International Journal of Forecasting 31 (2015) 33–50

Contents lists available at ScienceDirect

International Journal of Forecasting

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

Forecasting GDP growth using mixed-frequency models with switching regimes

Fady Barsoum, Sandra Stankiewicz *

University of Konstanz, Department of Economics, Chair of Statistics and Econometrics, Box 129, 78457 Konstanz, Germany

ARTICLE INFO

Keywords:
Markov-switching
Business cycle
Mixed-frequency data analysis
Forecasting

ABSTRACT

For modelling mixed-frequency data with a business cycle pattern, we introduce the Markov-switching Mixed Data Sampling model with unrestricted lag polynomial (MS-U-MIDAS). Usually, models of the MIDAS-class use lag polynomials of a specific function which impose some structure on the weights of the regressors included in the model. This may lead to a deterioration in the predictive power of the model if the structure imposed differs from the data generating process. When the difference between the available data frequencies is small and there is no risk of parameter proliferation, using an unrestricted lag polynomial might not only simplify the model estimation, but also improve its forecasting performance. We allow the parameters of the MIDAS model with an unrestricted lag polynomial to change according to a Markov-switching scheme in order to account for the business cycle pattern observed in many macroeconomic variables. Thus, we combine the unrestricted MIDAS with a Markov-switching approach and propose a new Markov-switching MIDAS model with unrestricted lag polynomial (MS-U-MIDAS). We apply this model to a large dataset with the help of factor analysis. Monte Carlo experiments and an empirical forecasting comparison carried out for the U.S. GDP growth show that the models of the MS-U-MIDAS class exhibit nowcasting and forecasting performances which are similar to or better than those of their counterparts with restricted lag polynomials.

1. Introduction

Forecasting GDP growth is important for the decision making process, at both the central administrative (central bank, government) and industry levels. Due to difficulties in the measurement of GDP, it is published with a delay of a couple of months and revised repeatedly. This creates an obstacle for policy makers and market participants, who need to be ahead of changes in the economy, or at least to adjust quickly. Thus, reliable predictions are needed badly, but most of the existing forecasting models do not perform satisfactorily. This might be due to the fact that these models often ignore the non-linearities in the data (e.g., business cycle patterns), and/or fail to explore the informational content of the data published more frequently than the GDP or with a shorter lag. In addition, many models fail to make use of the informational content of large datasets, due to the problem of parameter proliferation.

Many approaches fail to account for the fact that macroeconomic variables often behave differently in different phases of the business cycle. Thus, a model with constant parameters might not reflect the present situation well, let alone be useful for forecasting. Furthermore, most models cannot include time series of different frequencies within the same regression. Instead, they require the data to be transformed (through either aggregation or interpolation), so that left- and right-hand side variables are of the same frequency. However, that might lead to a loss of

* Corresponding author.
E-mail addresses: fady.nagy-barsoum@uni-konstanz.de (F. Barsoum), sandra.stankiewicz@uni-konstanz.de (S. Stankiewicz).

http://dx.doi.org/10.1016/j.ijforecast.2014.04.002
0169-2070/

information. In addition, many models are not suitable for dealing with large datasets, which forces the forecaster to limit the number of explanatory variables and ignore useful information from other potential regressors. Finally, in most models it is not possible to include the most recent observations of the higher-frequency variable when the corresponding data on the dependent variable are not yet available. Thus, these models cannot be used for nowcasting. This is a major drawback, as higher-frequency variables are useful indicators of the current state of the economy. They are also published relatively often and with a small delay, which makes them potentially very useful for forecasting lower-frequency variables, such as GDP growth.

Various different approaches can be used to solve the above-mentioned problems, including regime-switching models, introduced by Hamilton (1989), Mixed Data Sampling Regressions (MIDAS), recently developed by Ghysels, Santa-Clara, and Valkanov (2002), and dynamic factor analysis (see e.g. Stock & Watson, 2002a). Regime-switching models allow the parameters of the model to change according to the current state of the economy (e.g., different parameters for the expansion and recession periods), and accounting for business cycle patterns in macroeconomic variables might improve the forecasting performance of the model. MIDAS models, on the other hand, can include time series of different frequencies in the same regression without transforming them through aggregation or interpolation. They are also very useful in nowcasting (MIDAS with leads), as they can make use of the observations of higher-frequency variables even if the data from lower-frequency variables for the corresponding period are not yet available.[1] Finally, dynamic factor analysis helps to exploit the informational content of large datasets by summarizing the variation in the observed variables using just a few unobserved factors. Thus, using a single factor that explains a large part of the dataset variation, instead of a single observed variable, may capture more information from the available dataset and ensure parsimony of the model.

A vast body of literature on Markov-switching models is available, usually in the context of modelling the business cycle patterns of macroeconomic data. Anas, Billio, Ferrara, and Duca (2007) explore the use of multivariate Markov-switching models for analysing the relationship between the phases of the business cycle in the United States and the Euro zone. Krolzig (2000) investigates the forecasting performance of the multivariate Markov-switching processes through Monte Carlo experiments and an empirical application to the United States business cycle. In addition, Clements and Krolzig (1998) study the forecasting performances of Markov-switching models through Monte Carlo simulations and an empirical study for the US GNP. Lahiri and Wang (1994) use the Markov-switching framework to predict turning points in the US business cycle. Cheung and Erlandsson (2005), Engel (1994) and Frömmel, MacDonald, and Menkhoff (2005), use Markov-switching models to explain and predict the fluctuations in exchange rates, whereas Evans and Wachtel (1993), Pagliacci and Barraez (2010) and Simon (1996) use the Markov-switching framework to analyse the past dynamics of inflation in Venezuela, the United States and Australia respectively.

MIDAS models, which were introduced to the literature only recently, have already found a number of interesting applications in both macroeconomics and finance. Kuzin, Marcellino, and Schumacher (2011) investigate the performance of the MIDAS model for nowcasting and forecasting GDP in the Euro area, relative to a mixed-frequency VAR (with missing values of the lower frequency variables interpolated using Kalman filter). They conclude that the two approaches seem to be complementary, as MIDAS performs better for short forecast horizons, whereas the mixed-frequency VAR does better for longer ones. Marcellino and Schumacher (2010) undertake a similar study, investigating the abilities of factor MIDAS models vs. state space factor models for forecasting German GDP. They find that factor MIDAS models usually outperform their state-space counterparts in forecasting, and that the most parsimonious MIDAS regression performs best overall.

Bai, Ghysels, and Wright (2013) compare MIDAS regressions with state space models through Monte Carlo simulations and an empirical exercise focusing on predicting GDP growth in the United States. They conclude that the two approaches are comparable in terms of forecasting performances. Clements and Galvão (2008, 2009) use MIDAS regressions of monthly and quarterly data for forecasting the GDP growth of the United States and obtain promising results, especially for MIDAS with leads. Andreou, Ghysels, and Kourtellos (2013) test the suitability of MIDAS factor models with leads for forecasting quarterly GDP growth in the United States with a large dataset of daily financial and quarterly macroeconomic indicators. They find such models to perform relatively well, especially in the crisis periods. Barsoum (2011) carries out a similar analysis for the United Kingdom, comparing MIDAS and factor-augmented MIDAS (both with and without leads) with a bunch of benchmark models. He obtains mixed results on the performances of MIDAS models in general, but promising results for MIDAS with leads.

Although they are useful, Markov-switching and MIDAS models can only address one problem at a time: either the issue of business cycle patterns or the difference in data frequencies. Guérin and Marcellino (2013) therefore combine the two approaches, introducing a Markov-switching Mixed Data Sampling model (MS-MIDAS).[2] They assess its forecasting performance through Monte Carlo simulations and carry out empirical studies on forecasting GDP growth in the United States and the United Kingdom, and show that MS-MIDAS is a useful approach. In their version of the model, Guérin and Marcellino (2013) use the so-called restricted lag polynomial, which is based on a specific function (e.g., exponential function). Depending on this function, a particular structure is imposed on the weights of the regressors in the model. This prevents parameter proliferation, but at the same time restricts the values that

1 This feature also makes MIDAS models useful for dealing with ragged-edge data (though this is outside the scope of this paper).

2 More recently, Bessec and Bouabdallah (2013) used a factor-augmented MS-MIDAS model for forecasting.