Available online at www.sciencedirect.com
ScienceDirect
International Journal of Forecasting 27 (2011) 529–542
international journal of forecasting
www.elsevier.com/locate/ijforecast

MIDAS vs. mixed-frequency VAR: Nowcasting GDP in the euro area

Vladimir Kuzin[a], Massimiliano Marcellino[b,*], Christian Schumacher[c]

[a] *DIW, Berlin, Germany*
[b] *European University Institute, Università Bocconi and CEPR, Italy*
[c] *Deutsche Bundesbank, Germany*

Abstract

This paper compares the mixed-data sampling (MIDAS) and mixed-frequency VAR (MF-VAR) approaches to model specification in the presence of mixed-frequency data, e.g. monthly and quarterly series. MIDAS leads to parsimonious models which are based on exponential lag polynomials for the coefficients, whereas MF-VAR does not restrict the dynamics and can therefore suffer from the curse of dimensionality. However, if the restrictions imposed by MIDAS are too stringent, the MF-VAR can perform better. Hence, it is difficult to rank MIDAS and MF-VAR a priori, and their relative rankings are better evaluated empirically. In this paper, we compare their performances in a case which is relevant for policy making, namely nowcasting and forecasting quarterly GDP growth in the euro area on a monthly basis, using a set of about 20 monthly indicators. It turns out that the two approaches are more complements than substitutes, since MIDAS tends to perform better for horizons up to four to five months, whereas MF-VAR performs better for longer horizons, up to nine months.

Keywords: Nowcasting; Mixed-frequency data; Mixed-frequency VAR; MIDAS

1. Introduction

In recent times, the development of econometric models based on mixed frequency data has attracted a considerable amount of attention. In particular, the mixed-data sampling (MIDAS) approach proposed by Ghysels, Santa-Clara, and Valkanov (2004) and Ghysels, Sinko, and Valkanov (2007) has proven useful for various different forecasting purposes. MIDAS can be regarded as a time-series regression tool that allows the regressand and regressors to be sampled at different frequencies, and where distributed lag polynomials are used to ensure parsimonious specifications. Whereas MIDAS was initially used for financial applications (e.g., Ghysels, Santa-Clara, & Valkanov, 2005, 2006), it has also been employed for forecasting macroeconomic time

* Corresponding author.
E-mail addresses: vkuzin@diw.de (V. Kuzin), massimiliano.marcellino@eui.eu (M. Marcellino), christian.schumacher@bundesbank.de (C. Schumacher).

0169-2070/$ - see front matter
doi:10.1016/j.ijforecast.2010.02.006

series, and in particular quarterly GDP with monthly indicators, in recent applications by Clements and Galvão (2008, in press), Marcellino and Schumacher (in press), and Wohlrabe (2009).

In this paper, we compare the MIDAS approach to a mixed-frequency VAR (MF-VAR) model, as was proposed by Zadrozny (1988), Mittnik and Zadrozny (2005) and Mariano and Murasawa (2010). The MF-VAR is a VAR which is operating at the highest sampling frequency of the time series to be included in the model. Low-frequency variables are interpolated according to their stock-flow nature, implying specific time-aggregation schemes. The high-frequency VAR, together with the time-aggregation restriction, can be cast in state space form and estimated by maximum likelihood. In this framework, the Kalman filter can tackle missing values at the end of the sample, and can take into account the mixed-frequency nature of the data.

Compared to single-equation MIDAS, MF-VAR is a system approach that jointly explains indicators and target variables without imposing functional a priori restrictions on the dynamics. This can be an advantage when few variables are being modelled, their dynamics are limited, and the VAR provides a good approximation to the data generating process (DGP). Otherwise, MIDAS can represent a more robust forecasting device. In addition, due to its single equation specification, a direct forecasting approach is preferable for MIDAS, while an iterated scheme is a more natural choice for the MF-VAR, since it is cast in state space form and iterated forecasts are provided by the Kalman filter directly. For a discussion of direct versus iterated forecasting, see for example Marcellino, Stock, and Watson (2006) and Chevillon and Hendry (2005).

It is difficult to rank the MIDAS and MF-VAR approaches based purely on theoretical considerations, since, as was mentioned above, their relative merits depend on the DGP. Therefore, their performances are better assessed in specific economic applications, and in this paper we focus on nowcasting and forecasting quarterly euro area GDP growth using a set of monthly indicators, an issue which is also relevant from the economic policy perspective.

In our application, we compare various specifications of MIDAS and MF-VAR models with single indicators, as well as with combinations of these models. In addition, we take into account the differences in availability of monthly indicators that emerge from different statistical publication lags. The nowcast and forecast comparison is based on the relative mean squared errors (MSE) at different horizons, and the analysis is conducted recursively, in a pseudo real-time way.

Our main finding is that, in the case of euro area GDP growth, the two approaches are more complements than substitutes, since MF-VAR tends to perform better for longer horizons, whereas MIDAS is better for shorter horizons.

The paper proceeds as follows. Section 2 provides a description of the MIDAS and MF-VAR approaches, as well as a discussion of their relative advantages. Section 3 presents the empirical results on nowcasting and forecasting quarterly euro area GDP growth using a set of monthly indicators. Section 4 summarizes our main findings and concludes.

2. Nowcasting quarterly GDP with ragged-edge data

In this paper we focus on quarterly GDP growth, which is denoted as y_{t_q}, where t_q is the quarterly time index $t_q = 1, 2, 3, \ldots, T_q^y$, with T_q^y being the final quarter for which GDP data are available. GDP growth can also be expressed at the monthly frequency by setting $y_{t_m} = y_{t_q} \forall t_m = 3t_q$, with t_m being the monthly time index. Thus, GDP growth y_{t_m} is observed only in months $t_m = 3, 6, 9, \ldots, T_m^y$, with $T_m^y = 3T_q^y$. The aim is to nowcast or forecast the GDP h_q quarters ahead, or $h_m = 3h_q$ months ahead, yielding a value for $y_{T_m^y + h_m}$.

Nowcasting means that, in a particular calender month, we do not observe the GDP for the current quarter. It can even be the case that the GDP is only available with a delay of two quarters. In April, for example, the euro area GDP is only available for the fourth quarter of the previous year, and a nowcast for the second quarter GDP requires $h_q = 2$. Thus, if a decision-maker requests an estimate of the current quarter GDP, the forecast horizon has to be set to be sufficiently large in order to provide the appropriate figures. For a further discussion on nowcasting, see Giannone, Reichlin, and Small (2008), for example.

In this section we assume, for the sake of exposition, that the information set for now- and