



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.

© 2017 International Institute of Forecasters. Published by Elsevier B.V. All rights reserved.

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 hyperinflation 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: mgarci@econ.puc-rio.br (M.G.P. Garcia), mcm@econ.puc-rio.br (M.C. Medeiros), gabrielrvsc@yahoo.com.br (G.F.R. Vasconcelos).

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.

Download English Version:

<https://daneshyari.com/en/article/5106427>

Download Persian Version:

<https://daneshyari.com/article/5106427>

[Daneshyari.com](https://daneshyari.com)