Constructing a Forward-Looking Core Inflation Measure for Brazil

João Ricardo R. Moreira*

March 9, 2026

Abstract

Core inflation is widely used to guide monetary policy, yet most existing measures were not designed to forecast future inflation and often perform poorly in this task. This paper constructs a forward-looking core inflation measure for Brazil using disaggregated consumer price data and generalized non-negative ridge regressions that estimate the weights of price subcomponents so that the resulting index is maximally predictive of future headline inflation. Using monthly data from 1999 to 2025, we evaluate its out-of-sample forecasting performance and compare it with the official core measures used by the Central Bank of Brazil. The results show improved forecasts across different inflation regimes, generally matching or outperforming existing indicators. The estimated weights are relatively high for food and fuel prices, components typically excluded from traditional core measures. A supervised trimmed inflation version implies highly asymmetric trimming, also contrasting with conventional measures. Overall, the method proves to be robust and suggests that existing core measures may be poorly suited to signaling future inflation in emerging economies.

Keywords: Core inflation; Inflation forecasting; Monetary policy; Regularized regression; Brazil.
JEL Codes: E31, E52, C53.

1 Introduction

Core inflation measures are widely used in monetary policy and investment strategies. They are intended to capture the persistent component of inflation and to signal where headline inflation is likely to stabilize once temporary shocks dissipate. However, commonly used core inflation measures were not designed with forecasting performance in mind and therefore often perform poorly in predicting future inflation (Rich and Steindel, 2005; Pincheira et al., 2016).

This limitation is particularly relevant in emerging market economies, where inflation dynamics tend to be more volatile and exposed to external shocks. Differences in institutional frameworks, the presence of regulated prices and greater macroeconomic instability often make inflation developments more difficult to predict (Aizenman et al., 2011). Brazil provides a particularly informative case in this context. After the adoption of the inflation-targeting regime in 1999, the Central Bank of Brazil gradually expanded the set of core inflation indicators used to monitor underlying inflation trends. Machado (2024) provides a comprehensive overview of these measures and documents their growing role in monetary policy discussions. However, both Machado (2024) and Silva Filho and Figueiredo (2011) find that the predictive ability of these indicators remains limited, despite the importance of timely inflation assessments for policy decisions.

Traditional core measures typically rely on permanent heuristic exclusion or trimming rules that are not designed to maximize predictive accuracy. Recent work has proposed an alternative approach that

*PhD Candidate at Sao Paulo School of Economics (FGV/EESP). Contact: joao.r.r.moreira@gmail.com. The author acknowledge the partial financial support provided by FAPESP through grant 2025/17514-4. Also, the author is grateful to Pedro Luiz Valls Pereira and Marcelo Fernandes, who provided suggestions on an earlier draft.

explicitly constructs core inflation indicators to forecast future headline inflation. Goulet Coulombe et al. (2024) introduced the assemblage regression methodology, which uses generalized non-negative ridge regressions to estimate the weights of price subcomponents so that the resulting index is maximally predictive of future headline inflation. Moreover, by ordering price subcomponents according to their cross-sectional ranks, the methodology switches the algorithm to be learning a supervised trimmed inflation measure. Using disaggregated CPI data for the United States, the Euro Area and Canada, the authors show that their approach outperforms existing core inflation measures in forecasting exercises.

This paper applies the assemblage regression framework to Brazil and evaluates its performance in an emerging market environment. For the reasons cited above, the results found for advanced economies may not necessarily hold in this context. To our knowledge, this is the first application of this methodology to an emerging market. Using monthly CPI data from 1999 to 2025, we construct forward-looking core inflation measures and assess their out-of-sample forecasting performance relative to the official core indicators used by the Central Bank of Brazil.

Our results show that the proposed measures generally match or outperform benchmark combinations of official core indicators, although the forecasting gains are more concentrated at short horizons (three to six months ahead) than those reported for advanced economies. We also find relatively high weights for food and fuel prices, components that are typically excluded from traditional core inflation measures. In addition, the supervised trimmed inflation version implies a highly asymmetric trimming relative to conventional indicators.

This paper contributes to the literature in three main ways. First, it tests the assemblage regression methodology in a more challenging environment, demonstrating that it delivers reliable forecasts gains. Second, it provides a new forward-looking core inflation measure for Brazil that better signals future headline inflation than existing indicators. Third, by documenting the elevated weights of food goods and transport fuel components in this core measure, it suggests that traditional exclusion-based core indicators may not be well suited for emerging market contexts.

The remainder of this paper is organized as follows. Section 2 presents the methodology, Section 3 reports the main empirical results, and Section 4 concludes the paper.

2 Assemblage Regression Framework

The Assemblage Regression framework was proposed by Goulet Coulombe et al. (2024). As the authors describe it, this method is a generalized non-negative ridge regression, where the dependent variable is future headline inflation and the regressors are either inflation subcomponents or transformations of them.

There are two variations of their approach. The first is referred to as Albacore$_{Comps}$, which reweights the inflation subcomponents to form a basket that maximizes correlation with future inflation. The second is called Albacore$_{Ranks}$, which achieves optimal trimming of inflation subcomponents - via temporary exclusion or reweighting - by estimating the Assemblage Regression in the rank space of the component data. The term Albacore stands for **a**daptive **l**earning-**ba**sed **core** inflation.

In this section, we introduce both formulations of Assemblage Regression and subsequently discuss their connection to related methods[1]. To set the stage, however, we begin by reviewing the fundamentals of Ridge Regression.

2.1 Ridge Regression

A Ridge regression is a penalized linear regression problem. It is a traditional least squares problem with a penalty on the coefficients to regularize the outcome of minimization. The coefficients are obtained

[1] To do this, we heavily borrow from Goulet Coulombe et al. (2024) sections on the Assemblage Regression framework.

via

$$\hat{\boldsymbol{\beta}}_{\text{Ridge}} = \arg\min_{\boldsymbol{\beta}} \sum_{t=1}^{T-1} (y_{t+1} - \boldsymbol{\beta}' X_t)^2 + \lambda \|\boldsymbol{\beta}\|_2 \tag{1}$$

where $||.||_2$ is the l_2 norm, which is equivalent to $\Sigma_{k=1}^{K} \beta_k^2$, and k denotes the number of components.

The penalty term introduces regularization by incorporating the *a priori* belief that each coefficient should contribute to the fit, but only modestly. In other words, it shrinks the coefficients toward zero. With a suitable choice of $\lambda > 0$, ridge regression mitigates the overfitting that often affects OLS (i.e., $\lambda = 0$), particularly when the number of predictors K is large relative to the sample size T, and the signal-to-noise ratio is low.

There is a trade-off: the benefit of reduced variance comes at the cost of increased bias. The ridge estimator $\hat{\boldsymbol{\beta}}_{\text{Ridge}}$ blends information from the data with our prior belief about $\boldsymbol{\beta}$. Therefore, the form and strength of the penalty must be chosen carefully.

The regularization parameter λ is selected via cross-validation, which acts as a pseudo-out-of-sample performance evaluation metric. The degree of shrinkage directly affects both the forecasts and the estimated coefficients—and, by extension, the interpretability of the model. Ideally, λ is chosen to maximize the model's predictive accuracy on previously unseen data.

2.2 Albacore$_{Comps}$

The first Assemblage Regression corresponds to a permanent exclusion version, or a supervised weighting of basket components, and is obtained via

$$\hat{\boldsymbol{\omega}}_{\text{c}} = \arg\min_{\boldsymbol{\omega}} \sum_{t=1}^{T-h} (\pi_{t+1:t+h} - \boldsymbol{\omega}' \Pi_t)^2 + \lambda \|\boldsymbol{\omega} - \boldsymbol{\omega}_{\text{headline}}\|_2 \quad \text{s.t.} \quad \boldsymbol{\omega} \geq 0, \quad \boldsymbol{\omega}' \boldsymbol{l} = 1 \tag{2}$$

where h is the forecasting horizon, T is the last training observation, $\pi_{t+1:t+h}$ is the average headline inflation between $t+1$ and $t+h$, and Π_t is a matrix of component time series at a user-specified level of aggregation. The term $\lambda \|\boldsymbol{\omega} - \boldsymbol{\omega}_{\text{headline}}\|_2$ is a penalty that shrinks the solution toward $\boldsymbol{\omega}_{\text{headline}}$, the official headline inflation weights at that level. We impose that the weights be nonnegative ($\boldsymbol{\omega} \geq 0$) and sum to one ($\boldsymbol{\omega}' \boldsymbol{l} = 1$). The series in Π_t are expressed as price growth rates; in our application, this is the 3-month-over-3-month growth rate. As argued by Goulet Coulombe et al. (2024), this enables minimal temporal smoothing while preserving timeliness. In other words, it balances timeliness (as opposed to year-over-year changes) and noise reduction (as opposed to month-over-month growth rate).

Albacore$_{\text{Comps}}$ is defined as $\pi^*_{\text{comps},t} = \boldsymbol{\omega}_{\text{c}}' \Pi_t$, and is therefore a constrained forecast—both in terms of the information set and how that information is synthesized. It represents the linear aggregation (or weighted average) of components most closely related to future headline inflation, as measured by $\pi_{t+1:t+h}$. Its adaptiveness is tied to the forecasting objective, making it natural to expect core measures to vary with the horizon h and to depend on the loss function itself. Core measures relevant to investment bankers—concerned with market reactions to inflation releases—may differ from those of interest to central bankers dealing with the long and variable lags of monetary policy. This highlights the versatility of the approach, which allows flexibility in selecting components and forecasting horizons.

From another perspective, Assemblage Regression can be seen as an algorithm that searches the space of basket weights for which today's consumer inflation is most predictive of average future inflation. While the resulting assemblage is not explicitly designed to match the consumption basket of a representative agent, it would be surprising if the two were entirely unrelated. If predictive performance (evaluated out-of-sample) proves insufficient, the penalty term $\lambda \|\boldsymbol{\omega} - \boldsymbol{\omega}_{\text{headline}}\|_2$ ensures shrinkage toward official weights, rather than toward zero as in standard ridge regression.

Finally, the strength of the regularization, controlled by λ, must be carefully chosen. Although it can be selected automatically, it must be done appropriately. The target series exhibits persistence, as do many components. Thus, using standard cross-validation procedures—designed for independent observations—can lead to overly optimistic assessments of generalization error and a downward-biased choice of λ. Following the recommendation of Goulet Coulombe et al. (2024), we adopt a non-overlapping block cross-validation strategy. The training sample is divided into 10 contiguous blocks, which are then used in a 10-fold cross-validation. The block length depends on the sample size (e.g., a 20-year estimation sample yields blocks of two years). While longer blocks may be preferable for $h = 12$ or more, one must balance this with the risk that having too few blocks and limited randomness could outweigh the benefits.

2.3 Albacore$_{Ranks}$

The second approach is a supervised trimmed inflation. It requires additional consideration to be framed within the Assemblage Regression framework described above. Trimming, by nature, implies that components enter and exit the index each month, leading to time variation in their associated weights. These changes depend on how each component's growth rate ranks relative to others in a given period. Attempting to model this process directly, without further structure, would likely violate the principle of simplicity.

Instead, one can more generally consider retrieving the summary statistics of the current realized price growth distribution that best serve a specific statistical objective. Equivalently, this means assigning weights to components as a function of their position within the empirical distribution at time t. A closer examination of the implications of this time variation reveals that we can, in fact, run an Assemblage Regression very similar to the one described above, but using the empirical order statistics of Π_t as regressors. To understand this, we can simply move from the component space to the "rank space" by expressing the fitted values in summation notation:

$$\begin{aligned}
\pi^*_{\text{ranks},t} &= \sum_{k=1}^{K} \omega_{k,t} \Pi_{k,t} \\
&= \sum_{k=1}^{K} \sum_{r=1}^{K} \omega_r \, I(\text{rank}(\Pi_{k,t}) = r) \, \Pi_{k,t} \\
&= \sum_{r=1}^{K} \omega_r \sum_{k=1}^{K} I(\text{rank}(\Pi_{k,t}) = r) \, \Pi_{k,t} \\
&= \sum_{r=1}^{K} \omega_r O_{r,t}, \quad \text{where} \quad O_{r,t} = \sum_{k=1}^{K} I(\text{rank}(\Pi_{k,t}) = r) \, \Pi_{k,t}
\end{aligned}$$

where $O_{r,t}$ is the r^{th} order statistic (i.e., the value attached to rank r) ate time t. If K were to be equal to 100, $O_{1:100,t}$ would be an empirical approximation of the percentiles based on the ordered sample at time t. The above derivation informs us that this kind of time variation in the component space implies time-invariant coefficients in the rank space (and vice versa). Thus, ease of implementation can be restored from running the model in rank space, i.e., using O_t, which is effectively just sorting the components at each t and stacking them in a matrix. Precisely, $O_{1:K,t} = sort(\Pi_{1:K,t}) \quad \forall t$. The switch to rank space-based weighting can also be formalized in our preferred matrix notation, with $O_t = A_t \Pi_t$ where A_t is a $T \times T$ selection matrix whose entries are $0's$ and $1's$ given by $I(rank(\Pi_{k,t}) = r) \quad \forall r, k$. Thus, we run

$$\hat{\boldsymbol{\omega}}_{\text{r}} = \arg\min_{\boldsymbol{\omega}} \sum_{t=1}^{T-h} (\pi_{t+1:t+h} - \boldsymbol{\omega}' O_t)^2 + \lambda \|D\boldsymbol{\omega}\|_2 \quad \text{s.t.} \quad \boldsymbol{\omega} \geq 0, \quad \overline{\pi}_{t+1:t+h} = \overline{\pi}^*_{ranks,t} \tag{3}$$

where D is the difference operator and Albacore$_{Ranks}$ is defined as $\pi^*_{ranks,t} = \boldsymbol{\omega_r' O_t}$. Rather than learning which components to include, the problem will now be learning which ranks (and with which weight) to include or exclude. $\|D\boldsymbol{\omega}\|_2$ (or equivalently $\Sigma_{r=1}^{K}(\omega_r - \omega_{r-1})^2$ in summation notation) is a fused ridge penalty, as explained in Hastie et al. (2015). It favors a smooth weighting scheme, and in the limit $\lambda \to \infty$, pushes the model towards the sample mean solution where each rank receives a weight of $1/k$. In the other limit ($\lambda = 0$), the solution will be sparse for reasons that will be discussed below.

The equality constraint has been changed from $\boldsymbol{w}'l = 1$ to $\bar{\pi}_{t+1:t+h} = \bar{\pi}^*_{ranks,t}$. In the rank space, $\boldsymbol{w}'l = 1$ would imply that the average of all weights is $1/K$, which does not enforce symmetric trimming, but rather would impose equivalent masses above and below $1/K$, and thus blocking highly asymmetric outcomes which we will find to be successful in nearly all our experiments. The $\bar{\pi}_{t+1:t+h} = \bar{\pi}^*_{ranks,t}$ constraint brings back discipline by forcing residuals to have mean of 0. They could deviate marginally because the model has no intercept, and to minimize the sum of squared residuals, some bias could be exchanged for variance reduction. We block this possibility by imposing the constraint that Albacore$_{Ranks}$ has the same long-run mean as headline inflation (over the training sample).

This stands in contrast to most trimming approaches, which opt for a preselected band within which ranks are assigned their official (components space) weights, while those outside of it have a weight of 0. For instance, in the case of the United States, the FRB Cleveland trimmed mean CPI sets to 0 the first and last 16% of ranks, and then a weighted average of the center band is reported. In a similar vein, the FRB Dallas' Trimmed Mean PCE trims out 24% from the lower tail and 31% from the upper tail. It is important to note how our proposition differs from what is commonly done with trimmed measures. It does not reweight included components according to some rule based on their headline inflation weights, nor does it enforce that weights sum to 1 at every point in time.

2.4 Related Work

As Goulet Coulombe et al. (2024) put it, econometrically, Assemblage Regression's closest relatives are the regression-based forecast combination literature (Wang et al., 2023) and autoregression-based filtering (Hamilton, 2018). Albacore$_{Comps}$ shares various similarities with regularized forecast combination schemes such as that of Diebold and Shin (2019). However, a key distinction is that we are putting parts of something together. Forecast combination exercises, on the other hand, construct an optimal weighted average using noisy estimates of the same quantity.

Core inflation estimation can be framed as a signal extraction problem, which has led to the proposal of many variants of factor analysis (Stock and Watson, 2016). Compared to those methods, especially when relying on Principal Component Analysis (PCA), Assemblage Regression retains necessary shrinkage, while ensuring, through supervision, that the extracted signal is relevant. By discarding noise, the resulting factors are expected to incorporate signals relevant for forecasting headline inflation. However, they are not optimized for this purpose. Factors are extracted as the linear combination of components that best explain the variation in all components.

Conceptually, focusing on maximal predictability is not a radical departure from approaches that build on Gordon (1975)'s popular suggestion by either excluding or reweighting items according to their overall volatility, their cyclical volatility, their persistence, or even their sensitivity to the economic business cycle. Albacore$_{Comps}$ encompasses most of these objectives but does so within a framework explicitly focused on forecasting headline inflation.

Assemblage Regression in the component space can be seen as a constrained multivariate-to-univariate Hamilton filter, where, instead of regressing the target on its own lags, we regress it on its lagged subcomponents. Furthermore, the target is the average path between the forecast date and the forecast horizon. The usage of the average path, typical in forecasting studies and particularly natural for inflation, leads it to take an implicit mean over many "horizon tuning parameters", a strategy that has been documented to be successful in output gap extraction applications.

As argued in Goulet Coulombe et al. (2024), there have been rare instances in the literature where the predictability of headlines has played a more direct role in informing the weighting in the components space. Because of that, some papers focus on aggregation schemes. However, they generally overlook the covariance structure (Ravazzolo and Vahey, 2009) or lack, necessary nonnegativity restrictions (Gamber and Smith, 2019), among other things.

Albacore$_{Ranks}$ is Goulet Coulombe et al. (2024)'s proposal for a maximally forward-looking core inflation measure based on temporary exclusion. Traditionally, this is achieved by systematically ranking individual price changes from lowest to highest at each point in time and excluding the upper and lower tails of the distribution—either symmetrically or asymmetrically—with median inflation (Bryan and Pike, 1991) representing the extreme case of retaining only the midpoint rank. In the forecast combination literature involving trimming methods, most approaches focus on permanently eliminating the least accurate contributors (Diebold and Shin, 2019) or experimenting with different cutoff configurations for a trimmed mean estimator (Stock and Watson, 2004).

Albacore$_{Ranks}$ offers greater flexibility by assigning optimized weights to each rank rather than merely selecting trimming points. Mechanically, Albacore$_{Ranks}$ is a constrained functional regression (Morris, 2015), where the predictors (i.e., order statistics series) are functions of the empirical distribution of price growth rates.

Finally, it is worth noting that there have been numerous papers forecasting inflation using machine learning methods (Medeiros et al., 2021), Recently, there has been some attention on coupling these with components-level data (Boaretto and Medeiros, 2023). We differ from this strand of literature by our objective, to create a core inflation measure with desirable properties. This leads us to focus on constrained linear forecasts based on components, rather than on changes in functional form.

3 A Forward Looking Core Inflation for Brazil

3.1 Data

We use the IPCA[2] index, which is produced by the IBGE[3]. IPCA is the official CPI used in Brazil within the inflation-targeting framework. We have monthly data spanning from 1999m08 to 2025m12, comprising 317 observations before transformations.

For comparison purposes, we also use nine core inflation measures developed by the Central Bank of Brazil for the IPCA. These measures encompass a variety of methodologies commonly employed in the related literature. The measures are: EX0, which excludes all items from food-at-home and administered prices[4]; EX1, which excludes 10 out of 16 items from food-at-home and all administered prices; EX2, which excludes the same 10 food-at-home items and also vehicle and domestic fuels from administered prices; EX3, which builds on EX2 with further exclusions; P55, which reflects the exact variation of the 55th percentile of the items' distribution; EXFE, which excludes all items from food-at-home and energy prices; MA, a "trimmed mean" measure where the 20% of items with the largest and smallest variations are excluded; MS, which applies the same trimming procedure as MA but also smooths the variations of specific remaining items; and DP, a "double-weighting" measure in which the new weights are a combination of the original ones and each item's volatility.

[2] Portuguese acronym for Broad National Consumer Price Index

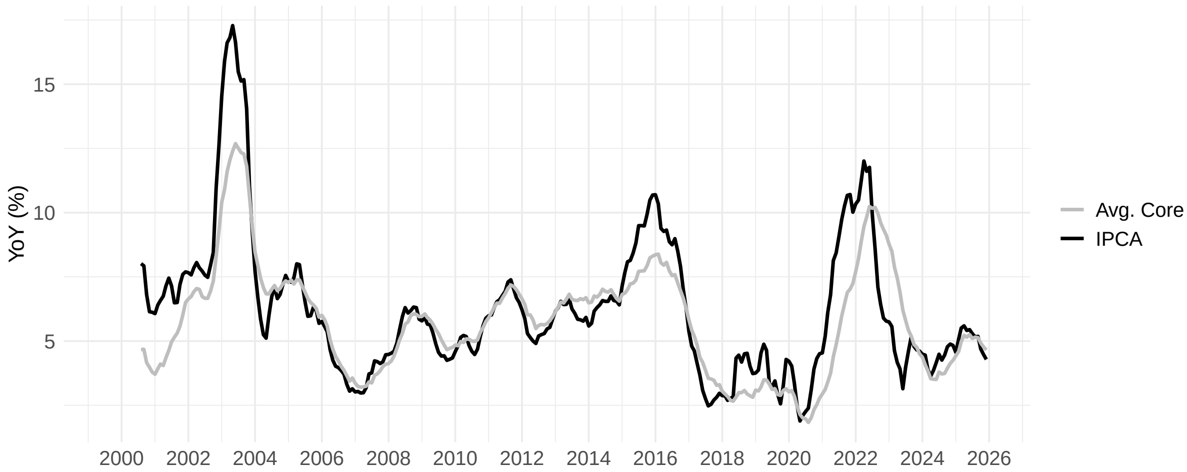
[3] Portuguese acronym for Brazilian Institute of Geography and Statistics

[4] All core inflation time series are taken from the Central Bank of Brazil website. The corresponding series codes in the Central Bank database are: EX0 (11427), EX1 (16121), EX2 (27838), EX3 (27839), P55 (28750), EXFE (28751), MA (11426), MS (4466), DP (16122).

All disaggregated series are seasonally adjusted and expressed as 3-month changes, which balance timeliness (compared to 12-month trailing averages) and noise reduction (relative to month-over-month growth rates). Rank positions are calculated using month-over-month growth rates (as done for trimmed mean inflation), and the resulting order statistics time series are then smoothed using a 3-month moving average. We use IBGE's seasonally adjusted series. For the Central Bank's core inflation measures, since the authority does not provide official seasonally adjusted series, we apply seasonal adjustment using the X-12 ARIMA method.

The plot below shows year-over-year inflation as measured by the IPCA and the average core measure, which is the arithmetic mean between all core measures produced by Central Bank of Brazil.

Figure 1: Year-Over-Year Inflation Rate (%)



Several visual insights emerge from the graph. The first is the high volatility of inflation in Brazil. As shown in the table 2 in the appendix, the sample mean is approximately 6.2%, with a standard deviation of 4.1%. This implies that, about 63% of the time, annual inflation falls between 2.2% and 10.3%, which is a rather wide range. Additionally, there are at least three periods of high inflation in the sample: 2002, 2015, and 2021, during which annual inflation exceeded 10%. Conversely, the only clear episode of low inflation appears to be the 2017–2019 window, following Brazil's economic crisis[5].

Another important visual insight from the graph is the lagging behavior of the average core inflation, specially after Covid-19 crisis. In this sense, it appears to function more as a lagged filtered series than a forecasting tool. Moreover, as shown in table 2 in the appendix, all core inflation measures display lower standard deviations than the IPCA, which is both expected and desirable. However, they also exhibit a downward bias, as their means are lower than that of the headline index. Likely due to this, we will show in the next section that linear combinations (without intercept) of the core inflation measures used by the Central Bank perform poorly in forecasting. However, we observe a significant improvement in predictive performance when a constant term is included.

3.2 Results and Forecasting Performance

Constructing Albacore for each level involves predicting the average path of headline inflation for h months ahead. We construct Albacore based on $h \in \{3, 6, 12, 24\}$, capturing short-term and medium-term developments of interest to policymakers.

[5] Brazilian GDP fell by 6% cumulatively over the 2015–2016 period.

We define two out-of-sample test sets: one covering the period prior to the COVID-19 crisis (2016m01 - 2019m12) and another encompassing the post-COVID period (2020m01 - 2025m12). As illustrated in Figure 1, these two periods display distinct characteristics. The former is characterized by a rapid decline followed by a phase of low inflation, while the latter captures the COVID-19 inflation shock, marked by a sharp rise and subsequent fall in inflation.

We evaluate the forecasting performance with root mean squared errors (RMSEs) and Model Confidence Sets, as proposed by Hansen et al. (2011). We base the estimation on an expanding window. We compare the forecasting accuracy to a set of benchmark models. This set includes all the officially reported core inflation and trimmed inflation series from the Central Bank of Brazil. As pointed out by Goulet Coulombe et al. (2024), given that the core inflation time-series address the problem from various angles, the information they convey may differ. Hence, instead of focusing on a single measure of core inflation we enhance our benchmarks by combining them and, in this way, performing an ex-ante weighting of the best currently available measures. Moreover, in 2025, some of the BCB core inflation measures were re-estimated, and their historical time series were revised. This introduces a form of data leakage in the BCB measures, creating look-ahead bias in the out-of-sample evaluation for the benchmark. Consequently, this makes it more difficult for alternative models to outperform them.

The benchmarks are constructed by including different series in a nonnegative regression. Our first benchmark, which we call B_{Int}, is comprised of headline, every core and trimmed mean inflation cited along with an intercept, which allows the model to include the long-run average. The second benchmark model, which we call B_{nch}, is a linear combination of the nine official core inflation measures, headline inflation itself and the average core inflation measure. To clarify, Table 3 in the appendix presents the benchmark weights using the 1999m10–2019m12 training sample as an example.

We construct Albacore for Brazil using monthly prices at different levels of disaggregation. Our core inflation measures are based on levels 2 (referred to as "subgrupos" in IBGE's classification), 4 ("itens") and 7 ("subitens"), which include 19, 51, and 377 subcomponents, respectively[6].

Table 1 presents the forecasting performance of Albacore for Brazil. The left panel reports results for the 2016m01–2019m12 test sample, while the right panel shows results for the 2020m01–2025m12 period. In both cases, we present results for Albacore$_{Comps}$ and Albacore$_{Ranks}$ with $K \in \{19, 51, 377\}$ and $h \in \{3, 6, 12, 24\}$, alongside our benchmarks B_{Int} and B_{nch}, as described earlier. The values in bold are in the superior Confidence Set according to the Range Statistic, as in Hansen et al. (2011), considering the forecast horizon and the out-of-sample period.

First, it is worth examining the performance of the benchmark models. As expected, in the relatively stable low-inflation period (2016m01 to 2019m12), the B_{nch} benchmark model performs fairly well, remaining within the confidence set for forecast horizons $h = 3$, $h = 6$ and $h = 12$.The opposite pattern emerges in the post-COVID period (2020m01–2025m12), during which B_{nch} performs clearly worse than B_{Int}. As discussed earlier, this outcome is likely driven by the downward bias present in the Central Bank of Brazil's core inflation measures, which makes a linear combination of all central bank core measures (B_{nch}) an effective forecasting tool primarily during periods of below-average inflation. This contrasts with the findings of Goulet Coulombe et al. (2024) for the United States, where the benchmark models perform best when excluding the intercept during highly volatile inflation periods, while the intercept version yields better results in stable inflation environments.

Turning to the Albacore models over the 2016m01–2019m12 test period, the best results are achieved by Albacore$_{Comps}$ at level 4 ($K = 51$) and level 7 ($K = 377$). The former belongs to the confidence set for $h = 3$, $h = 6$ and $h = 12$, while the latter is included in the superior set for $h = 6$, $h = 12$ and

[6]While the IBGE labels these levels as subgroups (subgrupos), items (itens), and subitems (subitens), we refer to them as levels 2, 4, and 7 for clarity. This naming convention is based on the number of digits used by IBGE to define each level of disaggregation.

$h = 24$. Both models exhibit consistent performance, conditional on the MCS specification, throughout this out-of-sample window.

Table 1: Forecasting Performance of Albacore for Brazil - Root Mean Square Error

	2016m01–2019m12				2020m01–2025m12			
	h = 3	h = 6	h = 12	h = 24	h = 3	h = 6	h = 12	h = 24
Level 2 (K=19)								
Albacore_{Comps}	0.223	0.208	0.264	0.259	**0.304**	**0.329**	0.350	0.400
Albacore_{Ranks}	0.203	**0.194**	**0.220**	0.233	**0.329**	**0.294**	**0.245**	0.270
Level 4 (K=51)								
Albacore_{Comps}	**0.189**	**0.171**	**0.205**	0.230	**0.328**	0.348	0.355	0.394
Albacore_{Ranks}	0.220	0.212	**0.231**	0.237	**0.331**	**0.286**	**0.234**	**0.242**
Level 7 (K=377)								
Albacore_{Comps}	0.228	**0.189**	**0.220**	**0.209**	**0.324**	**0.288**	**0.265**	0.323
Albacore_{Ranks}	**0.177**	0.321	0.448	**0.187**	0.399	0.408	0.365	0.347
Benchmark								
B_{Int}	0.218	0.207	**0.218**	**0.221**	0.332	**0.281**	**0.232**	**0.188**
B_{nch}	**0.199**	**0.171**	**0.189**	0.226	0.358	0.336	0.353	0.360

Table shows RMSE values considering four forecast horizons (3,6,12 and 24) and two out-of-sample tests (2016m01–2019m12 and 2020m01–2025m12). The values in bold are in the superior Confidence Set according to the Range Statistic ($\alpha = 0.05$), as in Hansen et al. (2011), considering the forecast horizon and the out-of-sample period.

In the post-pandemic period (2020m01–2025m12), by contrast, Albacore_{Ranks} shows stronger performance. In particular, Albacore_{Ranks} at level 4 ($K = 51$) is included in the superior set for every forecast horizon in this out-of-sample evaluation. Although it never attains the lowest RMSE, the model delivers stable results conditional on the MCS specification over this test period.

Overall, these results suggest that, even in shock-prone environments, Albacore can extract valuable predictive signals from disaggregated data, particularly at short and medium horizons. The short-run accuracy gains delivered by Albacore ($h = 3, 6, 12$) are broadly consistent with those reported by Goulet Coulombe et al. (2024) for the United States, Canada and the Euro Area, suggesting a more general pattern. However, for medium- to long-run forecasts ($h = 24$), the results for Brazil appear to be less robust than those documented by Goulet Coulombe et al. (2024) for the United States. While Albacore-based measures are often able to match the benchmark models, they struggle to systematically outperform them. This finding is perhaps unsurprising, as inflation in an emerging economy tends to be, from a modeling perspective, more heavily influenced by unpredictable shocks.

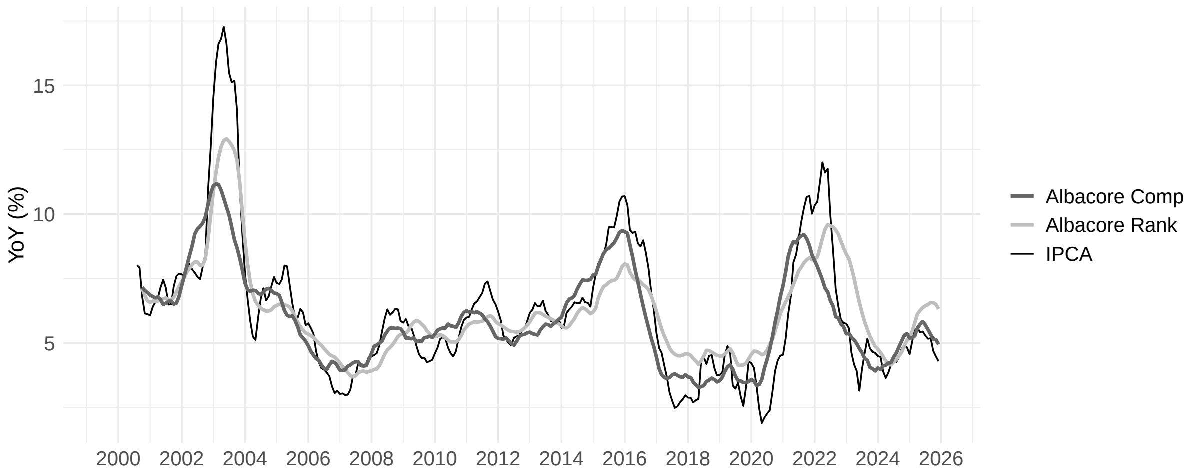
Comparing the two Albacore versions, Albacore_{Ranks} outperforms Albacore_{Comps} in the post-pandemic out-of-sample period, whereas the latter performs better in the pre-pandemic evaluation. This pattern suggests that Albacore_{Ranks} has greater potential to signal early inflationary pressures, while Albacore_{Comps} appears to be more precise in stable inflation environments. In this sense, both models seem suitable for practical applications because they complement each other. One might consider tracking both an Albacore*Comps* and an Albacore_{Ranks} model, depending on the specific objectives, such as forecast horizon and timeliness.

On the one hand, these results partially reaffirm the findings of Goulet Coulombe et al. (2024) regarding the strong performance of Albacore_{Ranks} in high inflation risk environments. On the other hand, they contrast with the evidence reported for the United States, where Albacore_{Ranks} emerges as the superior model regardless of forecasting horizon, level of disaggregation, or evaluation sample. Our findings are more closely aligned with those for Canada and the Euro Area, where no single model consistently dominates across all settings[7].

In the appendix, we present the same table using confidence sets constructed with semi-quadratic statistics rather than range statistics. As these statistics are less sensitive to outliers, the confidence sets reported in the appendix tend to be more parsimonious. Nevertheless, this alternative specification also indicates that Albacore$_{Ranks}$ (levels 2 and 4) performs better in the 2020m01–2025m12 out-of-sample period, while Albacore$_{Comps}$ (levels 4 and 7) delivers superior performance in the 2016m01–2019m12 evaluation period.

For a more intuitive understanding of our results, the graph below displays the year-over-year inflation rate as signaled by our Albacore measures alongside headline inflation. This Albacores were estimated for $h = 12$ using full sample. The Albacore$_{Comps}$ was estimated using $k = 377$ while Albacore$_{Ranks}$ was estimated for $k = 19$. We selected this combination to target a 12-month forecast horizon. The graph highlights a clear leading behavior during upward inflationary shocks, particularly for Albacore$_{Comps}$. It is worth noting, however, that this result is partly driven by construction, since the full sample is used to generate the plot below[8].

Figure 2: Year-Over-Year Inflation Rate (%)



These Albacores were estimated for $h = 12$ using full sample. The Albacore$_{Comps}$ was estimated using $k = 377$ while Albacore$_{Ranks}$ was estimated for $k = 19$.

In Table 5, in the appendix, we present descriptive statistics for the estimated Albacores. Beyond their forecasting advantages, these core inflation indicators also exhibit desirable properties such as lower bias and reduced volatility compared to the average core inflation measure.

3.3 Components and Ranks

In the graphs below, we illustrate the estimated coefficients (weights) of ranks and components. Figure 3 shows Albacore$_{Comps}$ weights for $h = 12$, considering full-sample estimation and level 7 ($K = 377$), which have been re-aggregated back to level 2 for ease of communication, compared to the headline inflation mean weights for 2025. Looking at it, some features stand out. First, *Transportation*, which includes

[7]The closer resemblance of our results to those reported for Canada and the Euro Area, relative to the United States, should not be interpreted as evidence of underlying structural similarities across these economies. A more parsimonious explanation is related to data availability and sample length. The United States benefits from a longer and more homogeneous inflation history, which may facilitate the identification of a single dominant forecasting model. In contrast, shorter or more limited samples—such as those available for Brazil, Canada, and the Euro Area in comparable exercises—may naturally lead to less clear-cut model dominance.

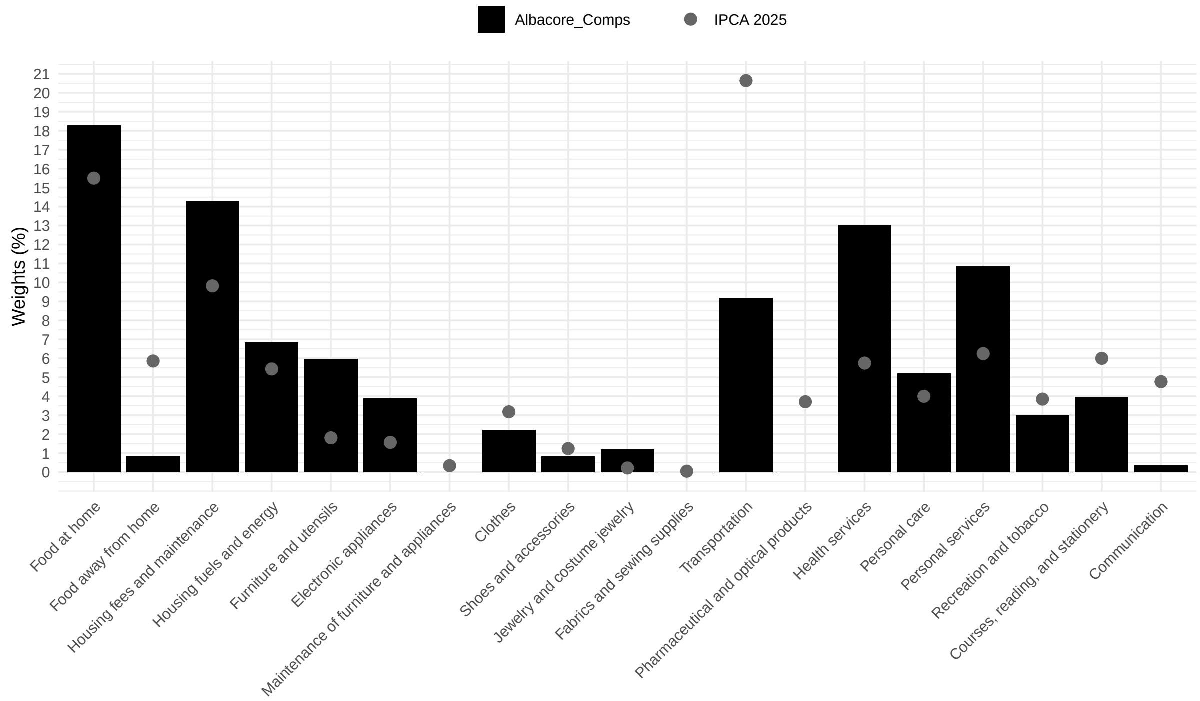
[8]To address this concern, the appendix presents the same graph in out-of-sample performances. The leading behavior is maintained

transportation fuel prices, receive significantly lower weights than in the headline inflation. This exclusion is similar to what many core measures implement, such as EX0 in the Brazilian case, although Albacore does not fully exclude this component. By contrast, *Food at home* receives a significantly higher weight than in the IPCA headline index, which differs from most exclusion-based core measures. In turn, *Food away from home* is assigned a lower weight relative to the headline index, which is also atypical from the perspective of traditional core measures. The algorithm may be capturing a leading relationship of *Food at home*[9] relative to *Food away from home* and to headline inflation itself. Accordingly, it suggests that placing excessive emphasis on the *Food away from home* component may not provide a strong forecasting signal.

Another noteworthy feature is that the model assigns higher weights to consumption goods such as *Electronic Appliances* and *Furniture*, while assigning lower weights to *Clothes* and *Shoes*, relative to the official headline inflation. This pattern may be related to product characteristics. The items receiving higher weights tend to be more expensive and are typically purchased with credit. IIn this sense, the result may carry meaningful economic interpretation, reflecting the role of credit conditions in shaping consumption-driven inflation dynamics.

Finally, it is worth noting the substantial contribution of three service categories: *Housing Fees*, *Health Services* and *Personal Services*. They receive nearly double its official weight. This finding is broadly consistent with Goulet Coulombe et al. (2024) for the United States, where the authors also show that health and other service components are informative indicators of trend inflation.

Figure 3: Albacore $_{Comps}$ Weights (%)



This Albacores$_{Comps}$ was estimated, using full sample, for $h = 12$ and $k = 377$, but reaggregated back to $k = 19$ for ease of communication.

In fact, the high-dimensional regression framework enables an investigation of component importance at the most disaggregated level, revealing several noteworthy patterns. The first relates to the IPCA methodology and the role of price regulation in Brazil. Many prices are imputed. Moreover, a significant

[9] We could call this component as Food Goods.

share of prices in Brazil are regulated, that is, their adjustments are determined by regulatory agencies or state-owned enterprises. Approximately 25% of the IPCA basket corresponds to regulated prices.

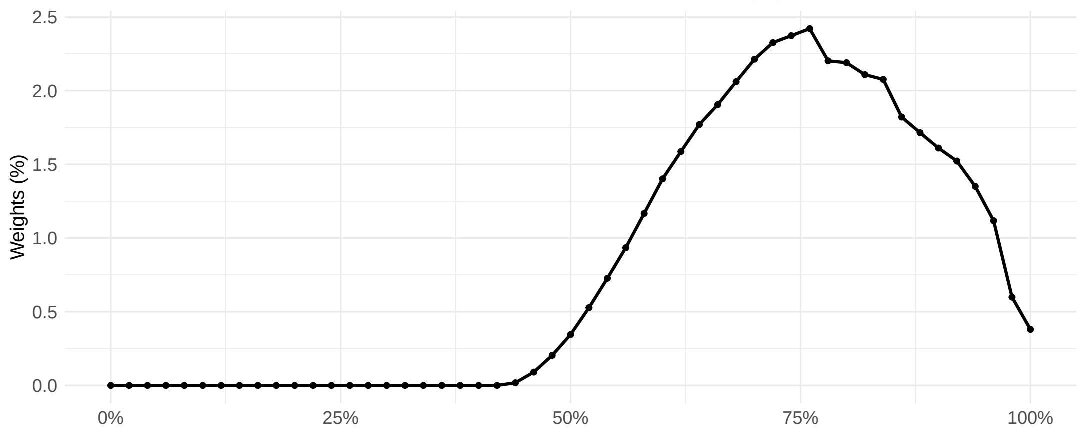
This combination of the IPCA's methodological features and price regulation may distort the economic interpretation of some of the weights obtained by Albacore. For instance, *Health Insurance* is one of the heaviest items in the IPCA, and the Albacore model assigns it an even higher weight (7.1% versus 4.1% in the headline index). However, it is difficult to determine whether this increase reflects the true economic relevance of this component to future inflation or the IBGE's methodological treatment. Health insurance prices are adjusted once a year by a regulatory agency, but the adjustment is incorporated linearly into the IPCA (one-twelfth each month). Since these adjustments are usually based on past inflation, *Health Insurance* effectively behaves like a moving intercept.

A similar anomalous pattern appears in the *Professional Council Fees* component, which represents roughly 0.05% of the IPCA. This item refers to the annual contributions that workers pay to unions and professional associations. Its value is also regulated and defined once a year, and the IBGE distributes the annual change evenly across the following twelve months, causing it to behave as another moving intercept. In this case, there is little room for an economic interpretation of the high weight (5.4%) assigned by the model, which likely stems from this intercept-like effect.

An interesting feature enabled by the high-dimensional regression framework is the potential for economic analysis. If we focus on non-regulated prices, several items stand out in the Albacore estimation at the most disaggregated level: *Used Vehicles* (5.0% vs 1.6% in IPCA healdline), *Dentist Services* (5.2% vs 0.6%), *Video Games* (3.9% vs 0.003%), and *Nail Polish* (3.7% vs 0.004% in the headline), *Parking* (3.6% vs 0.1%) and *Condominium Fees* (6.2% vs 2.2%). For example, the higher weights for *Dentist Services* may reflect increased demand for orthodontic treatments, whitening and other elective procedures due to higher income. Similarly, *Used Automobiles* and *Video Games* could be responding to domestic demand or external shocks, while *Nail Polish* may follow trends in non-essential consumption enabled by income growth. The elevated weights of *Parking* and *Condominium Fees* likely reflect their association with higher usage of private transportation during periods of income growth (in the case of *Parking*) and annualized adjustments (in the case of *Condominium Fees*). However, it is obvious that the model does not allow us to make precise causal claims regarding these mechanisms. It is only suggestive.

Figure 4 shows the weights of Albacore$_{Ranks}$. It exhibits a rather asymmetric trimming compared with standard core measures, with ranks below the median receiving almost zero weight. The highest weights lie in the fourth quartile, so the distribution is centered around the 75th percentile, which is almost the same as Goulet Coulombe et al. (2024) found for the United States.

Figure 4: Albacore$_{Ranks}$ Weights (%)



This

Albacore$_{Ranks}$ weights were estimated for $h = 12$ and $k = 51$ using full sample.

As noted by Rich et al. (2022), traditional trimming methods tend to underestimate early trends. In addition, upward price adjustments generally exhibit greater persistence than downward ones. Empirical evidence on firm pricing behavior indicates that prices respond more quickly to cost increases than to cost decreases (Nakamura and Steinsson, 2008). Accordingly, Albacore$_{Ranks}$ assigns relatively greater importance to price changes that are more likely to endure over time.

Last but not least, it is important to emphasize that the component weights vary substantially across samples, forecast horizons and levels of disaggregation. Providing a unified economic interpretation for all these changes is challenging, as no clear and systematic pattern emerges. In contrast, the rank weights are remarkably stable across aggregation levels and display a clear pattern. As the forecast horizon increases, the weights shift toward the right tail of the distribution. We show this pattern in the appendix.

4 Final Remarks

This paper constructs core inflation measures for Brazil that are explicitly designed to forecast headline inflation. We apply the assemblage regression methodology proposed by Goulet Coulombe et al. (2024), thereby addressing a key limitation of existing indicators and testing Goulet Coulombe et al. (2024)'s framework in a more challenging environment. To our knowledge, we are the first to do so. Our application to Brazilian inflation shows that forward-looking core measures can be adapted to an emerging market context, providing meaningful predictive signals at both short and medium horizons.

Although their forecasting performance generally matches or outperforms benchmark combinations of official measures, its predictive gains are generally lower than that observed in advanced economies, likely reflecting Brazil's more volatile inflation, regulated prices and numerous imputed CPI components. Future research could address these challenges by focusing on free-market prices or alternative sub-components. In addition, subsequent work could investigate whether this loss of accuracy, observed in the Brazilian case, also occurs in other emerging markets.

The forecasting accuracy gains of the Brazilian versions of the maximally forward-looking core inflation measures are relatively more concentrated in the short run (3 and 6 months ahead), whereas, as reported by Goulet Coulombe et al. (2024), in advanced economies they are relatively more accurate in the medium to long run (12 and 24 months ahead).

Another notable finding is the high weights estimated for *Food at home* (essentially food goods) and *Transportation* (essentially transport fuel prices). This result contrasts sharply with the findings of Goulet Coulombe et al. (2024) for advanced economies. It suggests that traditional core measures, which typically exclude these components, may not be well-suited for emerging economies that are more exposed to external shocks. Future research could explore whether this pattern is general across emerging markets.

Overall, the methodology proves to be robust and flexible, demonstrating that disaggregated price information can enhance predictive accuracy relative to traditional core measures while also improving the economic understanding of price dynamics. The paper provides a practical tool for policymakers and underscores the value of forecasting-oriented core inflation measures in emerging markets.

References

J. Aizenman, M. Hutchison, and I. Noy. Inflation targeting and real exchange rates in emerging markets. *World development*, 39(5):712–724, 2011.

G. Boaretto and M. C. Medeiros. Forecasting inflation using disaggregates and machine learning. *arXiv preprint arXiv:2308.11173*, 2023.

M. F. Bryan and C. Pike. Median price changes: an alternative approach to measuring current monetary inflation. *Economic Commentary*, (12/1/1991), 1991.

F. X. Diebold and M. Shin. Machine learning for regularized survey forecast combination: Partially-egalitarian lasso and its derivatives. *International Journal of Forecasting*, 35(4):1679–1691, 2019.

E. N. Gamber and J. K. Smith. Constructing and evaluating core inflation measures from component-level inflation data. *Journal of Forecasting*, 38(8):833–852, 2019.

R. J. Gordon. Alternative responses of policy to external supply shocks. *Brookings Papers on Economic Activity*, 1975(1):183–206, 1975.

P. Goulet Coulombe, K. Klieber, C. Barrette, and M. Göbel. Maximally forward-looking core inflation. *Available at SSRN 4758517*, 2024.

J. D. Hamilton. Why you should never use the hodrick-prescott filter. *Review of Economics and Statistics*, 100(5):831–843, 2018.

P. R. Hansen, A. Lunde, and J. M. Nason. The model confidence set. *Econometrica*, 79(2):453–497, 2011.

T. Hastie, R. Tibshirani, and M. Wainwright. Statistical learning with sparsity. *Monographs on statistics and applied probability*, 143(143):8, 2015.

V. Machado. Core inflation in brazil: past and present. Technical report, Central Bank of Brazil, Research Department, 2024.

M. C. Medeiros, G. F. Vasconcelos, Á. Veiga, and E. Zilberman. Forecasting inflation in a data-rich environment: the benefits of machine learning methods. *Journal of Business & Economic Statistics*, 39 (1):98–119, 2021.

J. S. Morris. Functional regression. *Annual Review of Statistics and Its Application*, 2(1):321–359, 2015.

E. Nakamura and J. Steinsson. Five facts about prices: A reevaluation of menu cost models. *The Quarterly Journal of Economics*, 123(4):1415–1464, 2008.

P. M. Pincheira, J. Selaive, and J. Nolazco. The evasive predictive ability of core inflation. *Available at SSRN 2712490*, 2016.

F. Ravazzolo and S. P. Vahey. Measuring core inflation in australia with disaggregate ensembles. *INFLATION IN AN ERA OF RELATIVE PRICE SHOCKS*, page 178, 2009.

R. Rich and C. Steindel. A review of core inflation and an evaluation of its measures. 2005.

R. W. Rich, R. J. Verbrugge, and S. Zaman. Adjusting median and trimmed-mean inflation rates for bias based on skewness. *Economic Commentary*, (2022-05), 2022.

T. N. T. d. Silva Filho and F. M. R. Figueiredo. Has core inflation been doing a good job in brazil? *Revista Brasileira de Economia*, 65:207–233, 2011.

J. H. Stock and M. W. Watson. Combination forecasts of output growth in a seven-country data set. *Journal of forecasting*, 23(6):405–430, 2004.

J. H. Stock and M. W. Watson. Core inflation and trend inflation. *Review of Economics and Statistics*, 98(4):770–784, 2016.

X. Wang, R. J. Hyndman, F. Li, and Y. Kang. Forecast combinations: An over 50-year review. *International Journal of Forecasting*, 39(4):1518–1547, 2023.

A Descriptive Statistics

Table 2: Descriptive Statistics - SA 3M Moving Avg Annualized (%)

	1999m10–2015m12		2016m01–2019m12		2020m01–2025m12		Full Sample	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
EX0	6.1	2.2	3.6	1.7	5.2	3.0	5.5	2.5
EX1	6.6	2.8	4.2	1.9	5.4	2.7	6.0	2.8
EX2	6.5	2.9	3.5	2.0	5.9	3.4	5.9	3.1
EX3	6.4	2.4	3.5	1.8	5.6	3.2	5.8	2.7
P55	5.9	2.5	4.2	1.8	5.5	2.4	5.6	2.5
EXFE	6.4	2.4	4.3	1.9	5.3	2.8	5.8	2.5
MS	6.5	2.2	4.4	1.9	5.5	2.8	6.0	2.4
MA	5.5	2.5	3.6	1.7	5.0	2.4	5.1	2.4
DP	6.6	2.8	4.3	2.0	5.5	2.6	6.0	2.8
IPCA	6.8	4.3	4.4	2.9	5.9	4.1	6.2	4.1
Avg. Core	6.3	2.4	4.0	1.8	5.4	2.7	5.7	2.5

Average and standard deviation values calculated within periods.

B Benchmark Design Example

To clarify, table 4 presents the benchmark weights using the 1999m10–2019m12 training sample. The left panel displays the B_{NCH} weights by forecast horizon. For the short-run ($h = 3$), the algorithm returns a simple average. As the horizon increases — for example, at $h = 12$ — the algorithm assigns non-zero weights only to the EX0 and MS measures. The right panel shows the B_{INT} weights, which include an intercept. A positive intercept emerges, likely correcting for the downward bias of core measures discussed earlier. In this setup, the $h = 12$ benchmark relies primarily on the constant and the EX1 measure, for example.

Table 3: Benchmark Weights Sample: 1999m10-2019m12

	B_{nch}				B_{int}			
	h = 3	h = 6	h = 12	h = 24	h = 3	h = 6	h = 12	h = 24
Intercept	–	–	–	–	0.14	0.25	0.32	0.41
IPCA	0.09	0.00	0.00	0.09	0.05	0.00	0.00	0.00
EX0	0.09	0.09	0.32	0.09	0.00	0.00	0.00	0.00
EX1	0.09	0.22	0.00	0.09	0.46	0.41	0.37	0.21
EX2	0.09	0.05	0.00	0.09	0.18	0.10	0.00	0.00
EX3	0.09	0.00	0.00	0.09	0.00	0.00	0.00	0.00
P55	0.09	0.21	0.00	0.09	0.04	0.00	0.00	0.00
EXFE	0.09	0.00	0.00	0.09	0.00	0.00	0.00	0.00
MS	0.09	0.37	0.68	0.09	0.00	0.00	0.00	0.00
MA	0.09	0.00	0.00	0.09	0.00	0.00	0.00	0.00
DP	0.09	0.00	0.00	0.09	0.00	0.00	0.00	0.00
Avg. Core	0.09	0.06	0.00	0.09	0.00	0.00	0.00	0.00

For clarification, this table shows benchmarks composition for this specific training sample.

C Albacore Performance With a Different Confidence Set

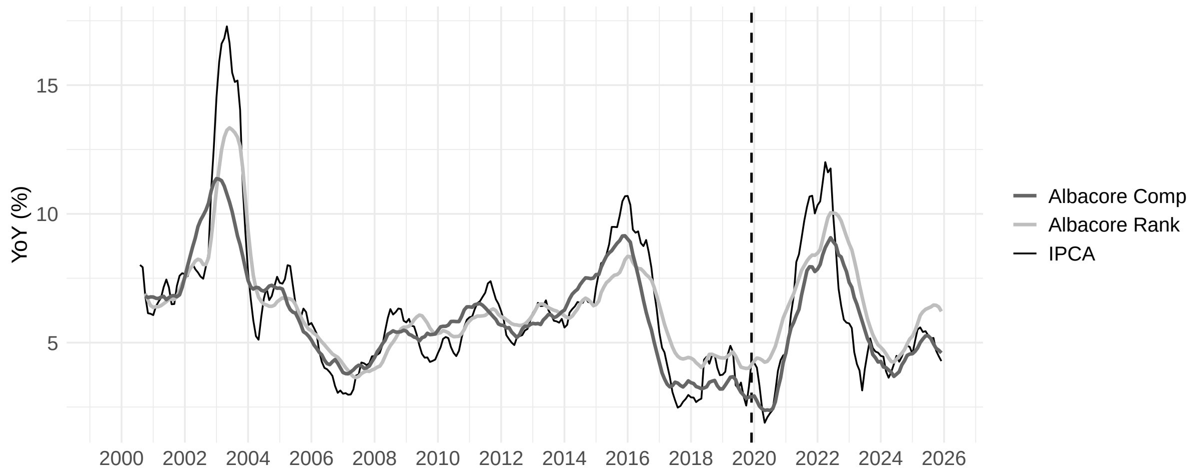
Table 4: Forecasting Performance of Albacore for Brazil - Root Mean Square Error

	2016m01–2019m12				2020m01–2025m12			
	h = 3	h = 6	h = 12	h = 24	h = 3	h = 6	h = 12	h = 24
Level 2 (K=19)								
Albacore_{Comps}	0.223	0.208	0.264	0.259	**0.304**	**0.329**	0.350	0.400
Albacore_{Ranks}	**0.203**	**0.194**	**0.220**	**0.233**	**0.329**	**0.294**	**0.245**	0.270
Level 4 (K=51)								
Albacore_{Comps}	**0.189**	**0.171**	**0.205**	**0.230**	**0.328**	0.348	0.355	0.394
Albacore_{Ranks}	0.220	**0.212**	**0.231**	**0.237**	**0.331**	**0.286**	**0.234**	**0.242**
Level 7 (K=377)								
Albacore_{Comps}	0.228	**0.189**	**0.220**	**0.209**	**0.324**	**0.288**	**0.265**	0.323
Albacore_{Ranks}	**0.177**	0.321	0.448	**0.187**	0.399	0.408	0.365	0.347
Benchmark								
B_{Int}	0.218	**0.207**	**0.218**	**0.221**	**0.332**	**0.281**	**0.232**	**0.188**
B_{nch}	**0.199**	**0.171**	**0.189**	**0.226**	**0.358**	**0.336**	0.353	0.360

Table shows RMSE values considering four forecast horizons (3,6,12 and 24) and two out-of-sample tests (2016m01–2019m12 and 2020m01–2025m12). The values in bold are in the superior Confidence Set according to the Semi-Quadratic statistics ($\alpha = 0.05$). In general, the Semi-Quadratic is less subject to outliers than the Range statistics. So, it often retain more models in the confidence set.

D Albacores Outlook in and Out-of-Sample Test Setting

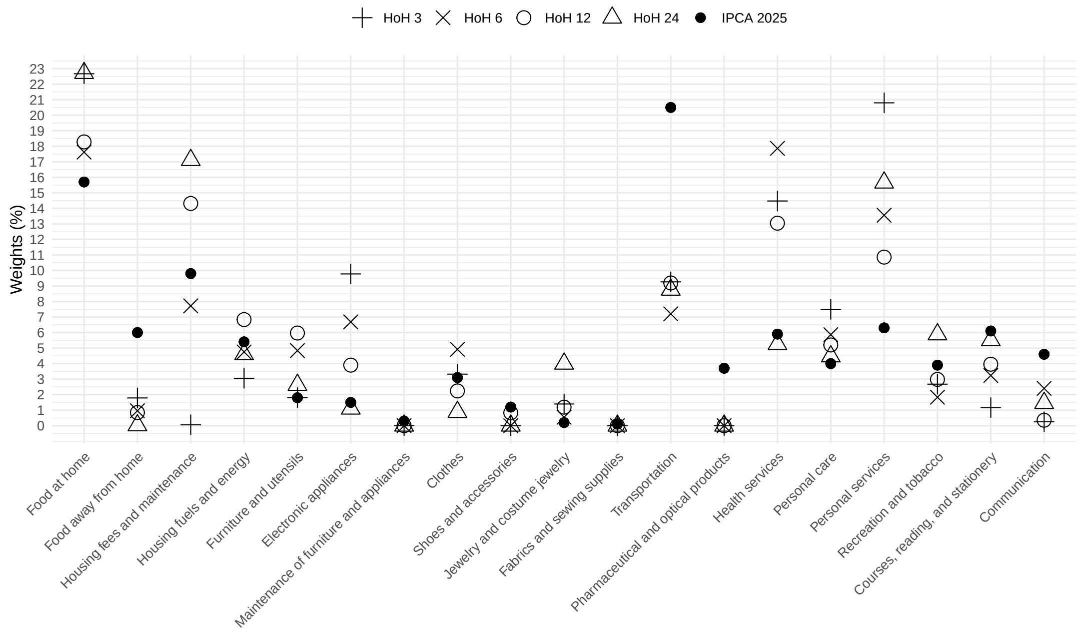
Figure 5: Year-Over-Year Inflation Rate (%) - 2020m01-2025m12 Out-of-Sample test



These Albacores were estimated for $h = 12$ using sample 1999m08-2019m12 sample. The out-of-sample test starts 2020m01. The Albacore_{Comps} was estimated using $k = 377$ while Albacore_{Ranks} was estimated for $k = 19$.

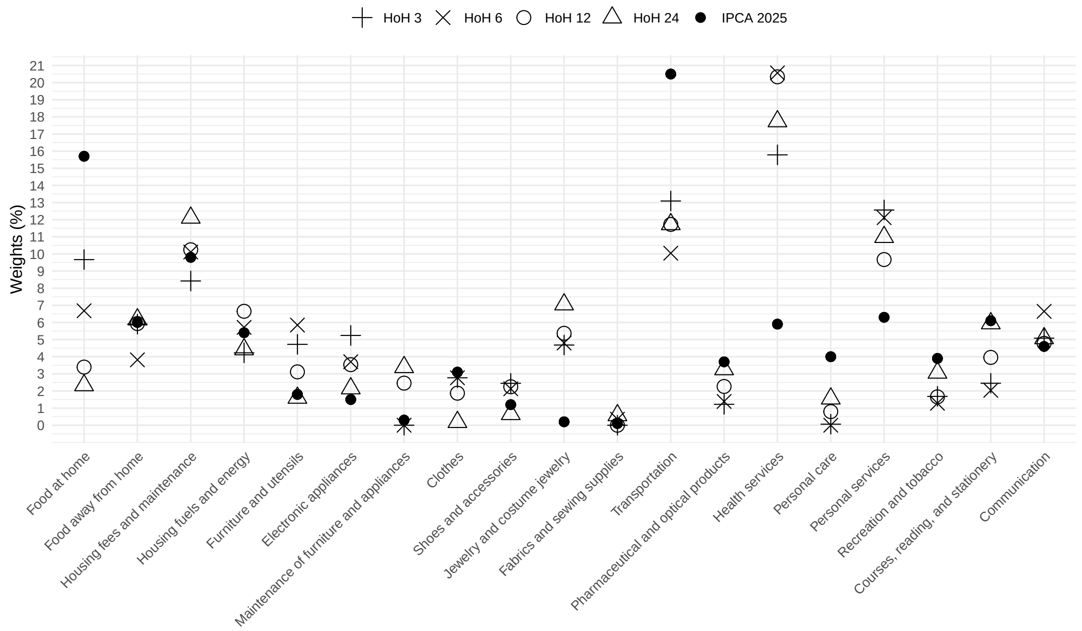
E Albacore$_{Comps}$ Weights According to Forecast Horizon

Figure 6: Albacore $_{Comps}$ Weights - ($K = 377$)



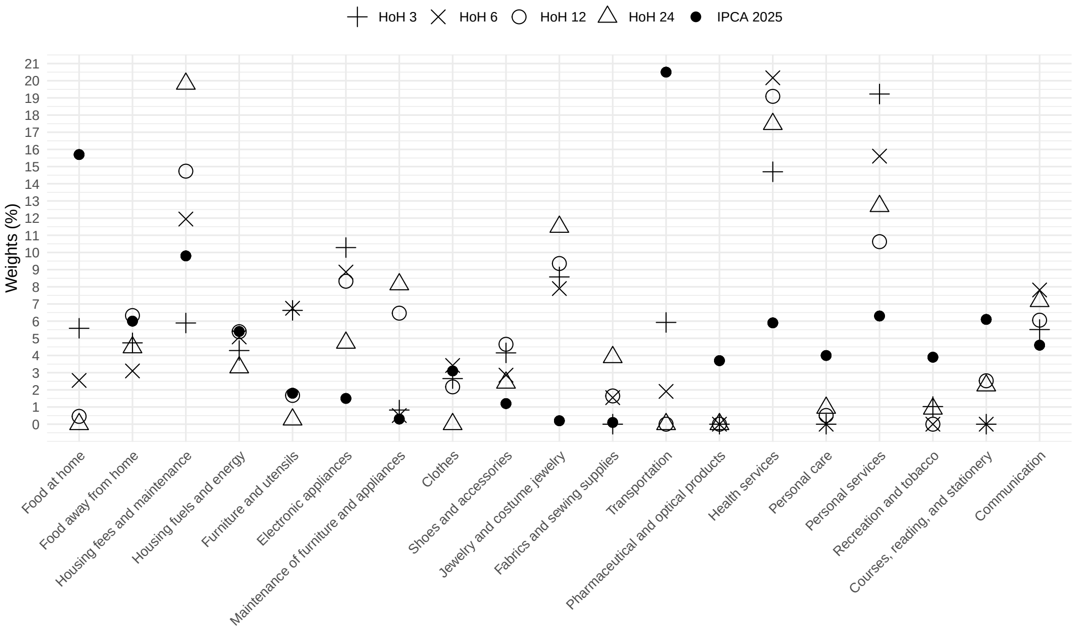
This Albacores$_{Comps}$ was estimated, using full sample, for $h = 12$ and $k = 377$, but reaggregated back to $k = 19$ for ease of communication.

Figure 7: Albacore $_{Comps}$ Weights - ($K = 51$)



This Albacores$_{Comps}$ was estimated, using full sample, for $h = 12$ and $k = 51$, but reaggregated back to $k = 19$ for ease of communication.

Figure 8: Albacore $_{Comps}$ Weights - ($K = 19$)

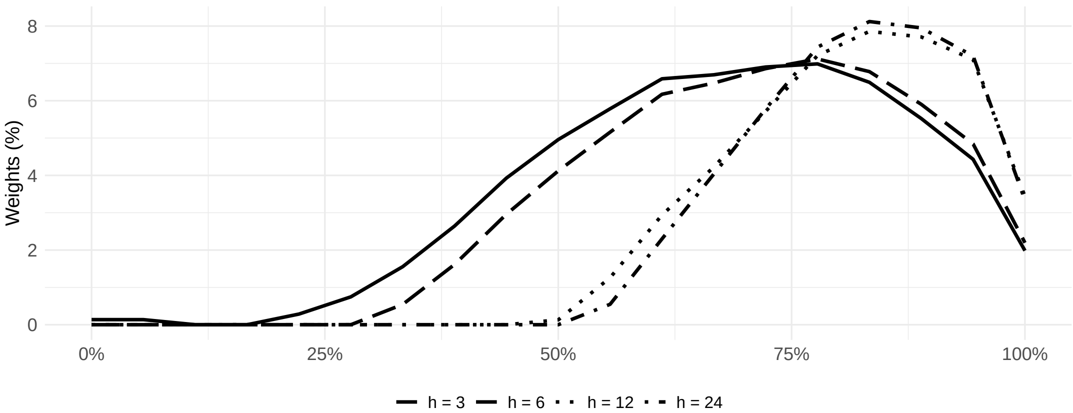


This

Albacores$_{Comps}$ was estimated, using full sample, for $h = 12$ and $k = 19$.

F Albacore$_{Ranks}$ Weights According to Forecast Horizon

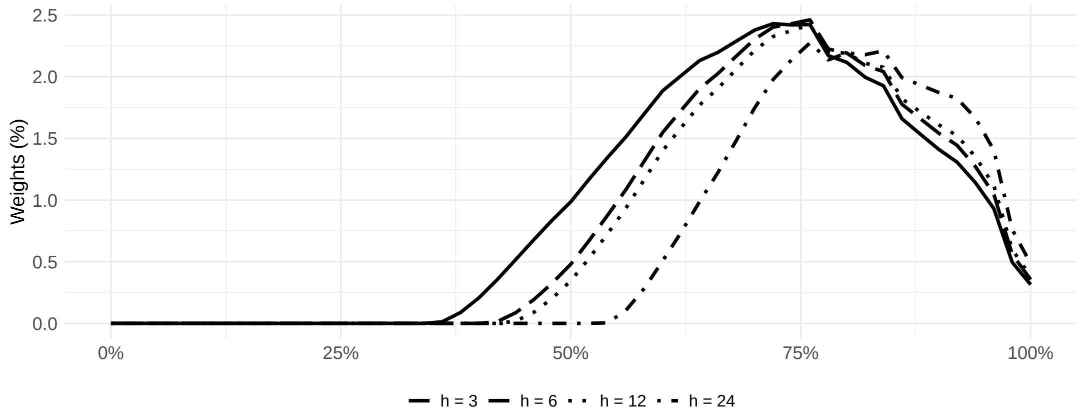
Figure 9: Albacore$_{Ranks}$ Weights According to Forecast Horizon ($K = 19$)



This

Albacore$_{Ranks}$ weights were estimated with $k = 51$ using full sample.

Figure 10: Albacore$_{Ranks}$ Weights According to Forecast Horizon ($K = 51$)



This Albacore$_{Ranks}$ weights were estimated with $k = 19$ using full sample.

G Other Underlying Properties of Estimated Albacores

Table 5: Descriptive Statistics - Albacores, Average Core Inflation Measures for Brazil and IPCA (%)

	$h = 3$		$h = 6$		$h = 12$		$h = 24$	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Level 2 (K=19)								
Albacore$_{Comps}$	6.2	2.8	6.2	2.5	6.2	2.3	6.3	2.2
Albacore$_{Ranks}$	6.1	2.4	6.3	2.4	6.2	2.2	6.2	2.2
Level 4 (K=51)								
Albacore$_{Comps}$	6.0	2.8	6.0	2.4	6.1	2.2	6.1	2.1
Albacore$_{Ranks}$	6.1	2.7	6.1	2.7	6.1	2.6	6.1	2.5
Level 7 (K=377)								
Albacore$_{Comps}$	6.0	3.2	6.0	2.5	5.9	2.1	5.9	1.7
Albacore$_{Ranks}$	6.1	2.0	6.1	1.6	6.1	1.5	6.1	1.5
Avg. Core	5.7	2.5	5.7	2.5	5.7	2.5	5.7	2.5
IPCA	6.2	4.1	6.2	4.1	6.2	4.1	6.2	4.1

Average and standard deviation values calculated within periods.