FORECASTING WITH MIXED FREQUENCY DATA:

MIDAS VERSUS STATE SPACE DYNAMIC FACTOR MODEL: “AN APPLICATION TO FORECASTING SWEDISH GDP GROWTH”

Author:

Chen Yu

1989/06/05
Abstract

Most macroeconomic activity series such as Swedish GDP growth are collected quarterly while an important proportion of time series are recorded at a higher frequency. Thus, policy and business decision makers are often confront with the problems of forecasting and assessing current business and economy state via incomplete statistical data due to publication lags. In this paper, we survey a few general methods and examine different models for mixed frequency issues. We mainly compare mixed data sampling regression (MIDAS) and state space dynamic factor model (SS-DFM) by the comparison experiments forecasting Swedish GDP growth with various economic indicators. We find that single-indicator MIDAS is a wise choice when the explanatory variable is coincident with the target series; that an AR term enables MIDAS more promising since it considers autoregressive behaviour of the target series and makes the dynamic construction more flexible; that SS-DFM and M-MIDAS are the most outstanding models and M-MIDAS dominates undoubtedly at short horizons up to 6 months, whereas SS-DFM is more reliable at long predictive horizons. And finally we conclude that there is no perfect winner because each model can dominate in a special situation.

Key words: mixed frequency data, MIDAS regression, state space model, dynamic factor model, Swedish GDP growth
# Table of Contents

1 Introduction .................................................................................................................. 1

2 Retrospect of Literature ............................................................................................... 2

3 The Model ...................................................................................................................... 4

3.1 Flat Aggregation Model .......................................................................................... 4

3.2 Mixed Data Sampling Regressions ........................................................................ 5

3.2.1 General MIDAS ............................................................................................... 5

3.2.2 Autoregressive Structure MIDAS .................................................................. 6

3.2.3 Multivariate MIDAS ...................................................................................... 7

3.3 State Space Model and Dynamic Factor Model ....................................................... 8

3.3.1 General State Space Model ............................................................................ 8

3.3.2 General Kalman Filter algorithm .................................................................... 9

3.3.3 Dynamic factor model ................................................................................. 10

3.3.4 DFM in State Space Form ............................................................................. 11

3.3.5 Estimation and Missing value ....................................................................... 12

4 Evaluation of the forecast accuracy ............................................................................. 12

5 Empirical applications ................................................................................................. 13

5.1 Data ..................................................................................................................... 13

5.2 Design of comparison experiments ...................................................................... 14

5.3 Results ............................................................................................................... 15

6 Conclusion ................................................................................................................ 18

References: ................................................................................................................. 19

Appendix ...................................................................................................................... 22
1 Introduction

With the enormous development and innovation in the field of information and computer technology, an important proportion of time series, especially financial data, are recorded on a daily or even intra-daily frequencies. On the other hand, most macroeconomic activity series are collected monthly or quarterly. For instance, Swedish real gross domestic product (GDP) is officially released quarterly in spite of the fact that many economic forecasters prefer monthly GDP forecasts for better and more flexible economic and business decision. It is published 60 days after the involving period (for example, the fourth quarter GDP of 2012 is released on the 1st of March) while many other coincident indicators like personal income (PI) and industrial production (IP) are sampled at a higher frequency. However, forecasting models are in general constructed using the same frequency data by time-aggregating higher frequency data to lower frequency data. As a result, substantial information might be loss, which is a common problem that many forecasters and researchers are confronted with. Thus, forecasting with mixed frequency data has become an interesting and important topic and it is getting more attention in recent years.

In practice, dealing with mixed frequency data, in the procedure of macroeconomic models’ construction, is such a common problem that a lot of methodologies have been explored to solve this kind of situation. The conventional and standard way is to aggregate all the high frequency variables to the same frequency by a flat aggregation scheme. Despite of its parsimony, information loss always happens during the process of time averaging, encouraging new possibility that using high frequency data in models for variables with lower frequency regressand.

One of the main approaches is to employ state space models together with Kalman filter, which are quietly common used in many dynamic time series models in economics and finance. According to Tsay (2010), general state space model is constructed by two system equations, an observation equation providing the connection between the data and the state, and a state equation that monitors the state evolution and transition. State space model is a flexible method to mixed frequency cases since Kalman filter can be applied conveniently to regard the lower frequency data as missing data.
For simplicity, Ghysels, Santa-Clara, and Valkanov (2004) invented another solution, which is so called mixed data sampling (MIDAS) regressions, preserving the past high frequencies and avoids information being discarded as well. What is more is that MIDAS regressions enable one to estimate a handful of parameters and solve the problem of parameter proliferation although the framework uses diverse frequency data.

The objective of this paper is to survey a few general methods when forecasting the mixed frequency time series. More specifically, in the first place, the paper will review relevant literature and specialize this to the forecasting context. Second, different models for mixed data frequencies will be examined. Finally, comparison among MIDAS and SS-DFM will be illustrated explicitly by the empirical study that forecasts Swedish GDP growth with various economic indicators.

The paper is constructed as follows: Section 2 is literature retrospect, in which we survey the state and latest achievements obtained by authors, in the area of mixed frequency forecasting. In section 3 we cover and specify the notation of the models mentioned above, namely, flat aggregation models, state space models and MIDAS regression models. Section 4 describes forecast evaluation. After that the empirical comparison experiments are carried out in Section 5. Section 6 provides summary and conclusion.

2 Retrospect of Literature

Recently, the interest of handling mixed frequency data in financial and macroeconomic time series analysis promotes the development of forecasting models. One of the methods is the infant promising MIDAS, invented by Ghysels, Santa-Clara, and Valkanov (2004). Ghysels and his colleagues examine the features of MIDAS regressions and find that the estimation of MIDAS models always perform more effective than the conventional flat aggregation models. Tay (2006), apply general MIDAS model and an AR model to predict real output growth of US. He finds that daily stock return is a useful index for forecasting quarterly GDP growth. Beside, his MIDAS model always outperforms the flat time-aggregation
model although his AR model even produce better accuracy than MIDAS regression during a specific period of time.

Andreou, Ghysels, Kourtellos (2010) review different MIDAS models involving in mixed frequency issues. They find that most research focus on using higher frequencies to improve the forecast of lower frequencies variable, whereas, in some cases, it is interesting and possible to do the reverse (see, for instance, Engle, Ghysels, and Sohn (2008) and Colacito, Engle, and Ghysels (2010)). Alper, Fendoglu, and Saltoglu (2008) use a linear univariate MIDAS regression based on square daily returns to evaluate the relative weekly stock market volatility forecasting performance of ten initial markets: BSE30 (India), HSI (Hong Kong), IBOVESPA (Brazil), IPC (Mexico), JKSE (Indonesia), KLSE (Malaysia), KS11 (South Korea), MERVAL (Argentina), STI (Singapore), and TWII (China-Taiwan) and four developed ones: S&P500 (the U.S.), FTSE (the U.K.), DAX (Germany), and NIKKEI (Japan). They find that MIDAS regressions, comparing to GARCH (1, 1), produce better forecasting accuracy for the ten emerging stock markets while there is no winner in terms of the developed stock markets which are less volatile. Similar empirical studies and conclusions can also be found in Clements, and Galvão (2008), Armesto, Engemann, and Owyang (2010), Armesto, Hernandez-Murillo, Owyang, and Piger (2009).

Comparing to MIDAS models, state space models together with Kalman filter is another more mature solution to mixed frequencies issues. So far, there are two mainstream models which are always being cast into state space representation, one is dynamic factor model (DFM) and another is so called mixed-frequency vector autoregressive model (MF-VAR). For instance, Mariano and Murasawa (2007) take advantage of American coincident indicators applied both two methods mentioned above to construct a new index to predict business cycles, namely US monthly real GDP. By means of SBIC selection criterion, they choose a two-factor coincident index as a new coincident index in forecasting US real GDP.

Aruoba, Diebold, and Scotti (2009) make a small-data dynamic factor model that forecasts economic activity like recession in real time. They used four variables, namely GDP, monthly employment, weekly initial jobless claims and daily yield curve term premium and cast the model into a state space form. They provide us a
standard example using Kalman filter to construct the missing values and illustrated that Kalman filter is one of the priority selections to handling missing observations. Kuzin, Marcellino, and Schumacher (2009) compare the MIDAS, AR-MIDAS and a set of mixed frequency VAR models written in a state space form by now- and forecasting euro quarterly GDP based on a set of 20 monthly indicators. They find that among all the three models, MF-VAR performs better than MIDAS and AR-MIDAS at longer horizons while AR-MIDAS has the highest accuracy at shorter horizons, which is similar to the conclusion of Bai, Ghysels, and Wright (2009).

Wohlrabe (2008) examines state-space MF-VAR and MIDAS and makes the comparison of the two methods by taking a Mote Carlo simulation forecasting study. He finds that small models which are chosen on BIC criterion are sufficiently accurate in forecasting; that restrictions and lags will influence the performance of forecasting; that for fixed target variables, the restrictions might improve the forecasting accuracy. In addition, under the condition of giving the structure and the length of the macroeconomic time series models, increasing sample size is a vain way to improve the forecasting accuracy. MF-VAR outperforms MIDAS if time series have a strong serially correlated component and GARCH effects will influence the performance of mixed-frequencies methods. There are also some other papers applying Kalman filter to treat missing information in mixed-frequencies forecasting but we don’t name them one by one here.

3 The Model

Under the circumstance of forecasting lower frequency variables via high frequencies and different ways of time aggregation, three main methods are provided in general. In this section, these three main approaches are described in detail.

3.1 Flat Aggregation Model

One of the simplest ways to solve the dilemma of mixed frequency data is to transform higher frequencies into the lower ones by means of taking simple average of higher frequency observations:
\[ \bar{X}_t = \frac{1}{m} \sum_{k=1}^{m} L_{HF}^k X_t. \]

Here \( m \) is the sampling times of the higher frequency observations \( X \). In addition, \( L_{HF} \) expresses the lag term of the higher frequencies. After being taken averaging, the higher frequencies \( X \) is converted to the same sampling rate as \( Y_t \). Thus, as what was mentioned by Armesto, Engermann, and Owyang (2010), we can make a simple Distributed Lag (DL) regression:

\[ Y_t = \alpha + \sum_{j=1}^{n} \gamma_j L^j X_t + \varepsilon_t, \quad (1) \]

where \( \gamma_j \)s are the slope of different \( X \)s after being taken time averaging. Besides, the second term means, for instance, using the first \( j \)-th quarter’s average of monthly \( X \)s.

### 3.2 Mixed Data Sampling Regressions

MIDAS regression, a new and promising time series approach that directly regressing on variables sampled at various frequencies, was initially proposed by Ghysels, Santa-Clara, and Valkanov (2004). It is a parsimonious and flexible method that challenges the status of state space models together with Kalman filter.

#### 3.2.1 General MIDAS

In order to prevent parameter proliferation, MIDAS regression, one of the most important issues in financial models, applies the succinct distributed lag polynomials. A general univariate MIDAS regression model for one-step ahead forecasting can be written as:

\[ y_t = \beta_0 + \beta_1 B(L^{1/m}; \theta)x_{t-1}^{(m)} + \varepsilon_t^{(m)} \quad (2) \]

where \( B(L^{1/m}; \theta) = \sum_{k=1}^{K} b(k; \theta)L^{(k-1)/m} \), the sum of exponential Almon lag and \( L^{(k-1)/m} x_{t-1}^{(m)} = x_{t-1-(k-1)/m}^{(m)} \). In this paper one step ahead means one quarter ahead. Notice that \( m \) is the times that higher frequency sampled between the lowers \( y_t \), \( t \) expresses the unit of time, \( b(k; \theta) \) is the exponential Almon lag proposed by Ghysels, Santa-Clara, and Valkanov (2004) with the specification as:
If we set $\theta = (\theta_1, \theta_2) = (0, 0)$, a time averaging model will be achieved.

Because of the complexity of the model, here we give a simple example, we set $m = 3$ (since the explanatory variables $x$ is monthly and $y$ is quarterly) and $K = 12$ (last 12 monthly information are used). Accordingly, $B(L^{1/3}; \theta) = \sum_{k=1}^{12} b(k; \theta) L^{(k-1)/3}$, and $L^{(k-1)/3} x_{t-1}^{(3)} = x_{t-1-(k-1)/3}^{(3)}$. Thus, for example, from equation (2), a one-step ahead univariate MIDAS regression can be represented as:

$$y_t = \beta_0 + \beta_1 \left[ b(1; \theta_1) x_{t-1}^{(3)} + b(2; \theta_1) x_{t-1-1/3}^{(3)} + \cdots + b(12; \theta_1) x_{t-4-2/3}^{(3)} \right]$$

(3)

Meanwhile, MIDAS model for $h$-steps ahead is also available via regressing $y_t$ on $x_{t-h}^{(m)}$ with nonlinear least squares (NLS):

$$y_t = \beta_0 + \beta_1 B(L^{1/m}; \theta) x_{t-h}^{(m)} + \varepsilon_t^{(m)}$$

(4)

To be more clarify, $h = 1/3$ means $2/3$ of the information of the current quarter is extracted. According to theoretical comparisons, generally, MIDAS performs worse as predictive horizons become longer. More details will be discussed in the following section.

### 3.2.2 Autoregressive Structure MIDAS

One elementary extension of the basic MIDAS model is to add an autoregressive dynamics factor, proposed by Clements and Galvão (2008), solving the problem raised by Ghysels, Santa-Clara, and Valkanov (2004). Ghysels and his partners conclude that efficiency loss will happen inevitably if lagged regressands are introduced into MIDAS models. In addition, adding an autoregressive term would lead to a “seasonal” polynomial, which can only be applied under the circumstances that seasonal patterns exist in explanatory variables. Thus, we follow the step of Clements and Galvão (2008) and an $h$-steps ahead MIDAS model including an AR term can be written as:

$$y_t = \beta_0 + \lambda y_{t-h} + \beta_1 B(L^{1/m}; \theta)(1 - \lambda L^h) x_{t-h}^{(m)} + \varepsilon_t^{(m)},$$

(5)
which is denoted as AR-MIDAS in this paper.

In order to obtain the estimation of equation (5), one could acquire the residuals of the standard MIDAS model, \( \hat{\varepsilon}_t \). Then an starting value for \( \lambda \), namely \( \hat{\lambda}_0 \), can be estimated from \( \hat{\lambda}_0 = (\sum \hat{\varepsilon}_{t-h}^2)^{-1} \sum \hat{\varepsilon}_{t-h} \). After constructing \( y_t^* = y_t - \hat{\lambda}_0 y_{t-h} \) and \( x_t^* = x_{t-h} - \hat{\lambda}_0 x_{t-2h} \) and implementing nonlinear least squares (NLS) to:

\[
y_t^* = \beta_0 + \beta_1 B(L^{1/m}; \theta)x_t^* + \varepsilon_t, \tag{6}
\]

then estimator of \( \hat{\theta}_1 \) is gained. Meanwhile, we can obtain a new value of \( \lambda, \hat{\lambda}_1 \), from the residuals of equation (6).

### 3.2.3 Multivariate MIDAS

In the further part of the paper, an empirical study about forecasting Swedish GDP growth will be implemented. In general, a handful of independent variables are prone to give rise to a result with less accuracy, especially for forecasting some crucial economic indices, like real output growth. An important merit of MIDAS is that it can be used to estimate quarterly GDP succinctly, including most of the leading indicators and some coincident indicator variables at a monthly interval. According to Clements and Galvão (2008), an \( h \)-steps ahead multivariate MIDAS model (denoted as M-MIDAS in this paper) to forecast GDP growth, combining the information composed of \( N \) monthly indicators, can be expressed as:

\[
y_t = \beta_0 + \sum_{i=1}^{N} \beta_{1,i} B_i(L^{1/m}; \theta)x_{t-h}^{(m)} + \varepsilon_t^{(m)}, \tag{7}
\]

where all the indexes are identified by \( i \). \( \beta_{1,i} \) is the weight to measure the influence of the indicators while \( \theta_i \) depicts the lagged effect of the indicators.

Also, we can obtain an example from equation (7), we set \( m = 3 \) because all the explanatory variables \( x_x \) are monthly and \( y \) is quarterly and \( K = 12 \). Thus, a one-step ahead MIDAS regression incorporates for instance, \( N = 10 \) indicators, can be represented as:
\[ y_t = \beta_{1,1} \left[ b(1; \theta_1) x_{1,t-1}^{(3)} + b(2; \theta_1) x_{1,t-1}^{(3)} + \cdots + b(12; \theta_1) x_{1,t-1}^{(3)} \right] + \\
\beta_{1,2} \left[ b(1; \theta_2) x_{2,t-1}^{(3)} + b(2; \theta_2) x_{2,t-1}^{(3)} + \cdots + b(12; \theta_2) x_{2,t-1}^{(3)} \right] + \cdots + \\
\beta_{1,10} \left[ b(1; \theta_{10}) x_{10,t-1}^{(3)} + b(2; \theta_{10}) x_{10,t-1}^{(3)} + \cdots + b(12; \theta_{10}) x_{10,t-1}^{(3)} \right] + \varepsilon_t^{(3)}, \quad (8) \]

### 3.3 State Space Model and Dynamic Factor Model

#### 3.3.1 General State Space Model

As being noted in a few literature, for example, Mergner (2009), Tsay (2010), when forecasters are confront with a dynamic system with missing components, state space models can provide a powerful and flexible support. More importantly, in terms of mixed frequency problems, Kalman filter is still valid by treating lower frequencies as missing or unobservable values. In general, a dynamic system can be expressed into a state space form and a linear Gaussian State Space model can be written as:

\[ s_{t+1} = T_t s_t + d_t + R_t \eta_t, \quad \eta_t \sim N(0, Q_t), \quad (9) \]
\[ y_t = Z_t s_t + c_t + e_t, \quad e_t \sim N(0, H_t), \quad (10) \]
\[ t = 1, \ldots, T, \]

where equation (10) is an observation equation that provide the connection between the data and the state, and equation (9) is a state equation that monitor the state evolution and transition. In addition, \( y_t \) denote a \( k \times 1 \) vector of observed data where missing values exist in huge numbers generally, \( s_t \) is an \( m \times 1 \) state vector, each \( s_t \) is a realization of the random variable \( S_t \) at time \( t \). \( T_t \) and \( Z_t \) are coefficient matrices with \( m \times m \) and \( k \times m \) dimension, \( R_t \) is an \( m \times n \) matrix which is usually composed of a section of the \( m \times m \) identity matrix, \( \eta_t \) and \( e_t \) are \( n \times 1 \) and \( k \times 1 \) Gaussian white noise series, and \( Q_t \) and \( H_t \) are positive-define covariance matrices of \( n \times n \) and \( k \times k \) dimension respectively.

The initial state vector \( s_1 \), is assumed to obey a normal distribution \( N(a_1, p_1) \), where \( a_1 \) and \( p_1 \) are of dimension \( m \times 1 \) and \( m \times m \). Besides, these two matrices are given initially.
3.3.2 General Kalman Filter algorithm

The purpose of Kalman filter is to use the given data $Y_t = \{y_1, \ldots, y_t\}$ and the model to acquire a conditional distribution of state variable $s_{t+1}$. From Equation (10), the conditional distribution of $s_{t+1}$ can be represented as:

$$a_{t+1} = E(s_{t+1} | Y_t),$$

$$P_{t+1} = Var(s_{t+1} | Y_t),$$

When the initial value $a_1$ and $p_1$ are available, updating the knowledge of the state vector recursively by the Kalman filter algorithm becomes realizable:

$$a_{t+1} = d_t + T_t a_t + K_t v_t,$$

$$P_{t+1} = T_t P_t L_t + R_t Q_t R_t^\prime,$$

with

$$v_t = y_t - c_t - Z_t a_t,$$

$$V_t = Z_t P_t Z_t^\prime + H_t,$$

$$K_t = T_t P_t Z_t^\prime V_t^{-1},$$

$$L_t = T_t - K_t Z_t,$$

where $K_t$ is known as Kalman gain, being of dimension $m \times k$, $v_t$ denotes one-step head error of $y_t$ when $Y_{t-1}$ is available.

When dealing with missing observations at $t = l +1, \ldots, l +h$, we set $v_t = 0, K_t = 0$ in Equations (11) and Kalman filter is still effective and valid:

$$a_{t+1} = d_t + T_t a_t, \quad (12)$$

$$P_{t+1} = T_t P_t T_t^\prime + R_t Q_t R_t^\prime,$$

More derivation is illustrated in Durbin and Koopman (2001).
3.3.3 Dynamic factor model

State space model is such a flexible method that it can cope with most kind of data especially for missing observations. In mixed frequency forecasting context, there are two main methods, one is mixed frequency vector autoregressive model (MF-VAR) while another promising approach is dynamic factor model (DFM). In the further case application, we mainly concentrate on the latter.

Dynamic factor model, proposed by Doz, Giannone, and Reichlin (2005), is designed to seek a set of common trends in a large panel of series. Unlike a Mixed-Frequency VAR system where all the variables needs to be endogenous, a dynamic factor model is prone to have better performance in predicting monthly real GDP growth, especially with less parameters. According to Mariano and Murasawa (2007) and Doz et al (2005), a DFM composed of a set of \( n \) standardized (mean zero and variance one) stationary monthly series \( x_t = (x_{1,t}, ..., x_{n,t})' \) can be expressed as:

\[
\begin{align*}
    x_t &= \Lambda f_t + u_t, \quad u_t \sim NID(0, \Sigma_u), \quad (13) \\
    \Phi_f(L)f_t &= v_t, \quad v_t \sim NID(0, \Sigma_v), \quad (14) \\
    \Phi_u(L)f_t &= w_t, \quad w_t \sim NID(0, \Sigma_w), \quad (15)
\end{align*}
\]

where \( f_t = (f_{1,t}, ..., f_{r,t})' \) is an \( r \times 1 \) latent component and \( u_t = (u_{1,t}, ..., u_{n,t})' \) is an \( n \times 1 \) idiosyncratic factors. In addition, \( \Lambda \) is the factor loading matrix being of dimension \( n \times r \). In Equation (14) and (15), \( p \) and \( q \) are the maximized order of polynomial \( \Phi_f(.) \) and \( \Phi_u(.) \) respectively. In our application, Equation (15), the so called moving average part, can be neglected.

In order to forecast Swedish quarterly GDP growth via DFM, we keep the core idea of Mariano and Murasawa (2007). In the first place, the forecast of monthly GDP growth \( \hat{y}_t \) is introduced as a latent variable:

\[
\hat{y}_t = \beta' f_t = \beta_1 f_{1,t} + \cdots + \beta_r f_{r,t} \quad (16)
\]

where \( \hat{y}_t \) can be also regarded as common components from the static factor models or principal component analysis.
Afterwards we introduce the evaluation of quarterly GDP growth $\hat{y}_t^Q$, which is in the 3-nd month of each quarter. Here we should notice that $\hat{y}_t^Q$ is the average of monthly series $\{\hat{y}_t\}$:

$$\hat{y}_t^Q = \frac{1}{3}(\hat{y}_t + \hat{y}_{t-1} + \hat{y}_{t-2})$$  \hspace{1cm} (17)

Besides, $\varepsilon_t^Q = y_t^Q - \hat{y}_t^Q$, which is the forecast error obeying a normal distribution with $\varepsilon_t^Q \sim N(0, \Sigma_e)$. We suppose all the innovations, namely $u_t$, $v_t$, $w_t$ and $\varepsilon_t^Q$, are mutually independent. Thus, the general nature about forecasting real GDP growth with monthly indicators via a dynamic factor model has been completed.

3.3.4 DFM in State Space Form

Despite the factor that there exists several different and mainstream ways to cast a dynamic factor model into a state space representation like Equation (9) and (10), our issue is forecasting with variables at different time intervals, which is a more specific application. Thus, the following exposition in state space transformation is mainly in the spirit of Bańbura and Rünstler (2007), which is also an extension to Zadrozny (1990).

In order to cast Equation (13) to (17) into state space form, we consider the case of $p = 1$ and $q = 0$. Also, we just consider one-factor situation because of the parsimonious rule. In addition, by means of changing the weight of Equation (17) and setting a new $\hat{y}_t^*$, where $\hat{y}_t^* = \frac{1}{3}\hat{y}_t$, we can simplify and develop the complicated state space form from Bańbura and Rünstler (2007). The new state vector $x'_t = (f_{1,t}, \hat{y}_{t-1}^*, \hat{y}_{t-2}^*, \hat{y}_t^Q)$, which greatly boosts the calculation of the state and observation equations described below:

$$\begin{bmatrix}
    x_{1,t} \\
    x_{2,t} \\
    x_{3,t} \\
    \vdots \\
    x_{N,t} \\
    y_{t}^Q
\end{bmatrix} =
\begin{bmatrix}
    Z_{1,1} & 0 & 0 & 0 \\
    Z_{2,1} & 0 & 0 & 0 \\
    Z_{3,1} & 0 & 0 & 0 \\
    \vdots & \vdots & \vdots & \vdots \\
    Z_{N,1} & 0 & 0 & 0 \\
    0 & 0 & 0 & 1
\end{bmatrix} \begin{bmatrix}
    f_{1,t} \\
    \hat{y}_{t-1}^* \\
    \hat{y}_{t-2}^* \\
    \hat{y}_t^Q
\end{bmatrix} +
\begin{bmatrix}
    u_{1,t} \\
    u_{2,t} \\
    u_{3,t} \\
    \vdots \\
    u_{N,t} \\
    \varepsilon_t^Q
\end{bmatrix}$$  \hspace{1cm} (18)
\[
\begin{bmatrix}
    f_{1,t} \\
    \hat{y}_{t-1} \\
    \hat{y}_{t-2} \\
    \hat{y}_{t-3} \\
    \hat{y}_{t-4}
\end{bmatrix}
= 
\begin{bmatrix}
    A_1 & 0 & 0 & 0 \\
    \beta_1 & 0 & 0 & 0 \\
    0 & 1 & 0 & 0 \\
    \beta_1 & 1 & 1 & 0 \\
    \beta_1 & 1 & 1 & 0
\end{bmatrix}
\begin{bmatrix}
    f_{1,t-1} \\
    \hat{y}_{t-2} \\
    \hat{y}_{t-3} \\
    \hat{y}_{t-4}
\end{bmatrix}
+ 
\begin{bmatrix}
    v_1 \\
    0 \\
    0 \\
    0 \\
    0
\end{bmatrix}
\]  
(19)

The measurement vector \( y_t' = (x_{t,1}, x_{z,t}, x_{3,t}, \ldots, x_{N,t}, y_{tQ}) \) includes \( N \) monthly variables and quarterly GDP growth rate. After obtaining the state space equations, we first estimate the model parameter \( \theta = (z_{1,1}, \ldots, z_{N,1}, A_1, \beta_1, \Sigma_u, \Sigma_e) \). Second, we can acquire the latest state vector \( x_t' \) by kalman filter or smoother and then substitute it into Equation (18). Finally, we could get the latest quarterly GDP growth \( y_{tQ} \) from the new observation vector. Unlike MIDAS, which is a direct multiple step forecasting methodology, state-space DFM produces iterative predictors. In this paper, we denote the dynamic factor model in a state space representation as SS-DFM.

### 3.3.5 Estimation and Missing value

The specifying state space model (18) and (19) can be estimated via expectation maximization (EM) algorithm or maximum likelihood (ML) method. Because of the nature of quarterly GDP, \( y_{tQ} \), the first two months of each quarter are skipped. In the empirical study of Mariano and Murasawa (2003, 2007), they replace all the missing observations with zeros and successfully realize this technique by some standard NID random variables. In this paper, we make the state space equations more specify and set all the missing values to be NAs. Then we can employ the powerful Kalman filter or smoother to estimate the model parameters \( \theta = (z_{1,1}, \ldots, z_{N,1}, A_1, \beta_1, \Sigma_u, \Sigma_e) \) via the algorithms mentioned above.

### 4 Evaluation of the forecast accuracy

When assessing and comparing two models, mean squared prediction errors (MSPEs) or mean absolute prediction errors (MAPEs) often can be used as a means of measurement. Since MSPEs is known to be more sensitive and strict to outliers than MAPEs. In our paper, we choose MSPEs as our measurement scale:

\[
MSPE = \frac{1}{T} \sum_{t=1}^{T} (\hat{y}_t - y_t)^2
\]  
(20)

Also, we want to test whether the discrepancies of the two competing models in
predictive accuracy are statistically significant. We wish to test the null hypothesis:

$$H_0: \sigma_{1,t+h}^2 - \sigma_{2,t+h}^2 = 0,$$

where $\sigma_{1,t}^2$ and $\sigma_{2,t}^2$ denote the second moments of $e_{1,t+h}$ and $e_{2,t+h}$. Here we should notice that $e_{1,t,h}$ and $e_{2,t,h}$ are the $h$ step ahead forecasting errors from two competitive models. According to the bootstrap simulation evidence in Clark and McCracken (2005), one-sided test is a better choice because the usual two-sided alternative hypothesis leads to poor power results. Thus, the alternative to $H_0$ is

$$H_A: \sigma_{1,t+h}^2 - \sigma_{2,t+h}^2 < 0.$$

In our comparisons, state space dynamic factor model and multiple MIDAS are non-nested models since the two systems operate quite differently. Henceforth, we employ the DM t-test proposed by Diebold and Mariano (1995) that $E\left[\sigma_{1,t}^2 - \sigma_{2,t}^2\right]$ obeys a standardized normal distribution under the null hypothesis. In general, we prefer parsimonious models to larger models when both have similar predictive power. Thus, the results of DM-test can be a valuable reference when we compare forecast evaluation. Interestingly, although DM-test is pretty useful in evaluating predictive accuracy, Deibold (2012) emphasizes that DM-test is intended in comparing forecasts, not models.

5 Empirical applications

In this part, two competing models cited in the preceding sections will be applied to forecast Swedish quarterly real GDP growth rate. In subsection 5.1, we describe the dataset. In subsection 5.2, the design of empirical comparison will be illustrated. Results and discussion will be presented in subsection 5.3.

5.1 Data

Our data is composed of quarterly Swedish real GDP and the six monthly coincident indicators. The six monthly indexes are Money Supply 3, 3-month treasury bills, 2-year government bonds, retail sale index, OMXS 30, and industrial production index. The exercise time span is from 1996M1 to 2012M12 and all the series are raw. As the last few years of the recessions do not seem to have common reasons, moreover, the performance of disparate indicators differs a lot for financial crisis (see, for instance,
Stock and Watson (2003)), the data is divided into in-sample (1996M1 to 2010M12) and out-of-sample (2011M1 to 2012M12). The in-sample data is adopted to estimate objective models while the out-of-sample data is used for forecasting evaluation. Since our forecast target is Swedish year on year GDP growth rate, we prefer to transform the candidate dependent variables into a year on year growth ($\Delta_4 \ln$), under the circumstance that it is stationary or first-difference stationary. More description about the dataset can be found in Table 1.

5.2 Design of comparison experiments

The principle idea of the exercises is to evaluate and compare the relative usefulness of MIDAS to state space DFM on the issue of predicting Swedish real GDP growth. In this situation, a benchmark model is helpful and necessary. There exists several benchmark models that can be used in a GDP growth forecasting context, for example, an ARIMA model is always helpful. Also, a random walk (RW) might work well when acting as a univariate benchmark. And an autoregressive distributed lag model (ADL) is also popular with many forecasters when they predict business cycles. In our experiments, we mainly use the ARIMA as benchmark.

To understand how real-time is operated in the forecast experiments, we first consider single-indicator MIDAS. The in-sample data set includes data from 1996Q1 up to 2010Q4. A 1-quarter-ahead prediction of 2011Q1 is generated from regressing real output up to 2010Q4 on monthly explanatory variable up to 2010M9. Then the monthly information updating to 2010M12 are extracted and we can forecast 2011Q1 under the help of estimated coefficients. Higher step predictions can be constructed in a similar way by recomputed every recursion. For the chosen of maximum number of lags of MIDAS, Kuzin, Marcellino, and Schumacher (2009) make $K=4$ while Clements and Galvão (2008) set $K=24$ in their applications. In our experiments, we collect the information of the last twelve months by setting $K$ equals to 12. For weight function, we choose exponential Almon lag rather than Beta weight lag although there is no difference in specification between these two weight functions according to Klaus Wohlrabe (2008). In addition, we do not restrict the two parameters $\theta_1$ and $\theta_2$ in exponential Almon lag weight. For state space DFM, we set the initial measurement vector from 1996M1 to 2010M12 which contains the six monthly
indexes and the quarterly real GDP growth rate. To make all the data at a monthly interval, we disaggregate lower-frequency output growth by setting the first two month of each quarter equal to NAs. Then we apply the state space system (Equation (18) and (19)) to get the estimated parameters. Meanwhile, we employ Kalman smoother to get the state vector of 2010M12. Finally, we substitute the state vector of 2010M12 into the observation equation repeatedly and we can get the estimated quarterly GDP growth for each three iterations.

For simplicity and clarity to the readers, we introduce basic idea and procedure of the application. Our experiment is mainly divided into three phases:

- First, we extract information from the six monthly indicators and estimate univariate MIDAS GDP forecasting models. Each univariate MIDAS is compared with our ARIMA benchmarks. The objective of this part is to check out whether we gain more prediction accuracy by means of using MIDAS rather than aggregating and averaging the monthly index to acquire a quarterly lower frequency.

- Second, we put in the autoregressive term and nest AR-MIDAS to see whether the forecasting accuracy of Swedish GDP growth improves.

- Last but not least, we fit the dynamic factor model in a state space form. Since state space system employs the powerful Kalman smoother and take advantage of all the six monthly indicators, to make the competition fair, we estimate a competitive Multivariate- MIDAS without any monthly information loss.

In detail, we compute MSPEs of each multi-step forecasts. MIDAS is estimated directly while state space DFM provides iterative predictions. All procedures adopted in this survey were written in R. For MIDAS, a MIDASR package was applied while a MARSS package for state space dynamic factor model.

5.3 Results

From Figure 1, we can see the general trends of Swedish quarterly real outputs growth. We can see obvious fluctuates due to the financial recession in 2008. This is a signal that we must be careful because generally most models have poor performance in
financial degeneration. Figure 1 also shows the forecasts of our benchmark, ARIMA (0, 0, 3), which is chosen by Akaike information criterion and Schwarz information criterion from Table 3. Figure 2-7 show the general real time movement of the six explanatory variables. Although the six chosen variables are monthly while GDP growth is at a lower quarterly interval, we can find some indicators exhibit synchronized appearance of GDP. For example, as can be seen in Figure 7, real GDP growth and IPI growth displays almost the same level of synchronized. In contrast, some financial index like OMXS30 growth appears to be irregular compared to real output growth.

In Table 4, we present MSPEs of the six single-indicator MIDAS models and the relative MSPEs, which is defined as MSPEs of MIDAS divided by the MSPEs of benchmark predictions. For each univariate MIDAS, the forecast horizon is constructed up to 12 months. It is obvious that, over the same horizon, the performance of the univariate MIDAS differs a lot between disparate indexes. For example, the MIDAS extracting M3 performs so unsatisfactorily that even our benchmark can surpass it breezily. In contrast, industrial production index acts excellently and it can be ranked in the list of best models. This is mainly because, as we cited above, the time series plot of IPI and GDP growth synchronize perfectly which have been an agreement by substantial materials. In addition, we can also find a valuable phenomenon that for some explanatory variables, MIDAS predictions within a monthly horizon outperforms those with a quarterly horizon. (e.g., TB3M for h=2/3 and h=1, IPI for h=1/3 and h=1) This might be a good suggestion that we can take advantage of the intra quarter monthly information which is more coincident with the explained variables. Anyway, no evident shows single-indicator MIDAS is that much more forecast capable than the quarterly benchmark MA model.

What is exciting is that, as can be seen from Table 5, MIDAS regressions beat the benchmark completely after adding an autoregressive term. That makes sense since AR-MIDAS is originally designed as an enhanced version of standard MIDAS in order to capture additional insights. An AR term considers autoregressive behaviour of the target series and makes the dynamic construction more flexible. As far as the relative performance of AR-MIDAS to MIDAS is concerned, in the third part of Table 5, most ratios MSPEs are smaller than 1 except the best index, IPI. Our
univariate MIDAS that exploits the information from IPI performs even better than the AR-MIDAS.

So far, all the prediction exercises presented above are based on the simple mixed data sampling regression model (MIDAS) and its autoregressive extension. And now it all comes down to a final faceoff between the two leading roles, M-MIDAS and SS-DFM, both of which exploit and extract the information from all the six indicators. The model specification is conducted as in Section 3. In Table 6, the MSPEs for horizons 1 to 4 for the M-MIDAS and SS-DFM are presented. We can find that sizeable gains is yielded from both two models. However, we should maintain clam since it is unfair and unbalanced to compare this two kind of models to the benchmark, even to MIDAS. Both two models are reasonable to have better accuracy because they grasp plenty of information from all the indicators. In Table 7, we can see how M-MIDAS versus SS-DFM by the relative performance. As can be seen, M-MIDAS dominates in short horzons (h=1, h=2) while SS-DFM takes over in longer forecast horizons (h=3, h=4). The ratios of MSPEs becomes lager reflects the truth that the predictive capability of M-MIDAS degenerates while SS-DFM tends to be stable as the forecast horizon enlarges. The second part of Table 7 shows the P value of the Diebold-Mariano test. For M-MIDAS with horizon h=1 to h=4, almost all the results of DM test are significantly at the level of a=0.2. This is a good sign because it implies that the forecast accuracy of M-MIDAS is higher than our benchmark. Similarly, SS-DFM has better predictive capability than the benchmark except for horizon h=1. Finally, we can find that for horizon h=1, the forecast evaluation of M-MIDAS is obviously superior to SS-DFM since P value is only 0.082. For longer horizon, we can not reject the null hyphothesis that the forecast accuracy of M-MIDAS is equal to SS-DFM. Also, we emphasis again that DM test is just used to compare forecasts, not models. Thus, neither of M-MIDAS and SS-DFM win a perfect victory.
6 Conclusion

This paper surveys a few general methods and examines different models for mixed frequency issues. The literature reveals that MIDAS is a simpler and parsimonious equation while state space models consisting of a system are prone to suffer more from parameter proliferation. Theoretical retrospects also indicate that one approach can never win completely, for example, Kuzin, Marcellino and Schumacher (2009), they compare the MIDAS, AR-MIDAS and a set of mixed frequency VAR(2) models written in a state space form while Mariano and Murasawa (2007) apply VAR and factor models for predicting monthly real GDP. It is unquestionable that there exist a number of similar empirical cases. However, to the best of our knowledge, none/few of them focus on M-MIDAS versus SS-DFM which is a fair and balance comparison. Hence, we compare the two models in forecasting Swedish output growth with a set of six monthly indicators.

The main finding is the following.

1. Single-index MIDAS is valuable when the explanatory variable is coincident with the target series. If we want to use the simplest model to gain considerable forecast accuracy, a univariate MIDAS is a good suggestion.

2. AR term enables MIDAS more promising especially under the circumstance that the indicator does not exhibit highly synchronized appearance of the target variable.

3. As a fair competitor to SS-DFM, M-MIDAS dominates undoubtedly at short horizons up to 6 months, whereas SS-DFM is more reliable at long predictive horizons. Although both of the two models defy the rule of parsimony, they are worth trying since sizeable forecast accuracy is gained in return.

In conclusion, our appraisal experiments proof that the approaches by exploiting the indicators which are available at a month frequency and using them directly rather than quarterly aggregating indeed facilitate the forecast accuracy. Both of the two models, MIDAS and SS-DFM, can dominate in a special situation.
References:


## Appendix

### Table 1: Description of Predictors/Indicators

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Description</th>
<th>Transformation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quarterly</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GDP</td>
<td>real GDP (SEK million)</td>
<td>$\Delta_4 \ln$</td>
</tr>
<tr>
<td>Monthly</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M3</td>
<td>real Money Supply 3</td>
<td>$\Delta_{12} \ln$</td>
</tr>
<tr>
<td>TB3M</td>
<td>3-month treasury bills</td>
<td>$\Delta \ln$</td>
</tr>
<tr>
<td>GB2Y</td>
<td>2-year government bonds</td>
<td>$\Delta \ln$</td>
</tr>
<tr>
<td>RSI</td>
<td>retail sale index</td>
<td>$\Delta_{12} \ln$</td>
</tr>
<tr>
<td>OMXS 30</td>
<td>OMX Stockholm 30 index</td>
<td>$\Delta \ln$</td>
</tr>
<tr>
<td>IPI</td>
<td>industrial production index</td>
<td>$\Delta_{12} \ln$</td>
</tr>
</tbody>
</table>

Notes: Swedish GDP, M3, retail sale index and industrial production index are available from Statistics Sweden's website (http://www.scb.se). 3-month treasury bills and 2-year government bonds comes from Sweden Riksbank website (http://www.riksbank.se). OMXS 30 can be found on Nasdaq OMX Nordic website (http://www.nasdaqomxnordic.com/).
**Table 2: Summary Statistics**

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Mean</th>
<th>S.D.</th>
<th>Min.</th>
<th>Max.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Quarterly</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GDP</td>
<td>0.02540</td>
<td>0.02870677</td>
<td>-0.06986</td>
<td>0.08027</td>
</tr>
<tr>
<td><strong>Monthly</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M3</td>
<td>0.07025</td>
<td>0.05705001</td>
<td>-0.02407</td>
<td>0.22352</td>
</tr>
<tr>
<td>TB3M</td>
<td>-0.010507</td>
<td>0.1257344</td>
<td>-0.762323</td>
<td>0.4496482</td>
</tr>
<tr>
<td>GB2Y</td>
<td>-0.011996</td>
<td>0.1013759</td>
<td>-0.337894</td>
<td>0.683436</td>
</tr>
<tr>
<td>RSI</td>
<td>0.03747</td>
<td>0.03042181</td>
<td>-0.07085</td>
<td>0.12604</td>
</tr>
<tr>
<td>OMXS 30</td>
<td>0.005847</td>
<td>0.05334825</td>
<td>-0.211484</td>
<td>0.137246</td>
</tr>
<tr>
<td>IPI</td>
<td>0.01901</td>
<td>0.08818266</td>
<td>-0.32950</td>
<td>0.25118</td>
</tr>
</tbody>
</table>

**Notes:** All the statistics are converted into the forms as being shown in Table 1 before summary.
Table 3: Model Selection of GDP growth (benchmark ARIMA Model)

The table presents the benchmark model selection. An MA (3) is chosen finally according to AIC and BIC selection criterion. Also, the model selection follows the principle of parsimony.

<table>
<thead>
<tr>
<th>Model</th>
<th>Log-likelihood</th>
<th>AIC</th>
<th>BIC</th>
<th>P-value of Ljung-Box test</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0, 0, 1)</td>
<td>170.39</td>
<td>-334.79</td>
<td>-328.13</td>
<td>0.0005702</td>
</tr>
<tr>
<td>(0, 0, 2)</td>
<td>173.36</td>
<td>-338.72</td>
<td>-329.84</td>
<td>0.0006717</td>
</tr>
<tr>
<td>(0, 0, 3)</td>
<td>193.66</td>
<td>-377.33</td>
<td>366.23</td>
<td>0.4424***</td>
</tr>
<tr>
<td>(0, 0, 4)</td>
<td>194.58</td>
<td>-377.16</td>
<td>-363.84</td>
<td>0.5946**</td>
</tr>
<tr>
<td>(1, 0, 0)</td>
<td>178.87</td>
<td>-351.75</td>
<td>-345.09</td>
<td>0.001907</td>
</tr>
<tr>
<td>(1, 0, 1)</td>
<td>179.31</td>
<td>-350.62</td>
<td>-341.74</td>
<td>0.0006989</td>
</tr>
<tr>
<td>(1, 0, 2)</td>
<td>186.39</td>
<td>-362.78</td>
<td>-351.68</td>
<td>0.3388</td>
</tr>
<tr>
<td>(1, 0, 3)</td>
<td>194.84</td>
<td>-377.68</td>
<td>-364.37</td>
<td>0.5614***</td>
</tr>
<tr>
<td>(1, 0, 4)</td>
<td>195.07</td>
<td>-376.15</td>
<td>-360.61</td>
<td>0.6441**</td>
</tr>
<tr>
<td>(2, 0, 0)</td>
<td>179.65</td>
<td>-351.3</td>
<td>-342.42</td>
<td>0.0002073</td>
</tr>
<tr>
<td>(2, 0, 1)</td>
<td>182.47</td>
<td>-354.94</td>
<td>-343.85</td>
<td>0.00509</td>
</tr>
<tr>
<td>(2, 0, 2)</td>
<td>189.23</td>
<td>-366.46</td>
<td>-353.14</td>
<td>0.3421</td>
</tr>
<tr>
<td>(2, 0, 3)</td>
<td>195.34</td>
<td>-376.68</td>
<td>-361.14</td>
<td>0.7462*</td>
</tr>
<tr>
<td>(2, 0, 4)</td>
<td>195.39</td>
<td>-374.79</td>
<td>-357.03</td>
<td>0.724</td>
</tr>
<tr>
<td>(3, 0, 0)</td>
<td>183.76</td>
<td>-357.51</td>
<td>-346.41</td>
<td>0.0008864</td>
</tr>
<tr>
<td>(3, 0, 1)</td>
<td>185.68</td>
<td>-359.37</td>
<td>-346.05</td>
<td>0.01632</td>
</tr>
<tr>
<td>(3, 0, 2)</td>
<td>189.62</td>
<td>-365.24</td>
<td>-349.7</td>
<td>0.2614</td>
</tr>
<tr>
<td>(3, 0, 3)</td>
<td>195.46</td>
<td>-374.93</td>
<td>-357.17</td>
<td>0.7027*</td>
</tr>
<tr>
<td>(3, 0, 4)</td>
<td>197.14</td>
<td>-376.28</td>
<td>-356.3</td>
<td>0.8658*</td>
</tr>
<tr>
<td>(4, 0, 0)</td>
<td>188.67</td>
<td>-365.34</td>
<td>-352.02</td>
<td>0.206</td>
</tr>
<tr>
<td>(4, 0, 1)</td>
<td>194.36</td>
<td>-374.73</td>
<td>-359.19</td>
<td>0.4615</td>
</tr>
<tr>
<td>(4, 0, 2)</td>
<td>192.72</td>
<td>-369.44</td>
<td>-351.69</td>
<td>0.2446</td>
</tr>
<tr>
<td>(4, 0, 3)</td>
<td>193.32</td>
<td>-368.64</td>
<td>-348.66</td>
<td>0.2181</td>
</tr>
<tr>
<td>(4, 0, 4)</td>
<td>197.14</td>
<td>-374.29</td>
<td>-352.09</td>
<td>0.8726</td>
</tr>
</tbody>
</table>
Table 4: Forecasting Swedish GDP growth: Univariate MIDAS

The table shows MSPEs of MIDAS and relative MSPEs, which is computed by the MSPEs of MIDAS divided by the MSPES of the benchmark. It is notable that forecast horizon $h = 1/3$ means 2/3 of the information of the current quarter is extracted while $h = 1$ indicates only the information update to the last quarter is adopted. For ratios MSPEs, the benchmark MA for horizon in $h = 1/3$, $h = 2/3$ and $h = 1$ are the same. Higher step relative MSPEs can are constructed in a similar way.

<table>
<thead>
<tr>
<th>Indicator</th>
<th>M3</th>
<th>TB3M</th>
<th>GB2Y</th>
<th>RSI</th>
<th>OMXS30</th>
<th>IPI</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MIDAS : MSPEs</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$h = 1/3$</td>
<td>0.0003787</td>
<td>0.0002268</td>
<td>0.0002483</td>
<td>0.0004810</td>
<td>0.0002886</td>
<td>0.0001453</td>
</tr>
<tr>
<td>$h = 2/3$</td>
<td>0.0003867</td>
<td>0.0001478</td>
<td>0.0001648</td>
<td>0.0004273</td>
<td>0.0002474</td>
<td>0.0001377</td>
</tr>
<tr>
<td>$h = 1$</td>
<td>0.0003700</td>
<td>0.0002270</td>
<td>0.0002582</td>
<td>0.0004638</td>
<td>0.0002546</td>
<td>0.0001521</td>
</tr>
<tr>
<td><strong>Ratio MSPEs: MIDAS/MA</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$h = 1/3$</td>
<td>1.614</td>
<td>0.967</td>
<td>1.058</td>
<td>2.050</td>
<td>1.230</td>
<td>0.620</td>
</tr>
<tr>
<td>$h = 2/3$</td>
<td>1.649</td>
<td>0.630</td>
<td>0.703</td>
<td>1.821</td>
<td>1.055</td>
<td>0.587</td>
</tr>
<tr>
<td>$h = 1$</td>
<td>1.577</td>
<td>0.968</td>
<td>1.101</td>
<td>1.977</td>
<td>1.085</td>
<td>0.648</td>
</tr>
<tr>
<td>$h = 4/3$</td>
<td>1.282</td>
<td>0.880</td>
<td>0.871</td>
<td>0.967</td>
<td>0.544</td>
<td>0.993</td>
</tr>
<tr>
<td>$h = 5/3$</td>
<td>1.547</td>
<td>0.574</td>
<td>0.996</td>
<td>0.896</td>
<td>0.471</td>
<td>0.797</td>
</tr>
<tr>
<td>$h = 2$</td>
<td>1.541</td>
<td>0.776</td>
<td>1.284</td>
<td>1.530</td>
<td>0.536</td>
<td>0.524</td>
</tr>
<tr>
<td>$h = 7/3$</td>
<td>1.746</td>
<td>1.119</td>
<td>0.964</td>
<td>0.809</td>
<td>0.625</td>
<td>0.723</td>
</tr>
<tr>
<td>$h = 8/3$</td>
<td>1.892</td>
<td>1.129</td>
<td>1.192</td>
<td>0.731</td>
<td>0.579</td>
<td>0.117</td>
</tr>
<tr>
<td>$h = 3$</td>
<td>1.885</td>
<td>1.491</td>
<td>1.215</td>
<td>2.189</td>
<td>0.648</td>
<td>0.223</td>
</tr>
<tr>
<td>$h = 10/3$</td>
<td>2.032</td>
<td>0.880</td>
<td>0.741</td>
<td>1.622</td>
<td>0.731</td>
<td>0.085</td>
</tr>
<tr>
<td>$h = 11/3$</td>
<td>2.120</td>
<td>1.051</td>
<td>0.631</td>
<td>1.098</td>
<td>0.935</td>
<td>0.155</td>
</tr>
<tr>
<td>$h = 4$</td>
<td>2.059</td>
<td>0.802</td>
<td>0.251</td>
<td>2.269</td>
<td>0.560</td>
<td>0.156</td>
</tr>
</tbody>
</table>
Table 5: Forecasting Swedish GDP growth: AR-MIDAS

<table>
<thead>
<tr>
<th>Indicator</th>
<th>M3</th>
<th>TB3M</th>
<th>GB2Y</th>
<th>RSI</th>
<th>OMXS30</th>
<th>IPI</th>
</tr>
</thead>
<tbody>
<tr>
<td>h=1</td>
<td>0.0002223</td>
<td>0.0001158</td>
<td>0.0001748</td>
<td>0.0001040</td>
<td>0.0001075</td>
<td>0.0001554</td>
</tr>
<tr>
<td>h=2</td>
<td>0.0002334</td>
<td>0.0001879</td>
<td>0.0001036</td>
<td>0.0001519</td>
<td>0.0001669</td>
<td>0.0001130</td>
</tr>
<tr>
<td>h=3</td>
<td>0.0001929</td>
<td>0.0002048</td>
<td>0.0002504</td>
<td>0.0002233</td>
<td>0.0001390</td>
<td>0.0001255</td>
</tr>
<tr>
<td>h=4</td>
<td>0.0002265</td>
<td>0.0001203</td>
<td>0.0002762</td>
<td>0.0002722</td>
<td>0.0001815</td>
<td>0.0001218</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ratios MSPEs: AR-MIDAS/MA</th>
</tr>
</thead>
<tbody>
<tr>
<td>h=1</td>
</tr>
<tr>
<td>h=2</td>
</tr>
<tr>
<td>h=3</td>
</tr>
<tr>
<td>h=4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ratios MSPEs: AR-MIDAS/MIDAS</th>
</tr>
</thead>
<tbody>
<tr>
<td>h=1</td>
</tr>
<tr>
<td>h=2</td>
</tr>
<tr>
<td>h=3</td>
</tr>
<tr>
<td>h=4</td>
</tr>
</tbody>
</table>
Table 6: Forecasting Swedish GDP growth: MSPEs of M-MIDAS vs SS-DFM

The table presents MSPEs for the multivariate MIDAS and state space dynamic factor models. Both two models extract the information from all the six indicators.

<table>
<thead>
<tr>
<th>Horizon\Model</th>
<th>M-MIDAS</th>
<th>SS-DFM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MSPEs</td>
<td></td>
</tr>
<tr>
<td>h=1</td>
<td>0.0000381</td>
<td>0.0002043</td>
</tr>
<tr>
<td>h=2</td>
<td>0.0000765</td>
<td>0.0001315</td>
</tr>
<tr>
<td>h=3</td>
<td>0.0001768</td>
<td>0.0001410</td>
</tr>
<tr>
<td>h=4</td>
<td>0.0001771</td>
<td>0.0001208</td>
</tr>
</tbody>
</table>

Table 7: Forecasting Swedish GDP growth:

Relative performance of M-MIDAS vs SS-DFM

The table presents ratios MSPEs for multivariate MIDAS and state space dynamic factor models. The second part of the table shows the p value of DM t-test, the null hypothesis is that Model 1 and Model 2 have equal forecast accuracy while the alternative is the forecast accuracy of Model 1 is less than those of Model 2.

<table>
<thead>
<tr>
<th>Horizon\Model</th>
<th>M-MIDAS/MA</th>
<th>SS-DFM/MA</th>
<th>M-MIDAS/ SS-DFM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ratios MSPEs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>h=1</td>
<td>0.1625204</td>
<td>0.8708619</td>
<td>0.1866202</td>
</tr>
<tr>
<td>h=2</td>
<td>0.2968430</td>
<td>0.5103637</td>
<td>0.5816302</td>
</tr>
<tr>
<td>h=3</td>
<td>0.6440848</td>
<td>0.5136480</td>
<td>1.2539420</td>
</tr>
<tr>
<td>h=4</td>
<td>0.5408519</td>
<td>0.3687471</td>
<td>1.4667288</td>
</tr>
<tr>
<td></td>
<td>P value of Diebold-Mariano Test</td>
<td></td>
<td></td>
</tr>
<tr>
<td>h=1</td>
<td>0.054</td>
<td>0.400</td>
<td>0.082</td>
</tr>
<tr>
<td>h=2</td>
<td>0.072</td>
<td>0.144</td>
<td>0.314</td>
</tr>
<tr>
<td>h=3</td>
<td>0.207</td>
<td>0.139</td>
<td>0.622</td>
</tr>
<tr>
<td>h=4</td>
<td>0.125</td>
<td>0.069</td>
<td>0.673</td>
</tr>
</tbody>
</table>
Figure 1: Quarterly Swedish GDP growth and forecasts of M-MIDAS, SS-DFM and benchmark MA (3).

Figure 2: Monthly Swedish M3 growth.

Figure 3: Monthly Swedish TB3M growth.
Figure 4: Monthly Swedish GM2Y growth.

Figure 5: Monthly Swedish RSI growth.

Figure 6: Monthly Swedish OMXS30 growth.
Figure 7: Monthly Swedish IPI growth.