International Journal of Forecasting 33 (2017) 121–131

Contents lists available at ScienceDirect

International Journal of Forecasting

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

Forecasting the Brazilian yield curve using forward-looking variables

Fausto Vieira [a,*], Marcelo Fernandes [b,a], Fernando Chague [c]

[a] Sao Paulo School of Economics, Fundação Getulio Vargas (FGV), Brazil

[b] Queen Mary University of London, United Kingdom

[c] Department of Economics, University of São Paulo, Brazil

ARTICLE INFO

Keywords:
Bonds
Factor-augmented VAR
Forecasting
Term structure
Yield curve

ABSTRACT

This paper proposes a forecasting model that combines a factor augmented VAR (FAVAR) methodology with the Nelson and Siegel (NS) parametrization of the yield curve in order to predict the Brazilian term structure of interest rates. Importantly, we extract the principal components for the FAVAR from a large data set containing a range of forward-looking macroeconomic and financial variables. Our forecasting model improves on the predictive accuracy of extant models in the literature significantly, particularly at short-term horizons. For instance, the mean absolute forecast errors are 15–40% lower than those of the random walk benchmark on predictions at the three-month horizon. The out-of-sample analysis shows that the inclusion of forward-looking indicators is the key to improving the predictive ability of the model.

1. Introduction

The yield curve of treasury bonds plays a central role in both pricing financial assets and shaping market expectations. As such, accurate forecasts of the yield curve are of great importance for the Treasury, central bankers and market participants in general. Unfortunately, none of the extant models in the literature are able to outperform the random walk benchmark consistently at short horizons, while at the same time providing good forecasts at longer horizons.

This paper proposes a forecasting strategy for the yield curve that achieves this. We provide out-of-sample evidence that our forecasting model improves on the random walk benchmark at short horizons (as early as one month ahead) while at the same time providing more accurate forecasts than the extant models at longer horizons. The key ingredient of our strategy is its reliance on a comprehensive data set of macroeconomic and financial variables that are mostly forward-looking. Specifically, we proceed in three steps. In the first, we estimate the entire yield curve using the Nelson and Siegel (1987, NS) parametrization of the yield curve. The NS parametrization successfully summarizes the variation in the yield curve using the level, slope and curvature factors. In the second stage, we predict the future paths of these factors by estimating a factor augmented VAR (FAVAR) model using a comprehensive dataset of macroeconomic and financial variables. Finally, we form forecasts of the yield curve for each maturity at different horizons, using the predicted evolution of the level, slope and curvature factors.

We ensure real time forecasts by repeating these three steps at each prediction point. As our forecasting model combines a Nelson–Siegel decomposition of the yield curve with a FAVAR specification, we denote it by NS-FAVAR. Our forecasts of the yield curve beat the random walk

* Corresponding author.
E-mail address: ytcfausto@yahoo.com.br (F. Vieira).

http://dx.doi.org/10.1016/j.ijforecast.2016.08.001
0169-2070/

benchmark as early as at the one-month horizon. This represents a significant improvement, given that the available models produce meaningful predictions only from the six-month horizon (see Diebold & Li, 2006; and Moench, 2008). At the one-month horizon, our model forecast errors are 5% lower than those of the random walk benchmark, whereas at longer horizons, they are 20%–40% lower than those of the random walk benchmark.

The use of a comprehensive dataset that contains a wide array of forward-looking macroeconomic and financial variables is critical to the superior short-horizon performance of our forecasting strategy. In this respect, Brazilian economic data sets provide a surprisingly rich array of variables. As a consequence of a high-inflationary past, Brazilian market participants consume a variety of price indexes and price expectations indexes, some of which are available at the weekly or even daily frequencies. Moreover, a large number of macroeconomic and financial expectations time series are available readily, in an effort to increase the transparency of the monetary policy and guide market expectations. Our dataset contains 142 macroeconomic and financial variables at the weekly frequency, of which 40% are forward-looking indicators. Examples of macroeconomic forward-looking variables that make an important contribution to our forecasts are the market expectations of GDP growth, of the federal government balance sheet and of the debt-to-GDP ratio.

Our forecasting strategy builds on the work of Diebold, Rudebusch, and Aruoba (2006). However, instead of including only a few macroeconomic variables, we use a comprehensive data set of 103 macroeconomic variables and 39 financial indicators. We deal with this large number of conditioning variables by implementing Bernanke, Boivin, and Eliasz (2005) FAVAR econometric model. The FAVAR model restricts attention to the dynamics of a few principal components that summarize the variation in the data set. We show that conditioning on a broader information set, with many forward-looking macro-financial indicators, is key to improving the predictability.

This is not the first paper to improve yield curve forecasts at shorter horizons. de Pooter, Ravazzolo, and van Dijk (2010) study models with and without arbitrage restrictions that use macroeconomic information. They find that autoregressive models with macroeconomic predictors achieve superior performances at shorter horizons, but fail to improve on the random walk benchmark at longer horizons. Exterkate, Van Dijk, Heij, Patrick, and Groenen (2013) discuss the importance of large data sets for improving yield curve forecasts at short horizons. The authors show that factor augmented Nelson and Siegel models are able to improve the short-term forecasting during volatile periods, but cannot improve on simpler models in periods of low volatility.

In addition, we are also not the first to advocate for the use of forward-looking variables. Altavilla, Giacomini, and Constantini (2014) and Altavilla, Giacomini, and Ragusa (2014) use market and survey expectations to produce lower short-term forecasting errors of the short-term yields at the three- and six-month horizons. However, they are unable to either improve forecasts at longer horizons or ameliorate longer-term yield predictions. In contrast, van Dijk, Koopman, van der Wel, and Wright (2014) improve the forecasting performance for long maturities and at longer horizons by allowing shifting endpoints in the yield curve factors, but their forecasts are weak at shorter horizons.

In summary, we contribute by ameliorating term structure forecasts for virtually every maturity, even at short horizons. We argue that the key is to condition on the information set spanned by a few principal components from a wide array of mainly forward-looking macro-financial indicators.

We organize the remainder of this paper as follows. Section 2 reviews the Nelson–Siegel approach to the modelling of the term structure of interest rates, and describes our forecasting strategy. Section 3 describes the data set, while Section 4 discusses the out-of-sample results of our forecasting strategy. Section 5 contains several robustness exercises, and Section 6 offers some concluding remarks.

2. The forecasting strategy

Our forecasting strategy has three steps. In the first, we estimate the entire yield curve using the Nelson and Siegel (NS) parametrization of the yield curve. In the second, we predict the evolution of the level, slope and curvature factors using a FAVAR approach. Finally, we back out yield forecasts for each maturity at different horizons using the predicted future path of the NS factors.

As such, our forecasting strategy is very similar to that of Diebold et al.'s (2006) VAR model for the level, slope and curvature factors, with the main difference being that they employ only a few macroeconomic variables, whilst we condition on a much broader information set. We do so by following Stock and Watson's (2002b) idea of conditioning on a small number of principal components from a wide array of macroeconomic and financial variables. In particular, we employ a FAVAR model for the level, slope, curvature factors of the yield curve, and for the principal components from a data set of 142 macroeconomic and financial variables.

The Nelson–Siegel decomposition of the yield curve posits that we may approximate the yield with maturity n by

$$\hat{y}_t^{(n)} = \hat{\beta}_{1t} + \hat{\beta}_{2t}\left(\frac{1-e^{-\lambda n}}{\lambda n}\right) + \hat{\beta}_{3t}\left(\frac{1-e^{-\lambda n}}{\lambda n} - e^{-\lambda n}\right), \tag{1}$$

where the betas may vary over time, capturing changes in the level, slope and curvature of the term structure, respectively. The NS decomposition allows one to form predictions of the entire yield curve by simply predicting the dynamics of the level, slope and curvature factors. Following Stock and Watson (2002b), we extract the principal components of a comprehensive data set of 142 macroeconomic and financial predictors at the weekly frequency, to proxy for the broad economic conditions.[1]

[1] Although the principal component analysis formally requires independent and identically distributed observations, Doz, Giannone, and Reichlin (2012) and Stock and Watson (2002a) show that its performance is similar to that of full maximum likelihood estimation for a large panel in the context of static and dynamic factor models, respectively.