International Journal of Forecasting 33 (2017) 679–693

Contents lists available at ScienceDirect

International Journal of Forecasting

journal homepage: www.elsevier.com/locate/ijforecast

Real-time inflation forecasting with high-dimensional models: The case of Brazil

Márcio G.P. Garcia [a], Marcelo C. Medeiros [a,*], Gabriel F.R. Vasconcelos [b]

[a] *Department of Economics, Pontifical Catholic University of Rio de Janeiro, Rua Marquês de São Vicente 225, Gávea, Rio de Janeiro, 22451-900, Brazil*

[b] *Department of Electrical Engineering, Pontifical Catholic University of Rio de Janeiro, Rua Marquês de São Vicente 225, Gávea, Rio de Janeiro, 22451-900, Brazil*

ARTICLE INFO

Keywords:
Real-time inflation forecasting
Emerging markets
Shrinkage
Factor models
LASSO
Regression trees
Random forests
Complete subset regression
Machine learning
Model confidence set
Forecast combination
Expert forecasts

ABSTRACT

We show that high-dimensional econometric models, such as shrinkage and complete subset regression, perform very well in the real-time forecasting of inflation in data-rich environments. We use Brazilian inflation as an application. It is ideal as an example because it exhibits a high short-term volatility, and several agents devote extensive resources to forecasting its short-term behavior. Thus, precise forecasts made by specialists are available both as a benchmark and as an important candidate regressor for the forecasting models. Furthermore, we combine forecasts based on model confidence sets and show that model combination can achieve superior predictive performances.

1. Introduction

Forecasting inflation in real-time is difficult and has been studied extensively in the literature. The forecasting of inflation has been crucial for both academics and practitioners at least since Fisher (1930) introduced the concept of real interest rates. We estimate models for forecasting inflation in real-time and in data-rich environments. By real-time we mean that the forecasts are computed based solely on the information that was available to the econometrician at the time when the forecasts were made. A data-rich environment is one in which the number of potential predictors is large, possibly larger than the sample size. We consider the case of an emerging economy with inflation targeting, where precise inflation forecasts are of the utmost importance for monetary policy and investment strategies (Iversen, Laséen, Lundvall, & Söderström, 2016).

Emerging markets usually exhibit higher and more volatile inflation, which tends to shorten the investment horizon. In Brazil, a country that only conquered hyper-inflation in 1994, most fixed-income assets are still very short. Therefore, the forecasting of short-term inflation is more important than in advanced economies, and financial institutions tend to devote more resources to the endeavor. Short-term inflation forecasting in Brazil is a difficult exercise, with lots of data, but it is also one in which extremely good expert forecasts exist and against which different econometric techniques may be compared.

The literature on inflation forecasting is vast, and there is substantial evidence that models based on the Philips curve do not provide good inflation forecasts. Although Stock and Watson (1999) showed that many production-related variables are useful predictors of US inflation, Atkeson and Ohanian (2001) showed that the Philips curve fails to beat even simple naïve models in many cases.

* Corresponding author.
E-mail addresses: mgarcia@econ.puc-rio.br (M.G.P. Garcia), mcm@econ.puc-rio.br (M.C. Medeiros), gabrielrvsc@yahoo.com.br (G.F.R. Vasconcelos).

http://dx.doi.org/10.1016/j.ijforecast.2017.02.002
0169-2070/

These results inspired researchers to investigate a range of different models and variables in order to improve inflation forecasts, with the variables used including financial variables (Forni, Hallin, Lippi, & Reichlin, 2003), commodity prices (Chen, Turnovsky, & Zivot, 2014) and expectation variables (Groen, Paap, & Ravazzolo, 2013).

Real-time inflation forecasting has been considered by several authors in recent years. Iversen et al. (2016) evaluated the forecasts made in real time to support monetary policy decisions at the Swedish Central Bank from 2007 to 2013. The authors compared dynamic stochastic general equilibrium (DSGE) models with Bayesian vector autoregressive (BVAR) models. Monteforte and Moretti (2013) proposed a mixed-frequency model for the daily forecasting of euro area inflation in real-time. The authors showed that the predictive performance of the mixed-frequency model is superior to those of forecasts based only on economic derivatives. Clements and Galvão (2013) considered real-time inflation forecasts from AR models and with revised data. Finally, Groen et al. (2013) evaluated the use of Bayesian model averaging (BMA) for forecasting inflation in real-time. However, none of these authors considered the use of large-dimensional machine learning models.

There is also a growing body of literature on inflation forecasting in Brazil. Arruda, Ferreira, and Castelar (2011) used several linear and nonlinear models and the Phillips curve to forecast inflation. The authors showed that some nonlinear models and the simple autoregressive (AR) model produced smaller forecast errors than the Phillips curve. Figueiredo and Marques (2009) used long-memory heteroskedastic models to show that Brazilian inflation has long-range dependence on both the mean and the variance. However, they did not exclude the importance of the short-term AR component. The relevance of past inflation was also pointed out by Kohlscheen (2012). More recently, Medeiros, Vasconcelos, and Freitas (2016) considered different high-dimensional models for forecasting Brazilian inflation. The authors showed that techniques based on the *least absolute shrinkage and selection operator* (LASSO) have the smallest forecasting errors for short horizon forecasts. For longer horizons, the AR benchmark is the best model for point forecasting, even though there are no significant differences between them. Factor models also produce good long-horizon forecasts in a few cases. However, none of these papers have considered real-time forecasts.

This paper makes use of the most important advances in econometric modeling to estimate real-time forecasts of the Brazilian CPI inflation (IPCA). This is not only the most widely used inflation measure in Brazil, but also the index that is used to set the inflation target for central bank policy.

As far as we know, this is the first paper to use high-dimensional and machine learning models to forecast inflation in real-time for an emerging economy, using expert survey forecasts as potential candidate predictors. The models used here may be classified as either shrinkage models, such as the LASSO (Tibshirani, 1996), the adaptive LASSO (Zou, 2006), or the post-ordinary least squares (Belloni & Chernozhukov, 2013), or models that combine information, such as target factors (Bai & Ng, 2008) and complete subset regression (Elliott, Gargano, & Timmermann, 2013, 2015). We also included AR models and random walk forecasts as benchmarks and the random forest model (Breiman, 2011) as a nonlinear alternative. As a robustness check, we compare the high-dimensional models with the unobserved component stochastic volatility (UC-SV) model advocated by Stock and Watson (2007) and a Bayesian vector autoregression with priors from Bańbura, Giannone, and Reichlin (2010). Furthermore, we use the Brazilian Central Bank's (BCB) compilation of forecasts by specialists to gauge the quality of our forecasts, and also include them as potential variables in our models. The specialists forecasts are obtained from the FOCUS report, which contains expectations for several variables regarding the Brazilian economy (Marques, 2013). The FOCUS is an online environment that collects projections about key Brazilian macroeconomic variables from more than a hundred professional forecasters. The report was created to support the inflation target regime, and is published by the Brazilian Central Bank weekly on Mondays. The information is collected from several agents in the market, such as banks, fund managers, and consulting companies. We use the median, mean and standard deviation of these market expectations in our models. In addition, the FOCUS report also publishes the Top5 expectations, which includes only the five agents who were the most accurate on previous periods. The expectations are collected daily, but many forecasters only update their forecasts on Fridays, since the survey is published on Mondays. In addition to inflation, the report also publishes expectations on GDP, industrial production, exchange rates and other variables. All of this information is used by the Brazilian Central Bank to gauge its monetary policy. Finally, following Samuels and Sekkel (2017), we use a forecast combination strategy based on the model confidence sets proposed by Hansen, Lunde, and Nason (2011). The idea is to compute the average of the forecasts from the models included in a given confidence set. We show that this delivers forecasts that are superior to those of all of the individual models, as well as to the simple average of all models.

We estimated forecasts for forecast horizons of between five days before the CPI index is published to 11 months plus five days (a total of 12 forecasts). For the five-day-ahead forecast, the LASSO and FOCUS (expert) forecasts are virtually the same. For the second horizon, the adaptive LASSO is superior than any other model. For the remaining horizons, the complete subset regression dominates all other alternatives. The results are the same if we either use the root mean squared error or the mean absolute error. In terms of accumulated inflation, the complete subset regression is the model which delivers the most precise forecasts. However, most of the forecasts from different models are not statistically different according the model confidence set. In light of this finding, we construct the final forecast as the average of the models included in the confidence set. This approach delivers the best forecasts among all the competing alternatives. Finally, we also compute density forecasts for each model based on bootstrap re-sampling. According to the log-score statistic, the CSR has superior performance for most of the forecasting horizons except the first two where LASSO based methods are ranked as the best models.

Following this introduction, this paper has four sections. Section 2 describes the models and empirical procedures that were used, while Section 3 explains the dataset. The main results are presented and discussed in Section 4. Finally, the main conclusions are summarized in Section 5. A more detailed description of the dataset is included in the Appendix.

2. Empirical methods

This section describes the methods used in this paper for forecasting future inflation. We consider a direct forecast approach where the inflation h periods ahead, π_{t+h}, is modeled as a function of a set of predictors measured at time t, such as:

$$\pi_{t+h} = T(\boldsymbol{x}_t) + u_{t+h}, \tag{1}$$

where $T(\boldsymbol{x}_t)$ is a possibly nonlinear mapping of a set of q predictors, u_{t+h} is the forecasting error, and $\boldsymbol{x}_t = (x_{1t}, \ldots, x_{qt})' \in \mathbb{X} \subseteq \mathbb{R}^q$ may include weakly exogenous predictors, lagged values of inflation and a number of factors computed from a large number of potential covariates. Importantly, our focus on real-time forecasts means that $\boldsymbol{x}_t$ contains only variables that are observed and available to the econometrician at time t. Many variables are published months after their period of reference, and these variables are not included in the dataset at time t. Note further that our consideration of direct forecast models for each horizon avoids the necessity of estimating a model for the evolution of $\boldsymbol{x}_t$.

For most of the methods considered in this paper, the mapping $T(\cdot)$ is linear, such that:

$$\pi_{t+h} = \boldsymbol{\beta}'\boldsymbol{x}_t + u_{t+h}, \tag{2}$$

where $\boldsymbol{\beta} \in \mathbb{R}^q$ is a vector of unknown parameters.

2.1. Factor models with targeted predictors

Factor models using principal components are a very popular approach for avoiding the curse of dimensionality when the number of predictions may be large. The idea is to extract common components from all variables, thus reducing the model dimension.

Consider Eq. (2). When the number of candidate predictors q is large, including potentially larger than the sample size T, ordinary least squares (OLS) is infeasible or has a very large variance. One way to circumvent this drawback is to use factors as predictors instead of $\boldsymbol{x}_t$. The factors can be observed as per Fama and French (1993, 1996) or unobserved as per Bernanke, Boivin, and Eliasz (2005) and Han (2015), but our focus is on unobserved factors. Consider the following forecasting model:

$$\pi_{t+h} = \sum_{i=1}^{p} \boldsymbol{\gamma}_i'\boldsymbol{f}_{t-i} + u_{t+h}, \tag{3}$$

where $\boldsymbol{f}_t$ is a vector of k common factors extracted from $\boldsymbol{x}_t$ and k is much smaller than q. Note that $\boldsymbol{f}_t$ is not observed and must be estimated by principal components. For a discussion of the assumptions and theory behind factor models and when we can treat factors as observed variables, see Bai and Ng (2002, 2006, 2008).

Bai and Ng (2008) argued that the forecasting performance of factor models could be improved by targeting the predictors. The idea is that if many variables in $\boldsymbol{x}_t$ are irrelevant predictors of π_{t+h}, a factor analysis using all of the variables may result in noisy factors with poor forecasting abilities. The target factors are regular factor models with a pre-testing procedure that selects only relevant variables to be included in the factor analysis. We list the steps of this procedure below, and point out where our methodology differs from that proposed by Bai and Ng (2008). Let $x_{i,t}, i = 1, \ldots, q$, be the candidate variables and $\boldsymbol{w}_t$ a set of fixed regressors that will be used as controls in the pre-testing. We follow Bai and Ng (2008) and use $\boldsymbol{w}_t$ as AR terms of π_t. The procedure is as follows.

1. For $i = 1, \ldots, q$, regress π_{t+h} on $\boldsymbol{w}_t$ and $x_{i,t}$ and compute the t-statistics for the coefficient corresponding to $x_{i,t}$. We include four lags of each candidate variable in the pre-testing. Bai and Ng (2008) use only the variables in t and select the lags later.
2. Sort all t-statistics calculated in Step 1 in descending order.
3. Choose a significance level α, and select all of the variables that are significant using the computed t-statistics.
4. Let $\boldsymbol{x}_t(\alpha)$ be the variables selected in Steps 1–3. Estimate the factors $\boldsymbol{F}_t$ from $\boldsymbol{x}_t(\alpha)$ by principal components.
5. Regress π_{t+h} on $\boldsymbol{w}_t$ and $\boldsymbol{f}_t \subset \boldsymbol{F}_t$. The number of factors in $\boldsymbol{f}_t$ is selected using the BIC. Bai and Ng (2008) also selected the number of lagged factors using the BIC. However, we did not use lagged factors because we use lagged variables as regressors in the pre-testing.

The same procedure was used by Medeiros and Vasconcelos (2016), who showed that target factors reduce the forecasting errors slightly in most cases when compared to factor models without targeting.

2.2. LASSO and adaptive-LASSO

When estimating parameters in large dimensions, *shrinkage* methods form a successful alternative to factor models. The idea is to shrink the parameters that correspond to irrelevant variables to zero. Under some conditions, it is possible to handle more variables than observations. Among shrinkage methods, the least absolute shrinkage and selection operator (LASSO), introduced by Tibshirani (1996), and the adaptive LASSO (adaLASSO) of Zou (2006) have received particular attention. It has been shown that the LASSO can handle more variables than observations, and the correct subset of relevant variables can be selected (Efron, Hastie, Johnstone, & Tibshirani, 2004; Meinshausen & Yu, 2009; Zhao & Yu, 2006). As was noted by Zhao and Yu (2006) and Zou (2006), the LASSO requires a rather strong condition denoted the "irrepresentable condition" in order to attain model selection consistency, and does not have the oracle property. Zou (2006) proposes the adaLASSO in order to escape these deficiencies. The adaLASSO is a two-step methodology which uses a first-step estimator, usually the LASSO, to weight the relative importance of the regressors.

The LASSO estimator is defined as

$$\hat{\boldsymbol{\beta}} = \arg\min_{\hat{\boldsymbol{\beta}}} \left[\sum_{t=1}^{T} \left(\pi_{t+h} - \boldsymbol{\beta}'\boldsymbol{x}_t\right)^2 + \lambda \sum_{j=1}^{q} |\beta_j| \right], \tag{4}$$

where λ controls the amount of shrinkage and is determined by data-driven techniques such as cross-validation or the use of information criteria.

The adaLASSO is defined as:

$$\hat{\boldsymbol{\beta}} = \arg\min_{\hat{\boldsymbol{\beta}}} \left[\sum_{t=1}^{T} \left(\pi_{t+h} - \boldsymbol{\beta}'\boldsymbol{x}_t\right)^2 + \lambda \sum_{j=1}^{q} w_j|\beta_j| \right], \tag{5}$$

where $w_j = |\widehat{\beta}_j^*|^{-\tau}$ represents different weights on the penalization of each variable, $\widehat{\beta}_j^*$ is the parameter estimated in the first step, and $\tau > 0$ determines how much we want to emphasize the difference in the weights. Medeiros and Mendes (2016) showed that the conditions that must be satisfied on the adaLASSO are very general. The model works even when the number of variables increases faster than the number of observations and when the errors are non-Gaussian and heteroskedastic.

The most common value used for τ is one. However, Medeiros and Vasconcelos (2016) showed that selecting τ using the BIC reduces the forecasting errors. They refer to this model as Flex-adaLASSO. The value of τ is not bounded on both sides like λ. If $\tau \to 0$, we have the traditional LASSO without weights, but we do not have an upper bound. Note that if $\tau \to \infty$, then $w_i \to 0$ and we have no penalty. Thus, selecting τ using an information criterion requires one to establish an upper bound, otherwise the problem becomes computationally infeasible. If we use the LASSO as the first model, some weights will be infinite. We deal with this issue computationally by summing $T^{-\frac{1}{2}}$ to all coefficients from the first model.

Belloni and Chernozhukov (2013) showed that estimating a linear regression using the variables selected by the LASSO (post-OLS) works at least as well as just using the LASSO itself in terms of the rate of convergence to the oracle, and it also has a smaller bias. We estimated the post-OLS regression for the Flex-adaLASSO in order to check whether it reduced the forecasting error.

2.3. Random forest

The random forest (RF) methodology was initially proposed by Breiman (2011) as a way to reduce the variance of regression trees, and is based on the bootstrap aggregation (bagging) of randomly constructed regression trees.

A regression tree is a nonparametric model based on the recursive binary partitioning of the covariate space $\mathbb{X}$, where the function $T(\cdot)$ is a sum of local models (usually just a constant), each of which is determined in $K \in \mathbb{N}$ different regions (partitions) of $\mathbb{X}$. The model is usually displayed in a graph which has the format of a binary decision tree with $N \in \mathbb{N}$ parent (or split) nodes and $K \in \mathbb{N}$ terminal nodes (also called leaves), and which grows from the root node to the terminal nodes. Usually, the partitions are defined by a set of hyperplanes, each of which is orthogonal to the axis of a given predictor variable, called the *split variable*. Hence, conditional on a knowledge of the subregions, the relationship between π_{t+h} and $\mathbf{x}_t$ in Eq. (1) is approximated by a piecewise constant model, where each leaf (or terminal node) represents a distinct regime.

We represent a complex regression-tree model mathematically by introducing the following notation. The root node is at position 0 and a parent node at position j generates left- and right-child nodes at positions $2j + 1$ and $2j + 2$, respectively. Every parent node has an associated split variable $x_{s_j t} \in \mathbf{x}_t$, where $s_j \in \mathbb{S} = \{1, 2, \dots, q\}$. Furthermore, if we let $\mathbb{J}$ and $\mathbb{T}$ be the sets of indexes of the parent and terminal nodes, respectively, a tree architecture can be determined fully from $\mathbb{J}$ and $\mathbb{T}$.

The forecasting model based on regression trees can be represented mathematically as

$$\pi_{t+h} = H_{\mathbb{J}\mathbb{T}}(\mathbf{x}_t; \boldsymbol{\psi}) + u_{t+h} = \sum_{i\in\mathbb{T}} \beta_i B_{\mathbb{J}i}(\mathbf{x}_t; \boldsymbol{\theta}_i) + u_{t+h}, \tag{6}$$

where

$$B_{\mathbb{J}i}(\mathbf{x}_t; \boldsymbol{\theta}_i) = \prod_{j\in\mathbb{J}} I(x_{s_j,t}; c_j)^{\frac{n_{i,j}(1+n_{i,j})}{2}} \times \left[1 - I(x_{s_j,t}; c_j)\right]^{(1-n_{i,j})(1+n_{i,j})}, \tag{7}$$

$$I(x_{s_j,t}; c_j) = \begin{cases} 1 & \text{if } x_{s_j,t} \le c_j \\ 0 & \text{otherwise,} \end{cases} \tag{8}$$

$$n_{i,j} = \begin{cases} -1 & \text{if the path to leaf } i \text{ does not include the parent node } j; \\ 0 & \text{if the path to leaf } i \text{ includes the right-child node of the parent node } j; \\ 1 & \text{if the path to leaf } i \text{ includes the left-child node of the parent node } j. \end{cases} \tag{9}$$

Let $\mathbb{J}_i$ be the subset of $\mathbb{J}$ that contains the indexes of the parent nodes that form the path to leaf i; then, $\boldsymbol{\theta}_i$ is the vector that contains all of the parameters c_k such that $k \in \mathbb{J}_i$, $i \in \mathbb{T}$. Note that $\sum_{j\in\mathbb{J}} B_{\mathbb{J}i}\left(\mathbf{x}_t; \boldsymbol{\theta}_j\right) = 1$, $\forall\, \mathbf{x}_t \in \mathbb{R}^{q+1}$.

A random forest is a collection of regression trees, each of which is specified in a bootstrapped sub-sample of the original data. Suppose that there are B bootstrapped sub-samples, and denote the estimated regression tree for each of the sub-samples by $H_{\mathbb{J}_b\mathbb{T}_b}(\cdot; \boldsymbol{\psi}_b)$. The final prediction is defined as:

$$\widehat{\pi}_{t+h} = \frac{1}{B} \sum_{b=1}^{B} H_{\mathbb{J}_b\mathbb{T}_b}(\boldsymbol{x}_t; \boldsymbol{\psi}_b). \tag{10}$$

A regression tree is estimated for each of the bootstrapped sub-samples by repeating the following steps recursively for each terminal node of the tree until the minimum number of observations at each node is achieved.

1. Randomly select m out of q covariates as possible split variables.
2. Pick the best variable/split point among the m candidates.
3. Split the node into two child nodes.

Random forests can deal with very large numbers of explanatory variables, and the predicted model is highly nonlinear. It is important to notice that bootstrap samples are calculated using block bootstraps, since we are dealing with time series.

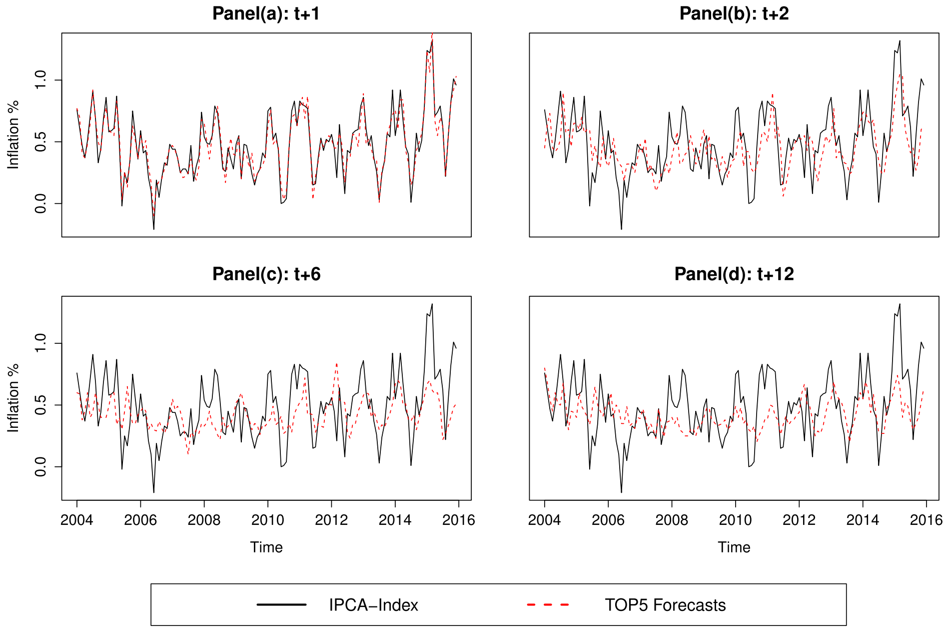

Fig. 1. Brazilian consumer prices index and focus Top5 forecasts.

2.4. Complete subset regression with targeted predictors

The complete subset regression (CSR) was developed by Elliott et al. (2013, 2015). The idea is that selecting the optimal subset of $\boldsymbol{x}_t$ for predicting π_{t+h} by testing all possible combinations of regressors is at least very demanding computationally, and often actually unfeasible. If we have q candidate variables, the CSR selects a number $n \leq q$ and computes all combinations of regressions using only n variables. The forecast of the model will be the average of all regressions in the subset.

The CSR works well for small numbers of candidate variables. However, the number of regressions to be estimated increases very quickly for large sets; for example, for $q = 25$ and $n = 4$, we need to estimate 12,650 regressions. As the number of candidate variables here is much larger, we adopt a pre-testing procedure which is similar to that used with the target factors. We start by fitting a linear regression of π_{t+h} to each of the candidate variables (including lags) and saving the t-statistic of each variable.[1] The t-statistics are ranked in absolute value and we selected the $\tilde{q}$ variables that were most relevant according to the ranking. The CSR forecast is calculated on these variables. We used $\tilde{q} = 25$ and $n = 4$.

3. The data

Inflation is measured using the Brazilian consumer price index (IPCA), which is the official inflation index in Brazil. Furthermore, a sizeable number of inflation-linked bonds use the IPCA as their reference. The dataset is obtained from Bloomberg and from the Central Bank of Brazil, and covers the period from January 2003 to December 2015, a total of 156 observations. We have 59 macroeconomic variables and 34 variables linked to specialist forecasts. The number of macroeconomic variables is smaller than that of Medeiros et al. (2016) because we are using only variables that were available in the period when the forecast was computed. The dataset also includes expert forecasts from the FOCUS survey produced by the Central Bank of Brazil. Our expectation variables include the median of the h-period-ahead specialist forecasts; the median of the top five (Top5) experts, i.e., the five experts who produced the best forecasts in the previous period; and, finally, the mean and the standard deviation of the Top5. The macroeconomic variables cover several inflation and industry indexes, unemployment and other variables related to labour, energy consumption, exchange rates, stock markets, government accounts, expenditure and debt, taxes, monetary variables and exchange of goods and services. Both the inflation series and the Top5 median are presented in Fig. 1, which shows that the Top5 delivers the smallest RMSE for $h = 1$ (five days ahead), but rapidly loses performance as h grows.

4. Main results

4.1. Forecasting errors

We estimate all models described in Section 2 for $h = 1, \ldots, 12$. Recall that $h = 1$ is five days before the IPCA inflation is published, $h = 2$ is one month and five days, and $h = 12$ is 11 months and five days. This section discusses the results and compares the forecasting errors of all models. We also include in the comparison forecasts estimated using autoregressive models with lags selected by the BIC

[1] We did not use a fixed set of controls, $\boldsymbol{w}_t$, in the pre-testing as we did on the target factors.

Table 1
Forecast mean absolute errors and root mean squared errors.

RMSE × 1000 (MAE × 1000)	Brazilian consumer price index Forecast horizon												Acc.
	$t+1$	$t+2$	$t+3$	$t+4$	$t+5$	$t+6$	$t+7$	$t+8$	$t+9$	$t+10$	$t+11$	$t+12$	
RW	2.41	3.23	3.68	4.10	4.40	4.62	4.76	4.33	3.76	3.40	3.03	2.75	33.94
	(1.99)	(2.63)	(3.01)	(3.38)	(3.44)	(3.64)	(3.71)	(3.41)	(3.04)	(2.73)	(2.58)	(2.13)	(26.11)
AR	2.30	2.89	3.26	3.31	3.23	3.18	3.04	2.82	2.72	2.70	2.67	2.64	20.75
	(1.93)	(2.21)	(2.47)	(2.60)	(2.54)	(2.49)	(2.37)	(2.16)	(2.13)	(2.07)	(2.06)	(2.01)	(16.14)
Factors	1.33	2.19	2.42	2.48	2.44	2.49	2.48	2.37	2.29	2.50	2.38	2.44	14.31
	(0.98)	(1.75)	(1.88)	(1.93)	(1.83)	(1.91)	(1.96)	(1.89)	(1.79)	(2.01)	(1.86)	(1.93)	(9.63)
LASSO	**0.95**	1.85	2.85	3.21	2.75	2.83	2.79	3.33	2.80	3.33	3.51	3.33	17.09
	(**0.74**)	(1.46)	(2.28)	(2.44)	(2.11)	(2.24)	(2.12)	(2.65)	(2.15)	(2.69)	(2.89)	(2.71)	(12.42)
F. aL	0.98	**1.58**	2.20	2.43	2.39	2.42	2.53	2.86	2.48	2.56	2.54	2.46	13.50
	(0.75)	(**1.30**)	(1.75)	(1.94)	(1.82)	(1.89)	(2.04)	(2.33)	(1.94)	(2.06)	(2.01)	(1.88)	(9.39)
P. OLS	0.98	1.62	2.23	2.23	2.49	2.52	2.53	3.08	2.52	2.66	2.61	2.46	14.02
	(0.75)	(1.34)	(1.80)	(1.80)	(1.89)	(1.97)	(2.02)	(2.48)	(1.94)	(2.11)	(2.06)	(1.89)	(9.58)
RF	1.43	1.95	2.56	2.54	2.66	2.88	2.82	2.85	2.71	2.65	2.64	2.46	15.67
	(0.97)	(1.45)	(1.93)	(1.93)	(2.06)	(2.30)	(2.21)	(2.25)	(2.09)	(1.96)	(1.99)	(1.82)	(12.36)
CSR	1.05	1.64	**2.04**	**2.23**	**2.25**	**2.29**	**2.29**	**2.26**	**2.26**	**2.27**	**2.25**	**2.26**	**11.93**
	(0.88)	(1.33)	(**1.69**)	(**1.75**)	(**1.79**)	(**1.80**)	(**1.80**)	(**1.80**)	(**1.81**)	(**1.79**)	(**1.77**)	(**1.78**)	(**8.41**)
FOCUS	**0.95**	1.83	2.39	2.48	2.53	2.57	2.56	2.53	2.55	2.57	2.58	2.60	16.82
	(0.76)	(1.50)	(1.87)	(1.91)	(1.93)	(1.97)	(1.94)	(1.91)	(1.93)	(1.93)	(1.94)	(1.96)	(12.51)
Top5	0.96	1.69	2.32	2.48	2.62	2.70	2.77	2.67	2.51	2.65	2.56	2.55	16.69
	(0.74)	(1.39)	(1.83)	(1.90)	(1.99)	(2.07)	(2.06)	(2.03)	(1.99)	(1.97)	(1.91)	(1.89)	(12.12)

The table shows the root mean squared errors and mean absolute deviations (in parentheses) of the forecasts. The values in bold represent the best model according to each measure of error and for each forecast horizon. All values are multiplied by 1000. The column Acc. shows the forecast errors accumulated over 11 months.

and random walk forecasts. All models are estimated in a nine-year rolling-window scheme, with the first forecast being for January 1, 2012. Thus, the models are evaluated based on 48 point forecasts, with the last forecast being for December 2015.[2] This period covers various different situations within the Brazilian economy. The Brazilian GDP increased by 1.9% and 3% in 2012 and 2013 respectively, 2014 had an increase of 0.1%, and 2015 had a decrease of 3.7%. Fig. 1 shows that the state of the economy does not affect the precision of the short-term forecasts. However, the errors for longer forecasting horizons were bigger in 2015, which was an year of 10.67% inflation. Note that the inflation target is 4.5% and its ceiling is 6.5%.

Table 1 shows the root mean squared error (RMSE) and the mean absolute error (MAE) for all forecasting models. The model with the smallest forecasting error for each horizon is displayed in bold. The last column of Table 1 shows the cumulative error for the 11-month inflation. The LASSO and the Flex-adaLASSO have the smallest errors for $h = 1$ and $h = 2$, while the CSR has the smallest errors for all other horizons. However, for $h = 1$, the LASSO forecasts are not statistically different from the expert forecasts. On the other hand, there is a substantial gain from using the CSR models for the longer horizons. The target factor models become more competitive as h increases. The random walk and the autoregressive models both perform poorly. The random forest was not the best model at any horizon, but its performance was not bad overall. In fact, its cumulative forecasting error was smaller than that of the FOCUS, the Top5 or the LASSO. In addition, the Post-OLS estimation with the variables selected using the Flex-adaLASSO delivers larger errors than the Flex-adaLASSO itself.

The reason why the LASSO and the Flex-adaLASSO are the best models for short horizons is that the expert forecasts are very precise for $h = 1$ and $h = 2$. As has been mentioned, market players devote considerable resources to inflation forecasting. Therefore, variable selection models such as the LASSO perform better than methods that combine information from many variables, such as target factors and CSR. However, the expert forecasts lose their predictive power as the forecast horizon increases and many variables become more relevant. Models that combine information can extract common information on all variables that are useful for forecasting inflation. Fig. 2 shows the average numbers of variables selected by the LASSO and the Flex-adaLASSO at all horizons. The number of variables selected is very small for both models for $h = 1$ and 2, but grows for longer horizons, especially in the case of the LASSO. For shorter horizons, the Flex-adaLASSO is mostly a combination of specialist forecasts.

Frequently, the model with the smallest average squared error is not the model with the smallest errors in most of the 48 rolling windows. Table 2 shows the ranking of models for each forecasting window. The table reports the proportion of cases where each model is in each position of the ranking. The results are aggregated for all horizons. Surprisingly, the random walk, which performed badly in terms of average errors, was the best model in 24%

[2] We start producing forecasts in 2012 in order to have a reasonable number of point forecasts for each forecasting horizon while still having enough observations for the in-sample estimation of the models. As the models are for direct forecasts, we have 108 observations from which to estimate models for $h = 1$, 107 for $h = 2$, and so on. We also show results for 24 rolling windows.

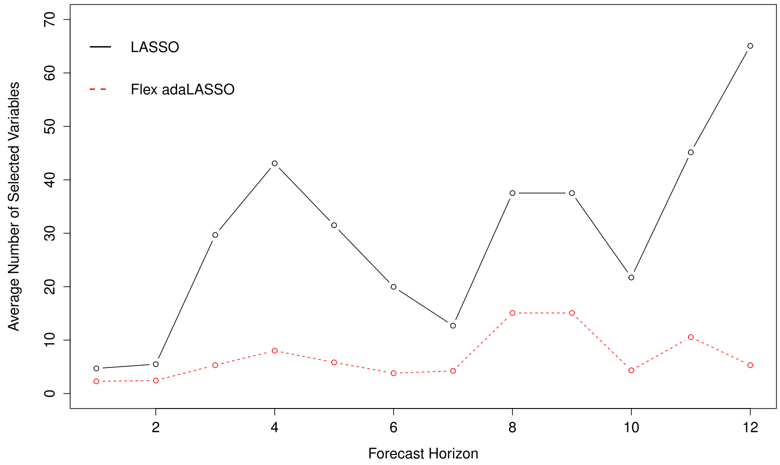

Fig. 2. Average number of variables selected by the shrinkage methods.

Table 2
Proportion of the time each model was in each position of the error ranking.

	Brazilian consumer price index Model position									
	1	2	3	4	5	6	7	8	9	10
RW	0.24	0.01	0.08	0.07	0.02	0.05	0.03	0.02	0.28	0.19
AR	0.04	0.06	0.06	0.06	0.03	0.07	0.06	0.06	0.24	0.31
Factors	0.14	0.05	0.18	0.07	0.09	0.15	0.09	0.10	0.06	0.08
LASSO	0.06	0.08	0.12	0.08	0.15	0.15	0.14	0.18	0.02	0.02
F. aL	0.03	0.07	0.10	0.14	0.18	0.15	0.09	0.20	0.02	0.02
P. OLS	0.05	0.10	0.11	0.12	0.18	0.12	0.11	0.17	0.03	0.02
RF	0.05	0.14	0.11	0.16	0.17	0.09	0.11	0.13	0.03	0.03
CSR	0.07	0.14	0.10	0.15	0.10	0.10	0.15	0.09	0.06	0.04
FOCUS	0.08	0.16	0.08	0.07	0.05	0.06	0.16	0.04	0.19	0.12
Top5	0.24	0.20	0.06	0.08	0.02	0.05	0.09	0.02	0.07	0.17

The table shows the proportion of the time that each model is in each ranking position, aggregated over all forecast horizons.

of cases, the same proportion as the Top5. However, the same two models delivered the worst forecasts in 19% and 17% of cases, respectively. The CSR model, which is the best model on average at most horizons, had the smallest errors for only 7% of forecasts, while the Flex-adaLASSO model, which is the second-best model when considering the cumulative inflation, is the best model in only 3% of cases. The models with the smallest average errors are those that perform well when most models are performing poorly. However, they are no longer the best models when all models are doing well.

We show the forecast error correlations in Fig. 3. The figure displays heat-maps for horizons 1, 2, 6, and 12. The pattern is very similar for all horizons. The FOCUS and the Top5 are positively correlated with each other, but their correlations with all other models are negative. The remaining forecasts are all positively correlated. The two best models, namely the Flex-adaLASSO and CSR models, have strong negative correlations with both of the expert forecasts considered in this paper. This shows that even though some models and the expert forecasts all have small forecast errors, their forecasts differ considerably. This in turn opens the possibility of improving the results using combinations of these forecasts, as will be discussed on the next section.

4.2. Model confidence sets and model combination

This section reports on the model confidence set (MCS) approach developed by Hansen et al. (2011). The MCS allows us to compare large numbers of models simultaneously. The test returns a confidence set that includes the best model with probability $(1 - \alpha)$. The set becomes wider (with more models) as we decrease α, while large values of α may result in a set with only a single model.

The MCS uses bootstrapped samples of a given loss function, in our case squared errors, to create the test statistics. The confidence set estimates p-values for all models using the bootstrapped samples, and uses α to select which models should be inside the set. Since models are removed from the set interactively, the MCS also generates a ranking. The best model has a p-value of 1 by definition, since it can only be as good as itself and there is no other model to compare. If model 1 is removed from the set with a p-value of k_1 and model 2 is later removed with a p-value of k_2, the test p-value if only models 1 and 2 are excluded will be $\max\{k_1, k_2\}$. Therefore, the p-value cannot decrease when a new model is excluded from the confidence set. We exclude models until the null hypotheses is no longer rejected. Hansen et al. (2011) propose two different statistics to be used as decision rules,

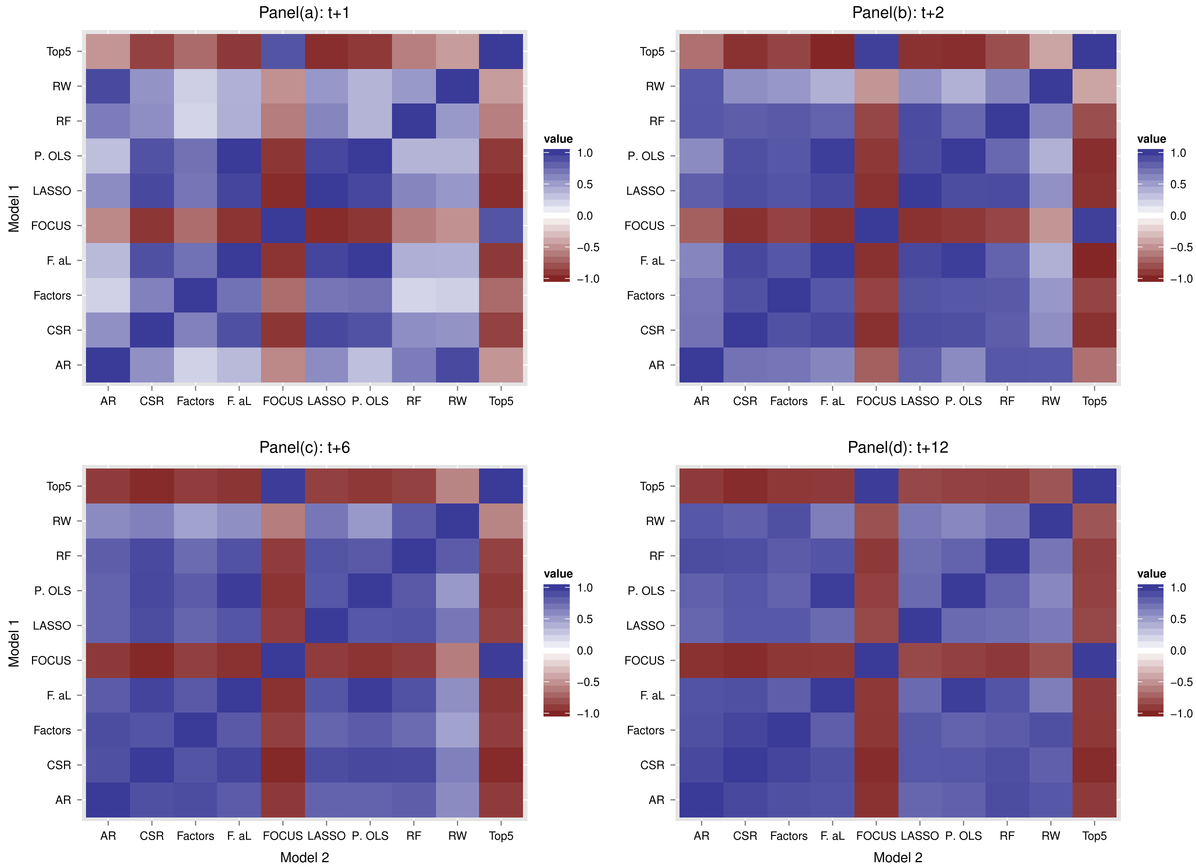

Fig. 3. Forecast error correlations.

the $T_{\max,M}$ and the $T_{R,M}$. We adopt the first, since it is simple and easy to compute. The second statistic creates the set by comparing all of the models two by two, which makes the procedure more intensive.

The MCS p-values are presented in Table 3. Values in bold represent models that remained in the confidence set with $\alpha = 20\%$. The autoregressive models and the random walk were removed from the set at most forecasting horizons. The only models which were in the confidence set for all horizons were the Flex-adaLASSO, the random forest, the complete subset regression and the FOCUS forecast. If we include the cumulative forecasts, we retain only the Flex-adaLASSO and the CSR as the models that are always in the set. If we look at the ranking, the CSR is the model with the most p-values of 1.

We use the results from Table 3 to generate combined forecasts from the models in the confidence set. These results are displayed in Table 4. The first row of the table shows the forecasting errors from averaging the forecasts from all models. The second row shows the forecasting errors from averaging the forecasts of the models in the confidence set, and the last row shows the forecasting error of the best model from Table 1 at each forecasting horizon.[3]

[3] The cumulative errors are calculated based on the 95% confidence set in order to include the specialist forecasts. This was done because of the results of Fig. 3, which show that the specialist forecasts are correlated negatively with the other forecasts.

The results in Table 4 show that the simple average of all models beats the results from the best individual model. The combined forecast from the MCS improves the results even more, especially when considering the shortest forecasting horizons. Even at horizon $h = 1$, which is only five days before the IPCA is published, the forecasting errors are considerably smaller when we combine forecasts. In many cases, the forecasting error is less than half that from the best individual model.

4.3. Look-ahead bias in the MCS combined forecasts

Our combined forecasts based on the MCS are contaminated with look-ahead bias, as we need to know the forecasting errors in order to estimate the confidence set. However, the selected models in the confidence set tend to be stable over the time period considered here. We tested the stability of the results and ensured that they were free of look-ahead bias by splitting the sample of 48 observations into two sub-samples: one with 36 observations, for estimating the confidence set, and the other with 12 observations, for estimating the combined forecasts. We also estimated the simple average forecast for this 12-month period.

The results are displayed in Table 5, and show that the combined MCS forecasting errors calculated without a look-ahead bias are still smaller than those calculated

Table 3
Model confidence set.

	Brazilian consumer price index Forecast horizon												Acc.
	$t+1$	$t+2$	$t+3$	$t+4$	$t+5$	$t+6$	$t+7$	$t+8$	$t+9$	$t+10$	$t+11$	$t+12$	
RW	0.00	0.00	0.00	0.00	0.01	0.01	0.00	0.03	0.01	0.05	0.15	**0.68**	0.01
AR	0.00	0.03	0.03	0.03	0.02	0.11	**0.42**	**0.56**	**0.43**	**0.75**	**0.74**	**0.70**	0.03
Factors	0.16	0.03	**0.48**	**0.80**	**0.68**	**0.71**	**0.62**	**0.52**	**0.79**	**0.54**	**0.35**	**0.75**	**0.35**
LASSO	**0.94**	**0.48**	0.07	**0.31**	**0.28**	**0.22**	**0.42**	0.13	**0.43**	0.05	0.02	0.01	0.06
F. aL.	**0.79**	**1.00**	**0.24**	**0.65**	**0.68**	**0.71**	**0.90**	**0.46**	**0.66**	**0.67**	**0.90**	**0.95**	**0.34**
P. OLS	**0.76**	**0.74**	0.19	**0.91**	**0.64**	**0.59**	**0.76**	**0.42**	**0.41**	**0.85**	**0.79**	**0.93**	**0.35**
RF	**0.27**	**0.48**	**0.28**	**0.91**	**0.61**	**0.22**	**0.27**	**0.56**	**0.39**	**0.76**	**0.81**	**0.93**	0.06
CSR	**0.31**	**0.74**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
FOCUS	**1.00**	**0.48**	**0.48**	**0.80**	**0.64**	**0.59**	**0.90**	**0.34**	**0.72**	**0.58**	**0.90**	**0.70**	0.05
Top5	**0.94**	**0.49**	**0.29**	**0.88**	**0.42**	**0.20**	**0.45**	0.19	**0.72**	**0.85**	**0.79**	**0.95**	0.05

The table shows the model confidence set p-values for all forecasting horizons, along with the 12-month accumulated inflation. Values in bold are included in the $\alpha = 20\%$ or 80% confidence set. The p-values can be used to rank the models. Models with p-values of 1 are the best models, or those that remain in all confidence sets.

Table 4
Combined forecast mean absolute errors and mean squared errors.

RMSE $\times$ 1000 MAE $\times$ 1000	Brazilian consumer price index Forecast horizon												Acc.
	$t+1$	$t+2$	$t+3$	$t+4$	$t+5$	$t+6$	$t+7$	$t+8$	$t+9$	$t+10$	$t+11$	$t+12$	
All models	0.74 (0.62)	1.18 (0.89)	1.48 (1.15)	1.53 (1.17)	1.52 (1.15)	1.58 (1.18)	1.58 (1.18)	1.63 (1.32)	1.49 (1.15)	1.55 (1.24)	1.52 (1.25)	1.42 (1.12)	9.72 (7.13)
MCS models	0.42 (0.33)	0.71 (0.58)	0.80 (0.63)	1.22 (0.97)	1.15 (0.85)	1.73 (1.35)	1.38 (1.08)	1.81 (1.47)	1.33 (1.05)	1.22 (0.97)	1.19 (0.95)	1.27 (0.99)	9.69 (6.72)
Best ind. model	0.96 (0.74)	1.58 (1.30)	2.04 (1.69)	2.23 (1.75)	2.25 (1.79)	2.29 (1.80)	2.29 (1.80)	2.26 (1.80)	2.26 (1.79)	2.27 (1.79)	2.25 (1.77)	2.26 (1.78)	11.93 (8.41)

The table shows the forecast errors of the average forecasts of all models and of those in the confidence set. The last row shows the best individual model, as a comparison with the combined forecasts. All values are multiplied by 1000.

Table 5
Combined forecast mean absolute errors and mean squared errors without look-ahead bias.

RMSE $\times$ 1000 (MAE $\times$ 1000)	Brazilian consumer price index Forecast horizon												Acc.
	$t+1$	$t+2$	$t+3$	$t+4$	$t+5$	$t+6$	$t+7$	$t+8$	$t+9$	$t+10$	$t+11$	$t+12$	
All models	0.75 (0.67)	1.62 (1.33)	2.23 (1.93)	2.24 (1.84)	2.32 (1.95)	2.39 (2.01)	2.29 (1.89)	2.27 (1.89)	2.20 (1.78)	2.10 (1.73)	2.03 (1.73)	2.09 (1.81)	15.67 (13.68)
MCS models	0.43 (0.35)	0.93 (0.83)	1.02 (0.86)	1.71 (1.43)	1.71 (1.43)	1.35 (1.10)	1.80 (1.51)	1.61 (1.39)	1.85 (1.48)	1.64 (1.32)	1.71 (1.40)	1.86 (1.61)	12.08 (9.97)

The table shows the forecast errors of the average forecasts of all models and of those in the confidence set. The last row shows the best individual model, as a comparison with the combined forecasts. All values are multiplies by 1000.

with a simple average across all models. Note that these results are only for the last 12 months in the sample (January–December, 2015), which was the worst year in our sample for the Brazilian economy, in terms of GDP growth.

4.4. Different window sizes

Given the length of the dataset, it is not viable to test the models on a completely different sample. However, we can check whether changing the size of the rolling window, and consequently the number of forecasts, has any significant impact on our results.

Increasing the window size from 9 to 10 years reduces the number of forecasts (windows) from 48 to 24. Table 6 shows the forecasting RMSEs and MAEs when the models are estimated on a larger window of observations. The results are similar to the case of 48 windows. However, the errors in Table 6 are generally larger because the forecasts are just for 2014 and 2015, years in which the Brazilian economy was more unstable (especially 2015). As was mentioned earlier, the forecasting errors for longer horizons are larger in 2015, and that shifted the errors up. The target factor model has the smallest errors for several forecasting horizons. The other models that deserve a mention are the LASSO and Flex-adaLASSO, which performed well on shorter horizons, and the complete subset regression, which has good results for longer horizons. We have already detected an improvement in the target factor model for longer horizons in the results for 48 rolling windows; however, the difference here is that factor models are able to beat the CSR in some cases for 24 rolling windows.

We show the model confidence set results for the 24-rolling-window analysis in Table 7. The results are similar to those for the 48 windows. However, the only

Table 6
Forecast mean absolute errors and root mean squared errors for 24 rolling windows.

RMSE ∗ 1000 *MAE* ∗ 1000	Brazilian consumer price index: 24 rolling windows Forecast horizon												Acc.
	$t+1$	$t+2$	$t+3$	$t+4$	$t+5$	$t+6$	$t+7$	$t+8$	$t+9$	$t+10$	$t+11$	$t+12$	
RW	2.69 (2.25)	3.80 (3.13)	4.37 (3.68)	4.81 (4.02)	5.14 (4.10)	5.46 (4.45)	5.72 (4.69)	5.29 (4.36)	4.40 (3.72)	3.77 (3.00)	3.45 (2.97)	3.17 (2.44)	43.16 (34.70)
AR	2.61 (2.24)	3.41 (2.66)	3.29 (2.68)	3.55 (2.81)	3.82 (3.12)	3.77 (3.06)	3.60 (2.90)	3.38 (2.68)	3.33 (2.71)	3.26 (2.62)	3.22 (2.61)	3.26 (2.54)	32.59 (29.57)
Factors	1.50 (1.15)	2.45 (1.98)	2.72 (2.32)	2.96 (2.40)	3.15 (2.49)	3.11 (2.41)	2.95 (2.20)	2.85 (2.33)	**2.56** (2.09)	**2.45** (2.01)	**2.45** (**1.97**)	**2.55** (2.07)	21.83 (18.04)
LASSO	**0.95** (**0.76**)	2.15 (1.75)	2.87 (2.35)	3.17 (2.48)	3.21 (2.45)	3.24 (2.56)	3.23 (2.54)	3.21 (2.51)	3.20 (2.48)	3.32 (2.64)	3.79 (2.91)	3.55 (2.97)	25.27 (23.04)
F. aL	1.03 (0.83)	**1.76** (**1.46**)	2.50 (**2.08**)	2.84 (2.32)	2.84 (2.24)	2.97 (2.32)	3.32 (2.64)	3.30 (2.60)	3.32 (2.59)	3.28 (2.56)	3.31 (2.60)	3.35 (2.70)	24.84 (22.83)
P. OLS	1.04 (0.83)	1.77 (1.48)	2.58 (2.08)	**2.58** (**2.08**)	2.95 (2.28)	3.06 (2.38)	3.36 (2.72)	3.27 (2.58)	2.96 (2.33)	2.93 (2.31)	2.87 (2.17)	2.81 (2.15)	22.74 (20.02)
RF	1.65 (1.08)	2.30 (1.74)	3.03 (2.36)	3.11 (2.46)	3.36 (2.60)	3.69 (2.94)	3.57 (2.80)	3.64 (2.79)	3.39 (2.52)	3.10 (2.27)	3.08 (2.31)	2.99 (2.22)	25.17 (23.00)
CSR	1.05 (0.91)	1.87 (1.55)	**2.44** (2.09)	2.71 (2.18)	**2.77** (**2.23**)	**2.75** (**2.20**)	**2.68** (**2.11**)	**2.71** (**2.11**)	2.69 (**2.05**)	2.59 (**1.93**)	2.62 (1.99)	2.63 (**2.02**)	**18.04** (**16.40**)
FOCUS	0.97 (0.83)	2.14 (1.74)	2.99 (2.38)	3.13 (2.48)	3.20 (2.51)	3.25 (2.60)	3.22 (2.56)	3.20 (2.52)	3.22 (2.55)	3.25 (2.58)	3.27 (2.58)	3.28 (2.60)	26.95 (24.86)
Top5	0.99 (0.78)	1.92 (1.55)	2.80 (2.21)	3.09 (2.42)	3.29 (2.57)	3.34 (2.64)	3.46 (2.64)	3.35 (2.63)	3.05 (2.48)	3.35 (2.58)	3.27 (2.54)	3.24 (2.48)	26.94 (24.72)

The table shows the root mean squared errors and mean absolute deviations (in parentheses) of the forecasts based on 24 rolling windows. The values in bold represent the best model according to each measure of error and for each forecasting horizon. All values are multiplied by 1000. The column Acc. shows the errors of the 12-month cumulative forecast, built using the monthly forecasts.

Table 7
Model confidence set: 24 rolling windows.

	Brazilian consumer price index: 24 rolling windows Forecast horizons												Acc.
	$t+1$	$t+2$	$t+3$	$t+4$	$t+5$	$t+6$	$t+7$	$t+8$	$t+9$	$t+10$	$t+11$	$t+12$	
RW	0.00	0.00	0.00	0.00	0.02	0.03	0.01	0.01	0.06	**0.44**	**0.50**	**0.64**	0.12
AR	0.00	0.08	**0.23**	**0.39**	0.04	0.03	0.14	**0.72**	**0.26**	**0.27**	**0.43**	**0.43**	0.03
Factors	0.09	0.10	**0.56**	**0.75**	**0.35**	**0.67**	**0.33**	**0.60**	**1.00**	**1.00**	**1.00**	**1.00**	0.06
LASSO	**1.00**	**0.40**	**0.43**	**0.63**	**0.60**	**0.38**	**0.44**	**0.59**	**0.54**	**0.27**	**0.63**	0.17	0.03
F. aL	**0.86**	**1.00**	**0.50**	**0.34**	**0.55**	**0.57**	**0.58**	**0.55**	**0.26**	**0.44**	**0.63**	**0.20**	0.05
P. OLS	**0.91**	**0.60**	**0.50**	**0.75**	**0.45**	**0.67**	**0.44**	**0.59**	**0.28**	**0.35**	**0.33**	**0.45**	0.06
RF	**0.34**	**0.53**	**0.48**	**0.69**	**0.67**	**0.38**	**0.58**	**0.72**	**0.54**	**0.35**	**0.33**	**0.45**	0.05
CSR	**0.91**	**0.29**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**0.71**	**0.66**	**0.59**	**0.78**	**1.00**
FOCUS	**0.79**	**0.40**	**0.41**	**0.63**	**0.67**	**0.23**	0.19	**0.23**	**0.49**	0.19	**0.50**	**0.29**	0.02
Top5	**0.79**	**0.53**	**0.56**	**0.59**	**0.39**	**0.38**	**0.37**	**0.55**	**0.51**	0.19	0.19	**0.64**	0.12

The table shows the model confidence set p-values for all forecast horizons, along with the 12-month cumulative inflation using 24 rolling windows. The values in bold are those included in the $\alpha = 20\%$ or 80% confidence set. The sizes of the p-values can be used to rank the models. Models with p-values of 1 are the best models, or those that remain in all confidence sets.

model in the confidence set on the accumulated inflation is the CSR. If we look at the monthly horizons individually, the models that are included in the 80% confidence set at all horizons are the Flex-adaLASSO, the Post-OLS estimated with the variables selected using the Flex-adaLASSO, the random forest and the CSR. The CSR was the last remaining model in six cases, against four of the target factors. The LASSO and the Flex-adaLASSO are the last remaining model in one case each.

4.5. Bayesian alternatives

This section shows the results using two alternative Bayesian models. First, the unobserved component stochastic volatility (UC-SV) model proposed by Stock and Watson (2007), which is a very popular model for the US inflation; and second, a large Bayesian vector autoregressive (BVAR) using all variables in the dataset and with priors defined as per Bańbura et al. (2010).

4.5.1. Unobserved component stochastic volatility model

The UC-SV model is described by the following equations:

$$\begin{aligned} \pi_t &= \tau_t + e^{h_t/2}\varepsilon_t, \\ \tau_t &= \tau_{t-1} + u_t, \\ h_t &= h_{t-1} + \upsilon_t, \end{aligned} \tag{11}$$

where $\{\varepsilon_t\}$ is a sequence of independent and normally distributed random variables with zero mean and unit variance, $\varepsilon_t \sim N(0, 1)$; u_t and υ_t are both normal with zero mean and variance given by inverse-gamma priors; $\tau_1 \sim$

Table 8
Forecast mean absolute errors and root mean squared errors for BVAR and UCSV.

RMSE × 1000 (MAE × 1000)	Brazilian consumer price index Forecast horizon												Acc.
	$t+1$	$t+2$	$t+3$	$t+4$	$t+5$	$t+6$	$t+7$	$t+8$	$t+9$	$t+10$	$t+11$	$t+12$	
UCSV	2.56 (2.14)	3.17 (2.66)	3.54 (2.91)	3.86 (3.11)	4.06 (3.19)	4.09 (3.24)	3.95 (3.06)	3.51 (2.75)	3.08 (2.47)	2.79 (2.28)	2.59 (2.10)	2.61 (2.00)	27.74 (21.79)
BVAR	1.19 (0.92)	2.14 (1.66)	2.53 (1.91)	2.73 (2.04)	2.88 (2.20)	2.98 (2.28)	3.05 (2.33)	3.09 (2.37)	3.11 (2.39)	3.12 (2.39)	3.13 (2.41)	3.14 (2.41)	19.27 (14.60)

The table shows the root mean squared errors and mean absolute deviations (in parentheses) of the forecasts. The values in bold represent the best model according to each measure of error and for each forecasting horizon. All values are multiplied by 1000. The column Acc. shows the errors of the 12-month accumulated forecasts. The order of the BVAR is 4 and the UCSV is estimated using MCMC.

$N(0, V_\tau)$; and $h_1 \sim N(0, V_h)$, where $V_\tau = V_h = 0.12$. The model is estimated by Markov chain Monte Carlo (MCMC) methods. The h-step-ahead forecast is computed as $\widehat{\pi}_{t+h} = \widehat{\tau}_{t|t}$. We computed forecasts for the same forecast horizons as in the previous sections, and the forecasting errors are calculated for 48 months out-of-sample as before.

4.5.2. Bayesian vector autoregressive model

Let $\boldsymbol{Y}_t = (y_{1,t}, y_{2,t}, \dots y_{n,t})'$ be described as the following VAR model:

$$\boldsymbol{Y}_t = \boldsymbol{c} + \boldsymbol{A}_1\boldsymbol{Y}_{t-1} + \cdots + \boldsymbol{A}_p\boldsymbol{Y}_{t-p} + \boldsymbol{u}_t, \tag{12}$$

where $\boldsymbol{c}$ is an n-dimensional vector of constants; $\boldsymbol{A}_i$, $i = 1 \dots p$, are $(n \times n)$ matrices of coefficients; and $\boldsymbol{u}_t$ is an n-dimensional error vector. The same model may be written as a system of equations:

$$\boldsymbol{Y} = \boldsymbol{XB} + \boldsymbol{U}, \tag{13}$$

where $\boldsymbol{Y} = (\boldsymbol{Y}_1, \dots, \boldsymbol{Y}_T)'$ is a $(T \times n)$ matrix, $\boldsymbol{X} = (\boldsymbol{X}_1, \dots, \boldsymbol{X}_T)'$ is a $(T \times k)$ matrix with $k = np + 1$, and $\boldsymbol{X}_t = (1, \boldsymbol{Y}_{t-1}, \dots, \boldsymbol{Y}_{t-p})'$, $\boldsymbol{U} = (\boldsymbol{u}_1, \dots, \boldsymbol{u}_T)'$, and $\boldsymbol{B} = (\boldsymbol{c}, \boldsymbol{A}_1, \dots, \boldsymbol{A}_p)'$.

The model is estimated using dummy observations $\boldsymbol{Y}_d$ and $\boldsymbol{X}_d$ of dimensions $T_d \times n$ and $T_d \times k$ respectively (for details on creating the dummy observations, see Bańbura et al., 2010). Using these dummies is equivalent to imposing the normal inverted Wishart prior on the covariance matrix of $\boldsymbol{B}$. The dummy observations are used to create $\boldsymbol{Y}^* = (\boldsymbol{Y}', \boldsymbol{Y}_d')'$ and $\boldsymbol{X}^* = (\boldsymbol{X}', \boldsymbol{X}_d')'$. The posterior mean of $\boldsymbol{B}$ is the same as the ordinary least-squares (OLS) estimates of the regression of $\boldsymbol{Y}^*$ on $\boldsymbol{X}^*$, and also the same as the Minnesota prior. In addition, the use of dummies ensures that the number of observations is larger than the number of variables for each regression of the VAR, which makes the OLS estimation of $\boldsymbol{B}$ feasible. Another important issue is the choice of the expected value of the priors for the diagonal of the $\boldsymbol{A}_1$ matrix. We choose a value of 0.5 for each element.

4.5.3. Results

The results for the UC-SV and the large BVAR are given in Table 8. The forecasting errors are calculated for the last 48 observations of the dataset using a rolling window scheme. The UC-SV does not use any information other than the past inflation, and thus, its forecasting errors are larger than those of the multivariate models that use the FOCUS and other macroeconomic variables as regressors. Nevertheless, the UC-SV is comparable to the other univariate models even though it cannot beat the AR model. The BVAR is much more accurate than the UC-SV and competes with the high-dimensional models directly, especially for short forecasting horizons. However, it is not the best model at any horizon. Note that an h-step-ahead forecast for the BVAR is iterated, while the models in Table 1 are estimated for the horizon of interest directly. Table 9 shows the (Giacomini & White, 2006) (GW) test p-values, where the null hypothesis is that the two models have the same forecasting accuracies. The null is rejected in most cases for the UC-SV, except when it is compared to the random walk and the AR models. It has the same forecasting ability as the random walk for short horizons and is similar to the AR model for long horizons. Note that these models are all univariate and very simple. Therefore, it is natural that the test fails to detect significant differences between them when considering only 48 rolling widows. The Bayesian VAR (second panel of Table 9) is just as accurate as any of the other multivariate models five days before the CPI is published ($h = 1$). This is because all of the models include forecasts by specialists, which are very accurate right before the CPI is published, and therefore, the forecasting errors are small for all of the multivariate models, and differences between them are not detected by the GW test. The performance of the BVAR is clearly inferior for longer forecasting horizons. The CSR, which is the best models in most cases, is statistically different from the BVAR for all horizons longer than one.

The main conclusion from the above results is that the performances of the two Bayesian alternatives considered in the paper are inferior to those of the machine learning methods.

4.6. Density forecasts

So far we have analyzed only point forecasts for several models. This section goes on to focus on density forecasts. The point forecasts in a rolling window scheme provide good information as to which model is most accurate on average, but do not tell us anything about the forecast uncertainty. We obtain the predictive densities by bootstrapping the in-sample residuals. For each model, in each rolling window, we randomly selected 100,000 observations of the in-sample residuals and added them to the point forecast, resulting on an empirical predictive density.[4]

[4] The procedures for the BVAR and the UCSV are slightly different, because we did not estimate direct forecasts with these models. The $t+1$

Table 9
Giacomini and White test *p*-values comparing the UCSV and the Bayesian VAR to all other models.

	Brazilian consumer price index: GW *p*-values												Acc.
	$t+1$	$t+2$	$t+3$	$t+4$	$t+5$	$t+6$	$t+7$	$t+8$	$t+9$	$t+10$	$t+11$	$t+12$	
	Forecast horizon: all models against UCSV												
RW	0.423	0.763	0.515	0.255	0.075	0.020	0.001	0.001	0.000	0.001	0.010	0.358	0.183
AR	0.020	0.107	0.275	0.079	0.037	0.017	0.024	0.053	0.124	0.683	0.566	0.854	0.178
Factors	0.000	0.003	0.000	0.001	0.001	0.000	0.001	0.005	0.004	0.183	0.082	0.282	0.073
LASSO	0.000	0.000	0.029	0.203	0.001	0.002	0.001	0.668	0.250	0.061	0.004	0.004	0.043
F. al.	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.106	0.006	0.185	0.841	0.532	0.079
P. OLS	0.000	0.000	0.000	0.001	0.001	0.001	0.001	0.342	0.043	0.553	0.940	0.577	0.083
RF	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.004	0.023	0.447	0.794	0.503	0.059
CSR	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.004	0.042	0.064	0.025
BVAR	0.000	0.000	0.000	0.000	0.001	0.000	0.002	0.085	0.912	0.223	0.059	0.042	0.002
	Forecast horizon: all models against BVAR												
RF	0.000	0.001	0.000	0.000	0.001	0.001	0.000	0.003	0.013	0.378	0.770	0.231	0.001
AR	0.000	0.001	0.001	0.001	0.045	0.232	0.968	0.188	0.051	0.035	0.023	0.008	0.003
Factors	0.518	0.810	0.453	0.311	0.061	0.043	0.027	0.014	0.009	0.035	0.010	0.025	0.092
LASSO	0.165	0.027	0.118	0.317	0.462	0.378	0.045	0.491	0.184	0.474	0.291	0.505	0.021
F. aL.	0.200	0.019	0.070	0.123	0.024	0.012	0.032	0.516	0.011	0.020	0.058	0.011	0.041
P. OLS	0.210	0.031	0.127	0.330	0.128	0.057	0.049	0.987	0.051	0.100	0.205	0.021	0.042
RF	0.202	0.065	0.737	0.116	0.095	0.432	0.089	0.068	0.016	0.018	0.012	0.004	0.002
CSR	0.399	0.027	0.018	0.011	0.005	0.003	0.002	0.001	0.001	0.001	0.001	0.001	0.145
UCSV	0.000	0.000	0.000	0.000	0.001	0.000	0.002	0.085	0.912	0.223	0.059	0.042	0.002

The table shows the Giacomini and White test *p*-values for all models compared to the UCSV (first block) and the Bayesian VAR (second block).

Table 10
Average log-scores for all models and all forecasting horizons.

	Brazilian consumer price index Forecast horizon: average log-scores											
	$t+1$	$t+2$	$t+3$	$t+4$	$t+5$	$t+6$	$t+7$	$t+8$	$t+9$	$t+10$	$t+11$	$t+12$
RW	−0.01	−0.99	−1.27	−1.79	−1.79	−1.74	−1.90	−1.78	−1.51	−1.17	−0.62	−0.61
AR	−0.05	−0.34	−0.50	−0.57	−0.48	−0.50	−0.49	−0.33	−0.28	−0.23	−0.18	−0.20
Factor	0.33	−0.01	−0.17	−0.29	−0.23	−0.21	−0.28	−0.18	−0.05	−0.18	−0.14	−0.16
LASSO	0.87	0.20	−0.79	−1.00	−0.44	−0.48	−0.33	−0.85	−0.47	−1.14	−1.62	−1.87
F. aL.	0.90	0.42	0.03	−0.02	−0.08	−0.06	−0.15	−0.46	−0.10	−0.30	−0.23	−0.04
P. OLS	0.93	0.39	0.02	−0.11	−0.15	−0.20	−0.30	−0.62	−0.06	−0.43	−0.35	−0.08
RF	0.14	−0.07	−0.24	−0.22	−0.20	−0.22	−0.18	−0.26	−0.28	−0.16	−0.20	−0.29
CSR	0.85	0.38	0.16	−0.04	−0.09	−0.10	−0.07	0.00	−0.02	0.00	−0.01	−0.02
UCSV	−0.85	−0.83	−0.83	−0.82	−0.80	−0.80	−0.80	−0.79	−0.78	−0.76	−0.76	−0.75
BVAR	0.62	−0.20	−0.40	−0.68	−0.69	−0.71	−0.81	−0.79	−0.82	−0.85	−0.99	−0.94

The table shows the average log-scores as described by Amisano and Giacomini (2007). The log-scores were calculated on the empirical densities generated using bootstraps. Each bootstrap forecast is the sum of the point forecast and a random observation of the in-sample residuals. Each empirical density was constructed using 100,000 bootstrap forecasts.

The predictive densities are used to estimate average log-scores following Amisano and Giacomini (2007). Suppose that we have estimated predictive densities $\hat{f}(\cdot)$ for a given model. Let Y be the observed value of the variable in the period that we aim to forecast. The log-scores are calculated as $S(\hat{f}, Y) = \log \hat{f}(Y)$, and will be larger when the probability of the observed Y is high. For each model, we compute average log-scores across all of the rolling windows for each forecasting horizon. In each case, the best model is the one with the highest average log-score. Amisano and Giacomini (2007) also propose a test for checking whether the predictive densities for two models are statistically equal, which will be referred as the AG test from here on. If we reject the null, the two models have different densities.

The average log-scores and AG test *p*-values are presented in Tables 10 and 11. The first interesting result in Table 10 is that even though the LASSO and the Flex-adaLASSO provide similar point forecasts, the Flex-adaLASSO densities have larger log-scores, especially for long forecasting horizons. The CSR, which is the best model for point forecasts for most forecasting horizons, also has the largest log-scores. However, the POLS and the Flex-adaLASSO are the best models for $t + 1$ and $t + 2$, even though the CSR also performs well at these horizons. The BVAR works well for $t + 1$, but its performance deteriorates very rapidly for longer horizons. Note that the BVAR calculates forecasts for horizons greater than one by iterating previous forecasts. As a result, its point forecasts and variances converge rapidly to their respective unconditional values. Table 11 shows that, in general, when there is a large difference between the average

densities were obtained in the same way as for the other models. We then used these densities to estimate bootstrap point forecasts, which were used to obtain the $t + 2$ densities, and kept iterating the bootstrap point forecasts until $t + 12$.

Table 11
Amisano and Giacomini test p-values comparing the forecasting densities of all models for all 12 forecast horizons.

Brazilian consumer price index

AG test p-values

$_{t+2}\backslash^{t+1}$	RW	AR	Factor	LASSO	F. aL.	P. OLS	RF	CSR	UCSV	BVAR
RW		0.798	0.215	0.002	0.001	0.001	0.699	0.000	0.009	0.001
AR	0.064		0.160	0.001	0.001	0.001	0.603	0.001	0.003	0.003
Factor	0.013	0.201		0.020	0.033	0.027	0.660	0.043	0.011	0.293
LASSO	0.004	0.003	0.210		0.768	0.513	0.066	0.829	0.001	0.141
F. aL.	0.002	0.011	0.032	0.230		0.212	0.053	0.608	0.001	0.046
P. OLS	0.003	0.017	0.050	0.335	0.277		0.049	0.414	0.001	0.034
RF	0.056	0.186	0.861	0.223	0.203	0.248		0.113	0.026	0.217
CSR	0.002	0.014	0.035	0.331	0.404	0.886	0.242		0.001	0.100
UCSV	0.612	0.051	0.013	0.002	0.004	0.005	0.021	0.003		0.001
BVAR	0.063	0.319	0.527	0.055	0.082	0.109	0.346	0.098	0.030	

$_{t+4}\backslash^{t+3}$	RW	AR	Factor	LASSO	F. aL.	P. OLS	RF	CSR	UCSV	BVAR
RW		0.029	0.007	0.244	0.004	0.005	0.012	0.007	0.163	0.031
AR	0.031		0.036	0.323	0.005	0.008	0.125	0.005	0.149	0.609
Factor	0.012	0.139		0.019	0.181	0.225	0.477	0.122	0.011	0.174
LASSO	0.194	0.196	0.049		0.016	0.018	0.037	0.016	0.874	0.142
F. aL.	0.005	0.024	0.265	0.013		0.863	0.090	0.315	0.005	0.081
P. OLS	0.007	0.051	0.449	0.020	0.045		0.131	0.325	0.007	0.091
RF	0.005	0.185	0.744	0.025	0.338	0.600		0.088	0.020	0.313
CSR	0.006	0.031	0.153	0.007	0.916	0.642	0.294		0.005	0.078
UCSV	0.039	0.389	0.071	0.624	0.014	0.027	0.024	0.010		0.152
BVAR	0.025	0.652	0.172	0.317	0.036	0.073	0.045	0.039	0.637	

$_{t+6}\backslash^{t+5}$	RW	AR	Factor	LASSO	F. aL.	P. OLS	RF	CSR	UCSV	BVAR
RW		0.022	0.009	0.017	0.004	0.007	0.004	0.006	0.022	0.054
AR	0.026		0.108	0.814	0.040	0.185	0.196	0.080	0.227	0.379
Factor	0.008	0.057		0.188	0.418	0.770	0.908	0.460	0.047	0.072
LASSO	0.023	0.914	0.169		0.051	0.222	0.245	0.046	0.158	0.331
F. aL.	0.005	0.062	0.415	0.031		0.610	0.401	0.929	0.011	0.067
P. OLS	0.008	0.249	0.975	0.183	0.127		0.814	0.753	0.039	0.190
RF	0.005	0.246	0.956	0.153	0.390	0.936		0.413	0.024	0.099
CSR	0.007	0.115	0.547	0.045	0.702	0.524	0.461		0.014	0.070
UCSV	0.046	0.293	0.024	0.282	0.013	0.042	0.024	0.014		0.767
BVAR	0.045	0.354	0.082	0.374	0.054	0.147	0.058	0.066	0.769	

$_{t+8}\backslash^{t+7}$	RW	AR	Factor	LASSO	F. aL.	P. OLS	RF	CSR	UCSV	BVAR
RW		0.015	0.005	0.005	0.003	0.006	0.003	0.003	0.022	0.034
AR	0.007		0.286	0.550	0.277	0.556	0.300	0.177	0.389	0.200
Factor	0.008	0.239		0.819	0.611	0.946	0.636	0.348	0.075	0.095
LASSO	0.100	0.063	0.009		0.240	0.816	0.293	0.097	0.056	0.122
F. aL.	0.030	0.659	0.181	0.120		0.309	0.883	0.410	0.018	0.073
P. OLS	0.061	0.383	0.075	0.318	0.053		0.413	0.121	0.060	0.150
RF	0.005	0.600	0.624	0.047	0.445	0.250		0.450	0.022	0.045
CSR	0.004	0.065	0.143	0.007	0.029	0.022	0.041		0.010	0.047
UCSV	0.045	0.092	0.037	0.861	0.297	0.626	0.040	0.008		0.950
BVAR	0.055	0.046	0.047	0.890	0.426	0.696	0.059	0.022	0.979	

$_{t+10}\backslash^{t+9}$	RW	AR	Factor	LASSO	F. aL.	P. OLS	RF	CSR	UCSV	BVAR
RW		0.009	0.008	0.021	0.005	0.007	0.005	0.004	0.090	0.104
AR	0.015		0.169	0.246	0.144	0.141	0.981	0.069	0.071	0.063
Factor	0.021	0.708		0.033	0.656	0.898	0.236	0.794	0.019	0.044
LASSO	0.936	0.009	0.010		0.062	0.061	0.343	0.014	0.264	0.222
F. aL.	0.013	0.549	0.382	0.017		0.443	0.241	0.423	0.014	0.038
P. OLS	0.030	0.259	0.179	0.045	0.117		0.252	0.714	0.013	0.040
RF	0.009	0.471	0.879	0.009	0.237	0.143		0.115	0.056	0.075
CSR	0.007	0.090	0.137	0.004	0.024	0.030	0.159		0.011	0.023
UCSV	0.319	0.042	0.041	0.368	0.074	0.233	0.015	0.010		0.882
BVAR	0.511	0.040	0.069	0.530	0.088	0.205	0.048	0.022	0.776	

$_{t+12}\backslash^{t+11}$	RW	AR	Factor	Lasso	F. aL.	P. OLS	RF	CSR	UCSV	BVAR
RW		0.068	0.029	0.021	0.094	0.211	0.087	0.026	0.617	0.437
AR	0.079		0.631	0.007	0.798	0.444	0.874	0.095	0.037	0.033
Factor	0.047	0.774		0.004	0.553	0.238	0.644	0.260	0.026	0.046
LASSO	0.005	0.002	0.001		0.006	0.006	0.007	0.004	0.089	0.290
F. aL.	0.069	0.317	0.484	0.001		0.085	0.850	0.110	0.075	0.115
P. OLS	0.088	0.482	0.668	0.001	0.539		0.469	0.063	0.174	0.197
RF	0.270	0.626	0.423	0.003	0.240	0.339		0.124	0.044	0.053
CSR	0.045	0.180	0.227	0.001	0.866	0.642	0.124		0.012	0.023
UCSV	0.670	0.062	0.034	0.016	0.020	0.028	0.088	0.015		0.610
BVAR	0.349	0.016	0.050	0.086	0.026	0.030	0.069	0.020	0.618	

The table shows p-values of the Amisano and Giacomini test. Each block shows the p-values for two different forecast horizons (gray and white cells). The test statistics were calculated on the log-scores estimated using bootstrap densities. Each bootstrap forecast is the sum of the point forecast and a random observation of the in-sample residuals. Each empirical density was constructed using 100,000 bootstrap forecasts.

log-scores of two models, we reject the null and obtain statistically different models.

5. Conclusion

We have tested several high-dimensional econometric models for forecasting inflation in real-time and using a large number of predictors. We have also considered a forecasting combination mechanism based on the model confidence sets. The methods discussed here have been evaluated using Brazilian inflation data (IPCA). The results can summarized as follows.

For five days ahead, the LASSO and FOCUS (expert) forecasts are virtually the same, and deliver the best forecasts. For the second horizon, the adaptive LASSO is superior to all other models considered. For the remaining horizons, the complete subset regression dominates all other alternatives. These results are the same regardless of whether we use the root mean squared error or the mean absolute error. In terms of accumulated inflation, the complete subset regression is the best model. However, most of the forecasts from the different models are not statistically different according to the model confidence set. We construct the final forecast as the average of the models included in the confidence set. This approach delivers the best forecasts from among all of the competing alternatives. The Bayesian VAR also produced accurate forecasts for shorter horizons, but not as good as those from some of the high-dimensional models.

Finally, we computed predictive densities for all individual models, using bootstrap and estimated log-scores to compare the models. The results are consistent with the point forecasts. Models from the LASSO family are better for $t+1$ and $t+2$ and the *CSR* is the best model for longer horizons.

Appendix. Data appendix

The dataset and the computer codes are available from https://github.com/gabrielrvsc/hdeconometrics. The "HD econometrics" repository is an R package that provides implementations of the models used in this paper and the data in a.rda file. The package contains a number of functions that are used in the paper, such as a function that selects the best LASSO model using the BIC, a function for the complete subset regression with several arguments to control for pre-testing, and a function for the Bayesian VAR model. Documentation is available in markdown and follows the same format as traditional R documentation. It can also be viewed in R if the package is installed. Each function has an example included in the documentation.

All of the variables included in the models are listed in Tables 12 and 13. The first table shows the macroeconomic variables, all of which obtained from Bloomberg.[5] Table 13 shows the variables from the expectations database of the Brazilian Central Bank.

Most of the variables in our dataset are published for month t before the Brazilian CPI, which is made public around the 10th day of month $t+1$. Some variables have some delay or are available only after the CPI is published. In such cases, we use the last available observation of such variables.

[5] The variable names in Table 12 are the same as those in the Bloomberg database.

Table 12
Macroeconomic variables.

	Prices and money		**Government and international transactions**
1	Brazil CPI IPCA	32	Brazil National Treasury Revenue Total
2	FGV Brazil General Prices IGP-M	33	Brazil Social Contribution over Net Profit Tax Income
3	FGV Brazil General Prices IGP-DI	34	Brazil PIS & PASEP Tax Income
4	FGV Brazil General Prices IGP-10	35	Brazil Central Government Net Revenue
5	Brazil CPI IPCA-15	36	Brazil Central Government Revenue from the Central Bank
55	Brazil Monetary Base	37	Brazil Central Government Total Expenditures
56	Brazil Money Supply M1 Brazil	38	Brazil National Treasury Gross Revenue
57	Brazil Money Supply M2 Brazil	39	Brazil Importing Tax Income
58	Brazil Money Supply M3 Brazil	40	BNDES Brazil Income Taxes
59	Brazil Money Supply M4 Brazil	41	Brazil National Treasury Revenue from Industrialized Products
		42	Brazil National Treasury Revenues from Other Taxes
	Employment	43	Brazil Central Government Revenue from the Social Security
14	IBGE Brazil Unemployment Rate	44	Brazil National Treasury Revenue from Import Tax
15	Brazil Unemployment Statistic Male	45	Brazil Current Account
16	Brazil Unemployment Statistic Total	46	Brazil Trade Balance FOB
17	IMF Brazil Unemployment Rate	47	Brazil Public Net Fiscal Debt as a percentage of GDP
18	CNI Brazil Manufacturing Industry Employment	48	Brazil Public Net Fiscal Debt
19	Brazil Industry Working Hours	49	Brazilian Federal Government Domestic Debt
		50	Brazil Public Net Government & Central Bank Domestic Debt
	Exchange rates and finance	51	Brazilian States Debt Total Consolidated Net Debt
22	USD-BRL X-RATE	52	Brazilian States Debt to Foreigners
23	USD-BRL X-RATE Tourism	53	Brazilian Cities Debt
24	EUR-BRL X-RATE	54	Brazilian Cities Debt to Foreigners
25	BRAZIL IBOVESPA INDEX		
26	Brazil Savings Accounts Deposits		
27	Brazil Total Savings Deposits		
28	Brazil BNDES Long Term Interest Rate		
29	Brazil Selic Target Rate		
30	Brazil Cetip DI Interbank Deposits		

Table 13
Focus expectation variables.

60	$t+1$ median	77	Top5 $t+5$ median
61	$t+2$ median	78	Top5 $t+6$ median
62	$t+3$ median	79	Top5 $t+7$ median
63	$t+4$ median	80	Top5 $t+8$ median
64	$t+5$ median	81	Top5 $t+9$ median
65	$t+6$ median	82	Top5 $t+10$ median
66	$t+7$ median	83	Top5 $t+11$ median
67	$t+8$ median	84	Top5 $t+12$ median
68	$t+9$ median	85	Top5 $t+13$ median
69	$t+10$ median	86	$t+1$ median^2
70	$t+11$ median	87	$t+1$ mean
71	$t+12$ median	88	$t+1$ mean^2
72	$t+13$ median	89	$t+1$ Std
73	Top5 $t+1$ median	90	$t+12$ median^2
74	Top5 $t+2$ median	91	$t+2$ mean
75	Top5 $t+3$ median	92	$t+2$ mean^2
76	Top5 $t+4$ median	93	$t+2$ Std

The first group of Table 12 covers prices and money. The CPI IPCA is the variable of interest, while the CPI IPCA-15 is another index that is released earlier and used as an indicator of the final CPI IPCA. These two indexes are adopted officially by the government. The FGV indexes are calculated by the Getlio Vargas Foundation (FGV). They are also important measures of inflation. The second group of the table covers employment variables, while the third group deals with exchange rates, financial variables, savings and interest rates. IBOVESPA is the most important Brazilian stock index, and BNDES is the national bank of investment, which lends money at lower rates and has a significant impact on national investment. The Selic is the target interest rate, and is published by the Central Bank. The last group of variables in Table 12 covers government and international transactions.

All variables in Table 13 are obtained from the Brazilian Central Bank expectations database. Recall that the forecasts for $h=1$ are made five days before the CPI was published, meaning that $t+13$ forecasts are for horizons of 12 months and five days. Our data include the FOCUS and the Top5 median forecasts for $h=1$ to $h=13$. We also include the average forecasts, the squared average and median forecasts, and their standard deviations for horizons 1 and 2.

References

Amisano, G., & Giacomini, R. (2007). Comparing density forecasts via weighted likelihood ratio tests. *Journal of Business and Economic Statistics*, *25*(2), 177–190.
Arruda, E., Ferreira, R., & Castelar, I. (2011). Modelos lineares e não lineares da curva de phillips para previsão da taxa de inflação no brasil. *Revista Brasileira de Economia*, *65*, 237–252.
Atkeson, A., & Ohanian, L. (2001). Are Phillips curves useful for forecasting inflation? *Federal Reserve Bank of Minneapolis Quarterly Review*, *25*, 2–11.
Bai, J., & Ng, S. (2002). Determine the number of factors in approximate factor models. *Econometrica*, *70*, 191–221.
Bai, J., & Ng, S. (2006). Confidence intervals for diffusion index forecasts and inference for factor augmented regressions. *Econometrica*, *74*, 1133–1155.
Bai, J., & Ng, S. (2008). Forecasting economic time series using targeted predictors. *Journal of Econometrics*, *146*, 304–317.
Bańbura, M., Giannone, D., & Reichlin, L. (2010). Large Bayesian vector autoregressions. *Journal of Applied Econometrics*, *25*, 71–92.
Belloni, A., & Chernozhukov, V. (2013). Least squares after model selection in high-dimensional sparse models. *Bernoulli*, *19*(2), 521–547.
Bernanke, B., Boivin, J., & Eliasz, P. (2005). Measuring the effects of monetary policy: a factor-augmented vector autoregressive (FAVAR) approach. *Quarterly Journal of Economics*, *120*, 387–422.
Breiman, L. (2011). Random forests. *Machine Learning*, *45*, 5–32.
Chen, Y.-C., Turnovsky, S., & Zivot, E. (2014). Forecasting inflation using commodity price aggregates. *Journal of Econometrics*, *183*, 117–134.
Clements, M., & Galvão, A. (2013). Real-time forecasting of inflation and output growth with autoregressive models in the presence of data revisions. *Journal of Applied Econometrics*, *28*, 458–477.
Efron, B., Hastie, T., Johnstone, I., & Tibshirani, R. (2004). Least angle regression. *The Annals of Statistics*, *32*, 407–499.
Elliott, G., Gargano, A., & Timmermann, A. (2013). Complete subset regressions. *Journal of Econometrics*, *177*(2), 357–373.
Elliott, G., Gargano, A., & Timmermann, A. (2015). Complete subset regressions with large-dimensional sets of predictors. *Journal of Economic Dynamics and Control*, *54*, 86–110.
Fama, E. F., & French, K. R. (1993). Common risk factors in the returns of stocks and bonds. *Journal of Financial Economics*, *33*, 3–53.
Fama, E. F., & French, K. R. (1996). Multifactor explanation of asset pricing anomalies. *Journal of Finance*, *51*, 55–84.
Figueiredo, E., & Marques, A. (2009). Inflação inercial como um processo de longa memória: análise a partir de um modelo ARFIMA-FIGARCH. *Estudos Econômicos*, *39*, 437–458.
Fisher, I. (1930). *The theory of interest*. New York: Macmillan.
Forni, M., Hallin, M., Lippi, M., & Reichlin, L. (2003). Do financial variables help forecasting inflation and real activity in the euro area? *Journal of Monetary Economics*, *50*, 1243–1255.
Giacomini, R., & White, H. (2006). Tests of conditional predictive ability. *Econometrica*, *74*(6), 1545–1578.
Groen, J., Paap, R., & Ravazzolo, F. (2013). Real-time inflation forecasting in a changing world. *Journal of Business and Economic Statistics*, *31*, 29–44.
Han, X. (2015). Tests for overidentifying restrictions in factor-augmented var models. *Journal of Econometrics*, *184*, 394–419.
Hansen, P., Lunde, A., & Nason, J. (2011). The model confidence set. *Econometrica*, *79*, 453–497.
Iversen, J., Laséen, S., Lundvall, H., & Söderström, U. (2016). Real-time forecasting for monetary policy analysis: the case of Sveriges Riksbank, Working Paper 318, Sveriges Riksbank Working Paper Series.
Kohlscheen, E. (2012). Uma nota sobre erros de previsão da inflação de curto-prazo. *Revista Brasileira de Economia*, *66*, 289–297.
Marques, A. (2013). Central Bank of Brazil's market expectations system: a tool for monetary policy. In *IFC bulletins. Vol. 36* (pp. 304–324). Bank for International Settlements.
Medeiros, M., & Mendes, E. (2016). ℓ_1-regularization of high-dimensional time-series models with flexible innovations. *Journal of Econometrics*, *191*, 255–271.
Medeiros, M., & Vasconcelos, G. (2016). Forecasting macroeconomic variables in data-rich environments. *Economics Letters*, *138*, 50–52.
Medeiros, M., Vasconcelos, G., & Freitas, E. (2016). Forecasting Brazilian inflation with high-dimensional models. *Brazilian Review of Econometrics*, *36*, 68–100.
Meinshausen, N., & Yu, B. (2009). LASSO-type recovery of sparse representations for high dimensional data. *The Annals of Statistics*, *37*, 246–270.
Monteforte, L., & Moretti, G. (2013). Real-time forecasts of inflation: the role of financial variables. *Journal of Forecasting*, *32*, 51–61.
Samuels, J., & Sekkel, R. (2017). Model confidence sets and forecast combination. *International Journal of Forecasting*, *33*, 48–60.
Stock, J., & Watson, M. (1999). Forecasting inflation. *Journal of Monetary Economics*, *44*, 293–335.
Stock, J., & Watson, M. (2007). Why has US inflation become harder to forecast? *Journal of Money, Credit and Banking*, *39*, 3–33.
Tibshirani, R. (1996). Regression shrinkage and selection via the LASSO. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, *58*, 267–288.
Zhao, P., & Yu, B. (2006). On model selection consistency of LASSO. *The Journal of Machine Learning Research*, *7*, 2541–2563.
Zou, H. (2006). The adaptive LASSO and its oracle properties. *Journal of the American Statistical Association*, *101*, 1418–1429.