BORRADORES DE ECONOMÍA

Forecasting Inflation from Disaggregated Data: The Colombian case

By:
Wilmer Martínez-Rivera
Eliana R. González-Molano
Edgar Caicedo-García

No. 1251
2023

BANCO DE LA REPUBLICA
COLOMBIA

Pronosticando la inflación a partir de datos desagregados: Caso Colombiano

Las opiniones contenidas en el presente documento son responsabilidad exclusiva de los autores y no comprometen al Banco de la República ni a su Junta Directiva

Wilmer Martínez-Rivera[1*], Eliana R. González-Molano[2], Edgar Caicedo-García[3]

[1*]Departamento de Inflación, Banco de la República, Bogotá, Colombia.
[2]Departamento de Información Económica, Banco de la República, Bogotá, Colombia.
[3]Departamento de Inflación, Banco de la República, Bogotá, Colombia.

*Autor corresponsal. E-mail(s): wmartiri@banrep.gov.co;
Coautores: egonzamo@banrep.gov.co; Ecaicega@banrep.gov.co;

Resumen

En este artículo se analiza la información mensual tanto agregada como desagregada del índice de precios al consumidor (IPC) en Colombia. Se explora las ventajas de pronosticar a nivel desagregado para luego agregar pronósticos y comparar con los pronósticos obtenidos al analizar la información agregada. El cálculo de pronósticos esta basado en el ajuste de modelos y técnicas que incluyen modelos de reducción de dimensión, modelos de selección de variables, modelos de Machine Learning así como modelos tradicionales de series de tiempo ARIMA. El periodo muestral de análisis es 2011 a 2022 cuyo cálculo de pronósticos fuera de muestra se da a partir de 2017 hasta 2022.

Keywords: Inflación, datos desagregados, pronósticos agregados

Clasificación JEL: C52 , E17 , E31

Forecasting Inflation from Disaggregated Data: The Colombian case

The opinions contained in this document are the sole responsibility of the authors and do not commit Banco de la República nor its Board of Directors

Wilmer Martínez-Rivera[1*], Eliana R. González-Molano[2], Edgar Caicedo-García[3]

[1*]Inflation Department, Banco de la República, Bogotá, Colombia.
[2]Economic Information Department, Banco de la República, Bogotá, Colombia.
[3]Inflation Department, Banco de la República, Bogotá, Colombia.

*Corresponding author(s). E-mail(s): wmartiri@banrep.gov.co;
Contributing authors: egonzamo@banrep.gov.co;
Ecaicega@banrep.gov.co;

Abstract

Based on monthly disaggregated Consumer Price Index (CPI) item series and macroeconomic series, we explore the advantages of forecast inflation from a disaggregated to an aggregated level by aggregating the forecasts. We compare the performance of this approach with the forecast obtained modeling aggregated inflation directly. For the aggregate level, we implement some of the techniques and models, helpful to work with many predictors, such as dimension reduction, shrinkage methods, and machine learning models. Also, we implement traditional time-series models. For the disaggregated data, we use its lags and a set of macroeconomic variables as explanatory variables. Direct and recursive forecast techniques are also explored. The sample period of the analysis is from 2011 to 2022, with forecasting and evaluation out of the sample from 2017. In addition, we evaluate the forecast accuracy during the COVID-19 period. We found a reduction in the forecast error from the disaggregate analysis over the aggregate one.

Keywords: Inflation, disaggregated data, forecast aggregation

JEL Classification: C52 , E17 , E31

1 Introduction

It is important for monetary policymakers to count on reliable forecasts for key macroeconomic variables, particularly inflation, to make appropriate decisions. This has led to a permanent effort of econometrists in trying to find ways of reducing forecasting errors, developing and implementing different methodologies for forecasting and forecast combination. On the other hand, most forecasting models are based on the aggregate variables and the relationship with some macroeconomic indicators that serve as explanatory variables, leaving out some disaggregated available information that might help to explain the dynamics of aggregate inflation and help to reduce forecasting errors.

In the literature on forecasting, there are two main arguments for aggregating forecasts of disaggregated variables instead of directly forecasting the aggregate variable of interest. On one hand, when forecasting the aggregate directly, the common or "average" trend of all the components is estimated, leaving out the properties of some items that might have distinct dynamics from the "average". Thus, modeling individual dynamics may generate a more accurate forecast for the aggregate. Modeling disaggregated variables may involve using a larger and more heterogenous information set, and specifications may vary across the disaggregate variables (Barker (2014)). A second argument favoring disaggregate is that forecast errors of disaggregated components might cancel partly, leading to more accurate aggregate predictions (also see Clements and Hendry (2002) for a discussion on forecast combination as bias correction). In contrast, some authors have argued that forecasting the aggregate directly is a better strategy because, in practice, the models for the disaggregate variables might be misspecified. Thus, aggregating forecasts from the misspecified models might not improve the forecast accuracy for the aggregate (Grunfeld and Griliches (1960)). On the other hand, a well-specified model does not necessarily imply higher forecast accuracy; unexpected shocks might affect the forecast errors of some of the disaggregate variables in the same direction so that forecast errors do not cancel out Hubrich (2005).

Some of the key works that support one of these approaches are the ones of Lütkepohl (1986, 1987), based on Monte Carlo simulation for small samples and different data generating processes - DGP´s. The results depend on the DGP. If the components are uncorrelated and for a short time forecast horizon, aggregating the component's forecasts may lead to lower forecasting error for certain DGPs.

Since there is no consensus, several empirical works have been performed in which forecasts for an aggregate are built by aggregating forecasts of components, finding mixed results. In particular, Hubrich (2002); Hubrich et al (2003) forecasts the annual inflation of the harmonized CPI for the Euro area aggregating forecasts of components of this indicator using different univariate and multivariate models, finding that this not necessarily reduces the forecast error in comparison with the forecasts obtained directly for the aggregate. The main reason for this result is that the forecasting errors of the components are not canceled out because some common shocks are not considered in the specification of the models and affect most components in the same direction leading to higher forecasting errors for the aggregate.

Benalal et al (2004) analyzes the forecasts of total and core inflation (excluding energy and non-processed food) based on the harmonized CPI for the eurozone obtained by aggregating forecasts of components. Additionally, they aggregate forecasts of the inflation of the four greatest countries of the euro area as a proxy of the total aggregated inflation of the Euro area. In this work also, different univariate and multivariate models are used. Some contributions of this work are the use of Bayesian models (BVAR), the model selection criterion is based on choosing the model with the smallest root mean squared error (RMSE) of the out-of-sample forecasts obtained recursively, and additionally, they check whether the variables and the signs of the coefficients included in the selected model satisfy the economic theory. The authors find that for long-term horizons, forecasting the aggregate directly produces smaller errors than aggregating components forecasts, but for the short-run horizons, there are no conclusive results. The results are more promising for core inflation, mainly because the components do not present important shocks. Therefore, there is minor volatility than in the case of total inflation. On the other hand, aggregating forecasts of the inflation of the four highest countries, as a forecast of synthetic total inflation of the eurozone, does not surpass the direct forecast of the aggregated synthetic inflation for any horizon. However, there are no big differences in the forecast errors between the two strategies.

Hendry and Hubrich (2006) generate different forecasts for the eurozone and the United States CPI from models that include not only its own lagged information but also lagged information of the components of the CPI. They find that this forecasting model produces less forecast error than the one obtained when directly forecasting the aggregate and/or aggregating forecasts of the components of the CPI for the case of the United States and, in some situations, for the eurozone. According to these authors, the benefit of including disaggregates to forecast the aggregate depends on the particular variable that is being analyzed, the sample, the forecasting model, and the forecast horizon.

Furthermore, Hendry and Hubrich (2011) analyzed the forecast ability to include disaggregate information in the model for the aggregate in comparison to the combination of disaggregated forecast to forecast the aggregate and a model for the aggregate based only on its own lagged information. Using Monte Carlo simulations, they checked the effect of the stochastic structure of disaggregates and their interdependencies, structural breaks, estimation uncertainty, and misspecification on the relative forecast accuracy of the three different approaches to forecasting the aggregate. They found that including disaggregate variables in the aggregate model helps forecast the aggregate if the disaggregates follow different stochastic structures, the components are interdependent, and only a selected number of components is included to reduce estimation uncertainty.

More recently, Joseph et al (2021) forecast CPI inflation in the United Kingdom using a large set of disaggregated CPI item series and 43 macroeconomic series. They implement a wide set of forecasting models such as dimensionality reduction models (principal component Analysis and Dynamic Factor), shrinkage models (Ridge, Lasso, and Elastic Net), and non-linear machine learning models (Support Vector Machines, Artificial Neural Networks, and Random Forest). With these forecasting tools, the authors evaluate the advantage of using a large dataset of predictors to improve the headline, core, and service inflation forecasts. They classified their results in forecast accuracy by remarking on the situations where each implemented model works better. Similarly, Botha et al (2022) explore statistical learning techniques and traditional time-series models over a big data set of predictors to improve the accuracy of inflation forecast in South Africa. They find that statistical learning models provide similar forecast performance over mid to long-term horizons compared to several benchmark models, such as traditional random walk, autoregressive, and Bayesian vector autoregressive, dimensionality reduction, and variable selection techniques. Also, they underline that statistical learning models can explain non-linear relationships and are particularly useful during periods of crisis, which is a similar conclusion found by Joseph et al (2021).

The main goal of this work is to empirically evaluate whether forecasts for aggregated inflation can be improved by reducing forecast error by aggregating forecasts of components of the CPI basket compared with forecasts obtained directly from models for the aggregate. The contribution of this paper is as follows. First, we use a disaggregated data set of 183-time series items of the CPI basket in Colombia that is measured and disclosed by the National Department of Statistics (DANE, its Spanish acronym) as predictors to forecast headline, core, goods, and services inflation, similarly as Joseph et al (2021) and Botha et al (2022) do. Also, we employ a set of 80 macroeconomic indicators as predictors. Second, given the dimension of this set of predictors, we implement dimensionality reduction techniques (principal component Analysis (PCA) and Dynamic Factor Models (DFM)), variable selection models ((Ridge (Rid) and Lasso (Las)), and the Machine Learning model Random Forest (RF). Furthermore, we use traditional time series models such as ARIMA, SARIMA, and TAR[1].

Third, we forecast each of the 183-time series items of the CPI basket using lags of each time series and the set of macroeconomic indicators. Then, by aggregating the forecast of components of the CPI basket, we compare it with forecasts obtained directly from models for the aggregate. Then, to achieve this, first, we use direct and recursive forecast methods for the traditional ARIMA and ARIMAX models; see, for instance, Taieb et al (2012) and Molano et al (2007).

The remaining section of this paper is organized as follows. Section 2 provides a literature review for the Colombian case related to forecast aggregation. Section 3 describes the forecasting methodology and a brief overview of the implemented models. The data description of the CPI index in Colombia is provided in section 4, and the results of the evaluation forecast performance are presented in section 5. Section 6 concludes. The appendix provides supplementary results.

2 Review of Inflation forecasting in Colombia

In Colombia, several empirical forecasting aggregations have been performed to forecast inflation. Here are some of the main conclusions found by Melo-Velandia (2006): i) the fewer CPI components are considered to forecast the aggregate, the smaller the forecasting error due to fewer sources of errors. ii) the more homogeneity within

[1] Threshold autoregressive models, see Tong (1990)

the disaggregates and the more heterogeneity between the disaggregates, the better aggregate forecasts are obtained. iii) for the short-term horizons, it works better to aggregate forecast from the disaggregates than forecast the aggregate directly; however, for mid-to-long-term horizons, the opposite produces better results. iv) one advantage of having a forecast for disaggregates is that it allows one to identify the items that present the most significant price changes, the direction of those changes, and explore some causes of this behavior.

The forecasting team of the Colombian Central Bank works with a set of forecasting models for the aggregates of inflation: headline, core inflation (non-food inflation), and food inflation. This set of forecasts includes both aggregate models directly and from aggregating forecasts of the components of each CPI aggregate. In particular, headline inflation forecasts have been built by aggregating individual forecasts of food, housing, clothing, and miscellaneous groups. These forecasts do not surpass others obtained with different methodologies and univariate models for the aggregate. For the case of non-food inflation, a forecast has been obtained by aggregating forecasts of tradable, non-tradable, and administrated goods. In this case, the results are more promising. This forecast produces minor RMSFE compared to other forecasts obtained from univariate models for non-food inflation currently used by the Bank. Finally, several exercises of aggregating forecasts of different predefined food groups have been carried out for food inflation. Those classifications are G6 (potatoes, milk, meat, fruits, vegetables, and other food items), G10 (meats, dairy products, edible oils, meals away from home, potatoes, tubercles, fruits, vegetables, cereals, and processed food) and the G3 (perishable food, and meals away from home). Of the forecasts obtained from these three aggregations, the one that historically produces minor forecast error is the one constructed from the G3 classification, see González-Molano (2008).

More recently, Alonso and Rivera (2017) used the hierarchy of the CPI index in Colombia to forecast the inflation at different levels of disaggregation and by regions to aggregate forecasts for headline inflation. The authors implement SARIMA models to forecast each CPI item series. They conclude that in periods with asymmetric effects, the strategy of aggregating forecast behaves better than forecasting headline inflation directly. Also, Peña Ordóñez et al (2019) forecast headline inflation using ARIMAX models for the 12 CPI groups of goods and services and compare them with those obtained by the random forest model. She forecasts monthly headline inflation one step ahead by aggregating the forecast for each group using CPI weights. She found smaller RMFSE with the random forest than with the disaggregated ARIMAX.

3 Methodology

We split the data analysis into two parts: the first part considers the headline inflation, core inflation, services, and goods, as the response variable; this set is called the aggregate inflation index. The second part uses each disaggregated CPI item series as the response. We include the disaggregated CPI item series and a set of economic indicators as predictors for the aggregate set. The economic indicators are included as predictors for the models for the CPI items. We will introduce some notation and details about the models we fit for aggregate and disaggregate inflation.

Let y_t^a an aggregate of inflation (headline, core, goods, etc.) and let $\{y_{1t}, y_{2t}, \ldots, y_{nt}\}$ the set of n disaggregated inflation index such that for each t

$$y_t^a = \sum_{i=1}^{n} w_i y_{it} \quad \text{where} \quad \sum_{i=1}^{n} w_i = 1. \tag{1}$$

We modeled both y_t^a and y_{it} for $t = 1, ..., T$, and $i = 1, \ldots, n$. Thus given observations through 1 to T (1:T) the h-step-ahead inflation forecast is given by

$$\hat{y}^a_{T+h|1:T} = f\left(y_1^a, \ldots, y_T^a, \hat{\theta}\right), \tag{2}$$

where $f(\cdot)$ is a fitted function that captures the relationship between the observations and the target $\hat{y}^a_{T+h|1:T}$, and $\hat{\theta}$ is a vector of parameter estimates. Similarly, we have each of the disaggregated prices.

$$\hat{y}_{i,T+h|1:T} = f_i\left(y_{i1}, \ldots, y_{iT}, \hat{\theta}_i\right) \quad \text{for } i = 1, \ldots, n. \tag{3}$$

Also, we include a set of economic indicators $X_t = \{x_{1t}, \ldots, x_{pt}\}$, for each $t = 1, ..., T$, and the disaggregated data as predictors such that

$$\hat{\hat{y}}^a_{T+h|1:T} = g\left(y_1^a, \ldots, y_T^a; X_1, \ldots, X_T; Y_1, \ldots, Y_T, \hat{\hat{\theta}}\right); \quad i = 1, 2, \ldots, n, \tag{4}$$

where $Y_t = \{y_{1t}, y_{2t}, \ldots, y_{nt}\}$, for $t = 1, ..., T$. Alternatively, we use dimension reduction techniques such as principal components analysis (PCA) and dynamic factor model (DFM) over the set of economic indicators and PCA over the disaggregated data set. For the disaggregated data, we also include the set of economic indicators in this way.

$$\hat{\hat{y}}_{i,T+h|1:T} = g_i\left(y_{i1}, \ldots, y_{iT}; X_1, \ldots, X_T, \hat{\hat{\theta}}_i\right). \tag{5}$$

Furthermore, we aggregate forecast from the disaggregated CPI items in (3) or (5) such that

$$\tilde{y}^a_{T+h|1:T} = \sum_{i=1}^{n} w_i \hat{y}_{i,T+h|1:T}. \tag{6}$$

Our goal is to forecast the aggregate as long as we minimize the root mean squared forecast error (RMSFE) under (2), (4) or (6). Thus we have

$$\begin{aligned} \text{RMSFE}_1^h &= \left[E\left(\left(y^a_{T+h} - \hat{y}^a_{T+h}\right)^2\right)\right]^{1/2} \quad \text{under (2) or (4), and} \\ \text{RMSFE}_2^h &= \left[E\left(\left(y^a_{T+h} - \tilde{y}^a_{T+h}\right)^2\right)\right]^{1/2} \quad \text{under (6).} \end{aligned} \tag{7}$$

The statistics used to evaluate the out-of-sample performance of every model is the RMSFE. Also, we support this evaluation by using the Diebold and Mariano (1995) test with the correction for short samples Harvey et al (1997).

3.1 Fitted models

We fit three classes of autoregressive models: ARIMA models (M0) and threshold autoregressive models (TAR) Tong (1990). Dimension reduction: principal component analysis (PCA) and dynamic factor analysis (DFA) Stock and Watson (2002a,b). Variable selection models: Ridge (Rid), Lasso (Las) Friedman et al (2010), and a nonlinear machine learning model, Random Forest (RF) Breiman (2001).

3.1.1 Autorregresive models

To fit the f and f_i functions in (2) and (3), we search for the best ARIMA model with a seasonal component according to the AIC criterion, see Hyndman and Khandakar (2008). This automatic procedure allows the rapid estimation of models for many time series variables. Also, we implement threshold autoregressive models (TAR) with two regimes Tong (1990) by following the procedure developed by Kung-Sik Chan, see Cryer and Chan (2008).

3.1.2 Dimension reduction models

The inclusion of a large set of predictors in the model (4) and (5) prevent us from using the standard ARIMA model with exogenous variable ARIMAX, see Shumway and Stoffer (2017) for instance. However, the dimension reduction techniques like PCA or DFM are an alternative to reduce the number of predictors and then use the classic ARIMAX approach to fit the functions g and g_i in (4) and (5), respectively. Nevertheless, when the dimension of the set of predictors is high in practice, DFM is not a good choice. That is the case of our set of disaggregated item series of the CPI.

3.1.3 Variable selection models and nonlinear learning models

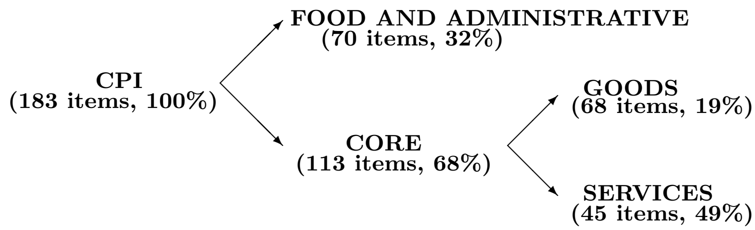
Alternatively, to tackle the dimension of the set predictor's problem, we appeal to the variable selection models, particularly the Ridge regression that imposes L_2 penalties to shrink the coefficients of predictors that contribute little to the predictive ability. Similarly, Lasso regression imposes L_1 penalties by adding the sum of absolute coefficients to the sum of squared residuals. Thus, some of the coefficients of the predictors are shrunken to exactly zero; see the detail of these two models in Friedman et al (2010). Finally, we implement the Random forest regression model of Breiman (2001) on the domain of nonlinear learning models. This model approximates an unknown nonlinear function partitioning the parameter space of the predictors recursively to minimize the squared errors in the fitted model.

4 Data

In Colombia, the Consumer Price Index (CPI) is measured and disclosed by the National Department of Statistics (DANE, its Spanish acronym). The last update to the items included in the CPI basket was in January 2019, letting as the reference base December of 2018[2]. The CPI tracks the behavior of prices of consumer goods

[2]Since 1978, in Colombia every ten years, the measurement methodology of the CPI is updated, see $https://www.dane.gov.co/files/investigaciones/fichas/metodologia_IPC-09.pdf$

Fig. 1 Structure of the CPI-2018, relationship among disaggregated CPI item series and aggregated components. Source: National Department of Statistics-DANE



and services that seek to represent average household expenditure. The selection of items included in the CPI basket considered the results of the National Household Budget Survey (ENPH, its Spanish acronym) developed between July 2016 and July 2017. The current CPI is composed of 443 units that are the source of 188 items, the minimum desegregated level available and released each month by the DANE. This information comes from different sources in 38 cities where the Colombian population buy goods and services[3].

In the analysis in the coming section, we are interested in headline inflation, and core inflation. Also, we emphasize two components of core inflation: Goods and Services. The core inflation in Colombia represents 68% of the full basket; see Figure 1. In Figure 1, we include the compositions of the CPI where we remark the number of items and their weight by aggregate (headline, core, goods, and services). We emphasize the number of items per aggregate because the analysis that we developed here includes forecast aggregation from the desegregated forecasts.

Several authors have found great utility and predictive power when using an extensive battery of economic variables to obtain inflation forecasts, mainly see Stock and Watson (2002b). In line with the above, in this document, in the inflation forecasting exercises, the information content of 80 variables of various types will be widely used by national and international analysts to follow and analyze the primary economic aggregates and price dynamics. Most of these variables are characterized because they are public, easy to follow, and available. The set of variables includes (See Table A1): consumer prices, producer prices, housing prices, production, the housing market, foreign trade, labor market, monetary aggregates, sector variables finance, trade sales, internal and external interest rates, exchange rates, international prices of basic goods, economic expectations, etc.

[3]See, DANE (2019), Metodología general del Índice de Precios al Consumidor -IPC-, Bogotá, Colombia (in Spanish).

5 Results and discussion

We evaluate the forecast performance of the models relative to the model with the minimum RMSFE per horizon (for the aggregate forecast group, see Table 2, for instance), named the "best one". We compare the RMSFE of the other models with the best one, per horizon, by using the Diebold and Mariano (1995) test for equal forecast ability. To evaluate the performance of the models before and after the beginning of COVID-19, using as reference January 2020, we split our out-of-sample forecasting period from March 2017 to November 2022 into two parts. Thus, since we use recursive out-of-sample forecasting from March 2017 to November 2022 (top panel on Table 2), we split this period from March 2017 to December 2019 and from January 2020 to November 2022 (middle and bottom panel on Table 2, respectively).

We forecast using the models described in section 3. We use direct and iterative forecast approaches, as summarized in Table 1. These models were used to forecast the aggregates of inflation for the Headline, Core, Goods, and Services. Also, we used the same models to forecast at the disaggregated level as we described in section 4 for the 183-time series items of the CPI basket (see Figure 1) to the subsequently forecast aggregation at the Headline, Core, Goods and Services level. Thus, we compare the performance of the aggregates (left panel on Table 2, for the Headline inflation) and the performance of the aggregates forecast from the disaggregates ones (right panel on Table 2). We only used it as a reference to compare the best model for the Aggregates because one of our goals is to determine whether the forecast aggregation (from the disaggregates) performs better than the Aggregates.

Variable selection models $RFYLX$, $RidYLX$, and $LasYLX$ (see Table 1) include the full data set of 183-time series items of the CPI basket (Ys) for the Aggregates while for disaggregates the Ys set correspond to the lags 1 to 12 of the fitted time-series itself.

5.1 Headline

In Table 2, we compare the forecast performance for all the implemented models for Headline inflation. The top panel in table 2 includes the RMSFE for each model for the horizons 1, 3, 6, 9, and 12; and for the out-of-sample forecasting period from March 2017 to November 2022. With bold format, we remark the model with minimum RMSFE for every horizon. Also, these results include the Diebold and Mariano (1995) test, where the * symbols represent the significance level for this test, compared to the model with minimum RMSFE. Thus, on the top panel for horizon 1, the model with the minimum RMSFE is the TAR model (see Table 1). Regarding the other models at the Aggregate forecast, only the $M0$, the PY, and the $PYFLX1$ models have a similar performance. In the case of the forecast aggregation (From disaggregates), all these models have worse performance compared with the model with the minimum RMSFE at horizon 1. Hence, there are no advantages in aggregating forecasts at this short-term horizon. For horizon 3, the model with the minimum RMSFE is the PY model. In this case, most models have the same forecast ability as the one with the minimum RMSFE, except for four models that include Xs variables and the ARp model, and for the disaggregates, the dynamic factors models have similar performance to the best aggregate model. For horizon 6, both the Aggregates and the disaggregates have similar behavior, i.e., most models have similar forecasting performance to the model with minimun RMSFE. Finally, for horizons 9 and 12, all the disaggregate models have a similar performance to the best Aggregate one.

When we split the evaluation period from March 2017 to December 2019 and from January 2020 to November 2022 (middle and bottom panel on Table 2, respectively), we identify the improvement from the disaggregates mainly for the horizons 6, 9 and 12. Even for some of the models, we notice that the RMSFE is smaller for the disaggregates than the best aggregate model[4], and according to DM test, the disaggregate models perform similarly to the best Aggregate model for horizons 6, 9, and 12. However, a remarkable advantage is that the forecast from the disaggregates has a better behavior in terms of variability, as we can see in Figures 2. For all the horizons, the forecasts from the disaggregate compared to the Aggregates are less volatile. As for the pandemic period, we found higher forecasting errors, as expected (bottom panel, Table 2).

The other important result is that for the whole period of out-sample forecasts, the best aggregate model for all the horizons is a univariate one, which includes only CPI items, not other additional explanatory variables. In particular, for the whole out-sample, in the short-run horizon ($h = 1$), it stands out the TAR model, and for horizons 9 and 12. While, for the pandemic period, the best forecasting model for horizons 6, 9, and 12 month-ahead is a multivariate dynamic factor model which contains additional explanatory variables ($PYFLX1$, $PYFLX7$, $PYFLXD$). This fact is emphasized in the pre-pandemic period, where the models with additional explanatory variables do not contribute to improving the forecast, whereas for the posterior period, the additional predictors help to improve the forecasts. As we found for the whole sample, also for $h = 1$, the best model is still the TAR, which may be able to capture some non-linearities.

[4] The big value for the TAR model for the disaggregates at horizon 12 happened due to outliers forecast for some of the time series items.

Model	Description	Forecast method
M0	Best ARIMA model according to AIC criterion	Iterative
TAR	Treshold autorregresive model (two regimes)	Iterative
ARp	AR(12) model	Direct
PY	PCA of the set of 181 CPI item series (Ys), fitted using ARIMAX	Iterative
M1	Best SARIMA model according to AIC criterion	Iterative
PYLX	PCA of the sets Ys and Log(Xs), fitted using ARIMAX	Iterative
PYLXD	PCA of the sets Ys and Log(Xs), fitted using ARIMAX	Direct
PYFLXD	PCA of the set Ys and DFM of the set Log(Xs), fitted using ARIMAX	Direct
PYFLX1[a]	PCA of the set Ys and DFM of the set Log(Xs), fitted using ARIMAX	Iterative
PYFLX7[b]	PCA of the set Ys and DFM of the set Log(Xs), fitted using ARIMAX	Iterative
RFYLX	Random Forest with the sets Ys[c] and Log(Xs)	Direct
RidYLX	Ridge with the sets Ys[c] and Log(Xs)	Direct
LasYLX	Lasso with the sets Ys[c] and Log(Xs)	Direct
PYXIPP	PCA of the sets Ys and Log(Xs), and disaggregate IPP data	Iterative
PYFXIPP	PCA of the set Ys, DFM of Log(Xs), and disaggregate IPP data	Iterative

Table 1 Fitted models: Principal components analysis (PCA), dynamic factor model (DFM). [a]one factor by following the suggestion of Stock and Watson (2002a), [b]seven factors by mimicking the explained variance obtained in the PCA analysis. [c] For the Aggregates correspond to the full set of 183 CPI item series, and the disaggregates correspond to the lags 1 to 12 of the fitted time series itself.

Headline	Aggregates					From disaggregates				
Model	1	3	6	9	12	1	3	6	9	12
M0	0.34	0.91	**1.68**	2.34	3.1	0.66***	1.26**	2.07	2.82	3.42
TAR	**0.34**	0.94	1.79	**2.26**	**2.78**	0.52***	1.09*	1.99	10.58	78.76
ARp	0.68***	1.6**	2.8	4.14	5.23	0.85***	1.34***	2.54*	4.27	5.49
PY	0.35	**0.88**	1.7	2.35	3.09	0.56***	1.15**	1.96	2.73	3.36
M1	1.46*	1.88	2.13	2.44	3.16	0.68***	1.26*	2.11	2.93	3.58
PYLX	0.49*	1.08*	2.09	3.21*	4.71	0.46***	0.92	1.54	2.17	2.94
PYLXD	1.45**	2.01***	4.08***	3.79***	7.38**	0.56***	1.01**	1.61	2.22	3.01
PYFLXD	0.71***	1.39***	3.15***	5.24**	5.49*	0.44***	0.87	1.43	2.27	3.2
PYFLX1	0.34	0.96	1.71	2.38	3.32	0.48***	0.96	1.62	2.32	3.13
PYFLX7	0.4*	1.07	2.08	2.85	3.92	0.45***	0.93	1.57	2.27	3.05
RFYLX	0.88**	1.51	2.21	2.89*	3.07	1***	1.57***	2.41	3.08	3.59
RidYLX	0.97**	1.47	2.16	2.48	3.13	1.08***	1.64***	2.36*	3.08	3.55
LasYLX	1.73**	2.22	2.58	2.75	3.24	1.35***	1.92**	2.58	3.14	3.58
PYXIPP	-	-	-	-	-	0.48***	0.92	1.5	2.16	2.91
PYFXIPP	-	-	-	-	-	0.49***	0.97	1.6	2.27	3
M0	0.23	0.47	0.66	0.89	1.13	0.41**	0.58	0.72	0.95	1.19
TAR	0.22	0.46	0.64	0.84	**1.05**	0.46***	0.63	1.02	1.42	3.44
ARp	0.4***	0.68**	1.72	3.36	5.23	0.61***	0.95***	1.26***	1.77***	2.32**
PY	0.25*	0.51**	0.73	0.95*	1.19	0.44**	0.65	0.83	1.09	1.37
M1	**0.16**	0.4	0.78	1.16***	1.41	0.44***	0.66**	1.01*	1.38	1.74
PYLX	0.25**	0.63**	1.29**	2.11**	2.81	0.4**	0.56	0.7	0.98	1.25
PYLXD	0.87***	2.09**	4.65***	3.26***	8.54	0.43**	0.6	0.78	1	1.31
PYFLXD	0.72***	1.14***	3.27***	5.68**	6.12	0.37**	0.51	0.62	0.94	1.24
PYFLX1	0.25*	0.65***	1.28***	1.88***	2.44**	0.4**	0.57	0.69	0.93	1.16
PYFLX7	0.26	0.73***	1.57***	2.41***	3.34**	0.37**	0.5	0.62	0.9	1.16
RFYLX	0.21	0.4	0.79	0.99	1.22	0.67***	0.9**	1.1	1.32	1.59
RidYLX	0.2	**0.38**	0.71	**0.76**	1.21	0.69***	0.89**	1.08*	1.29	1.56
LasYLX	0.21	0.43	**0.61**	0.94	1.35	0.67***	0.84*	1.01	1.31	1.68
PYXIPP	-	-	-	-	-	0.41**	0.59	0.72	0.99	1.28
PYFXIPP	-	-	-	-	-	0.4**	0.54	0.66	0.95	1.23
M0	0.42	1.21	2.36	3.41	4.69**	0.83***	1.7**	2.95	4.14*	5.19
TAR	0.42	1.27	2.54	3.29	**4.19**	0.58**	1.43	2.72	15.99	123.92
ARp	0.88***	2.18**	3.68	4.98	5.22	1.03***	1.65	3.48	6.16	8.17
PY	0.42	**1.14**	2.38	3.4	4.65*	0.66***	1.51	2.74	3.95*	5.03*
M1	2.05**	2.66	3.02	3.47	4.68	0.86***	1.69	2.92	4.16	5.23
PYLX	0.65	1.41	2.74	4.24	6.59	0.52**	1.19	2.13	3.1	4.38
PYLXD	1.86*	1.93*	3.27***	4.39	5.23	0.66**	1.32	2.22	3.17	4.46
PYFLXD	0.71***	1.62*	3.02*	4.61	4.41	0.49*	1.13	2	3.27	4.8
PYFLX1	**0.42**	1.21	**2.1**	**2.91**	4.31	0.55***	1.25	2.26	3.37	4.72
PYFLX7	0.51**	1.34	2.55*	3.34	4.65	0.53**	1.23	2.22	3.28	4.59
RFYLX	1.22**	2.13	3.15	4.24	4.6	1.25***	2.06**	3.34	4.42	5.3
RidYLX	1.36**	2.07	3.09	3.67	4.7	1.36***	2.17**	3.27	4.44	5.26
LasYLX	2.44**	3.16	3.74	4.04	4.82*	1.79***	2.61**	3.64	4.54	5.26
PYXIPP	-	-	-	-	-	0.54**	1.17	2.07	3.08	4.3
PYFXIPP	-	-	-	-	-	0.58***	1.28	2.24	3.28	4.49

Table 2 Root mean squared errors, in bold format, the minimum RMSFE per horizon. The list of *** symbols refers to the Diebold and Mariano test, which means statistical difference concerning the model with minimum RMSFE (* p <5%, * p <1%, and *** p <0.1%). We use a sample period 2011-2022, out-of-sample predictions from Mar2017 until Nov2022 (top panel), out-of-sample predictions from Mar2017 until Dec2019 (middle panel), and out-of-sample predictions from Jan2020 until Nov2022 (bottom panel). The description of the models is given in table 1.

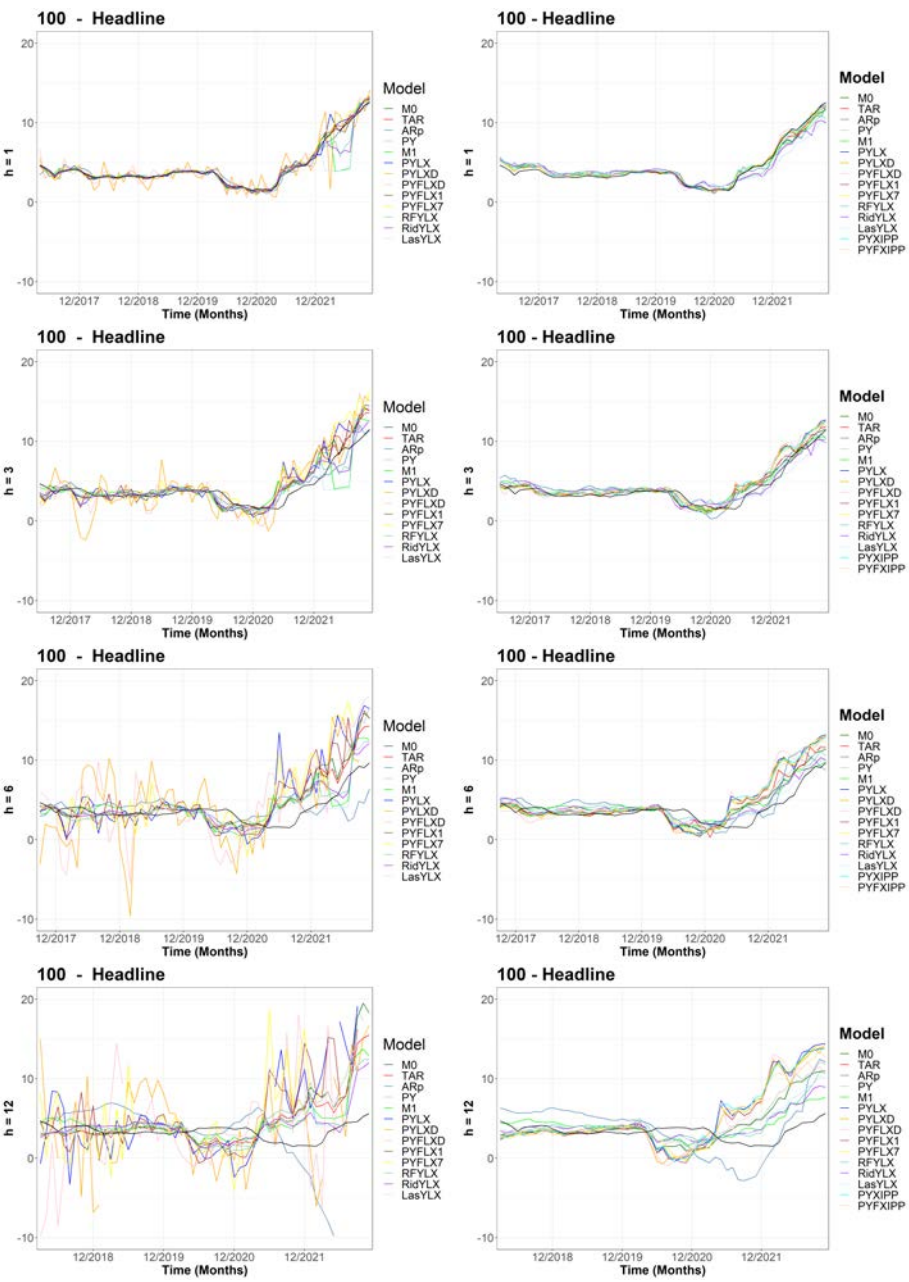


Fig. 2 Forecasts for the headline CPI inflation (coloring lines) compared to the observed headline inflation (black lines), lagged by h months. Forecast from aggregates models left panel and disaggregates models right panel. Out-of-sample predictions from Mar2017 to Nov2022 and for the horizons $h = 1, 3, 6, 9, 12$.

5.2 Core

For the complete out-sample evaluation period, see Table 3, we have found that for $h = 1$, the TAR model again has the minimum RMSFE and other models like the Random Forest (RF), Rid and Las have similar forecasting ability along with the univariate models $M0$ and $M1$. For all the disaggregate models but Rid and Las, the test of equal forecast ability is not rejected compared to the best aggregate model. For $h = 3$, the best forecasting model is $M1$, and all the other models seem to perform worse. Some univariate models compete in performance ($M0$, $M1$, TAR) for the disaggregated models, and others like PY and $PYFXIPP$. For $h = 6$, 9, and 12, the best model is the Random Forest (RF) and appears in models with equal forecasting ability like Rid and Las, along with $M1$ and TAR. Nevertheless, almost all models have similar performance for the disaggregated models. Again, we found that the forecasts from the disaggregated models are less volatile than the ones obtained by the aggregated models, see Figure 3.

The results for the pandemic period are quite similar to those for the whole sample. Same models with minimum RMSFE, except for the six-month-ahead horizon, the $M1$ model became better than the Random forest—higher forecasting errors than for the total and pre-COVID-19 samples. Less volatile forecasts are obtained from the disaggregated models, and most have equal forecasting ability to the best aggregate model for horizons further than six month-ahead.

In contrast with the results for the headline inflation, for the full out-sample evaluation, most forecasts obtained from the disaggregates models for horizon ($h = 1$) have a similar performance to the best aggregate one since the core inflation does not include the most volatile items in the CPI basket (food and administrated items). For horizons over the six months, the RF model performs best for both the Aggregates and the disaggregates models. This fact coincides with the results of the South African Study, see Botha et al (2022).

We also compared the performance of all models with traditional autoregressive and dimension reduction aggregates models by choosing one as a benchmark. The comparison is made by calculating the relative RMSFE regarding the benchmark. Relative RMSFE values below one imply a better forecast performance and values over the opposite. In addition, the comparison includes the Diebold and Mariano (1995) that allows us to determine if the forecast accuracy is significant. For instance, in Table A2 for the headline inflation, we use the aggregates AR(p) model as a benchmark (top panel). Likewise, we use the aggregates $M0$, $M1$, $PYFLXD$, and $PYFLX1$ models as benchmarks. In Table A3, we include these results for the Core inflation. From these results, we found that the traditional aggregate models that use the direct forecast method have poor forecast accuracy. Regarding the accuracy of the disaggregate models, we obtain a better forecast performance, although these results are not significant for the models $M0$, $M1$, and $PYFLX1$ for headline inflation. However, for the core inflation, we observe a significant improvement in the forecast accuracy except when we use the $M1$ as a benchmark.

Core	Aggregates					From disaggregates				
Model	1	3	6	9	12	1	3	6	9	12
M0	0.26	0.72**	1.35**	1.96*	2.56**	0.57	0.95	1.5	1.97	2.34
TAR	**0.25**	0.72**	1.43**	2.04	2.66	0.29	0.68	1.48	10.09	115.3
ARp	0.57***	1.15**	1.99	2.53	3.12**	0.55	1.02**	2.52	4.26	4.97
PY	0.27*	0.75**	1.36**	1.98*	2.6**	0.39	0.82	1.39	1.94	2.36
M1	0.57	**0.59**	1.2	1.81	2.4***	0.54	0.86	1.38	1.93*	2.42***
PYLX	0.32***	0.87**	1.81*	2.95*	4.13*	0.27	0.6***	1.05	1.61	2.17
PYLXD	0.83***	1.21***	2.15***	2.87*	4.53*	0.33	0.62*	1.1	1.66	2.15
PYFLXD	0.52***	1.06***	1.96**	3.21**	4.76**	0.28	0.6**	1.17	1.69	2.37
PYFLX1	0.29	0.85***	1.71**	2.59*	3.54*	0.29	0.62***	1.1	1.69	2.26
PYFLX7	0.32*	0.97***	2.04***	3.11**	4.17**	0.29	0.63*	1.07	1.6	2.14
RFYLX	0.48	0.69*	**1.18**	**1.7**	**2.11**	0.64	1.09**	1.78*	2.3	2.67
RidYLX	0.6	0.71**	1.31	1.77	2.32	0.75*	1.18**	1.81*	2.28	2.61
LasYLX	0.71	0.82**	1.44*	2.01	2.46	0.86**	1.29**	1.87	2.28	2.59
PYXIPP	-	-	-	-	-	0.28	0.62**	1.08	1.64	2.19
PYFXIPP	-	-	-	-	-	0.31	0.67	1.12	1.65	2.17
M0	0.16*	0.46	0.86***	1.21**	1.63	0.14**	0.36**	0.69	0.96	1.28
TAR	0.17*	0.47	0.75	0.96	1.32	0.18	0.39	0.81	1.26*	2.18**
ARp	0.27***	**0.37**	0.7	1.18	2.26***	0.26**	0.51	0.9	1.54	2.26
PY	0.18**	0.49	0.85***	1.2**	1.6	0.14**	0.36*	0.68	0.95*	1.29
M1	**0.13**	0.4	0.87	1.28*	1.69	0.14	0.38	0.83	1.24***	1.67***
PYLX	0.28***	0.9**	1.98*	3.24*	4.53*	0.14**	0.35*	0.66	0.93	1.24
PYLXD	0.72***	1.13***	1.75**	2.03**	5.11*	0.16	0.35*	0.7	0.98	1.24
PYFLXD	0.42***	1.02***	2.01**	2.55**	5.41***	0.14**	0.35	0.7	1.07	1.26
PYFLX1	0.24***	0.8**	1.71*	2.61*	3.56	0.13**	0.33*	0.63	0.86	1.12
PYFLX7	0.25***	0.82***	1.81*	2.85*	3.91	0.13**	0.33**	0.62	0.87	1.13
RFYLX	0.18**	0.43	0.7	1.05	**1.25**	0.26***	0.52**	0.91	1.21	1.73
RidYLX	0.18*	0.44	**0.7**	**0.89**	1.39	0.33***	0.53***	0.94	1.28	1.69
LasYLX	0.2**	0.5	0.88	1.23	1.65	0.32***	0.52***	0.91	1.27	1.68
PYXIPP	-	-	-	-	-	0.14**	0.37	0.68	0.95	1.27
PYFXIPP	-	-	-	-	-	0.14**	0.35*	0.64	0.89	1.18
M0	0.33	0.91**	1.76**	2.64	3.51*	0.79	1.32*	2.08	2.79	3.35*
TAR	**0.32**	0.91**	1.95***	2.9***	3.87	0.37	0.89	1.99	15.26	181.5
ARp	0.75***	1.6**	2.83	3.6	4.06	0.73	1.36**	3.58	6.22	7.33
PY	0.34	0.96**	1.78**	2.67*	3.6**	0.53	1.12*	1.91	2.75	3.37*
M1	0.8	**0.74**	**1.5**	2.32	3.16	0.75	1.18	1.83	2.57	3.22
PYLX	0.36	0.84	1.58	2.52	3.47	0.35	0.78**	1.38	2.21	3.07
PYLXD	0.92***	1.28*	2.53**	3.69	3.51	0.45	0.82	1.44	2.26	3.03
PYFLXD	0.6**	1.09**	1.89	3.9	3.59	0.37	0.78	1.54	2.26	3.4
PYFLX1	0.32	0.89**	1.7	2.56	3.5	0.39	0.83**	1.48	2.37	3.28
PYFLX7	0.38	1.11*	2.28	3.43	4.52	0.39	0.84	1.43	2.22	3.07
RFYLX	0.66	0.88*	1.56	**2.29**	**2.95**	0.87	1.48***	2.43*	3.21	3.63
RidYLX	0.82	0.92*	1.78	2.49	3.24	1	1.6**	2.46	3.14	3.55
LasYLX	0.99	1.06**	1.9*	2.71	3.32	1.17**	1.78**	2.57	3.14	3.53
PYXIPP	-	-	-	-	-	0.37	0.81*	1.41	2.24	3.08
PYFXIPP	-	-	-	-	-	0.41	0.89	1.49	2.3	3.11

Table 3 Root mean squared errors, in bold format, the minimum RMSFE per horizon. The list of *** symbols refers to the Diebold and Mariano test, which means statistical difference concerning the model with minimum RMSFE (* p <5%, * p <1%, and *** p <0.1%). We use a sample period 2011-2022, out-of-sample predictions from Mar2017 until Nov2022 (top panel), out-of-sample predictions from Mar2017 until Dec2019 (middle panel), and out-of-sample predictions from Jan2020 until Nov2022 (bottom panel). The description of the models is given in table 1.

5.3 Goods and Services

Similar results are obtained for goods and services CPI aggregates; for short-term horizons, the best model is a univariate $M1$ or TAR, and only a few other models, both aggregated and disaggregated, have equal forecast ability, see Tables 4 and 5. It seems that for the aggregate models is not necessary to use additional information, apart from the own lags and components of CPI, to improve the forecast, especially when using direct forecasting. Additionally, for the short-run, methodologies that make variable selections like Random Forest, Ridge, and Lasso do not contribute to improving the forecast for the models.

For further horizons, a multivariate model with additional explanatory variables has the minimum RMSFE, the RF, and almost all disaggregated models performed and the best aggregate model, even most with smaller RMSFE. During the pandemic period, most of the results obtained for the headline and core inflation still hold for goods and services. For services, no other aggregates model could beat the $M1$ for all horizons, and for $h = 1$, there is not a model, aggregate of disaggregated with equal forecast ability as the $M1$. Finally, as we found for the Headline and the Core inflation cases, the forecasts from the disaggregated models are less volatile than the ones obtained by the aggregated models; see Figures 4 and 5 for Goods and Services, respectively.

Goods	Aggregates					From disaggregates				
Model	1	3	6	9	12	1	3	6	9	12
M0	0.53	1.16	2.17*	3.13	4.34*	0.98**	1.51	2.17	2.92	3.61
TAR	0.51	1.12	1.95	2.98***	3.93**	0.53	1.39**	3.89	35.69	415.42
ARp	0.79**	1.37	3.27	4.25	4.63*	1.01***	1.98**	3.21	4.21	4.88
PY	0.54	1.15	2.15**	3.22**	4.44***	0.68**	1.31*	2.13	3	3.69
M1	**0.48**	**1.11**	2.1*	3.17**	4.31***	0.98**	1.43	2.04	3.01	3.79
PYLX	0.56	1.19	2.3	3.72	4.95	0.53	1.06	1.8	2.61	3.45
PYLXD	1.07***	1.64**	3.16***	5.02***	11.17	0.59	1.06	1.85	2.76	3.6
PYFLXD	0.89***	1.47*	3.09**	5.68**	6.86**	0.51	0.96	1.72	2.46	3.47
PYFLX1	0.58	1.33	2.36**	3.5**	4.44	0.56	1.12	1.93	2.9	3.79
PYFLX7	0.55	1.21	2.14	3.1	4.34	0.54	1.08	1.87	2.71	3.47
RFYLX	0.53	1.14	**1.8**	**2.47**	**3.27**	0.94***	1.48*	2.35*	3.21	3.77
RidYLX	0.51	1.18	1.95	2.65	3.37	1.22**	1.68**	2.41	3.14	3.71
LasYLX	0.56	1.26	2.12	2.95	3.59	1.32***	1.87*	2.52	3.12	3.66
PYXIPP	-	-	-	-	-	0.53	1.08	1.8	2.62	3.38
PYFXIPP	-	-	-	-	-	0.58	1.16	1.92	2.76	3.47
M0	0.19	**0.58**	1.08	1.54	2.29	0.29	0.47	0.8	1	1.35
TAR	0.22	0.64	1.33**	1.88**	2.37**	0.25	0.6	1.07	2.44	6.43
ARp	0.47***	0.9*	1.74	2.6	3.51**	0.32*	0.69	1.36	2.16	3.42***
PY	0.26**	0.76**	1.45**	2.01	2.81**	0.29	0.53	0.9	1.13	1.5
M1	**0.19**	0.66	1.44	2.1	2.76*	0.24	0.35*	0.9	1.38	1.87
PYLX	0.35***	1.05***	2.17	3.52	5.09	0.31	0.53	0.9	1.13	1.58
PYLXD	0.89***	1.69***	2.79**	5.19**	13.9	0.31	0.51	0.86	1.16	1.58
PYFLXD	0.56***	1.31**	1.82**	4.38	6.39	0.3	0.53	0.92	1.17	1.46
PYFLX1	0.3**	0.89***	1.75**	2.43*	3.52**	0.34*	0.54	0.9	1.17	1.63
PYFLX7	0.34**	1.04***	1.9**	2.75**	3.99**	0.32	0.51	0.86	1.1	1.51
RFYLX	0.24	0.62	0.99**	1.27	**1.56**	0.24	0.55	1.01	1.33	1.83
RidYLX	0.25	0.67	**0.9**	**1.22**	1.62	0.2	0.56**	1.06	1.4	1.88
LasYLX	0.28*	0.7	1.15	1.51	1.93	0.19	0.53**	1.01	1.39	1.84
PYXIPP	-	-	-	-	-	0.3	0.56	0.9	1.16	1.62
PYFXIPP	-	-	-	-	-	0.31	0.56	0.89	1.14	1.62
M0	0.72	1.55	2.97	4.42	6.24	1.36**	2.12	3.09	4.29	5.43
TAR	0.69	1.46	2.49	3.99**	5.47	0.7	1.9**	5.61	54.15	653.94
ARp	1.02*	1.73	4.44	5.73	5.91	1.39***	2.75**	4.5	5.9	6.47
PY	0.71	1.44	2.75	4.32*	6.1	0.92*	1.8	2.98	4.37	5.52
M1	**0.65**	1.45	2.68	4.16*	5.9	1.37**	2.03	2.84	4.29	5.52
PYLX	0.71	**1.33**	2.43	3.98	**4.73**	0.68	1.42	2.47	3.74	5.08
PYLXD	1.23***	1.59	3.54*	4.78	4.87	0.78	1.43	2.56	3.99	5.33
PYFLXD	1.13***	1.63	4.11**	7.03	7.49	0.66	1.27	2.33	3.49	5.17
PYFLX1	0.77	1.68	2.91	4.53	5.51	0.72	1.52	2.68	4.2	5.63
PYFLX7	0.71	1.36	**2.4**	3.51	4.81	0.7	1.46	2.59	3.92	5.15
RFYLX	0.71	1.51	2.42	**3.46**	4.78	1.3***	2.06	3.28	4.64	5.51
RidYLX	0.68	1.54	2.71	3.78	4.92	1.72**	2.34*	3.37	4.49	5.37
LasYLX	0.74	1.66	2.87	4.13	5.14	1.86***	2.63*	3.55	4.47	5.32
PYXIPP	-	-	-	-	-	0.69	1.43	2.46	3.75	4.95
PYFXIPP	-	-	-	-	-	0.76	1.56	2.66	3.98	5.09

Table 4 Root mean squared errors, in bold format, the minimum RMSFE per horizon. The list of *** symbols refers to the Diebold and Mariano test, which means statistical difference with respect to the model with minimum RMSFE (* p <5%, * p <1%, and *** p <0.1%). We use a sample period 2011-2022, out-of-sample predictions from Mar2017 until Nov2022 (top panel), out-of-sample predictions from Mar2017 until Dec2019 (middle panel), and out-of-sample predictions from Jan2020 until Nov2022 (bottom panel). The description of the models is given in table 1.

Services	Aggregates					From disaggregates				
Model	1	3	6	9	12	1	3	6	9	12
M0	0.21**	0.6*	1.11	1.62	2.1	0.53**	0.86	1.32	1.7	1.99
TAR	0.24**	0.7**	1.31	1.78	2.25	0.26	0.57	1	2.73	12.49
ARp	0.46***	0.88*	1.63	2.3	2.82	0.45***	0.78	2.45	4.34	5.11
PY	0.22***	0.6*	1.11	1.62	2.12	0.35**	0.7	1.19	1.64	2
M1	**0.18**	**0.52**	**1.02**	**1.51**	2.03	0.51*	0.78	1.22	1.63	2.07
PYLX	0.28***	0.72**	1.43**	2.27**	3.2*	0.24	0.5**	0.86	1.33	1.82
PYLXD	0.78***	1.07***	1.91***	2.7	3.5**	0.31**	0.56	0.94	1.37	1.76
PYFLXD	0.58***	0.97***	1.96**	2.86***	4.17*	0.25	0.54	1.02	1.55	2.06
PYFLX1	0.27***	0.79**	1.46*	1.89	2.58	0.25	0.51*	0.88	1.34	1.83
PYFLX7	0.34***	1.05**	2.08*	2.85	3.99	0.25	0.53	0.86	1.29	1.77
RFYLX	0.24***	0.58	1.07	1.58	**1.95**	0.61***	1.01**	1.61*	2	2.32*
RidYLX	0.23***	0.6*	1.1	1.54	2	0.69***	1.08**	1.62*	2*	2.26*
LasYLX	0.27***	0.68**	1.19	1.66	2.04**	0.77***	1.15**	1.66*	2.01*	2.25*
PYXIPP	-	-	-	-	-	0.24	0.51*	0.87	1.34	1.84
PYFXIPP	-	-	-	-	-	0.26	0.54	0.89	1.33	1.8
M0	0.18*	0.4	0.71	1	**1.32**	0.17	0.35	0.7	1.04	1.38
TAR	0.2**	0.46	0.81**	1.12	1.5	0.19	0.35	0.71	1.04	1.37
ARp	0.29***	**0.32**	**0.53**	**0.95**	1.54	0.25*	0.43	0.74	1.34	1.93
PY	0.19*	0.42	0.72	1.01	1.32	0.17	0.34	0.69	1.02	1.39
M1	**0.15**	0.38	0.81***	1.23	1.66	0.16	0.38	0.82	1.24	1.69
PYLX	0.26***	0.68***	1.39**	2.25*	3.16	0.17	0.33	0.62*	0.92	1.23
PYLXD	0.73***	0.79**	1.4**	1.68**	3.68**	0.2	0.32	0.7	1.01	1.22
PYFLXD	0.42***	0.93***	2.19**	2.69*	4.62	0.16	0.33	0.68	1.11	1.29
PYFLX1	0.23***	0.54***	0.98***	1.42***	1.89	0.15	0.33	0.62	0.9	1.16**
PYFLX7	0.29**	0.75**	1.48*	2.14	2.77	0.15	0.32	0.62*	0.9**	1.16**
RFYLX	0.17*	0.4	0.73	0.96	1.37	0.31***	0.5*	0.87*	1.18	1.7
RidYLX	0.18*	0.4	0.68	0.99	1.45	0.4***	0.54*	0.9	1.24	1.63
LasYLX	0.19**	0.42	0.74	1.07	1.46	0.38***	0.51*	0.87	1.23	1.62
PYXIPP	-	-	-	-	-	0.16	0.33	0.62*	0.92	1.23
PYFXIPP	-	-	-	-	-	0.15	0.33	0.63**	0.92**	1.19**
M0	0.25**	0.75*	1.45*	2.17*	2.88	0.73**	1.18	1.79	2.28	2.65
TAR	0.28	0.88**	1.72*	2.38*	3.04	0.32***	0.73	1.25	3.97	19.6
ARp	0.59***	1.22**	2.34*	3.32	4.03	0.58***	1.03*	3.52	6.41	7.7
PY	0.26**	0.75*	1.43	2.17*	2.92	0.47***	0.95*	1.58	2.2	2.67
M1	**0.2**	**0.63**	**1.22**	**1.82**	**2.48**	0.71**	1.05	1.56	2.03	2.53
PYLX	0.3**	0.76	1.49	2.3	3.25	0.3**	0.63	1.07	1.72	2.45
PYLXD	0.82***	1.3**	2.37**	3.63	3.22	0.39***	0.72	1.16	1.73	2.34
PYFLXD	0.7***	1.02**	1.65	3.08**	3.38	0.31***	0.7	1.31	1.99	2.83
PYFLX1	0.3**	0.99*	1.87	2.38	3.34	0.31***	0.64	1.1	1.75	2.51
PYFLX7	0.39**	1.29*	2.61	3.56	5.3	0.32***	0.68	1.08	1.66	2.41
RFYLX	0.28***	0.73	1.36	2.13	2.58	0.81***	1.36***	2.17*	2.71	3.01
RidYLX	0.27***	0.76**	1.45	2.05	2.61	0.88***	1.45**	2.19	2.69	2.96
LasYLX	0.33***	0.87**	1.56	2.2	2.68	1.02***	1.57**	2.27	2.71	2.95
PYXIPP	-	-	-	-	-	0.31**	0.65	1.1	1.75	2.49
PYFXIPP	-	-	-	-	-	0.33***	0.69	1.12	1.73	2.44

Table 5 Root mean squared errors, in bold format, the minimum RMSFE per horizon. The list of *** symbols refers to the Diebold and Mariano test, which means statistical difference concerning the model with minimum RMSFE (* p <5%, * p <1%, and *** p <0.1%). We use a sample period 2011-2022, out-of-sample predictions from Mar2017 until Nov2022 (top panel), out-of-sample predictions from Mar2017 until Dec2019 (middle panel), and out-of-sample predictions from Jan2020 until Nov2022 (bottom panel). The description of the models is given in table 1.

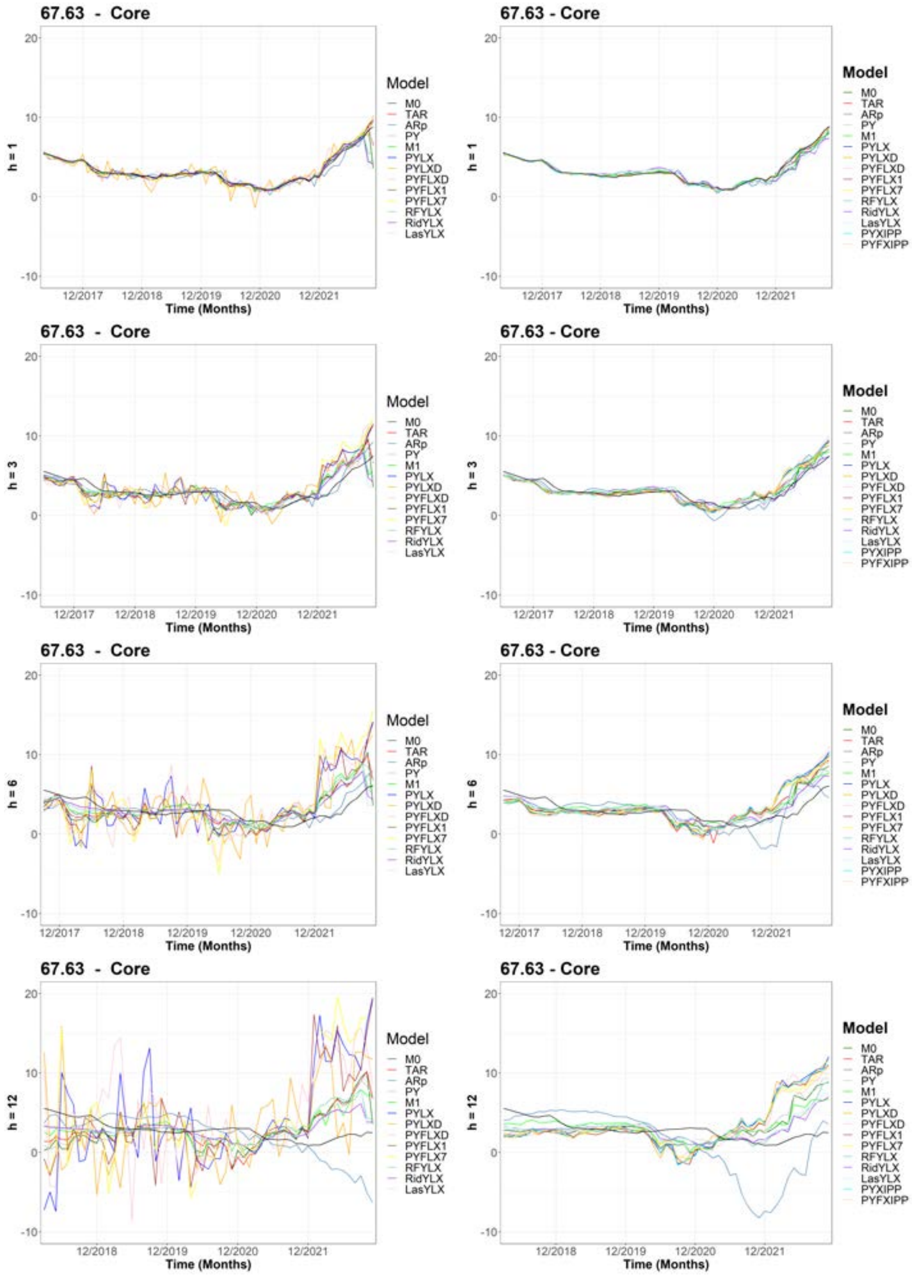


Fig. 3 Forecasts for the core CPI inflation (coloring lines) compared to the observed core inflation (black lines), lagged by h months. Forecast from aggregates models left panel and disaggregates models right panel. Out-of-sample predictions from Mar2017 to Nov2022 and for the horizons $h = 1, 3, 6, 9, 12$.

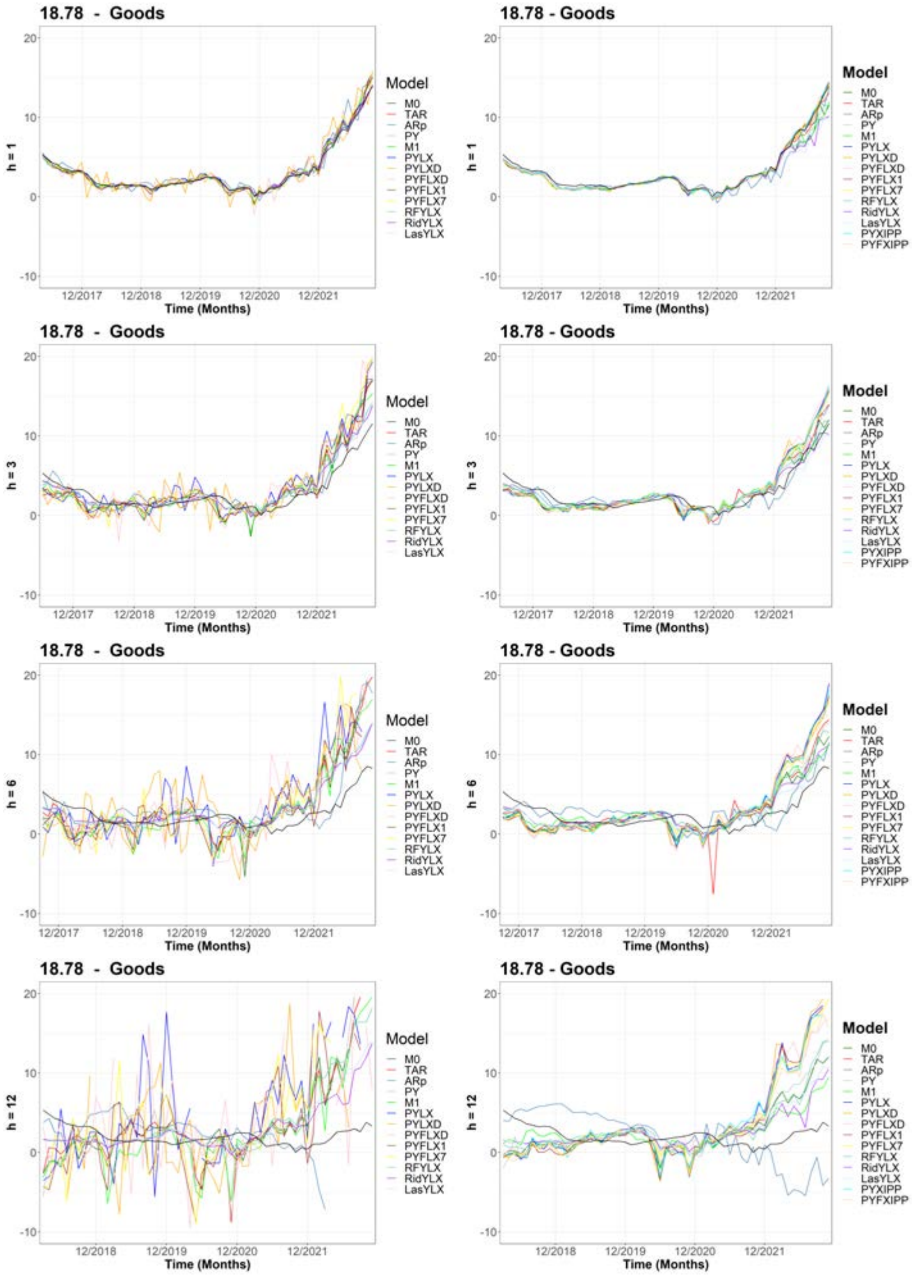


Fig. 4 Forecasts for the goods CPI inflation (coloring lines) compared to the observed goods inflation (black lines), lagged by h months. Forecast from aggregates models left panel and disaggregates models right panel. Out-of-sample predictions from Mar2017 to Nov2022 and for the horizons $h = 1, 3, 6, 9, 12$.

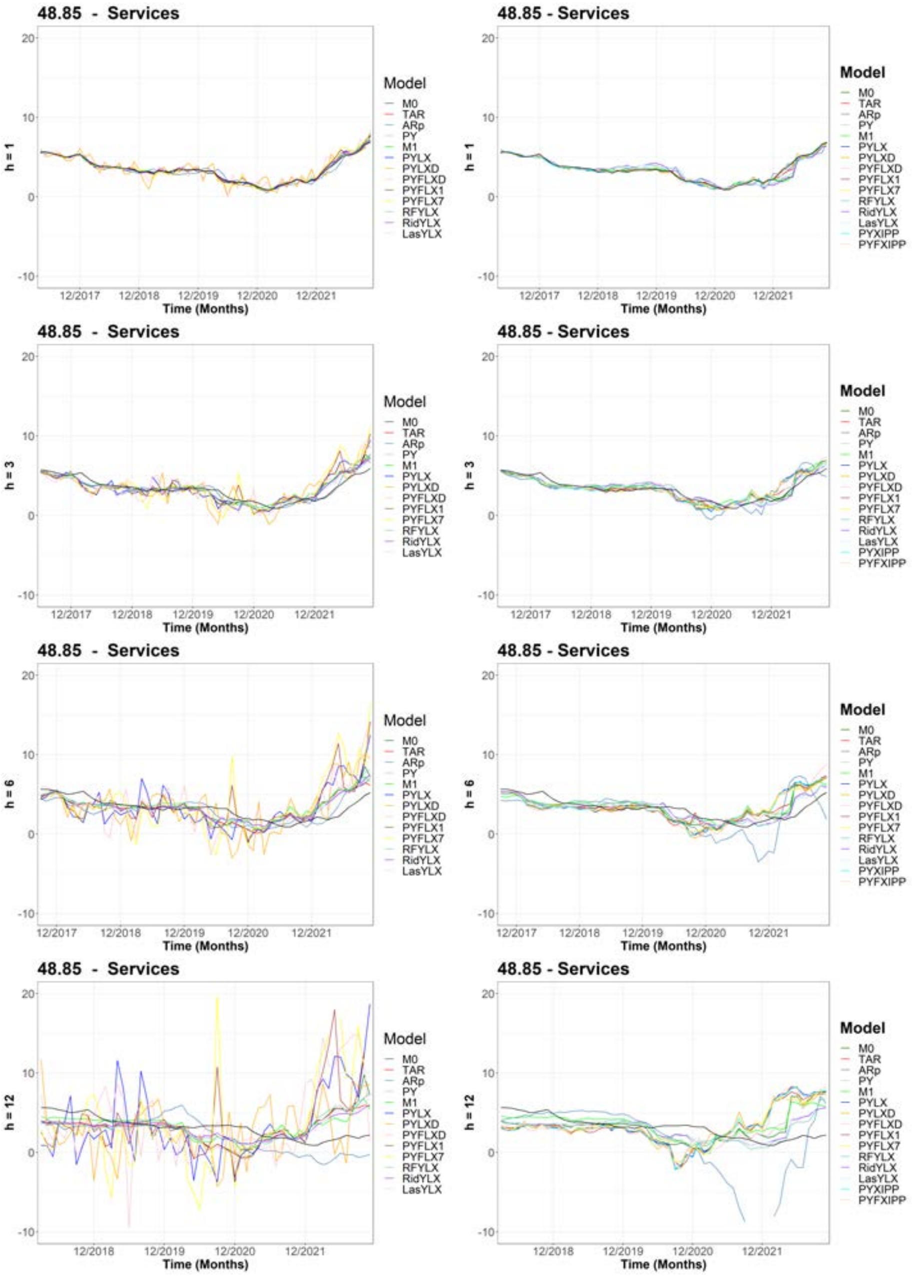


Fig. 5 Forecasts for the services CPI inflation (coloring lines) compared to the observed services inflation (black lines), lagged by h months. Forecast from aggregates models left panel and disaggregates models right panel. Out-of-sample predictions from Mar2017 to Nov2022 and for the horizons $h = 1, 3, 6, 9, 12$.

6 Conclusions

In this work, we compare the forecasting performance of different models and estimation methodologies that allow us to work with many predictors for both aggregates of inflation as a response variable and disaggregates to subsequently aggregate their forecasts using the CPI weights to obtain forecasts for the aggregates. The main conclusions are summarized as follows. We found that most disaggregated models are good competitors regarding forecast ability to the aggregated model with minimum RMSFE for horizons further than one month ahead. In general, for all the four CPI aggregates analyzed in this study, the disaggregated model generated a smaller forecasting error (RMSFE) than the aggregated for the mid-term horizons ($h = 6$, 9, 12). The models that include PPI components performed and the best-aggregated model for all CPI aggregates, sample periods, and horizons further than $h = 1$.

So, it is difficult to beat a simple autoregressive model for short-run horizons. Still, for further horizons, it seems that including a set of explanatory variables independent of the estimation methodology of the model helps to improve the forecasting ability. Also, a robust result obtained for all the CPI aggregates and sample periods analyzed is the good performance of disaggregated models to forecast an aggregate and the contribution of the PPI components to help to reduce forecasting errors or to compete with the best aggregate model.

Although it is difficult to beat a simple autoregressive model for short-run horizons, the TAR model, in some cases, improves the forecast performance without including the set of explanatory variables. In this sense, implementing threshold autoregressive models that include explanatory variables will be a potential application for future works. Also, other models from the domain of Machine Learning and time series models with Bayesian inference could be provided an additional improvement on the forecast performance. On the other hand, considering the good performance of aggregating disaggregated forecasts to obtain forecasts for aggregate inflation, it seems a good idea to evaluate different methodologies of forecast combination as a continuation of this study.

Regarding the evaluation period, for the pre-pandemic period, models for the aggregates without any additional explanatory variables performed well, but for the period when COVID-19 started and afterward, the forecasts improved by including additional variables into the models to help to explain the dynamics of both the disaggregate and the aggregates.

Declarations

- **Conflict of interest** The authors declare no conflicts of interest relevant to this work.
- **Availability of data and materials** The data that support the findings of this study are available from the corresponding author upon request.

References

Alonso JC, Rivera AF (2017) Pronosticando la inflación mensual en colombia un paso hacia delante: una aproximación “de abajo hacia arriba”. Revista de Métodos Cuantitativos para la Economía y la Empresa 23:98–118

Barker T (2014) Disaggregation in Econometric Modelling (Routledge Revivals). Routledge

Benalal N, Diaz del Hoyo JL, Landau B, et al (2004) To aggregate or not to aggregate? euro area inflation forecasting. Euro Area Inflation Forecasting (July 2004)

Botha B, Burger R, Kotzé K, et al (2022) Big data forecasting of south african inflation. Empirical Economics pp 1–40

Breiman L (2001) Random forests. Machine Learning 45(1):5–32

Clements MP, Hendry DF (2002) Modelling methodology and forecast failure. The Econometrics Journal 5(2):319–344

Cryer J, Chan K (2008) Time Series Analysis, with Applications in R. Springer-Verlag

Diebold FX, Mariano RS (1995) Comparing predictive accuracy. Journal of Business & Economic Statistics 13(3):253–263. https://doi.org/10.1080/07350015.1995.10524599

Friedman JH, Hastie T, Tibshirani R (2010) Regularization paths for generalized linear models via coordinate descent. Journal of Statistical Software 33(1):1–22. https://doi.org/10.18637/jss.v033.i01

González-Molano ER (2008) Pronósticos de agregados a partir de desagregados caso empírico: Inflación de alimentos en colombia. Borradores de Economía; No 504

Grunfeld Y, Griliches Z (1960) Is aggregation necessarily bad? The review of economics and statistics pp 1–13

Harvey D, Leybourne S, Newbold P (1997) Testing the equality of prediction mean squared errors. International Journal of Forecasting 13(2):281–291. https://doi.org/https://doi.org/10.1016/S0169-2070(96)00719-4

Hendry DF, Hubrich K (2006) Forecasting economic aggregates by disaggregates. ECB working paper

Hendry DF, Hubrich K (2011) Combining disaggregate forecasts or combining disaggregate information to forecast an aggregate. Journal of business & economic statistics 29(2):216–227

Hubrich K (2002) Forecasting euro area inflation: Does aggregating price component forecasts improve forecast accuracy? Tech. rep., Citeseer

Hubrich K (2005) Forecasting euro area inflation: Does aggregating forecasts by hicp component improve forecast accuracy? International Journal of Forecasting 21(1):119–136. https://doi.org/https://doi.org/10.1016/j.ijforecast.2004.04.005

Hubrich K, et al (2003) Forecasting euro area inflation: Does contemporaneous aggregation improve the forecasting performance? ECB Working Paper

Hyndman RJ, Khandakar Y (2008) Automatic time series forecasting: The forecast package for r. Journal of Statistical Software 27(3):1–22. https://doi.org/10.18637/jss.v027.i03

Joseph A, Kalamara E, Kapetanios G, et al (2021) Forecasting uk inflation bottom up. International Journal of Forecasting

Lütkepohl H (1986) Comparison of predictors for temporally and contemporaneously aggregated time series. International Journal of Forecasting 2(4):461–475. https://doi.org/https://doi.org/10.1016/S0169-2070(86)80001-2

Lütkepohl H (1987) Forecasting aggregated vector VARMA processes. Springer-Verlag

Melo-Velandia LF (2006) Forecasting food price inflation in developing countries with inflation targeting regimes: the colombian case. Borradores de Economía; No 409

Molano ERG, Velandia LFM, Olarte AG, et al (2007) Pronósticos directos de la inflación colombiana. Banco de la República

Peña Ordóñez AC, et al (2019) Pronóstico de la inflación colombiana: una aproximación desde un modelo arima desagregado y machine learning. Uniandes

Shumway RH, Stoffer DS (2017) Time Series Analysis and Its Applications. With R Examples. Springer Cham

Stock JH, Watson MW (2002a) Forecasting using principal components from a large number of predictors. Journal of the American Statistical Association 97(460):1167–1179. https://doi.org/10.1198/016214502388618960

Stock JH, Watson MW (2002b) Macroeconomic forecasting using diffusion indexes. Journal of Business & Economic Statistics 20(2):147–162. https://doi.org/10.1198/073500102317351921

Taieb SB, Hyndman RJ, et al (2012) Recursive and direct multi-step forecasting: the best of both worlds, vol 19. Department of Econometrics and Business Statistics, Monash Univ.

Tong H (1990) Non-linear Time Series, a Dynamical System Approach. Clarendon Press Oxford

Appendix A

Code	Variable Name	Source	Transf
1	CPI Services	DANE CPI: Services	LD
2	CPI services sin Comidas Fuera del Hogar	DANE CPI: Own Calculations	LD
3	PPI Produced and Consumed	DANE PPI	LD
4	CPI for goods excluding food and regulated goods subtracting tradable transport	DANE CPI: Own Calculations	LD
5	PPI P&C subtracting items related to vehicles	DANE PPI: Own Calculations	LD
6	PPI produced and consumed for items related to vehicles	DANE PPI: Own Calculations	LD
7	CPI for Health and Education	DANE CPI: Own Calculations	LD
8	DANE Industrial Production Index	DANE	L
9	DANE Manufacturing Production Index	DANE	L
10	PPI National Mining Production	DANE y Banco de la República: Historical PPI 2015	LD
11	CPI Housing regulated services	DANE CPI: Housing Regulated	LD
12	PPI Produced and Consumed from mining and quarrying	DANE PPI	LD
13	Economy Monitoring Indicator	DANE: ISE	L
14	VIS and Non-vis housing 12 metropolitan and 3 urban areas (completed units)	DANE: VIS and Non-VIS housing 15 areas of influence	L
15	Balance of payments Exports (goods)	Banco de la República: Balance of payments	LD
16	Balance of payments Imports (goods)	Banco de la República: Balance of payments	LD
17	PPI Internal offer	DANE	LD
18	PPI imports	DANE	LD
19	New home price index Banco de la Republica	Banco de la República: IPVNBR	LD
20	Bogotá Residential Property Price Index	DANE: Residential Property Price Index	LD
21	M3 + Deposits	Banrep: Own Calculations	LD
22	Number of disbursements for the purchase of new and used housing	Banco de la República: Housing data	L
23	Number of disbursements for the purchase of new and used housing, by financing entity: Fondo Nacional de Ahorro	Fondo Nacional del Ahorro: housing finance statistics	LD
24	Unemployment rate (%)	DANE:GEIH	LD
25	Employed (thousands)	DANE:GEIH	LD
26	Employment rate (%)	DANE:GEIH	LD
27	Monthly evolution of new vehicles (units)	Andemos	LD
28	International price of Brent oil	Investing.com	LD
29	Monetary Policy Interest Rate (%)	Banco de la República: Interest rates	LD
30	3M Libor	World Bank: External Rates	D
31	Financial Times Stock Exchange Index Colombia	Investing.com	LD
32	Nominal Exchange Rate Index	Banco de la República: ITCN	LD
33	EUR COP exchange rate	Banco de la República	LD
34	USD COP exchange rate	Banco de la República	LD
35	FTSE Mid-cap 250	Investig.com	LD
36	FTSE All Share (London)	Investing.com	LD
37	Colombian Stock Exchange Index COLCAP	Bolsa de valores de Colombia	LD
38	SP500	Yahoo Finance	LD
39	BOVESPA: Brazil Index	Investing.com	LD
40	VIX: CBOE Volatility Index	Investing.com	LD

Table A1 Macroeconomic Indicators used as predictors. Transformations over the variables (Transf): L = Logarithm, D = month-month difference, LD = Logarithm and difference, and OS = Original scaled.

Table A1 Continue...

Code	Variable Name	Source	Transf
41	CBOE Emerging Markets volatility index	Investing.com	LD
42	Total Imports	DANE and Banco de la República	LD
43	Non-durable goods imports	DANE and Banco de la República	LD
44	Durable goods imports	DANE and Banco de la República	LD
45	Fuels and Lubricants imports	DANE and Banco de la República	LD
46	Agricultural sector goods imports	DANE and Banco de la República	LD
47	Industrial sector goods imports	DANE and Banco de la República	LD
48	Construction materials imports	DANE and Banco de la República	LD
49	Imports of agricultural sector goods	DANE and Banco de la República	LD
50	Imports of industrial sector goods	DANE and Banco de la República	LD
51	Transportation equipment imports	DANE and Banco de la República	LD
52	Total Exports	DANE and Banco de la República	LD
53	Total non-traditional exports	DANE and Banco de la República	LD
54	Total Traditional exports	DANE and Banco de la República	LD
55	Coffee Exports	DANE and Banco de la República	L
56	Carbon Exports	DANE and Banco de la República	LD
57	Exports of Oil and derivatives	DANE and Banco de la República	LD
58	Ferroniquel exports	DANE and Banco de la República	D
59	Food CPI	Banco de la República: DANE data	LD
60	CPI Without Food	Banco de la República: DANE data	LD
61	Goods CPI	Banco de la República: DANE data	LD
62	services CPI	Banco de la República: DANE data	LD
63	Regulated goods CPI	Banco de la República: DANE data	LD
64	CPI excluding food and regulated goods	Banco de la República: DANE data	LD
65	Implicit Freight (%)	Banco de la República: DANE data	D
66	Implicit Freight	Banco de la República: DANE data	LD
67	Trade partners weighted average nominal exchange rate index	Banco de la República and DANE	LD
68	Quarterly Product Gap	Banco de la República: DANE data	D
69	Total energy demand GwH	XM Colombia	L
70	International Price Index for Agricultural Commodities	IndexMundi	LD
71	International food price index	IndexMundi	LD
72	International metal price index	Indexmundi	LD
73	Urea international price index	Indexmundi	LD
74	Monthly Inflation Expectations	Banco de la República	OS
75	Marketing Margin (CPI/PPI)	Banco de la República: DANE data	LD
76	Goods Marketing Margin (CPI/PPI)	Banco de la República: DANE data	LD
77	Food Marketing margin (CPI/PPI)	Banco de la República: DANE data	LD
78	Total demand / total intermediate consumption	Banco de la República: DANE data	LD
79	ONI climate index	NOAA	D

Headline	Aggregates					From disaggregates				
Model	1	3	6	9	12	1	3	6	9	12
M0	0.5***	0.57**	0.6*	0.57*	0.59	0.96	0.79	0.74	0.68	0.65
TAR	0.48***	0.55**	0.6*	0.55*	0.57	0.83	0.72**	0.69	1	5.16
ARp						1.24**	0.84**	0.91	1.03	1.05
PY	0.51***	0.55**	0.61*	0.57*	0.59	0.82**	0.72**	0.7	0.66	0.64
M1	2.13	1.17	0.76*	0.59*	0.61	1	0.79**	0.75*	0.71	0.69
PYLX	0.72***	0.68**	0.75	0.78	0.9	0.68***	0.58**	0.55*	0.52*	0.56
PYLXD	2.13	1.26	1.46	0.92	1.41	0.82**	0.63**	0.58*	0.54*	0.58
PYFLXD	1.05	0.87	1.13	1.27	1.05	0.64***	0.54**	0.51*	0.55	0.61
PYFLX1	0.5***	0.6**	0.61	0.57	0.64	0.7***	0.6**	0.58*	0.56	0.6
PYFLX7	0.59***	0.67*	0.74	0.69	0.75	0.67***	0.58**	0.56*	0.55	0.58
RFYLX	1.29	0.94	0.79	0.7	0.59	1.46	0.99	0.86	0.74	0.69
RidYLX	1.43	0.92	0.77	0.6*	0.6	1.58	1.03	0.84	0.74	0.68
LasYLX	2.54	1.39	0.92	0.66	0.62	1.98	1.2	0.92	0.76	0.69
PYXIPP	-	-	-	-	-	0.7***	0.57**	0.54*	0.52*	0.56
PYFXIPP	-	-	-	-	-	0.72***	0.61**	0.57*	0.55	0.57
M0						1.93	1.38	1.24	1.2	1.1
TAR	0.96*	0.97	1.01	0.98	0.96*	1.67	1.27	1.16	1.77	8.69
ARp	2.01	1.76	1.67	1.77	1.68	2.49	1.47	1.52	1.82	1.77
PY	1.02	0.97	1.02	1	1	1.64	1.27	1.17	1.16	1.08
M1	4.29	2.07	1.27	1.04	1.02	2.02	1.39	1.26	1.25	1.15
PYLX	1.44	1.19	1.25	1.37	1.52	1.36	1.02	0.92	0.93	0.95
PYLXD	4.28	2.22	2.43	1.62	2.38	1.65	1.11	0.96	0.95	0.97
PYFLXD	2.1	1.53	1.88	2.24	1.77	1.28	0.95	0.85*	0.97	1.03
PYFLX1	1.01	1.06	1.02	1.01	1.07	1.41	1.06	0.97	0.99	1.01
PYFLX7	1.18	1.18	1.24	1.22	1.26	1.34	1.02	0.94	0.97	0.98
RFYLX	2.58	1.66	1.32	1.23	0.99	2.94	1.73	1.44	1.31	1.16
RidYLX	2.86	1.62	1.29	1.06	1.01	3.18	1.81	1.41	1.31	1.15
LasYLX	5.1	2.45	1.54	1.17	1.04	3.99	2.11	1.54	1.34	1.15
PYXIPP	-	-	-	-	-	1.41	1.01	0.89	0.92	0.94
PYFXIPP	-	-	-	-	-	1.46	1.07	0.95	0.97	0.97
M0	0.23**	0.48	0.79	0.96	0.98	0.45**	0.67	0.97	1.15	1.08
TAR	0.22**	0.47	0.79	0.94	0.94	0.39**	0.61	0.91	1.7	8.52
ARp	0.47*	0.85	1.31	1.69	1.65	0.58**	0.71	1.19	1.75	1.73
PY	0.24**	0.47	0.8	0.96	0.98	0.38*	0.61	0.92	1.12	1.06
M1						0.47**	0.67	0.99	1.2	1.13
PYLX	0.34**	0.58	0.98	1.31	1.49	0.32**	0.49	0.72	0.89	0.93
PYLXD	1	1.07	1.91	1.55	2.33	0.38**	0.54	0.76	0.91	0.95
PYFLXD	0.49*	0.74	1.48	2.14	1.74	0.3**	0.46	0.67	0.93	1.01
PYFLX1	0.24**	0.51	0.8	0.97	1.05	0.33**	0.51	0.76	0.95	0.99
PYFLX7	0.28**	0.57	0.98	1.17	1.24	0.31**	0.49	0.74	0.93	0.96
RFYLX	0.6**	0.8	1.04	1.18	0.97	0.69	0.84	1.13	1.26	1.13
RidYLX	0.67**	0.78	1.01	1.02	0.99	0.74	0.88	1.11	1.26	1.12
LasYLX	1.19	1.19	1.21	1.12	1.02	0.93	1.02	1.21	1.29	1.13
PYXIPP	-	-	-	-	-	0.33**	0.49	0.7	0.88	0.92
PYFXIPP	-	-	-	-	-	0.34**	0.52	0.75	0.93	0.95

Table A2 Root mean squared forecast error, relative to AR(12) model (ARp, first panel), best ARIMA model (M0, second panel), best SARIMA model (M1, third panel), PYFLXD (fourth panel), and PYFLX (fifth panel). Sample period 2011-2022, out-of-sample predictions from Mar2017 until Nov2022. The description of the models is given in table 1. "-" for the models, PYXIPP and PYFXIPP for the aggregates means no applied.

Table A2 Continue...

Headline	Aggregates					From disaggregates				
Model	1	3	6	9	12	1	3	6	9	12
M0	0.48***	0.65***	0.53***	0.45***	0.56**	0.92***	0.9***	0.66***	0.54**	0.62**
TAR	0.46***	0.63***	0.54***	0.44***	0.54**	0.79	0.83	0.61**	0.79**	4.91*
ARp	0.96	1.15	0.89	0.79	0.95	1.19**	0.96	0.8**	0.81	1
PY	0.48***	0.63***	0.54***	0.45***	0.56**	0.78	0.83	0.62	0.52	0.61
M1	2.04	1.35	0.68**	0.47***	0.58**	0.96**	0.91	0.67**	0.56**	0.65*
PYLX	0.69**	0.77**	0.66**	0.61**	0.86	0.65	0.66	0.49**	0.41**	0.54*
PYLXD	2.04	1.45	1.29	0.72*	1.34	0.79***	0.73***	0.51***	0.42***	0.55**
PYFLXD						0.61**	0.62**	0.45***	0.43***	0.58**
PYFLX1	0.48***	0.69***	0.54***	0.45***	0.6**	0.67***	0.69***	0.51***	0.44**	0.57**
PYFLX7	0.56***	0.77**	0.66***	0.54**	0.71*	0.64***	0.67***	0.5***	0.43**	0.56**
RFYLX	1.23	1.08	0.7*	0.55**	0.56**	1.4	1.13	0.76	0.59**	0.65*
RidYLX	1.36	1.05	0.68**	0.47**	0.57**	1.51	1.18	0.75	0.59**	0.65*
LasYLX	2.43	1.6	0.82	0.52**	0.59*	1.9	1.38	0.82	0.6**	0.65*
PYXIPP	-	-	-	-	-	0.67***	0.66***	0.48***	0.41***	0.53**
PYFXIPP	-	-	-	-	-	0.69***	0.7***	0.51***	0.43**	0.55**
M0	0.99	0.94	0.98	0.99	0.93	1.91	1.3	1.21	1.19	1.03
TAR	0.95	0.92	0.99	0.96	0.9	1.66	1.19	1.13	1.75	8.11
ARp	1.99	1.66	1.64	1.74	1.57	2.47	1.39	1.49	1.8	1.65
PY	1.01	0.91	1	0.99	0.93	1.63	1.2	1.14	1.15	1.01
M1	4.25	1.95	1.25	1.03	0.95	2	1.31	1.24	1.23	1.08
PYLX	1.43	1.12	1.22	1.35	1.42	1.35	0.96	0.9	0.91	0.89
PYLXD	4.24	2.09	2.38	1.59	2.22	1.64	1.05	0.94	0.93	0.91
PYFLXD	2.08	1.45	1.85	2.2	1.65	1.27	0.9	0.84	0.95	0.96
PYFLX1						1.4	1	0.95**	0.98	0.94
PYFLX7	1.17	1.11	1.22	1.2	1.18	1.33	0.96	0.92	0.95	0.92
RFYLX	2.56	1.57	1.29	1.21	0.92	2.92	1.63	1.41	1.29	1.08
RidYLX	2.84	1.53	1.26	1.04	0.94	3.15	1.71	1.38	1.29	1.07
LasYLX	5.05	2.31	1.51	1.16	0.97	3.95	1.99	1.51	1.32	1.08
PYXIPP	-	-	-	-	-	1.39	0.95	0.88*	0.91	0.88*
PYFXIPP	-	-	-	-	-	1.44	1.01	0.93	0.96	0.9

Core	Aggregates					From disaggregates				
Model	1	3	6	9	12	1	3	6	9	12
M0	0.46***	0.62**	0.68	0.77	0.82	1.01	0.83	0.75	0.78	0.75
TAR	0.45***	0.62**	0.67	0.76	0.81	0.71	0.69**	0.68*	1.52	12.44**
ARp						0.97***	0.89**	1.27*	1.68	1.6
PY	0.48***	0.66**	0.68	0.78	0.83	0.69***	0.71**	0.7*	0.77	0.76**
M1	1.02	0.51**	0.6*	0.71	0.77**	0.96	0.75**	0.69*	0.76	0.78**
PYLX	0.57***	0.76	0.91	1.17	1.33	0.48***	0.52**	0.53*	0.64	0.7*
PYLXD	1.46	1.05	1.08	1.13	1.45	0.59***	0.55**	0.56*	0.66	0.69*
PYFLXD	0.92	0.92	0.98	1.27	1.53	0.49***	0.52**	0.59	0.67	0.76**
PYFLX1	0.51***	0.74	0.86	1.02	1.14	0.51***	0.54**	0.55*	0.67	0.73*
PYFLX7	0.57***	0.85	1.02	1.23	1.34	0.51***	0.55**	0.54*	0.63	0.69*
RFYLX	0.85	0.6**	0.59*	0.67	0.68**	1.14	0.95	0.89	0.91	0.86**
RidYLX	1.06	0.62**	0.66*	0.7*	0.74**	1.32	1.03	0.91	0.9	0.84**
LasYLX	1.26	0.72*	0.72	0.79	0.79**	1.52	1.13	0.94	0.9	0.83**
PYXIPP	-	-	-	-	-	0.49***	0.55**	0.54*	0.65	0.7*
PYFXIPP	-	-	-	-	-	0.55***	0.58**	0.56*	0.65	0.7*
M0						2.19	1.33	1.11	1.01	0.92**
TAR	0.97	0.99	0.99	0.98	0.98	1.54	1.11	1	1.97	15.15
ARp	2.17	1.6	1.48	1.29	1.22	2.11	1.42	1.87	2.17	1.94
PY	1.05	1.06	1.01	1.01	1.02	1.5	1.14	1.03	0.99	0.92*
M1	2.21	0.82***	0.89*	0.92	0.94	2.08	1.21	1.02	0.98	0.95
PYLX	1.25	1.22	1.34	1.51	1.62	1.04	0.84***	0.78**	0.82*	0.85*
PYLXD	3.18	1.69	1.59	1.46	1.77	1.29	0.87**	0.82**	0.85	0.84*
PYFLXD	2	1.48	1.45	1.63	1.86	1.07	0.83**	0.86*	0.86	0.93*
PYFLX1	1.1	1.18	1.27	1.32	1.38	1.11	0.87***	0.82**	0.86*	0.88**
PYFLX7	1.24	1.36	1.51	1.59	1.63	1.12	0.88***	0.8**	0.82**	0.83**
RFYLX	1.85	0.96	0.87***	0.87**	0.82**	2.48	1.53	1.32	1.17	1.04
RidYLX	2.3	1	0.97	0.9**	0.91	2.88	1.64	1.34	1.16	1.02
LasYLX	2.75	1.15	1.07	1.02	0.96	3.3	1.81	1.38	1.16	1.01
PYXIPP	-	-	-	-	-	1.07	0.87***	0.8**	0.84*	0.86*
PYFXIPP	-	-	-	-	-	1.19	0.93*	0.83**	0.84*	0.85**
M0	0.45	1.21	1.12	1.09	1.07	0.99	1.62	1.25	1.09	0.98
TAR	0.44	1.2	1.11	1.07	1.05	0.7	1.34	1.12	2.14	16.16
ARp	0.98	1.94	1.66	1.4	1.3	0.96	1.72	2.1	2.36	2.07
PY	0.47	1.28	1.13	1.1	1.08	0.68	1.39	1.16	1.08	0.99
M1						0.94	1.46	1.15	1.07	1.01
PYLX	0.56	1.48	1.5	1.64	1.72	0.47	1.02	0.88	0.89	0.91
PYLXD	1.44	2.05	1.79	1.59	1.89	0.58	1.06	0.92	0.92	0.9
PYFLXD	0.9	1.79	1.63	1.78	1.98	0.48	1.01	0.97	0.94	0.99
PYFLX1	0.5	1.43	1.42	1.43	1.48	0.5	1.06	0.92	0.93	0.94
PYFLX7	0.56	1.65	1.7	1.72	1.74	0.51	1.07	0.89	0.89	0.89
RFYLX	0.84	1.16	0.98	0.94*	0.88***	1.12	1.85	1.48	1.27	1.11
RidYLX	1.04	1.21	1.09	0.98	0.97	1.3	1.99	1.5	1.26	1.09
LasYLX	1.24	1.39	1.2	1.11	1.03	1.49	2.19	1.55	1.26	1.08
PYXIPP	-	-	-	-	-	0.48	1.06	0.9	0.91	0.91
PYFXIPP	-	-	-	-	-	0.54	1.13	0.93	0.92	0.91

Table A3 Root mean squared forecast error, relative to AR(12) model (ARp, first panel), best ARIMA model (M0, second panel), best SARIMA model (M1, third panel), PYFLXD (fourth panel), and PYFLX (fifth panel). Sample period 2011-2022, out-of-sample predictions from Mar2017 until Nov2022. The description of the models is given in table 1. “-” for the models, PYXIPP and PYFXIPP for the aggregates means no applied.

Table A3 Continue...

Core	Aggregates					From disaggregates				
Model	1	3	6	9	12	1	3	6	9	12
M0	0.5***	0.68***	0.69**	0.61***	0.54**	1.1***	0.9***	0.77***	0.62**	0.49**
TAR	0.48***	0.67***	0.68**	0.6***	0.53**	0.77	0.75	0.69	1.2***	8.15**
ARp	1.09	1.08	1.02	0.79**	0.66*	1.06*	0.96**	1.29**	1.33	1.05
PY	0.52***	0.71**	0.7**	0.62***	0.55**	0.75	0.77	0.71	0.61	0.5
M1	1.11	0.56***	0.61***	0.56***	0.5**	1.04**	0.82*	0.71*	0.6***	0.51**
PYLX	0.63***	0.83**	0.92	0.92	0.87	0.52	0.57	0.54**	0.5***	0.46**
PYLXD	1.6	1.14	1.1	0.89	0.95	0.65***	0.59***	0.56***	0.52**	0.45**
PYFLXD						0.53***	0.57***	0.6***	0.53**	0.5**
PYFLX1	0.55***	0.8**	0.87	0.81	0.74*	0.56***	0.59***	0.56***	0.53***	0.48**
PYFLX7	0.62***	0.92	1.04	0.97	0.88	0.56***	0.6***	0.55***	0.5***	0.45**
RFYLX	0.93	0.65***	0.6***	0.53***	0.44**	1.24	1.03	0.91	0.72***	0.56**
RidYLX	1.15	0.67***	0.67**	0.55***	0.49**	1.44	1.11	0.92	0.71***	0.55**
LasYLX	1.38	0.78*	0.74*	0.63***	0.52**	1.65	1.22	0.95	0.71***	0.54**
PYXIPP	-	-	-	-	-	0.54***	0.59***	0.55***	0.51**	0.46**
PYFXIPP	-	-	-	-	-	0.59***	0.63***	0.57***	0.52***	0.46**
M0	0.91	0.85*	0.79*	0.76*	0.72*	1.99	1.13***	0.88**	0.76**	0.66*
TAR	0.88*	0.83*	0.78*	0.74	0.71	1.4	0.93	0.79	1.49	10.96*
ARp	1.98	1.35	1.17	0.98	0.88	1.92	1.2	1.48	1.64	1.41
PY	0.95	0.89	0.8	0.76	0.74	1.37	0.97	0.82	0.75	0.67
M1	2.01	0.7***	0.7**	0.7*	0.68*	1.89	1.02	0.81	0.74	0.68*
PYLX	1.13	1.03	1.06	1.14	1.17	0.94	0.71	0.62	0.62*	0.61*
PYLXD	2.89	1.43	1.26	1.11	1.28	1.17	0.74***	0.65**	0.64**	0.61**
PYFLXD	1.81	1.25	1.15	1.24	1.35	0.97	0.71***	0.68**	0.65**	0.67**
PYFLX1						1.01	0.74***	0.65**	0.65**	0.64*
PYFLX7	1.12	1.15	1.19	1.2	1.18	1.02	0.74***	0.63**	0.62**	0.6**
RFYLX	1.68	0.81**	0.69**	0.66**	0.6**	2.25	1.29	1.04	0.89	0.75
RidYLX	2.09	0.84*	0.77*	0.68*	0.66*	2.61	1.39	1.06	0.88	0.74
LasYLX	2.49	0.97	0.84	0.78	0.7*	3	1.53	1.09	0.88	0.73
PYXIPP	-	-	-	-	-	0.97	0.74***	0.63**	0.63**	0.62**
PYFXIPP	-	-	-	-	-	1.08	0.79**	0.66**	0.64**	0.61**