests. These forecasts would also not be statistically superior to the average forecast without bias correction. Despite this, the results highlight the importance of considering the possibility of bias in the average forecast for short and intermediate horizons, as well as the possibility that this bias varies over time. The lack of statistical significance may be associated with estimation uncertainty, which is enlarged when considering fewer observations, which suggests a trade-off in defining the rolling window size: a smaller window can better capture variations in time but can increase the estimation instability.

Regarding the set of individual forecasts employed, the set comprising forecasts specifically generated for aggregate inflation, which includes the median of Focus' expectations (displayed in the initial rows of Panels B to I in Table 2.1), yields the most favorable forecast outcomes compared to alternative sets of individual forecasts. In certain instances, the use of indirect forecasts formed through the aggregation of forecasts for BCB disaggregates (namely, administered, tradable, and non-tradable items) produces combined forecasts exhibiting predictive performance that closely approximates those obtained when utilizing the set of direct forecasts to compute the combinations.

Table 2.1: Out-of-sample RMSE ratio for average forecast and BCAF and EBCAF models with respect to the available Focus: Jan/2015 to Jun/2022

Method/Forecasts	$h=0$	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$	$h=6$	$h=7$	$h=8$	$h=9$	$h=10$	$h=11$
A. Survey												
Focus (available)	**1.000**	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	*1.000*	*1.000*
Focus (*ex-post*)	**0.931*****	*0.972****	0.993***	1.001	1.000	1.000	1.000	0.999	0.999*	1.000	*1.001*	*1.001*
B. Average forecast												
Aggregates & Focus	1.215	*0.989*	0.953**	0.971*	0.982	0.970	*0.939***	**0.928*****	**0.930****	**0.935****	**0.968***	**0.986**
BCB & Focus	*1.208*	1.004	0.967	0.989	0.994	0.977	*0.945***	*0.951**	*0.948***	**0.950****	**0.973**	**0.985**
Groups & Focus	1.341	1.049	0.971	0.991	0.984	0.968	0.968	*0.950**	*0.953**	*0.976*	*0.975*	*0.992*
Subgroups & Focus	1.514	1.110	1.026	1.031	1.019	0.998	0.973	*0.960*	*0.977*	*0.996*	1.016	1.026
All & Focus	1.317	1.033	0.974	0.991	0.990	0.972	0.949**	**0.940****	**0.945***	*0.957**	*0.977*	*0.993*
C. Bias-corrected average forecast (BCAF) with rolling windows of 3 months												
Aggregates & Focus	1.340	1.116	1.087	1.096	1.083	1.041	0.982	0.972	0.995	1.047	1.122	1.183
BCB & Focus	1.378	1.132	1.077	1.100	1.067	1.028	0.967	0.967	0.993	1.045	1.118	1.146
Groups & Focus	1.499	1.198	1.106	1.138	1.093	1.049	1.012	0.977	0.984	1.068	1.130	1.163
Subgroups & Focus	1.627	1.240	1.152	1.142	1.097	1.055	1.010	0.979	1.008	1.079	1.161	1.206
All & Focus	1.463	1.175	1.107	1.121	1.084	1.041	0.989	0.972	0.994	1.058	1.130	1.173
D. Bias-corrected average forecast (BCAF) with rolling windows of 6 months												
Aggregates & Focus	1.261	1.029	0.979	0.969	*0.957*	*0.943*	**0.932**	*0.961*	0.994	1.040	1.096	1.135
BCB & Focus	1.283	1.056	0.974	0.977	*0.953*	**0.936**	**0.931**	0.977	1.009	1.049	1.106	1.116
Groups & Focus	1.407	1.108	1.001	1.006	0.968	**0.938**	*0.944*	0.969	1.008	1.091	1.121	1.136
Subgroups & Focus	1.562	1.157	1.045	1.021	0.983	0.957	0.951	0.974	1.035	1.103	1.157	1.183
All & Focus	1.380	1.091	1.000	0.992	0.962	*0.939*	*0.935*	0.968	1.010	1.069	1.118	1.142
E. Bias-corrected average forecast (BCAF) with rolling windows of 8 months												
Aggregates & Focus	1.235	*0.980*	*0.914**	**0.924**	**0.942**	*0.950*	*0.944*	*0.958*	0.991	1.035	1.089	1.128
BCB & Focus	1.266	1.003	*0.913**	*0.935*	**0.946**	0.955	0.952	0.983	1.012	1.050	1.100	1.114
Groups & Focus	1.380	1.053	0.937	0.958	*0.947*	*0.942*	0.961	0.975	1.016	1.094	1.117	1.139
Subgroups & Focus	1.531	1.098	0.982	0.975	0.965	0.965	0.963	0.976	1.043	1.108	1.157	1.186
All & Focus	1.354	1.035	*0.935*	*0.946*	*0.946*	*0.949*	0.951	0.971	1.014	1.070	1.114	1.141
F. Bias-corrected average forecast (BCAF) with rolling windows of 9 months												
Aggregates & Focus	1.209	**0.958**	**0.904***	**0.924**	*0.955*	0.958	*0.945*	0.964	0.993	1.039	1.093	1.127
BCB & Focus	1.239	*0.979*	**0.907***	*0.936*	0.961	0.966	0.955	0.991	1.015	1.052	1.103	1.116
Groups & Focus	1.355	1.030	*0.925*	*0.955*	*0.954*	*0.949*	0.961	0.983	1.019	1.097	1.120	1.140
Subgroups & Focus	1.500	1.074	0.972	0.973	0.973	0.972	0.963	0.985	1.047	1.113	1.163	1.185
All & Focus	1.327	1.011	*0.925*	*0.945*	0.957	0.958	0.951	0.978	1.016	1.074	1.118	1.141
G. Bias-corrected average forecast (BCAF) with rolling windows of 12 months												
Aggregates & Focus	*1.194*	*0.976*	*0.933*	*0.957*	0.983	0.979	0.968	0.988	1.019	1.058	1.109	1.144
BCB & Focus	1.215	0.995	0.935	0.967	0.991	0.992	0.981	1.016	1.040	1.071	1.122	1.137
Groups & Focus	1.331	1.043	0.953	0.976	0.973	0.967	0.986	1.009	1.044	1.113	1.137	1.158
Subgroups & Focus	1.468	1.084	0.998	0.996	0.994	0.993	0.991	1.012	1.074	1.133	1.179	1.196
All & Focus	1.303	1.025	0.954	0.972	0.982	0.980	0.977	1.004	1.042	1.092	1.135	1.158
H. Bias-corrected average forecast (BCAF) with expanding windows												
Aggregates & Focus	1.217	*0.987*	0.954*	0.972	0.986	0.976	0.953*	*0.948**	*0.957*	*0.969*	*1.005*	*1.025*
BCB & Focus	1.216	1.008	0.974	1.001	1.010	0.996	0.971	0.982	*0.981*	*0.986*	*1.012*	*1.024*
Groups & Focus	1.350	1.057	0.979	1.005	1.002	0.990	0.995	0.982	0.992	1.020	1.020	1.034
Subgroups & Focus	1.520	1.112	1.033	1.045	1.041	1.024	1.003	0.996	1.018	1.045	1.070	1.079
All & Focus	1.324	1.038	0.981	1.003	1.007	0.993	0.975	0.972	*0.982*	1.000	1.023	1.038
I. Extended bias-corrected average forecast (EBCAF) with expanding windows												
Aggregates & Focus	*1.074*	**0.967**	0.967	0.987	1.013	1.009	0.996	0.998	1.016	1.028	1.039	1.090
BCB & Focus	*1.139*	0.990	0.986	1.013	1.026	1.007	0.984	0.998	0.997	*0.996*	1.017	1.046
Groups & Focus	*1.208*	1.046	0.987	1.019	1.019	1.005	1.009	1.005	1.009	1.028	1.025	1.034
Subgroups & Focus	1.302	1.120	1.049	1.045	1.043	1.030	1.002	0.996	1.011	1.031	1.069	1.056
All & Focus	*1.153*	1.023	0.995	1.017	1.027	1.013	0.998	1.004	1.011	1.026	1.046	1.069

Notes: ***, **, and * indicate that for a specific forecast horizon, a forecast combination "comb" performed statistically better than the median of the available Focus at 1, 5, and 10% significance levels in a one-tailed Diebold-Mariano test with $\mathbb{H}_0 : \text{MSE}\big(\widehat{\pi}^{\text{comb}}_{t+h\,|\,t}\big) = \text{MSE}\big(\pi^{\text{Focus}}_{t+h\,|\,t}\big)$ *versus* $\mathbb{H}_1 : \text{MSE}\big(\widehat{\pi}^{\text{comb}}_{t+h\,|\,t}\big) < \text{MSE}\big(\pi^{\text{Focus}}_{t+h\,|\,t}\big)$. The two values highlighted in bold blue indicate the best and second-best methods for each horizon in terms of out-of-sample RMSE, while the six values in blue italics indicate the third- to eighth-best methods. All sets of individual forecasts included the median of the Focus survey's inflation expectations.

2.3 Time-varying bias correction for the average forecast

Analyzing the Brazilian Focus survey of forecasts for inflation, Carvalho & Minella (2012) note that there is empirical evidence for the existence of common forecast errors prevailing over idiosyncratic components among respondents. Moreover, they highlight the influence exerted by top-performing forecasters on other respondents, indicating a contamination phenomenon known as the epidemiology of the survey forecasts. Beyond that, the results of the preceding section reveal performance disparities in the bias-corrected average forecast when considering different sizes of rolling windows for bias estimation, as well as variations across forecast horizons. This prompts the hypothesis that biases vary over time, potentially possessing a certain degree of predictability. To verify this proposition, we introduce time-varying terms into the decomposition of individual forecasts. For the sake of simplicity and in line with Carvalho & Minella (2012), we assume that this term, which engenders temporal variations in biases, is common to all respondents and models. In practice, since we focus on the average forecast, questioning the adequacy of this hypothesis becomes a secondary concern.

2.3.1 Including a time-varying intercept bias

Error-component decomposition with time-varying intercept bias. Consider a new decomposition given by

$$\mathbb{E}_{t-h}(\pi_t) - \widehat{\pi}_{i,t\,|\,t-h} = \delta_i^h + \mu_t^h + \varepsilon_{i,t}^h, \qquad i = 1, \ldots, N \tag{2.11}$$

$$\mu_t^h = \mu_{t-h}^h + \nu_t^h \tag{2.12}$$

$$\pi_t - \mathbb{E}_{t-h}(\pi_t) = \theta_t^h \tag{2.13}$$

where we add a time-varying term μ_t common to all respondents or models in the individual forecast decomposition (Equation 2.11). Equation (2.12) indicates that this common term follows a random walk process – a common assumption in the literature on time-varying parameters –, and error term ν_t^h independent and identically distributed with zero mean. Lastly, notice that the decomposition of the conditional expectation using information available up to $t-h$ (Equation 2.13) is identical to the former (Equation 2.3). By combining the Equations (2.11) and (2.13), we obtain a forecast error given by

$$\pi_t - \widehat{\pi}_{i,t\,|\,t-h} = \delta_i^h + \mu_t^h + \theta_t^h + \varepsilon_{i,t}^h, \qquad i = 1, \ldots, N. \tag{2.14}$$

State-space representation. Averaging Equation (2.14) over i, we obtain a forecast error given by

$$\pi_t - \widehat{\pi}^{\text{av}}_{t\,|\,t-h} = \delta^h + \mu_t^h + \theta_t^h + \varepsilon_t^h, \tag{2.15}$$

where $\widehat{\pi}^{\text{av}}_{t\,|\,t-h} = \frac{1}{N}\sum_{i=1}^{N} \widehat{\pi}_{i,t\,|\,t-h}$, such as defined in Equation (2.1), $\delta^h = \frac{1}{N}\sum_{i=1}^{N} \delta_i^h$, and $\varepsilon_t^h = \frac{1}{N}\sum_{i=1}^{N} \varepsilon_{i,t}^h$. By defining $\alpha_t^h = \delta^h + \mu_t^h$ and $u_t^h = \theta_t^h + \varepsilon_t^h$, and isolating π_t, we can rewrite Equations (2.15) and (2.12) in a state-space representation given by

$$\pi_t = \alpha_t^h + \widehat{\pi}^{\text{av}}_{t\,|\,t-h} + u_t^h \tag{2.16}$$

$$\alpha_t^h = \alpha_{t-h}^h + \nu_t^h \tag{2.17}$$

Notice that α_t^h is a time-varying average intercept bias, and u_t^h and ν_t^h are error terms.

For identification purposes, we complete the state-space representation assuming the following distribution for disturbances u_t^h and ν_t^h:

$$\begin{pmatrix} u_t^h \\ \nu_t^h \end{pmatrix} \sim \mathcal{N}\left(\mathbf{0},\ \mathsf{diag}\left(\sigma_u^2,\ \sigma_\nu^2\right)\right), \tag{2.18}$$

where we omitted the indication of the forecast horizon h in the variances.

Time-varying BCAF (TV-BCAF). Considering a sample with T temporal observations, we can estimate the system composed by the Equations (2.16), (2.17), and (2.18) by maximum likelihood using the Kalman filter recursion. A time-varying bias-corrected average forecast is given by

$$\widehat{\pi}^{\text{TV-BCAF}}_{T+h\,|\,T} = \widehat{\alpha}^h_{T+h\,|\,T} + \widehat{\pi}^{\text{av}}_{T+h\,|\,T},$$

where $\widehat{\alpha}^h_{T+h\,|\,T}$ is a *predict value* for state variable α^h_{T+h}, recovered using the Kalman filter.

2.3.2
Adding a time-varying slope bias

Extended error-component decomposition with time-varying intercept and slope bias. Lastly, consider a full error-component decomposition in which besides time-varying intercept bias, we consider the possibility that there is a time-varying slope bias as well. Therefore, the model is able to capture both additive and multiplicative bias. In this setup, the decomposition system is

$$\mathbb{E}_{t-h}(\pi_t) - \beta_t^h \, \widehat{\pi}_{i,t\,|\,t-h} = \delta_i^h + \mu_t^h + \varepsilon_{i,t}^h, \qquad i = 1, \ldots, N \tag{2.19}$$

$$\mu_t^h = \mu_{t-h}^h + \nu_t \tag{2.20}$$

$$\beta_t^h = \beta_{t-h}^h + \eta_t \tag{2.21}$$

$$\pi_t - \mathbb{E}_{t-h}(\pi_t) = \theta_t^h \tag{2.22}$$

where ν_t and η_t are independent error terms. Notice that we already assume, by parsimony, that both time-varying intercept and slope bias follow random walk processes.

By combining the Equations (2.19) and (2.22), we obtain a forecast error

$$\pi_t - \beta_t^h \, \widehat{\pi}_{i,t\,|\,t-h} = \delta_i^h + \mu_t^h + \theta_t^h + \varepsilon_{i,t}^h, \qquad i = 1, \ldots, N. \tag{2.23}$$

State-space representation. Averaging Equation (2.23) over the cross-sectional dimension i and solving for π_t yields

$$\pi_t = \delta^h + \mu_t^h + \beta_t^h \, \widehat{\pi}_{t\,|\,t-h}^{\text{av}} + \theta_t^h + \varepsilon_t^h, \tag{2.24}$$

By defining $\alpha_t^h = \delta^h + \mu_t^h$ and $u_t^h = \theta_t^h + \varepsilon_t^h$, for $t = 1, \ldots, T$, we can write a state-space model given by

$$\begin{aligned}
\pi_t &= \alpha_t^h + \beta_t^h \, \widehat{\pi}_{t\,|\,t-h}^{\text{av}} + u_t^h \\
\alpha_t^h &= \alpha_{t-h}^h + \nu_t^h \\
\beta_t^h &= \beta_{t-h}^h + \eta_t^h \\
\begin{pmatrix} u_t^h \\ \nu_t^h \\ \eta_t^h \end{pmatrix} &\sim \mathcal{N}\left(\mathbf{0},\ \text{diag}\left(\sigma_u^2, \sigma_\nu^2, \sigma_\eta^2\right)\right),
\end{aligned}$$

where, likewise as before, we assume a multivariate normal distribution for (independent) disturbances u_t^h, ν_t^h, and η_t^h.

Time-varying EBCAF (TV-EBCAF). Just like before, considering a sample for $t = 1, \ldots, T$, we estimate $\{\alpha_t^h, \beta_t^h\}_t$ by maximum likelihood using the Kalman filter recursion. Thus, a version of time-varying extended bias-corrected average forecast is given by

$$\widehat{\pi}_{T+h\,|\,T}^{\text{TV-EBCAF}} = \widehat{\alpha}_{T+h\,|\,T}^{h} + \widehat{\beta}_{T+h\,|\,T}^{h}\,\widehat{\pi}_{T+h\,|\,T}^{\text{av}},$$

where $\widehat{\alpha}_{T+h\,|\,T}^{h}$ and $\widehat{\beta}_{T+h\,|\,T}^{h}$ are *predict values* for state variables α_{T+h}^h and β_{T+h}^h, respectively, both recovered using the Kalman filter.

2.3.3
Estimation

Just like before, we set January 2015 as the starting point for generating combined forecasts for all horizons. The sample for estimation begins in January 2012. To ensure the positivity of the error term variances, we impose an exponential form given by $\sigma^2 = \exp(\tau)$, where τ is a parameter to be estimated by maximum likelihood. In the TV-BCAF and TV-EBCAF models, there are two and three parameters to assess, respectively. To initialize the maximum likelihood estimator, we consider the OLS estimates generated from linear models with time-fixed parameters and the first 36 available observations, when applicable.[2] Thus, the initial values of the time-varying parameters (TPVs) are the OLS estimates for the coefficients of this initial stage. The starting variance-covariance matrix of TVPs corresponds to the conventional estimator for the variance-covariance matrix of parameters in the initial model. For the variance of error terms of the measurement equation, we use the variance of residuals from the initial model estimated by OLS. Finally, we set the initial variance of a state equation error to be $0.2^2 = 0.04$. We employ a quasi-Newton method (BFGS) as the optimization algorithm.

2.4
Results and discussion

Table 2.2 exhibits results of survey-based expectations, average forecast, BCAF based on both 9-month rolling windows and expanding windows, EBCAF based on expanding windows, and TV-BCAF and TV-EBCAF models for the period from January 2015 to June 2022. Figure 2.1 shows the evolution over

[2] As explained in the Footnote 1, for early periods, we may have fewer than 36 observations for some horizons. For example, in Jan/2015, only 25 stayed available for horizon $h = 12$. However, starting from Jan/2016, there are 36 starting observations available for all horizons.

time and by forecast horizon of actual inflation and forecasts resulting from some combination of forecasts considering the set of individual forecasts for the aggregate inflation (direct forecasting approach), including the available Focus consensus.

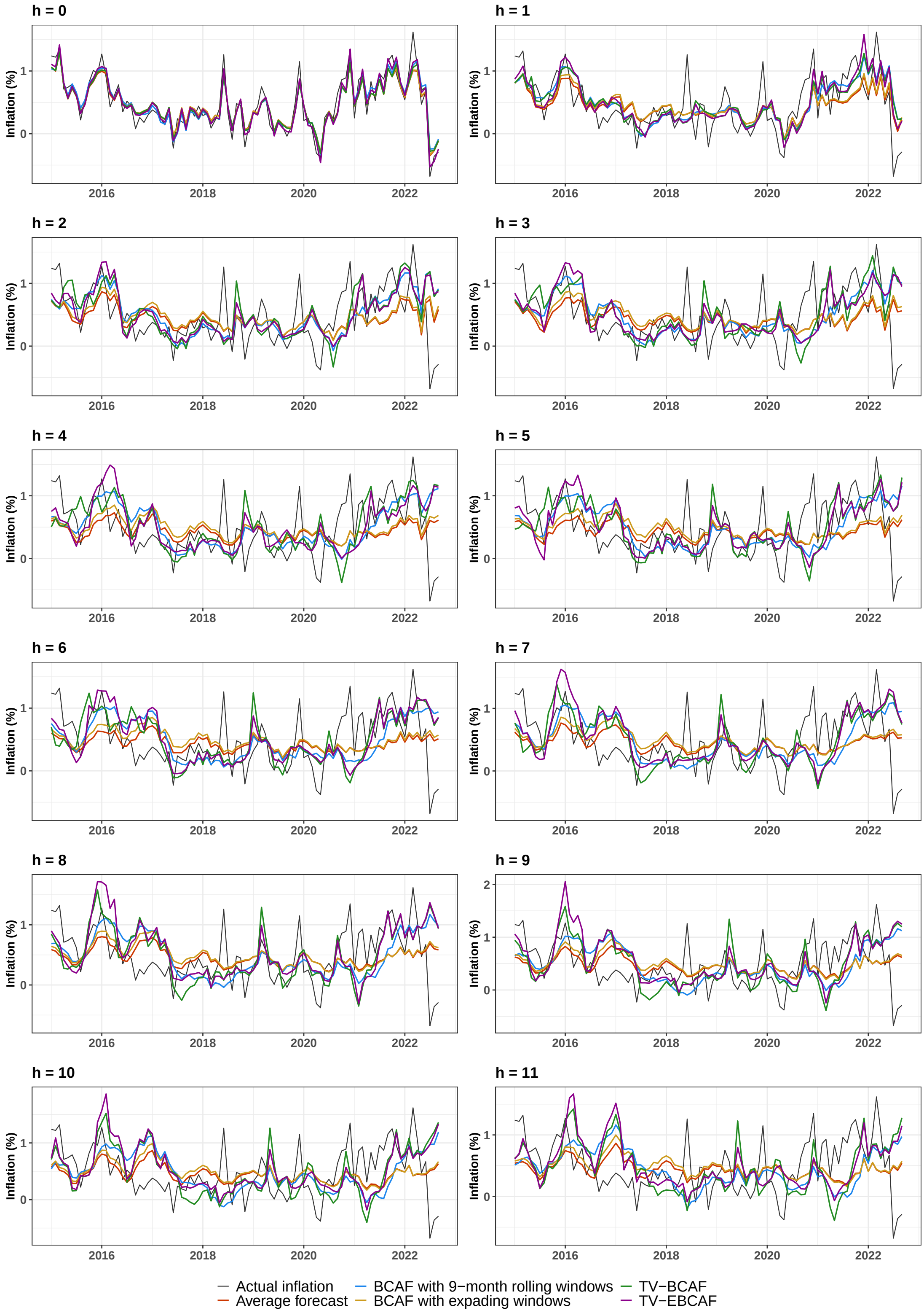
Firstly, it can be observed that the TV-BCAF models, which only include time-varying intercept bias (no slope bias), do not perform well with respect to their counterparts based on 9-month rolling windows or expanding windows and average forecast: almost all RMSE ratios are higher across forecast horizons and individual forecast sets addressed. Looking at the evolution of the forecasts obtained by the TV-BCAF, these models basically extrapolate the present forecast error into the future. The forecast error tends to propagate farther into the future as the forecast horizon extends.

Table 2.2: Out-of-sample RMSE ratio for average forecast and time-varying BCAF and EBCAF models with respect to available Focus: Jan/2015 to Jun/2022

Method/Forecasts	$h=0$	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$	$h=6$	$h=7$	$h=8$	$h=9$	$h=10$	$h=11$
A. Survey												
Focus (available)	**1.000**	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	*1.000*	*1.000*
Focus (*ex-post*)	**0.931*****	*0.972****	0.993***	1.001	1.000	1.000	1.000	0.999	0.999*	1.000	*1.001*	*1.001*
B. Average forecast												
Aggregates & Focus	1.215	*0.989*	*0.953***	*0.971**	*0.982*	*0.970*	**0.939****	**0.928*****	**0.930****	**0.935****	**0.968***	**0.986**
BCB & Focus	1.208	1.004	0.967	0.989	0.994	0.977	*0.945***	*0.951**	*0.948***	**0.950****	**0.973**	**0.985**
Groups & Focus	1.341	1.049	0.971	0.991	*0.984*	*0.968*	0.968	*0.950**	*0.953**	*0.976*	*0.975*	*0.992*
Subgroups & Focus	1.514	1.110	1.026	1.031	1.019	0.998	0.973	*0.960*	0.977	*0.996*	1.016	1.026
All & Focus	1.317	1.033	0.974	0.991	0.990	0.972	*0.949***	**0.940****	**0.945***	*0.957**	*0.977*	*0.993*
C. Bias-corrected average forecast (BCAF) with rolling windows of 9 months												
Aggregates & Focus	1.209	**0.958**	**0.904***	**0.924**	**0.955**	*0.958*	**0.945**	0.964	0.993	1.039	1.093	1.127
BCB & Focus	1.239	*0.979*	**0.907***	**0.936**	*0.961*	*0.966*	0.955	0.991	1.015	1.052	1.103	1.116
Groups & Focus	1.355	1.030	*0.925*	*0.955*	**0.954**	**0.949**	0.961	0.983	1.019	1.097	1.120	1.140
Subgroups & Focus	1.500	1.074	0.972	*0.973*	*0.973*	0.972	0.963	0.985	1.047	1.113	1.163	1.185
All & Focus	1.327	1.011	*0.925*	*0.945*	*0.957*	*0.958*	*0.951*	0.978	1.016	1.074	1.118	1.141
D. Bias-corrected average forecast (BCAF) with expanding windows												
Aggregates & Focus	1.217	*0.987*	*0.954**	*0.972*	0.986	0.976	*0.953**	*0.948**	*0.957*	*0.969*	*1.005*	*1.025*
BCB & Focus	1.216	1.008	0.974	1.001	1.010	0.996	0.971	0.982	0.981	*0.986*	*1.012*	*1.024*
Groups & Focus	1.350	1.057	0.979	1.005	1.002	0.990	0.995	0.982	0.992	1.020	1.020	1.034
Subgroups & Focus	1.520	1.112	1.033	1.045	1.041	1.024	1.003	0.996	1.018	1.045	1.070	1.079
All & Focus	1.324	1.038	0.981	1.003	1.007	0.993	0.975	0.972	0.982	1.000	1.023	1.038
E. Extended bias-corrected average forecast (EBCAF) with expanding windows												
Aggregates & Focus	*1.074*	**0.967**	*0.967*	0.987	1.013	1.009	0.996	0.998	1.016	1.028	1.039	1.090
BCB & Focus	*1.139*	*0.990*	0.986	1.013	1.026	1.007	0.984	0.998	0.997	*0.996*	1.017	1.046
Groups & Focus	1.208	1.046	0.987	1.019	1.019	1.005	1.009	1.005	1.009	1.028	1.025	1.034
Subgroups & Focus	1.302	1.120	1.049	1.045	1.043	1.030	1.002	0.996	1.011	1.031	1.069	1.056
All & Focus	*1.153*	1.023	0.995	1.017	1.027	1.013	0.998	1.004	1.011	1.026	1.046	1.069
F. Time-varying bias-corrected average forecast (TV-BCAF)												
Aggregates & Focus	1.233	1.005	1.018	1.040	1.070	1.049	0.995	0.989	1.001	1.052	1.121	1.184
BCB & Focus	1.226	1.026	1.005	1.034	1.040	1.018	*0.953*	*0.959*	*0.975*	1.030	1.104	1.134
Groups & Focus	1.364	1.103	1.044	1.079	1.067	1.019	1.004	0.991	0.975	1.056	1.110	1.149
Subgroups & Focus	1.512	1.157	1.122	1.114	1.096	1.025	0.993	0.993	0.996	1.068	1.142	1.194
All & Focus	1.336	1.079	1.062	1.072	1.071	1.026	0.987	0.987	0.986	1.049	1.120	1.166
G. Time-varying extended bias-corrected average forecast (TV-EBCAF)												
Aggregates & Focus	*1.087*	*0.998*	*0.948*	*0.973*	1.009	0.977	0.957	0.970	1.016	1.053	1.084	1.146
BCB & Focus	*1.145*	1.000	0.972	0.984	1.018	1.013	*0.947*	*0.960*	0.981	1.019	1.099	1.123
Groups & Focus	1.211	1.086	0.985	1.024	0.993	*0.964*	0.970	0.968	*0.969*	1.018	1.088	1.133
Subgroups & Focus	1.330	1.135	1.044	1.087	1.066	**0.952**	0.997	1.000	*0.964*	1.013	1.100	1.145
All & Focus	*1.156*	1.046	0.986	0.989	*0.979*	0.971	0.956	0.973	0.988	1.037	1.089	1.163

Notes: see Table 2.1.

Figure 2.1: Forecasts generated by selected combinations considering the set of individual forecasts for aggregate directly, by horizon



For shorter and intermediate horizons, the inclusion of the time-varying slope bias correction (TV-EBCAF model) produces improvements compared to the model that solely incorporates a time-varying intercept bias. Although the

TV-EBCAF model's numerical performance falls slightly behind the model that corrects only for intercept bias using a 9-month estimation window, there are marginal improvements when compared to the EBCAF model based on extended windows. This outcome suggests that a full correction for time-varying biases can be advantageous for inflation forecasting. When comparing the forecasts generated by TV-BCAF and TV-EBCAF models, we observe that the latter carries fewer ex-post errors forward, except for a noticeable overprediction of inflation in 2016 and 2017. It is worth noting that for more distant horizons, the prominent positive bias, in early 2019, and negative bias, in early 2020, are mitigated to some extent in the case of the TV-EBCAF model compared to the TV-BCAF model. This highlights the importance of incorporating a slope bias in the time-varying bias-correction approach.

By assuming a random walk process for the time-varying bias, we expect the forward loading of ex-post forecast errors. In this regard, we are investigating alternative specifications that incorporate an autoregressive (AR) model with a non-zero intercept and an AR term smaller than one for the states. However, the results obtained so far are not satisfactory. We are considering specifications that deviate from the traditional local-level model or assumptions regarding the random walk process for latent variables. It is important to preserve parsimony in specifying a state-space model since the estimation of these models may suffer from severe instabilities, along with difficulties associated with identification.

Finally, an important finding of our essay is the lack of statistical superiority of most forecasts generated by bias correction models. Although, on average, some models present a lower RMSE ratio than the average forecast, traditional models with fixed biases or models based on time-varying bias hardly demonstrate statistically superior performance compared to the Focus consensus. This outcome may be attributed to various factors. Firstly, the improved performance of bias-corrected procedures may be closely linked to the period of the COVID-19 pandemic, characterized by higher and more volatile inflation. Secondly, the uncertainty associated with parameter estimation in the less parsimonious corrected-bias models could offer another potential explanation. More specifically, models incorporating time-varying bias display a notable divergence of estimators for these latent variables, as evidenced by the variance-covariance matrix of the states. In practical terms, this instability in estimation over time leads to the occurrence of atypical forecasts, thereby hindering the achievement of statistical significance.

2.5 Final considerations

In this essay, we introduce a model for correcting the bias in the average forecast, which allows both intercept and slope biases to vary over time. Initially, we proceed with an estimation based on rolling windows of different lengths. Such models allow the parameters to oscillate over time to a greater degree than the variation that occurs in a procedure based on extended windows. Applying the procedure to different sets of individual inflation forecasts, we find good predictive results for these rolling-window-based models, particularly for windows ranging from 6 to 12 months in the case of intermediate forecast horizons (one to six months ahead). Based on this result, we suggest a state-space model that allows for obtaining time-varying bias components using all available information, that is, without the need to define *ad-hoc* a window size. Overall, the model that includes corrections for intercept and slope bias varying over time tends to perform slightly worse than rolling-window-based procedures. However, it is worth investigating other specifications for the state-space model and alternatives for reducing the variance of the estimated time-varying biases.

2.A Kalman filter

Following Hamilton (1994, Chapter 13) and Elliott & Timmermann (2016, Appendix A), let $\boldsymbol{y}_t$ be a n-dimensional vector of (observable) variable observed at period t, and $\boldsymbol{\xi}$ be a r-dimensional vector of state (unobservable) variables. Consider a generic state-space model consisting of a measurement equation, a state equation, and a perturbation equation as follows:

$$\boldsymbol{y}_t = \boldsymbol{H}\,\boldsymbol{\xi}_t + \boldsymbol{\varepsilon}_t$$

$$\boldsymbol{\xi}_t = \boldsymbol{F}\,\boldsymbol{\xi}_{t-h} + \boldsymbol{\nu}_t$$

$$\begin{pmatrix} \boldsymbol{\varepsilon}_t \\ \boldsymbol{\nu}_t \end{pmatrix} \sim \mathcal{N}\left(\boldsymbol{0}, \begin{bmatrix} \boldsymbol{Q} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{P} \end{bmatrix}\right),$$

where $\boldsymbol{H}$ and $\boldsymbol{F}$ are matrices of parameters, $\boldsymbol{Q}$ and $\boldsymbol{P}$ are covariance matrices, and $\boldsymbol{\varepsilon}_t$ and $\boldsymbol{\nu}_t$ are independent vectors of white noise. Given starting values $\widehat{\boldsymbol{\xi}}_{h\,|\,0}$ and $\boldsymbol{P}_{h\,|\,0}$, we recursively compute the following values until we get to $\widehat{\boldsymbol{\xi}}_{t+h\,|\,t}$ and $\boldsymbol{P}_{t+h\,|\,t}$, and then compute $\widehat{\boldsymbol{y}}_{t+h\,|\,t}$. We estimate the parameters in the matrices $\boldsymbol{F}$, $\boldsymbol{H}$, $\boldsymbol{Q}$, and $\boldsymbol{P}$ employing the maximum likelihood estimator combined with the recursiveness of the Kalman filter.

2.B
Forecast error and estimated intercept bias

Figure 2.B.1: Forecast error and estimate intercept bias over time: from $h = 0$ to $h = 3$, by set of forecasts

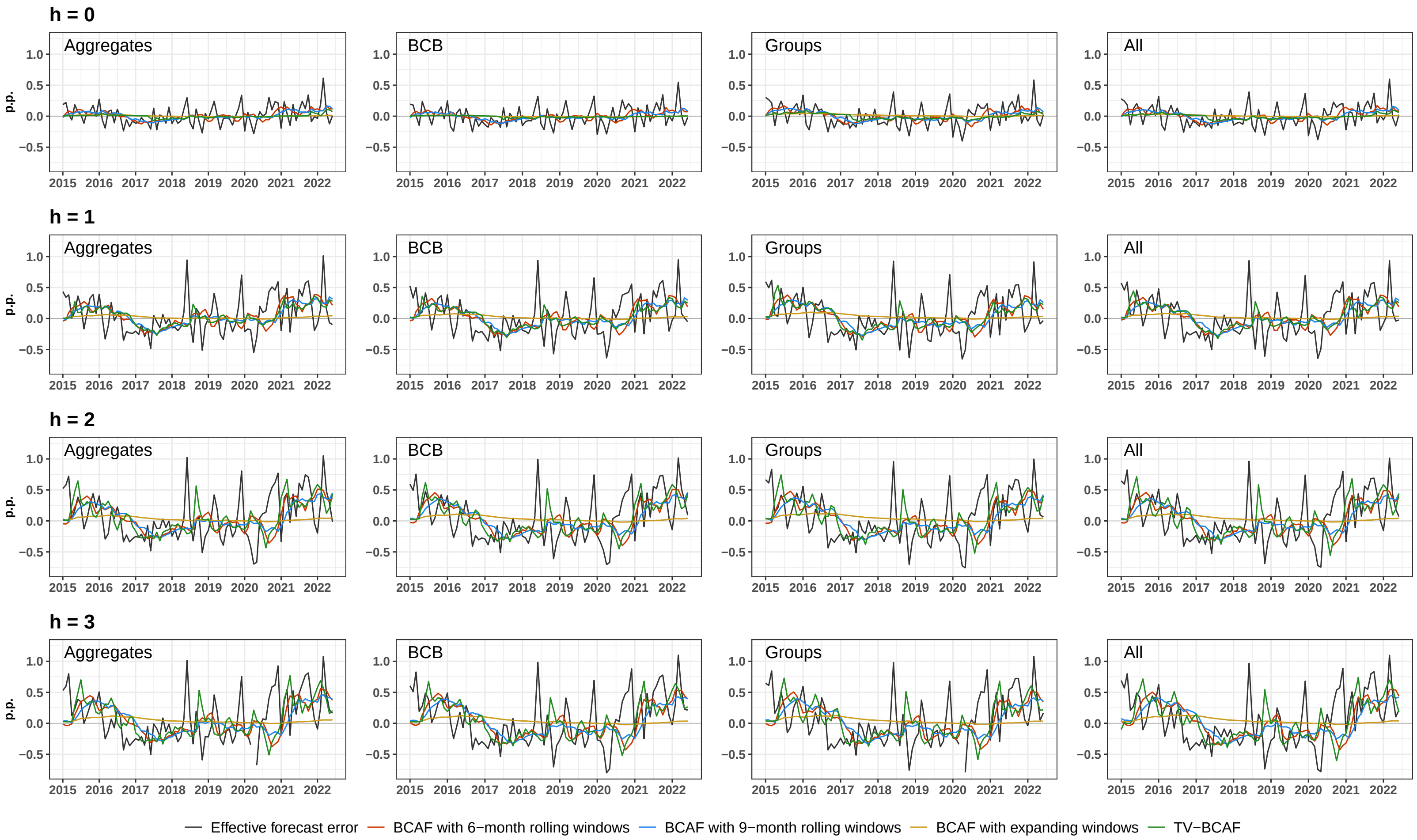


Figure 2.B.2: Forecast error and estimate intercept bias over time: from $h = 4$ to $h = 7$, by set of forecasts

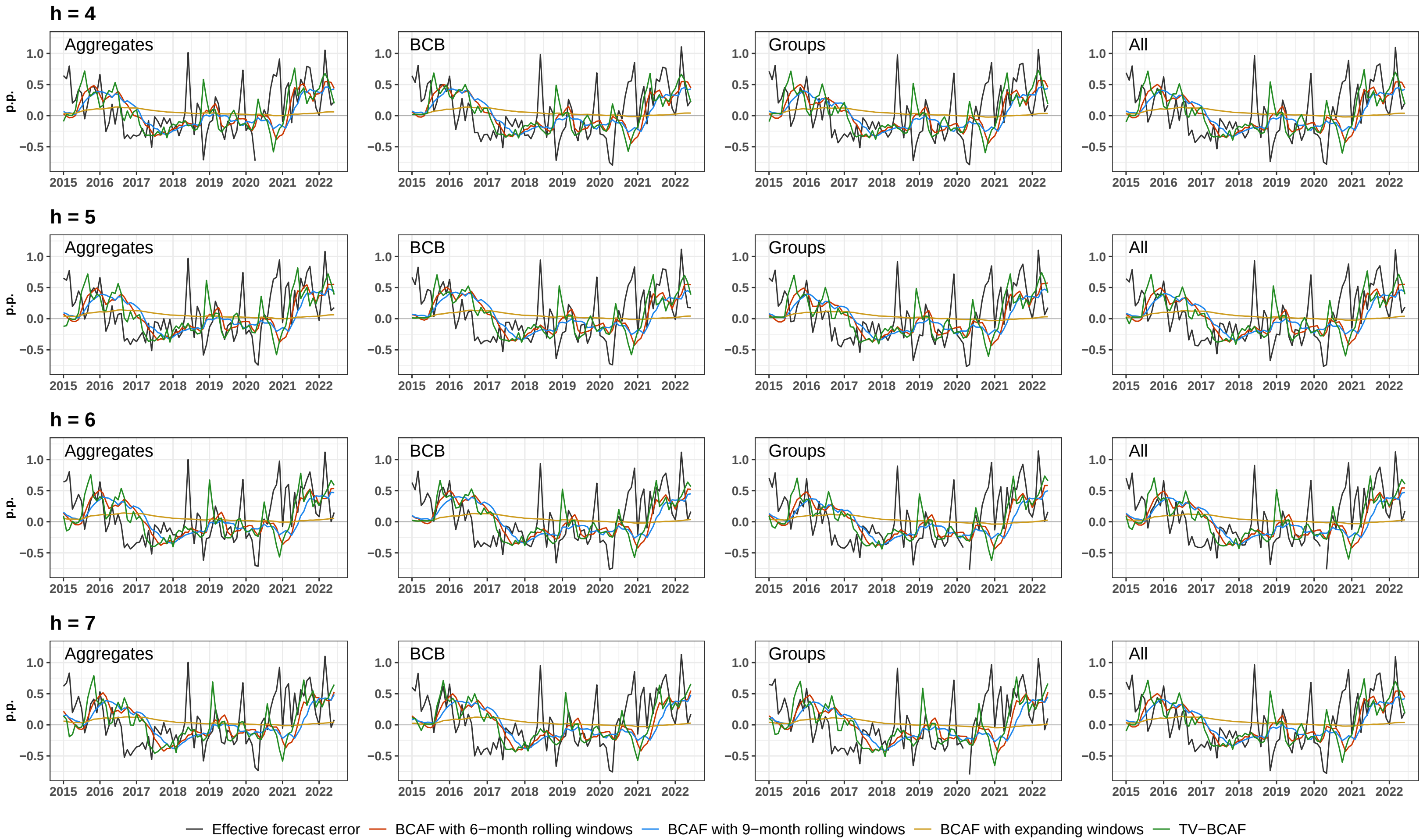
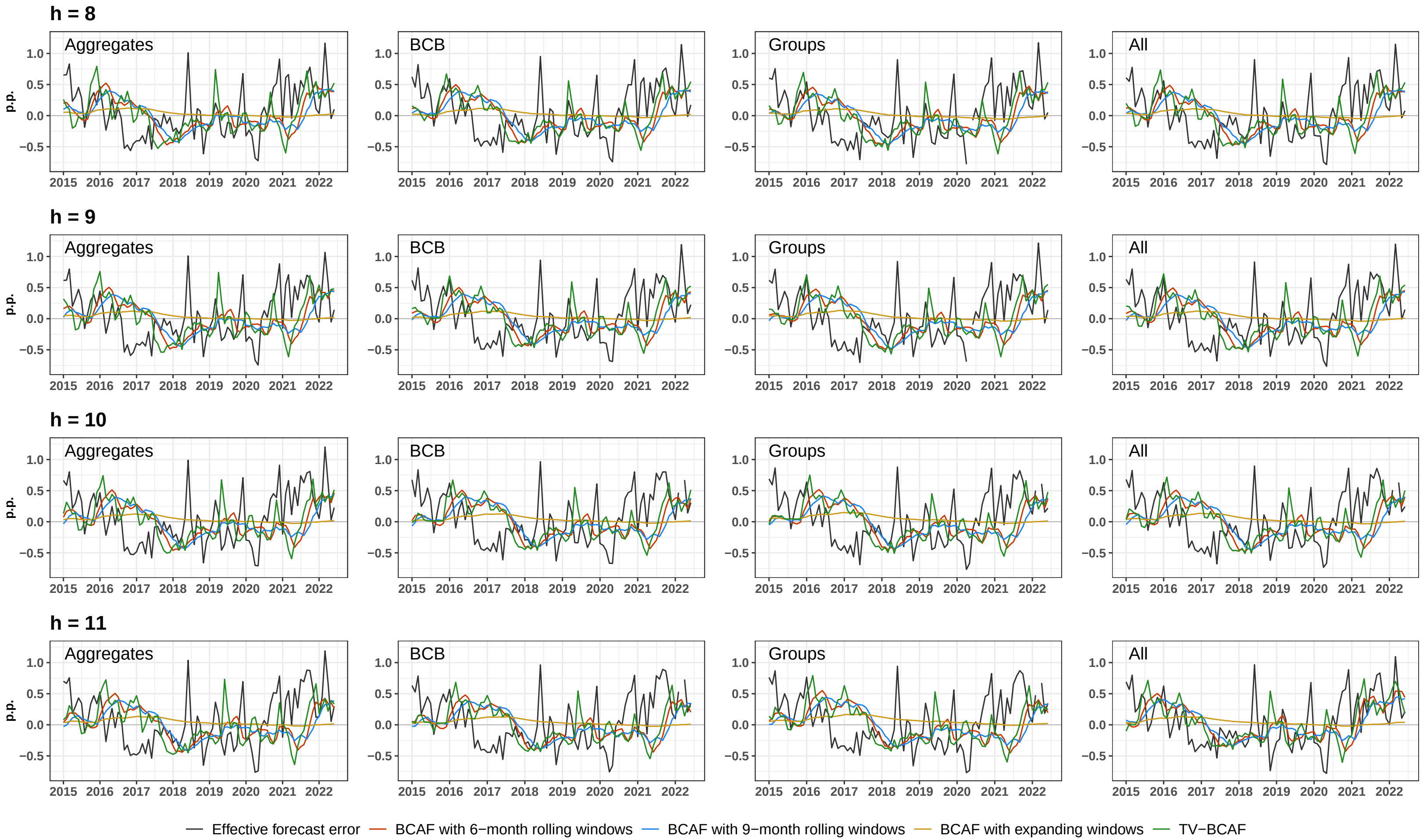


Figure 2.B.3: Forecast error and estimate intercept bias over time: from $h = 8$ to $h = 11$, by set of forecasts



3
What news and social media tell us about future inflation?

Abstract. We construct forward-looking indexes for inflation based on tweets and newspaper articles employing a supervised machine-learning approach. Using Brazilian data, we verify that the news-based indexes are able to anticipate long-term trends as well as capture short-term movements of the accumulated inflation over 3, 6, and 12 months ahead at various periods. Furthermore, the proposed indexes could improve inflation forecast performance. More specifically, for short horizons (3 and 6 months ahead), a bias correction model for the median of available survey-based expectations benefits from including news-based indexes. On the other hand, when considering longer-term inflation forecasts (12 months ahead), models that incorporate a large number of predictors can benefit from the inclusion of the indexes. Thus, considering indexes from social media and news sources can improve inflation forecasting. The intuition for the result is that it pays to consider a broader set of information than solely that resulting from survey-based expectations that account only for experts' opinions.

Keywords: inflation forecasting; unstructured data; Twitter; newspapers; elastic net; adaLASSO.

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

3.1 Introduction

Unstructured data are becoming very popular in economic modeling and forecasting. Newspapers and social networks such as Twitter produce a considerable volume of unstructured data that to some extent reflects the information flow. This essay investigates whether indexes constructed from tweets and newspaper articles can help us anticipate future movements in inflation. In particular, inflation forecasting is an old and relevant research topic that presents new perspectives when considering unstructured data. The literature has been expanding by employing both new econometric techniques (Inoue & Kilian, 2008; Garcia *et al.*, 2017; Medeiros *et al.*, 2021) and new databases such as Google Trends (Guzman, 2011; Li *et al.*, 2015; Niesert *et al.*, 2020), newspaper articles (Rambaccussing & Kwiatkowski, 2020; Larsen *et al.*, 2021), and Twitter (Angelico *et al.*, 2022).

Experts write articles and opinion pieces in newspapers about economics, politics, social questions, and the international scene. Several economists, politicians, consumers, and entrepreneurs share their thoughts on social media about inflation, prices, and related topics. Could this information be used to obtain more accurate inflation forecasts than available expectations? This essay aims to address this question and explore whether non-traditional data sources remain relevant and informative even in the presence of several macroeconomic and financial variables commonly used as predictors for inflation. Our application will address the Brazilian case. The Central Bank of Brazil manages the Focus Survey, a daily collection of inflation expectations provided by market specialists in the country. It is challenging to outperform these expectations, especially for shorter forecast horizons (Ang *et al.*, 2007; Faust & Wright, 2013; Garcia *et al.*, 2017).

In this essay, we use a supervised machine learning procedure via the elastic net to construct *forward-looking* indexes for inflation using information gathered from Twitter and newspapers. This procedure can be interpreted as a version of the time-varying dictionary approach (see Lima *et al.*, 2021, for example). The methodology rests on the occurrence counts of terms appearing in tweets or articles, with a broad set of predefined terms collected from Twitter and preselected n-grams from three relevant Brazilian newspapers used for articles. After selecting relevant terms for different inflation horizons employing an elastic net estimator, we construct two versions of indexes from three distinct information sets. The non-standardized version predicts inflation based on the latest available counts. In a standardized version, we divide the previous predicted value by the sum of the absolute values of each term in the linear model. The information sets consist of only Twitter, only newspapers, or both. Throughout the chapter, we detail the advantages and challenges of each version. Finally, in addition to visually verifying the adherence of the indexes to future inflation, we also conduct pseudo-out-of-sample forecasting exercises in which we compare models that include or disregard the indexes. We evaluate a simple historical bias correction model for available survey-based expectation, as well as a data-rich model that incorporates several predictors typically used in inflation forecasting.

Results overview. The news-based indexes are able to anticipate long-term trends and captured short-term movements in 3-, 6-, and 12-month-ahead cumulative inflation at various periods. Considering the benefits in forecasting inflation accumulated over 3 and 6 months ahead, the indexes contribute to a reduction in the root mean squared forecast error (RMSE) of a bias correction for available Focus' inflation expectations. The model including an index based solely on articles achieves the best predictive performance for 3-month cumulative inflation, delivering a reduction of 26% of RMSE compared to the median of the available Focus expectations – the Focus consensus. For 6-month-ahead inflation, the reductions are more modest, ranging from 7% to almost 13%, while the model that does not include any index registers a reduction of only 4%. In turn, for inflation accumulated over 12 months, the inclusion of an index based solely on tweets improves the already good result of a high-dimensional model. More specifically, there is an extra reduction of 11 percentage points in terms of RMSE, totaling almost 50% reduction in this metric compared to the Focus consensus. Our findings indicate that news-based indexes are particularly helpful from the beginning of the COVID-19 pandemic in Brazil, i.e., from 2020 onwards, a period of great economic and social instability.

Literature and contributions. Researchers extensively investigate the predictive power of Central Bank statements in forecasting a wide range of economic variables, including interest rates (Hubert & Labondance, 2021), output growth (Lima *et al.*, 2021), inflation (Dräger *et al.*, 2016), and multiple macroeconomic variables (Lin *et al.*, 2022). They also use newspapers articles to analyze economic fluctuations and growth (Larsen & Thorsrud, 2019; Thorsrud, 2020), inflation and inflation expectations (Larsen *et al.*, 2021), output growth (Martins & Medeiros, 2022), as well as several macroeconomic variables (Rambaccussing & Kwiatkowski, 2020; Kalamara *et al.*, 2022; Barbaglia *et al.*, 2022). Furthermore, using Twitter data, Angelico *et al.* (2022) build a daily indicator of expected inflation for Italy, a country that only possesses a monthly survey-based expectation. The resulting index is a good proxy for daily inflation expectations, with Twitter timely reflecting the beliefs of economic agents.

Similar to Lima *et al.* (2021), we compute indexes and incorporate them into forecast models. In contrast, Kalamara *et al.* (2022) directly employ time series of counts of terms, alongside other predictors, in forecasting. The literature points out the benefits of both approaches in enhancing predictive accuracy. Indexes offer the benefit of expanding possibilities beyond forecasting alone. For instance, practitioners may be interested in identifying patterns, anticipating trends, or detecting turning points. Hence, the use of a news-based index may be useful if it successfully captures relevant and informative aspects. Concerning the construction of the news-based index, a time-varying dictionary approach via supervised machine learning presents the advantage of the simplicity of implementation and interpretation since it involves a procedure with a target variable. These features distinguish this approach from more complex topic modeling techniques, such as those based on Latent Dirichlet Allocation employed by Larsen & Thorsrud (2019), Thorsrud (2020), Larsen *et al.* (2021), and Martins & Medeiros (2022).

We can summarize the main contributions of this essay in the following four points. First, we propose alterations to the time-varying dictionary approach explored by Lima *et al.* (2021) for constructing our indexes for inflation using Twitter and newspapers. Modifications involve an alternative way of computing the indexes that employ the parameter estimates of the linear model used in selecting terms, smoothing through more recent fits as well as normalizing for stability over time. We can naturally consider these changes to obtain indexes for other economic variables. Second, our essay innovates by considering news-based indexes to forecast inflation *directly*, taking the Brazilian case as an application, thus extending the use of such an index compared to Angelico *et al.* (2022), where the aim is obtaining a proxy for inflation expectations. Brazilian inflation expectations from the Focus survey consist of expert opinions, linked to the financial market. By showing that indexes based on tweets and articles help forecast inflation, a third contribution of our essay is to point out that information from a broader audience can be relevant to inflation prediction, as argued by Angelico *et al.* (2022) for Italy, for example. Fourth, we suggest a procedure for dealing with the secular increase in tweets over time to avoid artificially inflating the count of terms independent of the state of affairs.

Outline. This chapter has four more sections in addition to this Introduction. Section 3.2 details the Brazilian case and describes news data as well as the construction of the news-based indexes for inflation. Section 3.3 describes the forecasting methodology employed to evaluate the contribution of the indexes to inflation forecasting. Section 3.4 analyses the adherence of the indexes to future inflation as well as presents and discusses the results of the forecasting exercises. Finally, Section 3.5 concludes. Appendix 3.A explains terms from tweets and other predictors for inflation, while Appendix 3.B describes the adaptive LASSO that we employ for the evaluation of news-based indexes in inflation forecasting.

3.2
News-based indexes for the Brazilian inflation

3.2.1
The Brazilian context and the database for indexes

The Brazilian context. The *Instituto Brasileiro de Geografia e Estatística* (IBGE) computes the official Brazilian consumer price index (IPCA) from which we compute the monthly inflation. The Central Bank of Brazil (BCB) manages the Focus survey, a daily-frequency expectation system based on expert opinion. The Focus survey collects expectations for several variables, including inflation, for multiple horizons. Although this survey has a daily periodicity, the current week's expectations are released to the public by the BCB only at the beginning of next week. Consequently, it is important to differentiate between the *available* Focus, which the econometrician observes when they compute their forecast, and the *ex-post* Focus, which is from the current day but will only be available days later. Thus, it is pertinent to know whether additional information generates more accurate forecasts for inflation at several horizons than Focus-based expectations. Furthermore, a survey-based expectation may reveal information unavailable to the econometrician and include signals not contained in other variables. In this context, it may be useful to use the available expectation as a predictor in a forecast model as well as to control for it to select which variables contribute at the margin to forecasting inflation.

Multi-horizon forecasts. We consider three forecast horizons: inflation accumulated over 3, 6, and 12 months ahead, which we indicate by `inf3m`, `inf6m`, and `inf12m`, respectively. These horizons can be relevant for managing monetary policy as well as pricing and investment make-decision.

Overview of indexes and data from Twitter and newspapers. The news-based indexes considered in this essay are developed in partnership with Vox Radar, a Brazilian technology company focused on monitoring social networks (social listening). We have daily data for both tweets and articles. For Twitter data, we count mentions of various terms related to inflation in all tweets in Portuguese from 2010 onwards, disregarding tweets with terms about other economies such as "europa", "eua" (US), "fed", "alemanha" (Germany), "argentina", among other. The list of terms includes expressions about commodities, employment, exchange rate, expectations concerning prices and inflation, inflationary pressure, interest rates, investment, loans, costs, demand, supply shocks, taxes, and other macroeconomic-related terms. Some terms are similar to those in Angelico *et al.* (2022). We treat the data to control for Twitter usage over time. Twitter experienced substantial growth in recent years. Consequently, there is a secular increase in tweets over time, which, if not accounted for, could artificially inflate the count of terms irrespective of the prevailing economic context. To mitigate this, we construct a series of counts for generic terms such as "oi" (hi), "olá" (hello), "bom dia" (good morning), among others, and normalized each count of inflation-related terms by dividing it by the sum of counts of these generic terms. Table 3.A.1 in Appendix 3.A presents the list of generic terms.

For newspapers, we count n-grams with n up to 3 related to inflation after proceeding with tokenization, cleaning, and lemmatization of the articles obtained from three of the most relevant newspapers in Brazil (*Folha de São Paulo*, *Valor Econômico*, and *Estadão*), as in Martins & Medeiros (2022). Tokenization divides the text into smaller units called tokens, usually comprising words and punctuation. Cleaning involves removing irrelevant elements such as stopwords, rare words, digits, and punctuation. Lemmatization reduces words to their base form. These procedures are widely used in the pre-processing of textual data. After this pre-processing, there are more than 36,000 n-grams. To reduce the universe of terms, we select those n-grams that contain specific words (or parts of words).[1] Although the construction of the index employs a supervised machine learning method, which at first allows us to deal with the problem of dimensionality, including all this information would be counterproductive, besides the fact that many terms do not provide relevant information about future inflation.

[1] List of words (or parts of words), accompanied by the respective translations: "*preço*" (price), "*inflaç*" and "*inflac*" (root for inflation), "*ipca*" (Brazilian consumer price index), "*juro*" (interest), "*selic*" (Brazilian interest rate), "*demanda*" (demand), "*petróleo*" (oil), "*gasolina*" (gasoline), "*bacen*" and "*BC*" (Central Bank), "*commodit*" (root for commodities), "*camb*" and "*câmb*" (root for exchange rate), "*pib*" (GDP), and "*empreg*" (root for employment). We also include the 1-grams "*caged*" (a recording of hiring and firing employees in Brazil) and "*caro*" (expensive).

We smooth the series of counts by applying 132-day moving averages. This moving average aims to mitigate the effects that a great repercussion or unexpected increase of mentions of a certain term could have on obtaining the index. We investigate other sizes of moving averages, but overall, 132 days produce good results. We also apply the transformation $\log(\text{count}_{i,t} + 1)$, where $\text{count}_{i,t}$ is the resulting moving average, with i indexing the n-grams, and t indicating the period. This transformation aims to mitigate possible asymmetries in the distribution of counts. We are now ready to proceed with constructing the indexes for inflation.

3.2.2 Construction of the indexes

Let π_t be the inflation rate at period t. Let $\mathbf{news}_{t-h}$ be a p-dimensional vector of the counts of terms of tweets and n-grams of newspaper articles observed at period $t-h$. The construction of these counts follows the steps described in the previous subsection. At each period t and for each forecast horizon h, we estimate the linear model

$$\pi_t = \mu + \eta \,\text{Focus}^{\text{available}}_{t-h\,|\,t} + \boldsymbol{\phi}\,\mathbf{news}_{t-h} + \varepsilon_t, \tag{3.1}$$

where $\text{Focus}^{\text{available}}_{t-h\,|\,t}$ is the median of inflation expectations for period t from Focus survey *observed* at period $t-h$ (Focus consensus), ε_t is a projection error, and $(\mu, \eta, \boldsymbol{\phi}) \in \mathbb{R}^{p+2}$ is a vector of parameters. We estimate the model (3.1) employing the elastic net estimator. The estimator $(\widehat{\mu}, \widehat{\eta}, \widehat{\boldsymbol{\phi}})$ for $(\mu, \eta, \boldsymbol{\phi})$ is the result of the problem

$$\min_{\mu,\eta,\boldsymbol{\phi}} \left\{ \sum_t \left(\pi_t - \mu - \eta \,\text{Focus}^{\text{available}}_{t-h\,|\,t} - \boldsymbol{\phi}\,\mathbf{news}_{t-h}\right)^2 + \lambda \left(\frac{1-\gamma}{2} \|\boldsymbol{\xi}\|_2^2 + \gamma \|\boldsymbol{\xi}\|_1 \right) \right\} \tag{3.2}$$

where $\boldsymbol{\xi} = (\eta, \boldsymbol{\phi})$, and λ and γ are hyperparameters.

The presence of the available Focus improves the "stability" of the indexes. It is a guarantee that the selected terms may contribute in some way to predicting inflation beyond what is summarized by the Focus consensus. Moreover, employing the elastic net increases the probability that two relevant and highly correlated terms will be selected – compared to the LASSO, for example. For more advantages of using the elastic net, see Lima *et al.* (2021).

Finally, we compute two *updated* indexes from the most recent vector of news ($\mathbf{news}_t$) in "standardized" and "non-standardized" versions:

$$\texttt{index}_t^1 = \sum_{i=1}^{p} \widehat{\phi}_i \, \text{news}_{it} \;\in \mathbb{R}, \tag{3.3}$$

$$\texttt{index}_t^2 = \frac{\sum_{i=1}^{p} \widehat{\phi}_i \, \text{news}_{it}}{\sum_{i=1}^{p} |\widehat{\phi}_i \, \text{news}_{it}|} \;\in\; [-1, 1]. \tag{3.4}$$

Pros. Following, we list five benefits of the proposed methodology:

(i) Flexibility and adaptability for any variable of interest (with due care);

(ii) The past values of the index do not change, i.e., a new update in time does not modify the previous values of the index;

(iii) There is the automation of the selection of relevant terms, despite the need for pre-selection of n-grams of articles;

(iv) Possibility of relevant terms changing over time; the sign of the coefficient associated with a term can include change over time;

(v) The standardized version of the index is limited to the range of -1 to 1, which avoids significant instabilities over time.

Potential disadvantages or difficulties. Following, we highlight four potential complications of the methodology:

(i) Since the index is based on estimates, the model may take time to capture new relevant terms or exclude terms that are no longer relevant;

(ii) Need to set the size of the rolling window used in the estimation. A smaller window can make it possible to enter new terms more quickly at the cost of estimation uncertainty (instability);

(iii) It requires care so that the index is not unstable over time, especially in the non-standard version, which may show strange behavior at times. A time-varying intercept can generate significant instability, for example;

(iv) Need to condition on the available survey-based expectation to ensure "stability". In the absence of something like the Focus expectation, one could consider an autoregressive (AR) term, for example. Along the same lines, the inclusion of monthly dummies could also contribute to obtaining the index, for example. Nonetheless, conditioning on the available Focus has an economic interpretation – as will be argued further on.

3.2.3
Setup and important considerations

Selection of hyperparameters. We pick the λ from a grid of one hundred values with exponential decay whose definition follows the default of the package `glmnet` for R. For γ, we choose it from a grid of ten values that grows logarithmically according to the sequence

$$\Big\{ \big(\log(1.01 + j \cdot 0.2)\big)^{0.25} : j = 0, 1, \ldots, 8 \Big\} \cup \{1\}.$$

Then, both hyperparameters are selected via Bayesian Information Criterion (BIC).

Sensitivity to the pre-selection of terms and number of terms. There is a certain instability of the index concerning the pre-selection of terms. To avoid increasing the possibilities, we consider the same (broad) pre-selection of article terms for all horizons. By taking 16 (pieces of) terms and adding two more specific terms (see previous Footnote 1), we count 762 n-grams of newspaper articles. Regarding tweets, we considered the count of 397 terms. Therefore, we consider 1,159 terms in the estimation that originates the indexes. We consider only tweets, only articles, or both in the information set for constructing news-based indexes for inflation.

Intercept zero and instability. The indexes set $\mu = 0$ (intercept zero) in model (3.1) for the three horizons considered. Since the intercept varies (considerably) over time, it causes an increase in the "instability" of the indexes, which deteriorates the indexes visually and in terms of contribution to forecast performance. In a way, conditioning the model to some variable that generates stability in the estimation (such as controlling for the available Focus expectation or AR terms, for example) makes the requirement of the intercept dispensable.

Controlling for the available Focus. As previously mentioned, the presence of the available Focus expectation is necessary to guarantee the "adherence" of the indexes to future inflation rates. Furthermore, controlling for the available Focus survey generates an interesting economic interpretation: we manage to make the method include terms that generate "marginal gain" for the inflation adjustment after considering relevant available information from a survey. In other words, controlling for the Focus allows the estimator to select terms that capture the *inflationary surprise*, which may contribute to the relevance of the indexes in forecasting inflation.

Smoothing via averaging of fits of several models. A potential source of instability for the index is abrupt changes in the selection of terms by elastic net across the rolling windows. To alleviate this difficulty, we consider predictions from fits of models estimated in previous windows (with all models being evaluated in the most recent news vector). Formally, with $\widehat{\mathcal{M}}^{t-j}$ being the estimated model considering the period ending in $t-j$, we compute a "smoothed version" of the index via a simple average of the adjustments generated by evaluating each estimated model in the most recent vector of terms:

$$\texttt{index}_t^{i,s} = \frac{1}{J}\sum_{j=0}^{J-1} \widehat{\mathcal{M}}^{t-j}\big(\mathbf{news}_t\big), \qquad i \in \{1,2\},$$

where J is the number of fits we consider. Note that if we consider only the most recent fits, we will be left with $\widehat{\mathcal{M}}^{t}\big(\mathbf{news}_t\big)$, that is, one of the original versions presented in (3.3) and (3.4). Smoothing is necessary mainly for the 12-month cumulative inflation index. We consider the mean of the six most recent adjustments for all horizons. In general, this is the choice that generates the best results.

3.3 Evaluation of the relevance of news-based indexes

In addition to visually inspecting news-based indexes and comparing them to actual inflation, we conduct pseudo-out-of-sample forecasting exercises with models that include or exclude them. Besides the natural benchmark given by the Focus survey, we consider the four models to verify the usefulness of news-based indexes in forecasting. For the presentation of the models, consider the following variable definitions:

- π_t is the cumulative inflation over h periods (months) at the period t;
- $\text{Focus}_{t+h\,|\,t}^{\text{available}}$ is the median of the Focus survey inflation expectations accumulated for h periods ahead and available at the period t – the Focus consensus;
- u_t is a forecast error;
- $\widehat{\pi}_{T+h\,|\,T}$ is a forecast for h-period-ahead cumulative inflation based on information *available* at T.

Model 1 – Bias correction via OLS. Following Mincer & Zarnowitz (1969), we take a linear model that considers both intercept (α) and slope (β) historical bias for a forecast. In particular, we are interested in the *available* Focus-based inflation expectation. Thus, we have the following model:

$$\pi_t = \alpha + \beta\,\text{Focus}^{\text{available}}_{t\,|\,t-h} + u_t, \qquad t = 1, \ldots, T-h.$$

After the estimation of the parameters employing least squares, we are able to obtain a forecast that corrects for historical bias for the period T by computing

$$\widehat{\pi}_{T+h\,|\,T} = \widehat{\alpha} + \widehat{\beta}\,\text{Focus}^{\text{available}}_{T+h\,|\,T},$$

in which $\widehat{\alpha}$ and $\widehat{\beta}$ are OLS estimates.

Model 2 – Bias correction including news-based indexes. We can augment the previous simple bias correction model by adding indexes based on tweets and newspaper articles to test the forecasting performance. Thus, for each index in $\{\texttt{index}^{i,s}_t : i \in \{1,2\},\, s \in \{\text{smooth}, \text{not smooth}\}\}$, we define the model

$$\pi_t = \alpha + \beta\,\text{Focus}^{\text{available}}_{t\,|\,t-h} + \theta\,\texttt{index}^{i,s}_{t-h} + u_t, \qquad t = 1, \ldots, T-h.$$

As before, we compute a forecast via

$$\widehat{\pi}_{T+h\,|\,T} = \widehat{\alpha} + \widehat{\beta}\,\text{Focus}^{\text{available}}_{T+h\,|\,T} + \widehat{\theta}\,\texttt{index}^{i,s}_{t-h},$$

in which the coefficients with hat are least squares estimates.

Model 3 – Data-rich environment and estimation via adaptive LASSO (adaLASSO). Above models have the limitation of not including other potential predictors for inflation (macroeconomic variables, for example). Thus, we can consider including a large number of predictors, including their lags (about this, see Inoue & Kilian, 2008; Garcia *et al.*, 2017; Medeiros *et al.*, 2021). Defining $\boldsymbol{x}_{t-h}$ to be a p-dimensional vector with such variables *available* at period $t-h$, we can write a general linear model as follows:

$$\pi_t = \alpha + \beta\,\text{Focus}^{\text{available}}_{t\,|\,t-h} + \gamma\,\boldsymbol{x}_{t-h} + u_t, \qquad t = 1, \ldots, T-h,$$

in which γ is a p-dimensional vector of parameters.

However, when the number of predictors exceeds the number of temporal observations, we can resort to machine learning methods. We choose to display results from the adaptive LASSO (adaLASSO), a model that deals with the curse of dimensionality by selecting predictors.[2] After estimating the model, we calculate the forecast based on the latest available information, as previously done. Appendix 3.B provides a description of the adaLASSO.

Model 4 – adaLASSO including news-based indexes. Finally, we also add a news-based index in a linear model estimated via adaLASSO to verify the potential gains in forecast performance. Variable selection properties of the adaLASSO play an important role since it empirically determines whether or not indexes should be selected. Combined with evaluating forecasts based on a metric – e.g., root mean squared error (RMSE) – this will attest to the relevance (or not) of considering the indexes in a data-rich environment. In this case, for each $\{\texttt{index}_t^{i,s} : i \in \{1,2\},\, s \in \{\text{smooth}, \text{not smooth}\}\}$, the model is given by

$$\pi_t = \alpha + \beta\,\text{Focus}^{\text{available}}_{t\,|\,t-h} + \gamma\,\boldsymbol{x}_{t-h} + \theta\,\texttt{index}^{i,s}_{t-h} + u_t, \qquad t = 1, \ldots, T-h.$$

Finally, we compute our forecasts based on the most recent data set.

Pseudo-out-of-sample exercise (setup). We set expanding windows to compute multi-horizon inflation forecasts starting in Jan/2019 and ending in Jul/2022. Thus, we compute 43 forecasts for each horizon. In the case of linear models estimated via the adaLASSO, we consider three lags for each time-varying predictor and include monthly dummies. Finally, we use root mean square error (RMSE) as a metric and the Diebold-Mariano test to assess the forecast performance of the models that include or do not include a news-based index. We consider a one-tailed Diebold-Mariano test with null and alternative hypothesis given by $\mathbb{H}_0 : \text{MSE}(\widehat{\pi}^1_{t+h\,|\,t}) = \text{MSE}(\widehat{\pi}^2_{t+h\,|\,t})$ and $\mathbb{H}_1 : \text{MSE}(\widehat{\pi}^1_{t+h\,|\,t}) < \text{MSE}(\widehat{\pi}^2_{t+h\,|\,t})$, where $\widehat{\pi}^1_{t+h\,|\,t}$ indicates the model does not include an index and $\widehat{\pi}^2_{t+h\,|\,t}$ indicates the model including an index.

3.4 Results

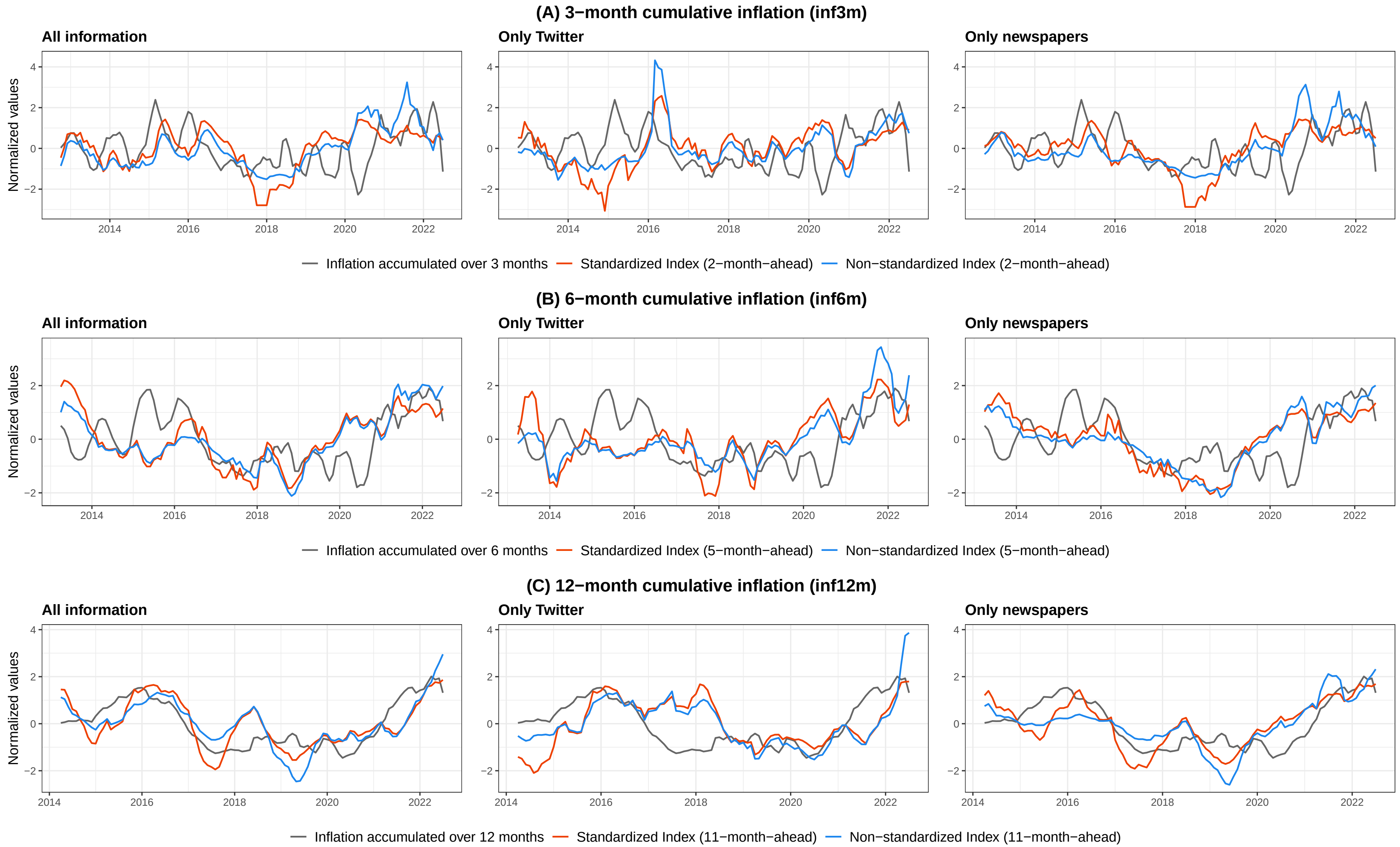
[2]We also consider other models such as LASSO, complete subset regression (CSR), and Random Forest (that admits nonlinearities). However, the adaLASSO performs better than the LASSO for inflation accumulated in 12 months, and both obtain similar performance in other horizons. Concerning the others, the adaLASSO is superior.

3.4.1
Visual inspection of the indexes

Figure 3.1 presents the actual inflation accumulated over 3, 6, and 12 months and news-based indexes in their different versions: standardized and non-standardized, and considering different sets of information – all information, only Twitter, and only newspapers. Each horizon is displayed in a row, and each information set is in a column. To facilitate the comparison, actual inflation (gray lines), as well as both standardized (red lines) and non-standardized (blue lines) indexes, are normalized over each period. We advance the indexes in time according to the horizon to compare them with the respective inflation. The start date differs for different horizons because we need more initial information in index construction (estimation) for longer horizons. Notice that standardized and non-standardized versions of the indexes often exhibit dissonant movements, underscoring the importance of considering both construction approaches and determining the one that suits each situation best. Despite indexes sometimes presenting discrepant magnitudes when compared to the respective actual inflation, we should verify whether the indexes can capture trends and track inflation's fluctuations over time.

For the inflation accumulated over 3 months, the indexes based on all information better capture inflation movements. The non-standardized index that uses all information tends to better track the ups and downs of inflation, especially from 2020 onwards. Even considering the smoothing via the average of different fits, we note that it is not possible to completely mitigate the noisy behavior that persists over time in virtually all indexes for this horizon. In addition, the isolated peak of the non-standardized index based only on tweets in 2016 is a negative highlight. For 6-month cumulative inflation, all indexes adhere reasonably to the long-term inflation trend. However, they do not capture shorter-term cycles well. For this horizon, all indexes show trends that differ from the inflation realized in 2013. Finally, regarding inflation accumulated during 12 months, indexes show more abrupt fluctuations than inflation, but most of them capture the smooth ups and downs of the serie. A considerable divergence occurred in the magnitude and trend of the standardized index based solely on Twitter over 2014. Additionally, the non-standardized index considering only Twitter shows a considerable increase throughout 2022, which may indicate the relevance of the standardized version of the index to attenuate such situations. Visually, the best fit belongs to indexes that consider all information, i.e., join tweets and articles.

Figure 3.1: News-based indexes and inflation both normalized, by horizon and information set



Despite difficulties anticipating some movements of inflation, news-based indexes have potential, and their consideration can contribute to decision-making regarding the prognostic of future inflation dynamics. Some movements not captured by the indexes, such as the sharp decline in 3- and 6-month cumulative inflation at the start of 2020 (at the beginning of the COVID-19 pandemic), are difficult to anticipate. On the other hand, it is worth highlighting that most of the indexes captures well the trend of increasing accumulated inflation over 6 and 12 months from 2021 onwards. From this period, the median of inflation expectations collected by the Central Bank of Brazil began to underestimate future inflation significantly (see Chapter 1). Additionally, note that indexes based on all information or only on articles for accumulated inflation in the next 12 months efficiently anticipate fast disinflation during the second half of 2016 and the first half of 2017. The Focus consensus does not reasonably anticipate this rapid decline in inflation. Econometric models also do not easily anticipate it, even in an information-rich environment, as pointed out in Chapter 1. Following, we investigate the benefits of the employ of news-based indexes in pseudo-out-of-sample forecasting exercises.

3.4.2 Evaluation of the predictive contribution of indexes

We generate 45 out-of-sample predictions for the 3-, 6-, and 12-month cumulative inflation, covering January 2019 to July 2022. Table 3.1 displays the forecast performance in terms of RMSE for *available* Focus consensus (last available median expectation when we compute our forecasts), *ex-post* Focus (median expectation of the reference day, but released only days later), and models that include or not a news-based index for inflation. We report the RMSE ratio using the available Focus RMSE as a reference point. If the RMSE ratio is less than 1, then the model performs better than the available Focus and, if greater than 1, worse than the available Focus. From Panel A of Table 3.1, we notice that the *ex-post* Focus improves the predictive performance slightly compared to the available Focus for all horizons. We expect this result since the experts have more updated information on the *ex-post* Focus. We also expect that the performance improvement would drop with the horizon increase since there is little relevant informational gain between a few days when looking at a longer horizon.

Table 3.2 reports the relative frequency in which the adaLASSO (high-dimensional model) automatically selects a news-based index. Figure 3.2 exhibits actual inflation and forecasts by the horizon (figure on the left) as well as the squared forecast errors (figure on the right) of main models/expectations. Each

Table 3.1: Out-of-sample RMSE with respect to available Focus

		`inf3m`	`inf6m`	`inf12m`
A. Survey				
Focus	Available	1.000	1.000	1.000
	Ex-post	0.960	0.984	0.996
B. Bias correction				
OLS (no index)		0.910	0.960	1.136
Including a non-std index	All information	*0.805****	0.859***	1.148
	Only tweets	0.920	0.906**	1.174
	Only articles	**0.740*****	0.863***	1.254
Including a std index	All information	0.881***	0.887***	0.901***
	Only tweets	0.917	0.928***	1.150
	Only articles	*0.843****	0.874***	0.813***
C. High-dimensional model				
adaLASSO (no index)		0.939	*0.761*	*0.614*
Including a non-std index	All information	0.939	*0.761*	0.721
	Only tweets	0.939	**0.758**	0.615
	Only articles	0.917	*0.761*	*0.614*
Including a std index	All information	0.939	*0.761*	0.623
	Only tweets	0.939	*0.760*	**0.504****
	Only articles	0.939	*0.761*	0.659

Notes: Forecasts covering the period from January/2019 to July/2022. The value highlighted in bold blue indicates the best result for each forecast horizon in terms of out-of-sample RMSE, while blue italics indicate the second- and third-best results. ***, **, and * indicate that a specific model that includes a news-based index performed statistically better than the corresponding model that did not include the index in a one-tailed Diebold-Mariano test at a significance level of 10, 5, and 1%, respectively.

horizon appears in a different panel (from A to C). For the inflation accumulated over 3 months (`inf3m`), bias correction models for available Focus estimated by OLS, adding or not a news-based index as an extra predictor, register the best performances. The RMSE reductions in comparison to the available Focus consensus range from 9% to 36%. For these low-dimensional models, including a news-based index contributes to a further reduction of up to 17 percentage points in relative RMSE, considering the non-standardized index based only on articles. The good performance of the low-dimensional model that includes indexes is mainly due to the reduction of (squared) forecast errors from the second half of 2021 (see Figure 3.2, Panel A). In high-dimensional models, we notice that the adaLASSO hardly selects news-based indexes (see Table 3.2). The method chooses the non-standardized index solely from articles only approximately 14% of the time, leading to a reduction of 2 p.p. on relative RMSE, but statically not

Table 3.2: Selection of news-based indexes by the adaLASSO (%)

	inf3m		inf6m		inf12m	
	Non-std	Std	Non-std	Std	Non-std	Std
All information	–	–	–	–	60.47	100.00
Only tweets	–	–	9.30	2.33	39.53	100.00
Only Articles	13.95	–	–	–	–	93.02

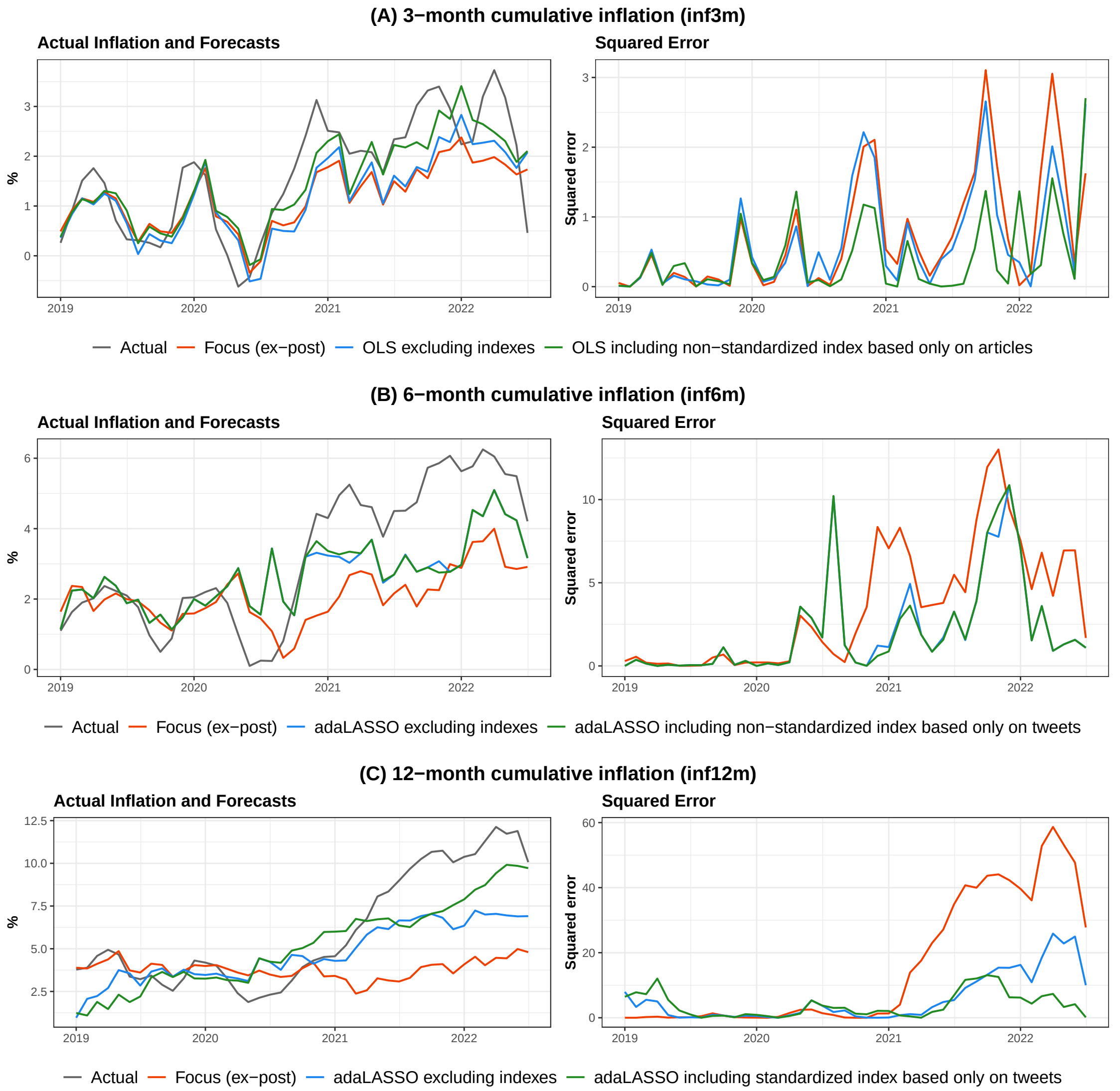
Notes: "–" indicates that the adaLASSO did not select a specific index for any period. $T = 43$.

significant according to a one-tailed Diebold-Mariano test.

For the inflation accumulated during 6 months (`inf6m`), the bias correction model not including any news-based index leads to a small reduction in RMSE (4%), which increases when we include an index (maximum reduction of almost 13%). This result highlights the contribution of a news-based index in a low-dimensional case. Intuitively, our news-based indexes still have predictive power conditional on the experts' available expectations. In contrast, high-dimensional models exhibit superior forecast performance, resulting in a significant decrease in RMSE of at least 24% relative to the available Focus. However, news-based indexes do not exhibit robust predictive power due to their infrequent selection, except for indexes based solely on tweets. Among these, adaLASSO chooses the non-standardized index 9.3% of the time, while the standardized version appears only once out of 43 time periods. Despite this, including these indexes do not result in a significant reduction in RMSE compared to the high-dimensional model that excluded them. Thus, when we control for a more extensive information set, news-based indexes lose their relevance for the analyzed horizon.

The accuracy of the 12-month cumulative inflation forecast (`inf12m`) deteriorates when we apply a historical bias correction to the available expectation. The results in Table 3.1 indicate an increase of more than 13% in RMSE compared to the available Focus. The situation worsens when we consider each of the three non-standardized indexes. However, standardized indexes based on all information or only on newspaper articles lead to a substantial reduction in RMSE, ranging from 10% to 19%, compared to the available Focus. These improvements are statistically significant, as attested by a one-tailed DM test. These findings underscore the benefits of implementing discipline through standardized versions of the indexes, given that the series of the count of terms can be volatile even with some smoothing applied. When we consider a large number of predictors in a linear model estimated employing adaLASSO, there is an expressive reduction of almost 39% in terms of RMSE. In this context, only the standardized index based solely on Twitter information can deliver an even better result: a reduction of al-

Figure 3.2: Inflation forecasts and squared forecast errors, by horizon



most 50% in RMSE, which is a decrease of 11 percentage points compared to the model that did not include any index. Notably, this index is automatically picked in 100% of the opportunities, as shown in Table 3.2. Moreover, according to Panel C of Figure 3.2, the predictive improvements come from better forecasts starting from 2021.

The results of the pseudo-out-of-sample forecasting exercises indicate that news-based indexes are particularly useful during periods of high instability[3], such as the onset of the COVID-19 pandemic in 2020 and onwards. As shown in Figure 3.2, except for the accumulated inflation in 6 months, the indexes significantly reduce the squared forecast error from the second half of 2020. Panel C of Figure 3.2 also suggests that disregarding the inaccurate forecasts

[3]This result is similar to the finding by Kalamara *et al.* (2022). They use newspaper data to forecast GDP, inflation, and unemployment in the United Kingdom.

generated for early 2019, the high-dimensional models, including or not news-based indexes, would deliver an even greater RMSE reduction. Another result indicates that smaller models perform better for a shorter horizon (3 months ahead), whereas models incorporating several predictors perform better for a longer horizon (12 months ahead). This result may occur because there is little room to improve the predictive performance of a survey-based expectation as we shorten the forecast horizon. Moreover, while a restricted model using "right variables" still generates some improvement concerning the inflation expectation in shorter horizons, a more extensive model is more susceptible to specification errors and estimation uncertainty. However, in long horizons, there is room for the effective contribution of other predictors – including news-based indexes, even considering an information-rich environment.

3.5 Conclusion

This essay presents novel approaches to constructing forward-looking inflation indexes using data from Twitter and newspapers through a supervised learning method shown as a time-varying dictionary approach. Considering the Brazilian case and different horizons for cumulative inflation, our news-based indexes are able to anticipate long-term trends. Furthermore, they capture short-term movements in inflation at various periods. We also highlight the benefits of news-based indexes for inflation forecasting by conducting pseudo-out-of-sample exercises. News-based indexes can improve forecast performance for different horizons. For short ones (3 and 6 months ahead), a low-dimensional model that considers the median of expectations from a survey as the unique predictor benefits from including news-based indexes. On the other hand, for larger horizons (12 months ahead), high-dimensional models can improve the forecast accuracy by incorporating these indexes, at least marginally. Thus, incorporating news-based indexes from social media and news sources can improve inflation forecasting.

There are several possibilities for extending the results of this essay that can be investigated in future research. The most natural extension is to look at sub-components of a price index and predict their variations individually. Since different disaggregates have specific characteristics and some are more difficult to predict, indexes based on tweets and articles can be interesting in predicting future values of these disaggregations. Moreover, one potential avenue of exploration beyond inflation forecasting is to use news-based indexes to model and predict demand for various goods and services.

3.A
Terms and variables

Generic terms on Twitter. Table 3.A.1 contains the generic terms whose count is used to normalize the count of terms related to inflation over time in order to control for the secular trend in the number of tweets. The translations to English are also presented.

Table 3.A.1: List of generic terms and their translations

Generic term	Translation	Generic term	Translation
oi	hi	ok	okay
olá	hello	sim	yes
bom dia	good morning	não	no
boa noite	good night	galera	folks
boa tarde	good afternoon	bora	let's go (slang)
escrever	to write	fazer	to do, to make
ler	to read	valeu	thanks (slang)
vamos	let's go	obrigado	thanks, thank you

Other predictors. In addition to news-based indexes, we consider the *available* Focus-based inflation expectation, seasonal dummies, and eighty more time-varying variables and their respective lags as predictors for inflation. These variables can be divided into ten categories: prices and money (19), commodities prices (4), economic activity (9), employment (5), electricity (4), confidence (3), finance (12), credit (4), government (12), and exchange and international transactions (9). The choice of the variables is similar to the variables used in Garcia *et al.* (2017).

Table 3.A.2 presents a description of all variables as well as the transformations implemented to guarantee the stationarity of the series. To get as close as possible to a real-time database, we consider the average disclosure delay of each variable. The penultimate column of Table 3.A.2 contains this information. We consider the last day of each month as the reference day on which multi-period forecasts are computed.

Table 3.A.2: Description of predictive variables

#	Variable	Description	Unit	Source	Lag	Transformation
		A. Prices and Money				
1	inf	Consumer Price Index (IPCA)	index	IBGE	1	% change
2	expec	Focus-based inflation expectations (*available*)	% per month	BCB	0	–
3	ipca15	Consumer Price Index - 15 (IPCA-15)	index	IBGE	0	% change
4	inpc	Consumer Price Index (INPC)	index	BCB	1	% change
5	ipc	Consumer Price Index - Brazil (IPC-Br)	index	FGV	1	% change
6	igpm	General Price Index - M (IGP-M)	index	FGV	1	% change
7	igpdi	General Price Index - DI (IGP-DI)	index	FGV	1	% change
8	igp10	General Price Index - 10 (IGP-10)	index	FGV	1	% change
9	ipc_fipe	Fipe Consumer Price Index (IPC-Fipe)	index	Fipe	1	% change
10	ipa	Wholesale Price Index (IPA)	index	FGV	1	% change
11	ipa_ind	IPA – industrial Products	index	FGV	1	% change
12	ipa_agr	IPA – agricultural Products	index	FGV	1	% change
13	incc	National Index of Building Costs (INCC)	index	FGV	1	% change
14	bm_broad	Broad Monetary Base – end-of-period balance	index	BCB	2	% change
15	bm	Monetary Base – working day balance average	Index	BCB	2	% change
16	m1	Money supply M1 – working day balance average	Index	BCB	2	% change
17	m2	Money supply M2 – end-of-period balance	Index	BCB	2	% change
18	m3	Money supply M3 – end-of-period balance	Index	BCB	2	% change
19	m4	Money supply M4 – end-of-period balance	Index	BCB	2	% change
		B. Commodities prices				
20	icbr	Brazilian Commodity index – all	index	BCB	1	% change
21	icbr_agr	Brazilian Commodity index – agriculture	index	BCB	1	% change
22	icbr_metal	Brazilian Commodity index – metal	index	BCB	1	% change
23	icbr_energy	Brazilian Commodity index – energy	index	BCB	1	% change
		C. Economic Activity				
24	ibcbr	Brazilian IBC-Br Economic Activity index	index	BCB	3	% change
25	month_gdp	GDP monthly – current prices	R$ million	BCB	1	% change
26	tcu	Use of installed capacity – manufacturing industry	%	FGV	1	first difference
27	pimpf	Industrial Production – general	index	IBGE	2	% change
28	pmc	Retail sales volume – total	index	IBGE	2	% change
29	steel	Steel production	index	BCB	1	–
30	prod_vehicles	Vehicle production – total	units	Anfavea	1	% change
31	prod_agr_mach	Production of agricultural machinery – total	units	Anfavea	1	% change
32	vehicle_sales	Vehicle sales by dealerships – total	units	Fenabrave	1	% change
		D. Labor Market				
33	unem	Unemployment (combination of PME and PNADC)	%	IBGE	3	first difference
34	employment	Registered employess by economic activity - Total	units	IBGE	1	first difference
35	aggreg_wage	Overall Earnings (broad wage income)	R$ (million)	IBGE	2	% change
36	min_wage	Federal Minimum Wage	R$	MTb	0	% change
37	income	Households gross disposable national income	R$ (million)	BCB	2	% change
		E. Electricity				
38	elec	Electricity consumption - total	GWh	Eletrobrás	3	% change
39	elec_res	Electricity consumption - residential	GWh	Eletrobrás	3	% change
40	elec_com	Electricity consumption - commercial	GWh	Eletrobrás	3	% change
41	elec_ind	Electricity consumption - industry	GWh	Eletrobrás	3	% change
		F. Confidence				
42	cons_confidence	Consumer Confidence index	index	Fecomercio	1	% change
43	future_expec	Future expectations index	index	Fecomercio	1	% change
44	conditions	Current economic conditions index	index	Fecomercio	1	% change
		G. Finance				
45	ibovespa	Ibovespa index	% per month	BM&FBOVESPA	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	Savings 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	0	–
55	embi	Emerging Markets Bond Index Plus – Brazil	b.p. acc. month	JP Morgan	0	first difference
56	vix	CBOE Volatility Index (VIX)	index	CBOE	0	
		H. Credit				
57	cred_total	Credit outstanding - total	R$ (million)	BCB	2	% change
58	cred_dgp	Credit outstanding as a percentage of GDP	% of GDP	BCB	2	first difference
59	indebt_house1	Household debt to income ratio – all	% of 12m income	BCB	2	first difference
60	indebt_house2	Household debt to income ratio without mortgage loans	% of 12m income	BCB	2	first difference
		I. Government				
61	net_debt_gdp	Net public debt (% GDP) - Consolidated public sector	% of GDP	BCB	2	first difference
62	net_debt	Net public debt - Total - Consolidated public sector	R$ (million)	BCB	2	first difference
63	net_debt_fedgov_bcb	Net public debt - Federal Government and Central Bank	R$ (million)	BCB	2	first difference
64	net_debt_states	Net public debt - State governments	R$ (million)	BCB	2	first difference
65	net_debt_cities	Net public debt - Municipal governments	R$ (million)	BCB	2	first difference
66	primary_result	Primary result - Consolidated public sector	R$ (million)	BCB	2	first difference
67	debt_fedgov_old	Gross general government debt - Method used until 2007	R$ (million)	BCB	2	% change
68	debt_fedgov_new	Gross general government debt - Method used since 2008	R$ (million)	BCB	2	% change
69	treasury_emit	National Treasury domestic securities - Total issued	R$ (million)	BCB	2	% change
70	treasury_mkt	National Treasury domestic securities - Total on market	R$ (million)	BCB	2	% change
71	treasury_term	National Treasury securities debt - medium term	months	BCB	2	first difference
72	treasury_dur	National Treasury securities debt - medium duration	months	BCB	2	first difference
		J. Exchange and International Transactions				
73	reer	Real Effective Exchange Rate	R$/other	BIS	2	% change
74	usd_brl_end	USD-BRL rate – end of period	R$/US$	BCB	0	% change
75	usd_brl_av	USD-BRL rate – monthly average	R$/US$	BCB	0	% change
76	eur_brl_end	EUR-BRL rate – end of period	R$/€	Bloomberg	0	% change
77	eur_brl_av	EUR-BRL rate – monthly average	R$/€	Bloomberg	0	% change
78	current_account	Current account – net	US$ (million)	BCB	2	% change
79	trade_balance	Balance on goods and services - net (Brazilian trade balance)	US$ (million)	BCB	2	% change
80	exports	Imports	US$ (million)	BCB	2	% change
81	imports	Exports	US$ (million)	BCB	2	% change

3.B
Adaptive LASSO (adaLASSO)

Consider a predictive linear model given by $\pi_t = \boldsymbol{\beta}\,\boldsymbol{x}_{t-h} + \varepsilon_t$, in which π_t is inflation at period t, $\boldsymbol{x}_{t-h}$ is a J-dimensional vector of predictors (and their lags) *observed* at period $t-h$, and ε_t is a forecast error. We can estimate the parameter vector $\boldsymbol{\beta}$ via adaptive LASSO (adaLASSO). Introduced by Zou (2006), this method solves

$$\widehat{\boldsymbol{\beta}}_{\text{adaLASSO}}(\lambda, \boldsymbol{\omega}) = \underset{\boldsymbol{\beta}}{\text{argmin}} \left\{ \frac{1}{T-h} \sum_{t=1}^{T-h} (\pi_t - \alpha - \boldsymbol{\beta}\,\boldsymbol{x}_{t-h})^2 + \lambda \sum_{j=1}^{J} \omega_j \,|\beta_j| \right\}$$

in which λ is a regularization parameter, and $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_J)$ is a vector of weights obtained previously via LASSO, an estimator that assumes $\omega_j = 1$ for all j (see Tibshirani, 1996). More precisely, we compute the adaLASSO weights via

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

in which the presence of $T^{-1/2}$ makes possible a variable that the LASSO had not selected in the first stage, i.e., the case in which $\beta_{\text{LASSO},j} = 0$, can be selected by the adaLASSO.

Finally, we get an h-periods-ahead forecast by computing $\widehat{\pi}_{t+h\,|\,t} = \widehat{\alpha} + \widehat{\boldsymbol{\beta}}_{\text{adaLASSO}}\,\boldsymbol{x}_t$.

Conclusion

In this dissertation, we address the following three topics related to inflation forecasting, focusing on the Brazilian case: (i) the aggregation of disaggregated forecasts, (ii) the combination of forecasts, and (iii) the employment of textual data for inflation prediction. The findings encompass the benefits of aggregating disaggregated forecasts using machine learning techniques, which outperform traditional time series models, particularly during volatile periods. A time-varying bias correction approach for individual inflation forecasts is promising for small and intermediate forecast horizons. While the state-space model slightly underperforms compared to rolling window procedures, short rolling windows demonstrate good forecast performance. Lastly, forward-looking indexes based on social media and newspaper articles are valuable in capturing short-term cycles and long-term trends in inflation. Considering these indexes tends to improve the inflation forecast performance. Overall, our findings highlight the importance of employing diverse forecasting methods and alternative data sources to enhance the accuracy of inflation forecasts.

Future research has several avenues to extend the findings of this dissertation. Replicating the procedures in different countries would enable the assessment of the generalizability of the results. Developing a methodology that combines several models to predict disaggregates and exploring various levels of disaggregation could improve the accuracy of inflation forecasts. Additionally, integrating unstructured data, via news-based indexes or sentiment analysis, into the disaggregated analysis to predict variations in price index sub-components presents an intriguing area for investigation. Another pertinent topic is dealing with the challenge of the slow-reacting intercept in models for inflation forecasting. Exploring the inclusion of a time-varying intercept into a high-dimensional model could offer a viable solution. Lastly, in terms of forecast combination, delving into alternative specifications for the state-space model and exploring methods to reduce the variance of time-varying biases would refine the approaches used to correct biased forecasts.

Bibliography

Altug, Sumru, & Çakmaklı, Cem. 2016. Forecasting inflation using survey expectations and target inflation: Evidence for Brazil and Turkey. *International Journal of Forecasting*, **32**(1), 138–153.

Ang, Andrew, Bekaert, Geert, & Wei, Min. 2007. Do macro variables, asset markets, or surveys forecast inflation better? *Journal of Monetary Economics*, **54**(4), 1163–1212.

Angelico, Cristina, Marcucci, Juri, Miccoli, Marcello, & Quarta, Filippo. 2022. Can we measure inflation expectations using Twitter? *Journal of Econometrics*, **228**(2), 259–277.

Araujo, Gustavo Silva, & Gaglianone, Wagner Piazza. 2023. Machine learning methods for inflation forecasting in Brazil: New contenders versus classical models. *Latin American Journal of Central Banking*, **4**(2), 1–29.

Aron, Janine, & Muellbauer, John. 2012. Improving forecasting in an emerging economy, South Africa: Changing trends, long run restrictions and disaggregation. *International Journal of Forecasting*, **28**(2), 456–476.

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

Bai, Jushan, & Ng, Serena. 2008. Forecasting economic time series using targeted predictors. *Journal of Econometrics*, **146**(2), 304–317.

Bai, Jushan, & Ng, Serena. 2009. Boosting diffusion indices. *Journal of Applied Econometrics*, **24**(4), 607–629.

Bańbura, Marta, Leiva-León, Danilo, & Menz, Jan-Oliver. 2021. *Do inflation expectations improve model-based inflation forecasts?* ECB Working Paper n. 2604.

Barbaglia, Luca, Consoli, Sergio, & Manzan, Sebastiano. 2022. Forecasting with economic news. *Journal of Business & Economic Statistics*, 1–12.

Baştürk, Nalan, Çakmakli, Cem, Ceyhan, S Pinar, & Van Dijk, Herman K. 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.

Bates, John M, & Granger, Clive WJ. 1969. The Combination of Forecasts. *Journal of the Operational Research Society*, **20**(4), 451–468.

Baybuza, Ivan. 2018. Inflation forecasting using machine learning methods. *Russian Journal of Money and Finance*, **77**(4), 42–59.

Benalal, Nicholai, Diaz del Hoyo, Juan Luis, Landau, Bettina, Roma, Moreno, & Skudelny, Frauke. 2004. *To aggregate or not to aggregate? Euro area inflation forecasting*. ECB Working Paper n. 374.

Bermingham, Colin, & D'Agostino, Antonello. 2014. Understanding and forecasting aggregate and disaggregate price dynamics. *Empirical Economics*, **46**(2), 765–788.

Blanchard, Olivier, & Galí, Jordi. 2007. Real wage rigidities and the New Keynesian model. *Journal of Money, Credit and Banking*, **39**, 35–65.

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

Bruneau, Catherine, De Bandt, Olivier, Flageollet, Alexis, & Michaux, Emmanuel. 2007. Forecasting inflation using economic indicators: the case of France. *Journal of Forecasting*, **26**(1), 1–22.

Capistrán, Carlos, & Timmermann, Allan. 2009. Forecast combination with entry and exit of experts. *Journal of Business & Economic Statistics*, **27**(4), 428–440.

Capistrán, Carlos, Constandse, Christian, & Ramos-Francia, Manuel. 2010. Multi-horizon inflation forecasts using disaggregated data. *Economic Modelling*, **27**(3), 666–677.

Carlo, Thiago Carlomagno, & Marçal, Emerson Fernandes. 2016. Forecasting Brazilian inflation by its aggregate and disaggregated data: a test of predictive power by forecast horizon. *Applied Economics*, **48**(50), 4846–4860.

Carvalho, Fabia A, & Minella, André. 2012. Survey forecasts in Brazil: A prismatic assessment of epidemiology, performance, and determinants. *Journal of International Money and Finance*, **31**(6), 1371–1391.

Choudhary, M Ali, & Haider, Adnan. 2012. Neural Network Models for Inflation Forecasting: An Appraisal. *Applied Economics*, **44**(20), 2631–2635.

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

Del Negro, Marco, & Eusepi, Stefano. 2011. Fitting observed inflation expectations. *Journal of Economic Dynamics and Control*, **35**(12), 2105–2131.

Diebold, Francis X, & Mariano, Roberto S. 1995. Comparing Predictive Accuracy. *Journal of Business & Economic Statistics*, **13**(3).

Dräger, Lena, Lamla, Michael J, & Pfajfar, Damjan. 2016. Are survey expectations theory-consistent? The role of central bank communication and news. *European Economic Review*, **85**, 84–111.

Duarte, Cláudia, & Rua, António. 2007. Forecasting inflation through a bottom-up approach: How bottom is bottom? *Economic Modelling*, **24**(6), 941–953.

Elliott, Graham, & Timmermann, Alan. 2016. *Economic Forecasting*. Princeton, New Jersey: Princeton University Press.

Elliott, Graham, Gargano, Antonio, & Timmermann, Allan. 2013. Complete subset regressions. *Journal of Econometrics*, **177**(2), 357–373.

Elliott, Graham, Gargano, Antonio, & Timmermann, Allan. 2015. Complete subset regressions with large-dimensional sets of predictors. *Journal of Economic Dynamics and Control*, **54**, 86–110.

Espasa, Antoni, & Albacete, Rebeca. 2007. Econometric modelling for short-term inflation forecasting in the euro area. *Journal of Forecasting*, **26**(5), 303–316.

Espasa, Antoni, Senra, Eva, & Albacete, Rebeca. 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, Jianqing, Masini, Ricardo, & Medeiros, Marcelo C. 2021. *Bridging factor and sparse models*. Available at SSRN 3789141.

Faust, Jon, & Wright, Jonathan H. 2013. Forecasting Inflation. *Chap. 1, pages 3–56 of:* Elliott, Graham, & Timmermann, Allan (eds), *Handbook of Economic Forecasting*, vol. 2A. North Holland: Elsevier.

Florido, Gustavo Porto. 2021. *Essays on Inflation Forecasting*. Ph.D. thesis, Sao Paulo School of Economics, FGV.

Forni, Mario, Hallin, Marc, Lippi, Marco, & Reichlin, Lucrezia. 2003. Do financial variables help forecasting inflation and real activity in the euro area? *Journal of Monetary Economics*, **50**(6), 1243–1255.

Fulton, Chad, & Hubrich, Kirstin. 2021. Forecasting US inflation in real time. *Econometrics*, **9**(36), 1–20.

Gaglianone, Wagner Piazza, & Issler, João Victor. 2023. *Microfounded Forecasting*. Mimeo., FGV EPGE.

Gaglianone, Wagner Piazza, Issler, João Victor, & Matos, Silvia Maria. 2017. Applying a microfounded-forecasting approach to predict Brazilian inflation. *Empirical Economics*, **53**(1), 137–163.

Galí, Jordi, & Gertler, Mark. 1999. Inflation dynamics: A structural econometric analysis. *Journal of Monetary Economics*, **44**(2), 195–222.

Garcia, Márcio GP, Medeiros, Marcelo C, & Vasconcelos, Gabriel FR. 2017. Real-time inflation forecasting with high-dimensional models: The case of Brazil. *International Journal of Forecasting*, **33**(3), 679–693.

Granger, Clive WJ, & Ramanathan, Ramu. 1984. Improved methods of combining forecasts. *Journal of forecasting*, **3**(2), 197–204.

Guzman, Giselle. 2011. Internet search behavior as an economic forecasting tool: The case of inflation expectations. *Journal of Economic and Social Measurement*, **36**(3), 119–167.

Hamilton, James Douglas. 1994. *Time Series Analysis*. Princeton University Press.

Hendry, David F, & Hubrich, Kirstin. 2011. Combining disaggregate forecasts or combining disaggregate information to forecast an aggregate. *Journal of Business & Economic Statistics*, **29**(2), 216–227.

Hubert, Paul, & Labondance, Fabien. 2021. The signaling effects of central bank tone. *European Economic Review*, **133**, 1–27.

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

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

Inoue, Atsushi, & Kilian, Lutz. 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.

Issler, Joao Victor, & Lima, Luiz Renato. 2009. A panel data approach to economic forecasting: The bias-corrected average forecast. *Journal of Econometrics*, **152**(2), 153–164.

Kalamara, Eleni, Turrell, Arthur, Redl, Chris, Kapetanios, George, & Kapadia, Sujit. 2022. Making text count: economic forecasting using newspaper text. *Journal of Applied Econometrics*, **37**, 896–919.

Larsen, Vegard H, & Thorsrud, Leif A. 2019. The value of news for economic developments. *Journal of Econometrics*, **210**(1), 203–218.

Larsen, Vegard H, Thorsrud, Leif Anders, & Zhulanova, Julia. 2021. News-driven inflation expectations and information rigidities. *Journal of Monetary Economics*, **117**, 507–520.

Li, Xin, Shang, Wei, Wang, Shouyang, & Ma, Jian. 2015. A MIDAS modelling framework for Chinese inflation index forecast incorporating Google search data. *Electronic Commerce Research and Applications*, **14**(2), 112–125.

Lima, Luiz Renato, Godeiro, Lucas Lúcio, & Mohsin, Mohammed. 2021. Time-varying dictionary and the predictive power of FED minutes. *Computational Economics*, **57**(1), 149–181.

Lin, Jianhao, Fan, Jiacheng, Zhang, Yifan, & Chen, Liangyuan. 2022. Real-time macroeconomic projection using narrative central bank communication. *Journal of Applied Econometrics*, 1–20.

Lütkepohl, Helmut. 1984. Linear transformations of vector ARMA processes. *Journal of Econometrics*, **26**(3), 283–293.

Martins, Leonardo Caio Ladalardo, & Medeiros, Marcelo Cunha. 2022. *Nowcasting GDP with News and Google Trends*. Unpublished manuscript.

Medeiros, Marcelo C, Vasconcelos, Gabriel, & Freitas, Eduardo. 2016. Forecasting Brazilian inflation with high-dimensional models. *Brazilian Review of Econometrics*, **36**(2), 223–254.

Medeiros, Marcelo C, Vasconcelos, Gabriel FR, Veiga, Álvaro, & Zilberman, Eduardo. 2021. Forecasting inflation in a data-rich environment: the benefits of machine learning methods. *Journal of Business & Economic Statistics*, **39**(1), 98–119.

Mincer, Jacob A, & Zarnowitz, Victor. 1969. The Evaluation of Economic Forecasts. *Pages 3–46 of: Economic forecasts and expectations: Analysis of forecasting behavior and performance*. NBER.

Moser, Gabriel, Rumler, Fabio, & Scharler, Johann. 2007. Forecasting Austrian Inflation. *Economic Modelling*, **24**(3), 470–480.

Moshiri, Saeed, & Cameron, Norman. 2000. Neural Network versus Econometric Models in Forecasting Inflation. *Journal of Forecasting*, **19**(3), 201–217.

Nakamura, Emi. 2005. Inflation forecasting using a neural network. *Economics Letters*, **86**(3), 373–378.

Newbold, Paul, & Granger, Clive WJ. 1974. Experience with forecasting univariate time series and the combination of forecasts. *Journal of the Royal Statistical Society: Series A (General)*, **137**(2), 131–146.

Niesert, Robin F, Oorschot, Jochem A, Veldhuisen, Christian P, Brons, Kester, & Lange, Rutger-Jan. 2020. Can Google search data help predict macroeconomic series? *International Journal of Forecasting*, **36**(3), 1163–1172.

Palm, Franz C, & Zellner, Arnold. 1992. To combine or not to combine? Issues of combining forecasts. *Journal of Forecasting*, **11**(8), 687–701.

Rambaccussing, Dooruj, & Kwiatkowski, Andrzej. 2020. Forecasting with news sentiment: Evidence with UK newspapers. *International Journal of Forecasting*, **36**(4), 1501–1516.

Reid, David J. 1969. *A comparative study of time series prediction techniques on economic data*. Ph.D. thesis, Dept. of Mathematics, University of Nottingham.

Stock, James H, & Watson, Mark W. 1999. Forecasting inflation. *Journal of Monetary Economics*, **44**(2), 293–335.

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

Thomas, Lloyd B. 1999. Survey measures of expected US inflation. *Journal of Economic Perspectives*, **13**(4), 125–144.

Thorsrud, Leif Anders. 2020. Words are the New Numbers: A Newsy Coincident Index of the Business Cycle. *Journal of Business & Economic Statistics*, **38**(2), 393–409.

Tibshirani, Robert. 1996. Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, **58**(1), 267–288.

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

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