Using the MIDAS approach for now- and forecasting Colombian GDP

Master Thesis
Econometrics

Author:
Gabriel Camilo Pérez Castañeda

Supervisor:
Prof. Dr. Dick van Dijk

May 11, 2009

MSc in Econometrics and Management Science
Erasmus School of Economics
Erasmus Universiteit Rotterdam

Abstract

This study applies the Factor - MIDAS approach (Marcellino and Schumacher, 2007) in the forecasting of Colombian GDP. The main objective is to test the performance of the predictions generated under this framework by means of Mean Squared Error values and forecast evaluation tests. Two forms of MIDAS (Mixed Data Sampling) projections were studied, MIDAS with exponential almon and MIDAS with unrestricted coefficients. Also, two methods for factor were used, one based on the EM algorithm and the other based on the state-space model with the Kalman Filter. Both methods are able to handle missing values at the end of the sample due lags of publication. In addition, the factors were calculated using a large dataset of macroeconomic variables and a subset of it. The regressions were estimated using fixed factor lags along with an automatic lag selection. The nowcast and forecast performance of these regressions were compared with a simple benchmark model AR(1) model. The empirical findings show in general, that the MIDAS projections do not outperform the benchmark when the forecast tests are applied. There is only slight evidence that the MIDAS projections do better in the nowcast horizon. In terms of lower Mean Squared Error values, the better results are achieved when the number of factor lags is at most 3. Moreover, in this case there is no difference in the performance of these two projections. The automatic factor lag selection did not show any improvement compared to the use of very few fixed factor lags.

Keywords: MIDAS projections, factor models, ragged edge data, Kalman Filter, EM-algorithm, nowcasting

Acknowledgements

I want to thank my supervisor Prof. Dr. Dick van Dick for his patience and useful advices. Also, I would like to give a special thank María Fernanda Cortés for her advices with the insights of the Colombian macroeconomic series and Andrés Perdomo at Departamento Nacional de Planeación (National Planning Bureau) for kindly provide the data for this research.

To my parents, sister and all the people who made possible for me to have more than a second chance to live this life.

Contents

List of Tables

Introduction

In every country or region in the world it is of vital importance for economic agents such as policy makers and investors, to have more certainty about how the economy is going to behave in the near and long future. Central Banks and various local and global agencies give particular attention to the forecasts of key indicators of macroeconomic activity, which are crucial to support further policy decisions. The frequency of publication for some of these indicators is lower compared to other macroeconomic variables and also has delays. That is the case of the Gross Domestic Output (GDP). The GDP is usually calculated with quarterly frequency and has a release lag of several weeks after the end of the period. These issues have motivated researchers to develop techniques that first, take in account data with higher frequency, which can improve the predictions of the current (Nowcast) and next (Forecast) periods and second, deal with the problem of "ragged-edge" data (Marcellino and Schumacher, 2007), which refers to unbalanced data samples due different publication dates for all the variables.

A recent methodology that deals with the different frequencies present in macroeconomic data was developed by Ghysels et. al. (2004), where the standard single frequency regression models are extended to incorporate data with low and high sampling frequencies. This scheme is denominated the Mixed-data sampling (MIDAS) approach. In a following study, Ghysels et. al.(2007) extend this approach to a general linear case and other case that includes nonlinearities as well.

Another relevant aspect in the literature regarding forecasting models, is the use of use of Dynamic Factor Models instead of the series themselves to predict key macroeconomic indicators. Since the work of Sargent and Sims (1977), there has been a great interest in the development of methodologies that obtain latent processes that are common components of macroeconomic datasets. These processes represent a gain in efficiency since they are able capture most of the variance and underlying dynamics of the whole set of series with a lower dimension. Usually, these components can be seen as the so-called *state of the economy*. Following this direction, Stock and Watson (2002) established a methodology that uses Principal Components Analysis (PCA) along with the Expectation-Maximization (EM) algorithm to estimate factors from macroeconomic series, which is able to deal with the ragged edge data mentioned above. Other methodologies, such as the one in Doz et. al. (2006) are based in different estimation frameworks such as the state-space model to calculate these factors in presence of ragged data as well.

In order to have a better forecast model, Marcellino and Schumacher (2007) combine the MIDAS projections with factors obtained using Dynamic Factor Models creating a in a framework denoted as Factor-MIDAS approach, which is simply high frequency factors as regressors in the MIDAS model.

In the Colombian case, the literature about models for forecasting the GDP is scarce. In the study made by Castro (2003a) a system of univariate and multivariate time series models to forecast Colombian GDP is shown. Nieto (1998) presented a methodology that estimates the higher frequency values of a certain variable as a latent process. This study is composed of two parts, first defines the ex-post estimation of unobserved processes and second, presents the ex-ante estimation when the lower frequency data is it not available yet. An application with the Colombian Gross National Product (GNP) is made in this study. In a following article Nieto (2007) extended the methodology in a multivariate setting using structural time series models. In this study an application to the Colombian industrial GDP is carried out. Other studies have focused on the construction of monthly coincident and leading indexes for the state of the economy which are the case of Nieto and Melo (2001), Nieto et al (2003) and Castro (2003b).

Up till now there is an open path to explore more methodologies to forecast the Colombian GDP, more specifically using a framework that allows the use the MIDAS regressions and Dynamic Factor Models. The purpose of this research is to apply the Factor MIDAS approach to the Colombian case and test whether it shows more forecasting accuracy than a simple time series model. The remainder of this paper is organized as follows. Chapter 1 describes the data, both monthly and quarterly. Chapter 2 discusses the factor estimation methods and the MIDAS projections and gives the explanation of the forecast experiment. Chapter 3 presents the empirical results. Chapter 4 concludes.

Chapter 1

Data

1.1 Monthly dataset of macroeconomic series

Table A.5 contains the list of the monthly series used in this study divided by economic sector. This dataset contains most of the monthly indicators that were used in Castro (2003a), comprising 60 series in total. The data span starts from 1982:01 to 2008:03, being the one available on 2008:03. It is not a real-time data set and does not contain vintages of data as they are not officially available in the Colombian economy. For this study, only the data since 1991:01 it is going to be used.

Stock and Watson (2002) point out that the factor estimation procedure that is going to be explained below, requires that the series are seasonally adjusted and I(0). First the seasonal adjustment this dataset was made using TRAMO-SEATS. Then, three unit root tests are calculated for each series to check their order of integration. The tests used are ADF, KPSS and ERS, which are implemented in EVIEWS® and the selection of the lags was chosen to be automatic depending on information criteria. For each series the three tests show fairly the same integration order. For series where one of the tests differs from the other two the final order of integration was the one that repeated the most. The results are shown in the table A.6. Natural logarithm was applied to all the series except the interest rate series.

Figure 1.1 presents six of the monthly macroeconomic series that are going to be used in the study. Most of the monthly series show a trend break in 1999, where there was a deep recession in the country. Some series, like the interest rates and inflation, present structural changes due changes in monetary policy derived from the inflation targeting scheme that was implemented in 1995 (Gómez et. al. 2002 p6). These facts are the base for the forecasting scheme that is going to be applied, which will be explained below in section 2.3.

1.2 Quarterly GDP

In Colombia, the gross domestic product (GDP) is calculated and released by the National Department of Statistics (DANE). Before 1994 it was released on an annual basis, however, starting at 1994:Q1 it has been published with quarterly frequency. Since 2008:Q1 the DANE incorporated a new methodology for the calculation of GDP. Nonetheless, for this research the span of the GDP growth goes from 1995:Q1 until 2007:Q4. With respect to the lag of publication, actually the release of the GDP is 10 to 12 weeks after the end of the referenced quarter.

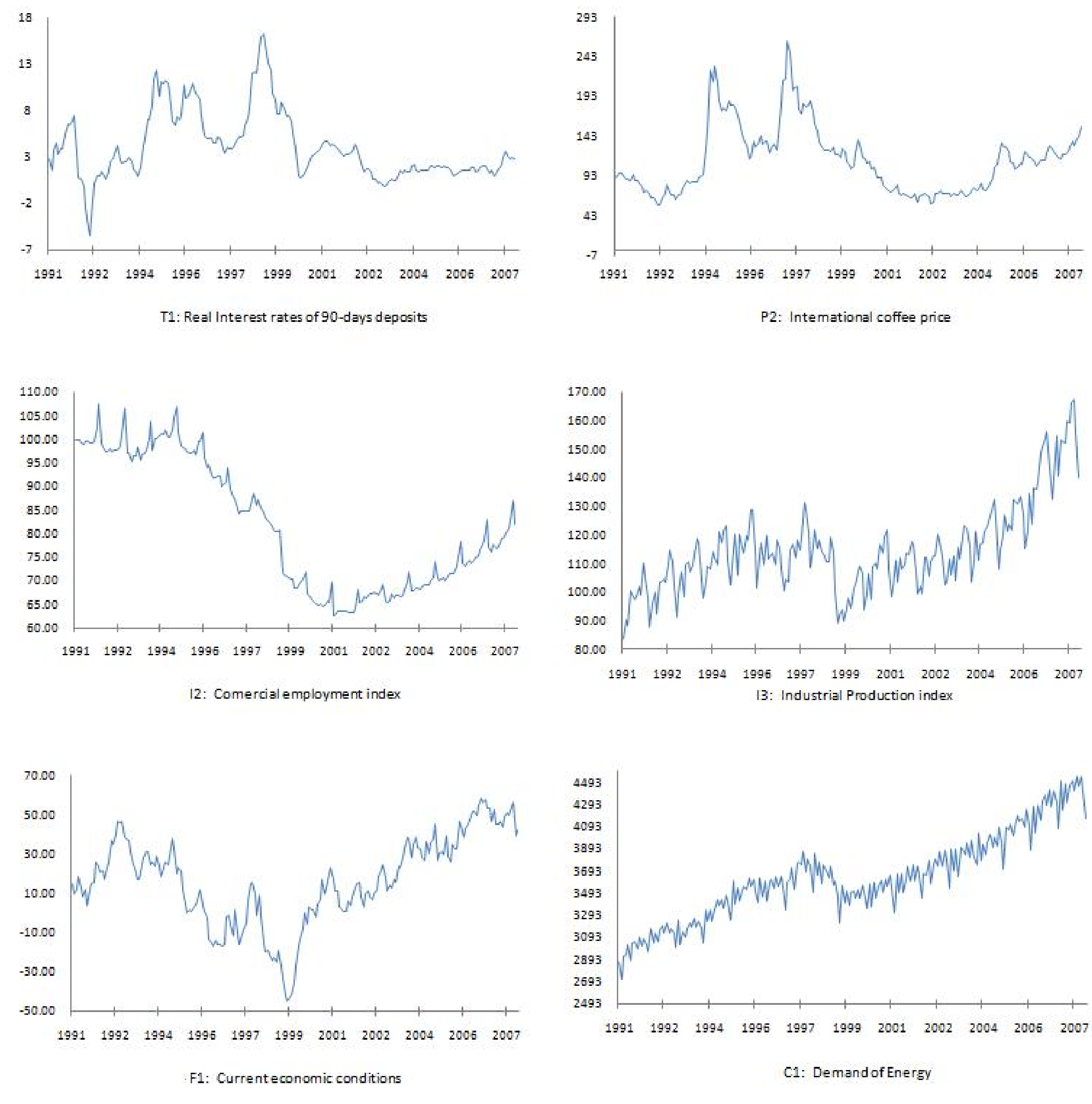

Figure 1.1: Six of the Colombian macroeconomic series included used for the estimation if factors.

Chapter 2

Econometric Methodology

2.1 Factor estimation with ragged-edge data

The objective of this section is to describe the factor estimation framework that is going to be used in this study. Marcellino and Schumacher (2007) denote GDP growth as y_{t_q} where t_q is the quarterly time index $t_q = 1, 2, \ldots, T_q$. It can also be expressed in a monthly frequency setting $y_{t_m} = y_{t_q}, \forall t_m = 3t_q$ with t_m as the monthly time index meaning that GDP y_{t_m} it is observed only at months $t_m = 3, 6, 9, \ldots, T_m$ with $T_m = 3T_q$. Using the information up to one specific month, the aim is to nowcast the current quarter GDP and to forecast it h_q quarters ahead, or $h_m = 3h_q$ months ahead using monthly factors. This prediction is denoted as $y_{T_m+h_m|T_m}$. In every month during the quarter of reference, a nowcast can be produced and forecasts for the next quarters can also be produced depending on the desired horizon.

On the other hand, the information set to create the factors has to contain stationary monthly indicators, which are comprised in a N-dimensional vector $\mathbf{X}_{t_m}$, where t_m denotes the frequency in months. In addition, some of the final observations might not be available at the end of the sample due to publication lags, leaving an unbalanced sample of $\mathbf{X}_{t_m}$. This is what has been called *ragged edge data*.

Assuming a structure on $\mathbf{X}_{t_m}$ the objective is to estimate few factors that hold almost the same information of it. Marcellino and Schumacher (2007) apply a two stage methodology that is based on earlier works on singe-frequency datasets. The first step is to estimate the factors using a technique that is capable to handle ragged-edge data, followed by a second step, where a method that can deal with mixed frequency data is applied to generate the forecast. In general, the scheme involves a low-frequency target variable that is augmented by high-frequency factors.

In order to estimate the factors the following structure is assumed:

$$\mathbf{X}_{t_m} = \Lambda F_{t_m} + \epsilon_{t_m} \tag{2.1}$$

where the r-dimensional factor is denoted as $\mathbf{F_{t_m}} = (f'_{1,t_m}, \ldots, f'_{r,t_m})'$. The loadings matrix Λ has dimensions $(N \times r)$. This matrix multiplied with the factors is the common component among the variables in $\mathbf{X}_{t_m}$.

To get the factors with ragged-edge data Marcellino and Schumacher (2007) describe three approaches based on the structure for the monthly data shown in (2.1). The first approach is to realign each time series in the sample in order to get a balanced dataset (Altissimo et al., 2006). The procedure is simply set $\widetilde{x}_{I,T_m} = x_{I,T_m - k_i}$, for $t_m = k_i + 1, \ldots, T_m$. This method is applied to all the series and gives a balanced data set $\widetilde{X}_{t_m}$ for $t_m = max(\{k_i\}_{i=1}^{N}) = 1, \ldots, T_m$. Having the balanced dataset, Altissimo et. al. (2006) propose to use dynamic PCA to estimate the factors. This procedure has the disadvantage that it affects the dynamic cross-correlations when the publication date for some series is not fixed due revisions. The dynamic PCA explained in Forni (2005) exploits the dynamic cross-correlations on the frequency domain and might be able to take these realignments into account. The next approach is the one considered in Stock and Watson (2002), where they use the EM and the PCA to obtain the factors when there are missing values. The third approach is followed by Doz et. al. (2006) and Kapetanios and Marcellino (2006) who express the factor model in the state space form and use quasi-ML to estimate the factors. In this research the Stock and Watson approach and the State - space approach are going to be applied.

2.1.1 Stock and Watson factor estimation

Stock and Watson (2002) define the statistical framework for Dynamic Factor Models. Let y_{t+1} denote the univariate time series to be forecasted and let X_t be a N -dimensional time series array of macroeconomic indicators, observed for $t = 1, ...T$, where y_t and X_t are both taken to have mean 0. The model used has the form

$$y_{t+1} = \beta' F_t + \gamma(L) y_t + \epsilon_{t+1}, \tag{2.2}$$

$$X_t = \Lambda F_t + e_t \tag{2.3}$$

where $F_t = (f'_t, \ldots, f'_{t-q},)\prime$ is $r \times 1$, where $r \leq (q+1)\bar{r}$, the ith row of Λ is $(\lambda_{i0}, \ldots, \lambda_{iq})$, and $\beta = (\beta_0, \ldots, \beta_q)$.

Stock and Watson focus on the multistep forecasting and rewrite the equation (2.2) as

$$y_{t+h}^{h} = \alpha + \beta_h(L)F_t + \gamma_h(L)y_t + \epsilon_{t+h}^{h}, \tag{2.4}$$

where y_{t+h}^{h} is the h-step ahead variable to be forecasted, the constant term is introduced explicitly, and the subscripts on the coefficients show the dependence of the projection on the forecast horizon.

Several points have to be addressed for this specification. In equations (2.2) - (2.3) the factors can be estimated using standard principal components. Stock and Watson (2002) advise that this specification is inconsistent with infinite distributed lags of the factors. However, they highlight the benefits in terms of forecasting accuracy in practical applications.

The estimation procedure is made in two steps, first, the sample data $\{X_t\}_{t=1}^{T}$ are used to estimate a time series of factors (the diffusion indexes), $\{\hat{F}_t\}_{t=1}^{T}$, then the estimators $\hat{\alpha}_h$, $\hat{\beta}_h(L)$ and $\hat{\gamma}_h(L)$ are obtained by regressing y_{t+1} onto a constant, $\hat{F}_t$ and y_t (and lags). The forecast of y_{T+h}^{h} is then formed as $\hat{\alpha}_h + \hat{\beta}_h(L)\hat{F}_T$ and $\hat{\gamma}_h(L)y_T$. The factors are estimated using principal components.

Stock and Watson consider missing values in the data for estimating factors using the Expectation-Maximization (EM) algorithm joint with the standard Principal Components Analysis (PCA). Let i be a variable from $\mathbf{X}_{t_m}$ which is a full data column vector of the observations until time T_m such that $\mathbf{X}_i = (x_{i,1}, \ldots, x_{i,T_m})'$. Now assume that there are some missing observations at the end of the sample due the ragged-edge data. Let $\mathbf{X}_i^{obs}$ be a vector that is a subset of $\mathbf{X}_i$ which takes into account only the observations that are available. A relation of these two vectors can be established in the following way by

$$\mathbf{X}_i^{obs} = \mathbf{A}_i\mathbf{X}_i, \tag{2.5}$$

where $\mathbf{A}_i$ is a matrix that handles missing values or missing frequencies. For example, in the case of missing values $\mathbf{A}_i$ is an identity matrix whose rows corresponding to the missing values are removed.

The EM procedure described in Marcellino and Schumacher (2007) is explained as follows. First, provide initial (naive) guesses for the missing values. This guesses along with the observed values provide a balance dataset $\hat{\mathbf{X}}^{(0)}$. Then standard PCA is applied to obtain the initial monthly factors $\hat{\mathbf{F}}^{(0)}$ and loadings $\hat{\mathbf{\Lambda}}^{(0)}$. The E-step is performed doing expectation of $\mathbf{X}_i$ conditional on $\mathbf{X}_i^{obs}$, $\hat{\mathbf{F}}^{(j-1)}$ and $\hat{\mathbf{\Lambda}}_i^{(j-1)}$. Basically it is an update from the preceding iteration

$$\hat{\mathbf{X}}_i^{(j)} = \hat{\mathbf{F}}^{(j-1)}\hat{\mathbf{\Lambda}}_i^{(j-1)} + \mathbf{A}_i'(\mathbf{A}_i\mathbf{A}_i')^{-1}(\mathbf{X}_i^{obs} - \mathbf{A}_i\hat{\mathbf{F}})^{(j-1)}\hat{\mathbf{\Lambda}}_i^{(j-1)}). \tag{2.6}$$

When this procedure is applied to all the variables i the result is a balanced dataset. Here is where the M-step starts which consists on reestimate the factors $\hat{\mathbf{F}}^{(j)}$ and loadings $\hat{\mathbf{\Lambda}}^{(j)}$ using PCA. Then the E-step is repeated and so on, until convergence is achieved. From now on this approach is going to be denoted as SW-EM.

2.1.2 State - Space factor estimation

Based on Doz et, al. (2006), Marcellino and Schumacher (2007) show how the factor model can be cast in a state-space form. The specification of this approach has the following form

$$\mathbf{X}_{t_m} = \mathbf{\Lambda}\mathbf{F}_{t_m} + \xi_{t_m} \tag{2.7}$$

$$\Psi(L_m)\mathbf{F}_{t_m} = \mathbf{B}\eta_{t_m} \tag{2.8}$$

where equation (2.7) is the static representation as it was shown in (2.1). Equation (2.8) describe the dynamic evolution of the factors using a Vector Autoregressive (VAR) representation with lag polynomial $\Psi(L_m) = \Sigma_{i=1}^{p}\Psi_i L_m^i$ being L_m^i the lag operator with $L_m^i x_{t_m} = x_{t_m - i}$. The vector η_{t_m} is q-dimensional and has the orthogonal dynamic shocks which drive the r factors, where $\mathbf{B}$ has dimensions $(r \times q)$. If the monthly lag p equals one, the model is already in state-space form. For lags higher than one, equation (2.8) has to be set into the VAR(1) form. Marcellino and Schumacher (2007) stress that for small dimensions of $\mathbf{X}_{t_m}$ the estimation of the model can be done using ML but for large datasets is infeasible. A quasi-ML estimation scheme is proposed by Doz et al. (2006). The matrices needed to specify the system are obtained by applying the following steps:

1. $\hat{\mathbf{F}}_{t_m}$ is estimated initially using PCA.
2. $\hat{\mathbf{\Lambda}}$ is calculated doing the regression of $\mathbf{X}_{t_m}$ on $\hat{\mathbf{F}}_{t_m}$.
3. $\hat{\mathbf{\Sigma}}_\xi$ is the covariance matrix of the errors $\hat{\xi}_{t_m} = \hat{\mathbf{\Lambda}}\hat{\mathbf{X}}_{t_m}$.
4. $\Psi(L_m)$ comes from estimating a VAR(p) on the factors $\hat{\mathbf{F}}_{t_m}$.
5. $\hat{\mathbf{\Sigma}}_\zeta$ is the residual covariance matrix of $\hat{\zeta} = \hat{\Psi}(L_m)\hat{\mathbf{F}}_{t_m}$ which follows from the estimation of the VAR(p)

6. $\hat{\mathbf{B}} = \mathbf{M}\mathbf{P}^{-1/2}$ is obtained by applying an eigenvalue decomposition on $\hat{\mathbf{\Sigma}}_\zeta$ where $\mathbf{M}$ is a $(r \times q)$ matrix that contains the q largest corresponding eigenvectors of the decomposition. $\mathbf{P}$ contains the largest q eigenvalues on the main diagonal.

The system now is cast into the state-space form (Harvey, 1991). Now, applying the Kalman filter it is possible to get new estimates for the monthly factors. Marcellino and Schumacher (2007) stress that the coefficients on the matrices have to be estimated from a balanced dataset, as it is needed in the first step. However, factor estimation using the Kalman filter applies to the unbalanced data and is suitable to the ragged edge problem.

The difference with the Stock and Watson approach is the incorporation of a dynamic structure inside the factor estimation. In the applications below this method will be denoted as KAL-SS.

2.2 MIDAS Approach

Ghysels et. al. (2004) and (2007) set the framework of the MIDAS regressions in the following way. Let y_t be a variable that is sampled on a quarterly interval and let $x_t^{(m)}$ be another variable that is sampled monthly. $x_t^{(m)}$ is observed three times in each period of y_t (m=3). The interest is on project y_t on the span of lagged observations of $x_{t-j/m}^{(m)}$. The superscript on $x_{t-j/m}^{(m)}$ denotes the higher sampling frequency where the lag is a fraction of the unit interval between $t-1$ and t.

$$y_t = \beta_0 + \beta_1 B\left(L^{(1/m)};\theta\right) x_t^{(m)} + \epsilon_t^{(m)} \tag{2.9}$$

for $t = 1, \ldots, T$ and where $B(L^{1/m};\theta) = \sum_{k=0}^{K} B(k;\theta) L^{k/m}$ and $L^{1/m}$ is a lag operator such that $L^{1/m} x_t^{(m)} = x_{t-1/m}^{(m)}$ and the lag coefficients in $B(k;\theta)$ of the corresponding lag operator $L^{1/m}$ are parameterized as a function of a small-dimensional vector of parameters θ.

A more general MIDAS model has the form

$$y_{t+1} = \beta_0 + f\Big(\sum_{i=1}^{K}\sum_{j=1}^{L} B_{ij}(L^{1/m_i}) g(x_t^{(m_i)})\Big) + \epsilon_{t+1} \tag{2.10}$$

Where the functions f and g can have unknown parameters. As estimation methods Ghysels et. al. (2004) consider maximum likelihood (MLE), non-linear squares (NLS) and generalized method of moments (GMM).

2.2.1 Basic Factor-MIDAS approach

Marcellino and Schumacher (2007), introduced the Factor-Midas approach, which is a generalization of the MIDAS regressions presented above using factors obtained from several macroeconomic variables instead of the variables themselves. Supposing there is only one factor $r = 1$ the forecast model for horizon h_q quarters with $h_q = h_m/3$ is

$$y_{t_q+h_q} = y_{t_m+h_m} = \beta_0 + \beta_1 b(L_m, \theta)\hat{f}_{t_m}^{(3)} + \epsilon_{t_m+h_m} \tag{2.11}$$

where the polynomial $b(L_m, \theta)$ is the exponential Almon lag with

$$b(L_m, \theta) = \sum_{k=0}^{K} c(k, \theta)L_m^k, c(k, \theta) = \frac{exp(\theta_1 k + \theta_2 k^2)}{\sum_{k=0}^{K} exp(\theta_1 k + \theta_2 k^2)}. \tag{2.12}$$

The factor $\hat{f}_{t_m}^{(3)}$ is a skip-sampled version of the monthly factor $\hat{f}_{t_m}$. The superscript (3) indicates that every third observation starting from the t_m-th is included in the regressor $\hat{f}_{t_m}^{(3)}$.

For the case of $r > 1$ with $\mathbf{F}'_{t_m} = (f'_{1,t_m}, \ldots, f'_{r,t_m})'$, the model generalizes to

$$y_{t_q+h_q} = y_{t_m+h_m} = \beta_0 + \Sigma_{i=1}^{r}\beta_{1,i} b(L_m, \theta_i)\hat{f}_{i,t_m}^{(3)} + \epsilon_{t_m+h_m} \tag{2.13}$$

The estimation of the MIDAS-basic regression is performed by non-linear least squares. From now on this approach will be named MIDAS-Basic.

2.2.2 The Unrestricted MIDAS

As an alternative to the MIDAS-basic approach, the Unrestricted MIDAS approach is a simple but less parsimonious way to include several lags of the high frequency into the regression. The equation has the form

$$y_{t_m+h_m} = \beta_0 + \mathbf{D}(L_m)\hat{\mathbf{F}}_{t_m}^{(3)} + \epsilon_{t_m+h_m} \tag{2.14}$$

where $\mathbf{D}(L_m) = \Sigma_{k=0}^{K}\mathbf{D}_k L_m^k$ is an unrestricted lag polynomial of order K. The estimation of the polynomial $\mathbf{D}(L_m)$ and β_0 is performed by OLS. From now on this setting will be denoted as MIDAS-U.

Even thought, this setting is easy to estimate, it lacks from the parsimony that is present in the MIDAS-basic approach.

2.2.3 MIDAS extensions and benchmark model

The MIDAS projections can be extended to allow for autoregressive lags in the following way

$$y_{t+h} = \beta_0 + \lambda y_t + \beta_1 B\left(L^{(1/m)};\theta\right) x_t^{(m)} + \epsilon_t^{(m)}. \tag{2.15}$$

Gyshels et al. (2007) highlight that this specification should be use with caution as it has the disadvantage that it includes implicitly a seasonal pattern in the model. In spite of that, this specification is implemented to check whether there is a gain in forecast power or not.

As a benchmark the AR(1) model is going to be used. The dynamic of a variable y_t which follows an AR(1) model is

$$(1-\phi_1 L)y_t = \epsilon_t \tag{2.16}$$

where L is the lag operator such that $Ly_t = y_{t-1}$, ϵ_t is *iid* and $E(\epsilon_{t+1}|y_t) = 0$.

2.3 Forecast experimental design

This section will describe the empirical application of the MIDAS regressions in order to test their performance when applied to forecast Colombian GDP. A recursive experiment is performed to calculate the factors using the methods described in section 2.1 and to estimate the MIDAS projections showed on sections 2.2.1 and 2.2.2 prior to the generating of the forecasts. The monthly data was seasonally adjusted and outlier corrected as it was described in Chapter 1. To conduct the experiment, the ragged-edge pattern shown at the end of the final dataset was replicated. For simplicity it was assumed this patter is stable over the evaluation period, even though the real patter of publication lag of the Colombian series is not fixed and it can change from one month to the other. Only few series have fixed publication dates. Although this assumption might not reflect the real pattern that was present in the months during the evaluation period, is a plausible solution to carry on the study due the lack of this information.

Another setback found in the gathering of data is change in the measure methodology of the GDP. There is a new vintage that starts from 2000:Q1 and at this moment there is no series available that splice the two vintages in a consistent way. For this reason the evaluation period does not go beyond 2007:Q4.

The study is performed applying a rolling window scheme. The motivation for this choice is the structural changes present in the Colombian macroeconomic data as they were described above. For each window the factors are estimated first and subsequently they are used to estimate the MIDAS projections.

For each quarter in the evaluation period, three nowcasts can be calculated, first using the data up to the first month of the quarter, then up to the second month and finally up to the third month. Forecasts can also be calculated using the information that is available up to the first, second and third month of the previous quarter for the case of the forecast one period ahead. Using the same scheme, the two periods ahead forecast is computed only using the monthly information corresponding the months two quarters before. Backcasts can also be calculated depending on the lag of publication of the GDP. For this study, the lag of publication allows to have two backcasts after the end of the quarter. These backcasts can be calculated in the first and second month of the next quarter. It total there are eleven projections made for in each quarter in the evaluation period.

2.3.1 Specification of the number of factors

In order to set the appropriate number of factors the criteria Bai and Ng (2002) proposed an information criteria to establish the number of static factors which are going to be used in the regressions. This information criteria has the following

$$IC_{p2} = ln(V(r, \mathbf{F})) + r(\frac{N + T_m}{NT_m})ln(min\{N, T_m\}) \tag{2.17}$$

where

$$V(r, \mathbf{F}) = \frac{1}{NT_m}\Sigma_{i=1}^{N}\Sigma_{t_m+1}^{T_m}(x_{i,t_m} - \mathbf{\Lambda}_i\mathbf{F}_{t_m}) \tag{2.18}$$

is the residual sum of squares. This information criterion has to be minimized with the purpose of determining the number of factors.

2.3.2 Number of series in factor estimation

The number of macroeconomic indicators included in the factor estimation has become a relevant issue in recent studies. It was believed the more data and series are incorporated into the factor estimation procedure the better and efficient the estimated factors will be. However, Boivin and Ng (2006) find out that the use of too many series in the estimation of the factors do not yield to improvements in their forecast ability. This study shows how the presence of cross-correlation in the idiosyncratic elements in ϵ in the equation (2.1) affects the factors, yielding to less efficient to forecasts.

Following this approach, in this research there are going to be used two sets of series to estimate the factors. The first set is composed of all the series described in Table A.5 and it is going to be denoted as ALL. The second set is constituted the series included in the coincident and leading indices, with additional series that have a high R^2 with the factors and its idiosyncratic term is not correlated with the rest of the series. This set of series includes: Business Survey question 1: current economic conditions, Business Survey question 6: number of orders, industrial production index excluding coffee threshing, currency in circulation in real terms, demand of energy and gas, total imports excluding capital and durable goods, loan portfolio of the financial system, Money Supply, real interest rate and approved building area. This set is named CL from now on.

2.4 Out of sample forecasting evaluation

One of the main concerns in this study is the forecast performance of the MIDAS projections. The way this is examined is by means of the out-of-sample forecast for the different prediction horizons defined above.

One way to measure the forecast performance of the MIDAS projections is the Mean Square Error (MSE). This quantity is defined as

$$MSE_j^h = \frac{1}{P}\Sigma_{i=1}^{P}(y_{i+h} - \hat{y}_{i+h|i}^{j}) \tag{2.19}$$

where $\hat{y}_{i+h}$ is the forecasted value of the model j at the horizon h and y_{i+h} is the observed value.

2.4.1 Forecast evaluation statistics

Using MSE values alone to state if a model holds more accuracy than a benchmark or any other model can yield to mislead results, due the fact that the variability of the forecasts is not taken into account. In the literature there are several tests designed to evaluate if two competing models differ statistically in their accuracy. Among the most used, is the test developed by Diebold and Mariano (1995) which is described below. The basic idea of this test is to state if on average the differences of the distances between the forecasts and the true realized values are statistically equal to zero or not.

To put it in a formal way, let y_t be the series to be forecasted and let $\hat{y}_{t+h|t}^1$ and $\hat{y}_{t+h|t}^2$ be the two competing models. The forecast errors of this models are $e_{t+h|t}^1 = y_t - \hat{y}_{t+h|t}^1$ and $e_{t+h|t}^2 = y_t - \hat{y}_{t+h|t}^2$ respectively. The base of the test is the differential

$$d_t = L(e_{t+h|t}^1) - L(e_{t+h|t}^2) \tag{2.20}$$

where $L(e_{t+h|t}^i) = (e_{t+h|t}^i)^2$ is the square loss function. The purpose of the test is to prove whether the null hypothesis of equal predictive accuracy

$$H_0 : E[d_t] = 0 \tag{2.21}$$

is true or not. The Diebold-Mariano test statistic has the following function

$$S = \frac{\bar{d}}{(\hat{\Omega}/n)^{1/2}} \tag{2.22}$$

where

$$\hat{d} = \frac{1}{n}\Sigma_{t=1}^{n} d_t \hat{\Omega} = \gamma_0 + 2\Sigma_{j=1}^{\infty}\gamma_j, \quad \gamma_j = cov(d_t, d_{t-j}). \tag{2.23}$$

In sufficient large samples the Diebold-Mariano statistic is approximately a standard normal ($S \sim N(0, 1)$).

In practice, the variance estimator proposed by Newey and West (1987) can be used instead. This estimator has the advantage to be consistent in the presence of autocorrelation and heteroscedasticity. This variance estimator has the following form

$$\hat{\Omega}^{NW} = \hat{\Psi}_0 + \Sigma_{j=1}^{m}(1 - j/(m + 1))[\hat{\Psi}_j + \hat{\Psi}_j'] \tag{2.24}$$

$$\hat{\Psi}_j = \sum_{n}^{j=1} d_t d_{t-j}$$

where $m = floor[4(n/100)^(2/9)]$ was suggested by Newey and West (1994). A variant of the Diebold-Mariano test is the one proposed by Harvey et.al. (1997) where a small sample correction to this was performed. The statistic has the following form

$$S^* = (\frac{n + 1 - 2h + n^{-1}h(h - 1)}{n})^{1/2} S \tag{2.25}$$

where n is the size of the evaluation period, h is the number of periods ahead and S is the Diebold-Mariano statistic shown in (2.22). The Modified Diebold-Mariano test has a Student's t distribution with n-1 degrees of freedom. The Modified Diebold Mariano statistic can also be calculated with the Newey-West variance estimator, which is denoted as MDM-NW.

As an alternative to the Diebold-Mariano(1995) and Harvey et.al. (1997) tests, Giacomini and White (2006) introduced a test that can handle misspecified models and is suitable for recursive parameter updating and nested models. They propose the following chi-square distributed test statistic

$$S_{GW} = n(n^{-1}\Sigma_{t=1}^{n} h_t \hat{d}_{t+h})\hat{\Omega}^{-1}(n^{-1}\Sigma_{t=1}^{n} h_t \hat{d}_{t+h}) \sim \chi_{\nu}^2 \tag{2.26}$$

where $h_t = [1, \hat{d}_{t+h|t}]'$ is a $\nu \times 1$ test function which is a constant and lagged difference as instruments. The variance matrix Ω is estimated using the Newey and West (1987) covariance estimator. This test will be denoted as GW henceforth.

Chapter 3

Estimation and Results

3.1 Factor Estimation

The first part of the study is focused on the estimation of the factors. In order to evaluate the forecast of the models shown in Chapter 2 the sample is divided in an estimation and evaluation periods . The estimation period starts from 1995:Q1 to 1999:Q4 and the evaluation period goes from 2000:Q1 to 2007:Q4. The sequence of regression estimates used to get predictions is the *rolling window* scheme. This scheme was preferred due the presence of trend breaks and structural changes in the Colombian macroeconomic series as it can be seen in Figure 1.1.

The estimation of the monthly factors is done according to this scheme keeping in mind that the number of series must not be higher than the window size. Specifically, the estimation window comprises 60 months (20 quarters), so it was necessary to use more monthly data before the start point of the window. Once the factors are estimated, only the periods that comprehend the window are used. The factors are estimated recursively for every rolling window keeping the ragged edge pattern described in section (2.1).

For each model the out-of-sample forecasts were computed over the evaluation period 2000:Q1 to 2007:Q4. These models are combinations of the two MIDAS projections, the two factor estimation methods (SW-EM, KAL-SS), the series used to estimate the factors (ALL,CL), presence of one lag of GDP , lags of the monthly factor and the number of factors. The information criterion of Bai and Ng (2002) shown in section was applied. The maximum number of factors that were tested was $r = 3$. The criterion showed an optimal value of at most $r = 2$ which was the maximum number of factors taken into account in the MIDAS projections.

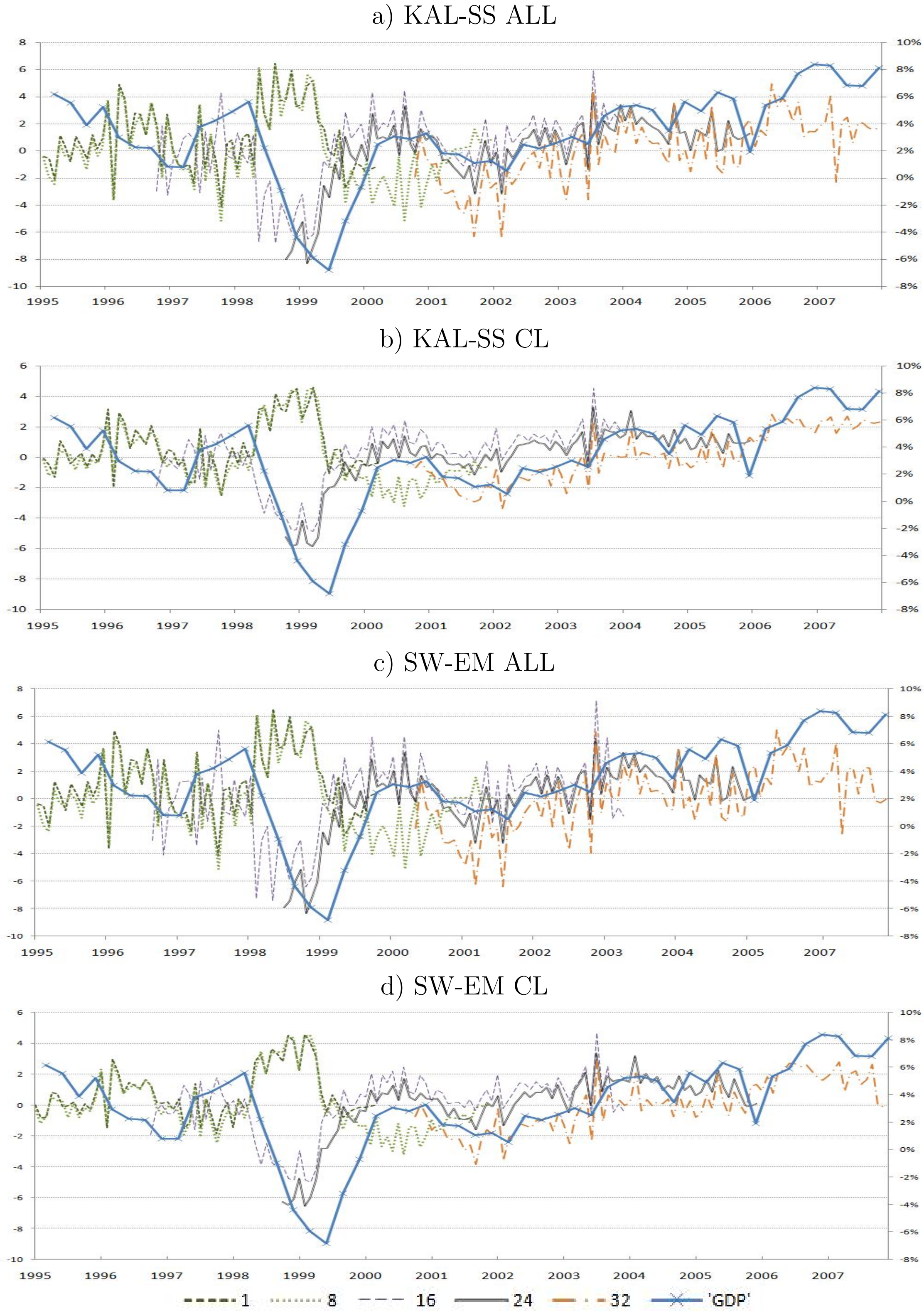

Figure 3.1: Comparison of the first factor using the KAL-SS method for a) the ALL set and b) the CL set and the SW-EM for c) ALL set and d) CL set, for different windows with the real GDP growth.

At this stage, the factor estimation procedure SW-EM evidenced sensitivity to the initial (naive) guesses for the missing values present in the ragged edge pattern at the end of the sample. Different initial values were used to estimate the factors, finding that the estimated missing values are not robust to the changes in the initial guesses. First, the mean of the transformed series was used, then a moving average of the three past values before the missing value and finally, an ARMA model was estimated for each series to get predictions for the last observations where the missing values were. These three approaches yield different results, yet an early experiment of the influence in the forecast ability showed better results when the last two approaches were applied. In spite of that, the third approach was too computational expensive. In the end, for practical issues the second approach was used. Figure 3.1 shows the first factor estimated across the evaluation period for the two factor estimation methods, using the ALL set and the CL set. It is important to notice how this factor exhibits comovement with the GDP growth for both factor estimation methods, moving direct or inverse depending on the window. In addition, the factors estimated with the ALL set look rougher than the ones estimated with the CL series.

3.2 Estimation of the MIDAS projections

The second stage of the procedure was the estimation of the MIDAS-U and MIDAS-basic projections. The first estimations of the MIDAS-basic projection showed problems of convergence in the non-linear squares estimation despite the restrictions on the θ_2 parameter. To circumvent local minima problems, different initial values were used and the results did not make a difference. Furthermore, for some windows the θ parameter estimates were very large values ($> 10^6$). Another problem was the close to singularity of the Hessian matrix for some windows, being this more frequent in the case of $r = 2$. This is an issue for further research to check possible presence of identification issues in the implementation of MIDAS projections with this data. To overcome this problems in a practical way there were implemented two solutions, the first one was to place more restrictions on the θ parameters. In case of failure, the second solution implemented was to use the estimated parameters from the previous window.

3.3 MIDAS projections vs. benchmark model

Marcellino and Schumacher (2007) compare their nowcasts and forecast ranking the MSEs from the lowest to the highest . However, they do not include forecast evaluation tests in their analysis.

The back-, now- and forecast results for the different combinations of MIDAS projections, factor estimation methods, set of variables used in the factor estimation and GDP lags in the equation are found in Tables A.1 to A.4. Each table contains the relative MSEs for one projection and presence of GDP lags in the equation. The results are divided by MIDAS projection and factor estimation method only for models with $r = 1$. The reason for this choice will be exposed in section 3.5.

In Tables A.1 and A.2 it can be seen that MIDAS projections with no GDP lags show very few MSEs which are lower than one. Most of them are in the MIDAS-U projection with KAL-SS factors for $K \leq 3$ in the prediction horizons $h_m = 2, \ldots, 9$. In general, looking only at the MSE values, the MIDAS projections alone do not seem to do better than the AR(1) benchmark model.

In contrast to the regressions with no GDP lags, the results of MIDAS projections estimated using the equation (2.15) give a different picture. Just by looking at the MSE values, from Tables A.3 and A.4 it can be seen that relative MSEs of this models with one lag show lower MSE than the models with no GDP lags. Moreover, some of these combinations of models seem to perform much better than the AR(1) benchmark based only on the comparison of MSE values. Among this are the MIDAS-U projection with less than three factor lags and almost all the MIDAS-basic projections.

To check if these comparisons are statistically significant, the forecast evaluation tests explained in section 2.4.1 were applied to all the model combinations. For these comparisons against the benchmark, the alternative hypothesis was set to be $H_A : \sigma^2_{AR} - \sigma^2_{MIDAS} > 0$ to show whether or not there is evidence in favor of the MIDAS projections in case the null hypothesis is rejected.

The MSEs that are significant are shown in with superscripts in Tables A.1 to A.4, being (a) MDM-NW test significant at 5%, (b) MDM-NW test significant at 10% , (c) GW test significant at 5% and (d) GW significant at 10%.

The number of rejections of the null hypothesis is low compared to the amount of models that were estimated, nevertheless there MIDAS-basic projection showed more rejections that the MIDAS-U. Notice that monthly forecast horizons $h_m = -1, 1$ did not report any rejection for none of the test performed. Among the test statistics used, MDM-NW shows very few rejections, still GW did not show any rejection. It is hard to state that there is one projection that outperforms the benchmark. It has to be kept in mind that due the large amount of comparisons one can incur in data snooping by accepting a model that rejects the null hypothesis.

3.4 Influence of the factor lags

Marcellino and Schumacher (2007) set a boundary for all their MIDAS projections to $K \leq 12$ and they employ the BIC criterion to select automatically the number of monthly factors lags K present in the MIDAS-U projections. The MIDAS-U projection with $K = 0$ is denoted for them as MIDAS-U0 and they set it as a benchmark for the MIDAS regressions.

In this study the BIC criterion was used also to select K but in a different fashion. The automatic selection is performed for K, first up to six lags and then up to twelve lags. These approaches are denoted as BIC_6 and BIC_{12} respectively. The motivation behind this choice is to check if the use of the automatic lag selection helps to improve forecast performance for certain windows in the evaluation period compared to the models with fixed K. For MIDAS-U projections with $r = 2$ the monthly lags taken into account where up to $K \leq 6$ due the short window size.

From the results for MIDAS-U shown in Table A.1 and Table A.3 it can be seen how as K increases the MSE increases too. This might arise because a consistency problem in the estimated parameters due the lack of degrees of freedom to estimate the coefficients as the factor lags increase. The smallest MSEs are found in the models with low values of K. Moreover, the MSEs of MIDAS-U projections with automatic lag selection BIC_6 and BIC_{12} are equal or higher than the MIDAS-U projections with fixed K.

For MIDAS-basic, Table A.2 and Table A.4 show mixed results when comparing the number of factor lags in the equation. In the case of no lags of GDP in the equation (Table A.2), increasing K in the exponential almon in equation (2.12) leads to increases in the MSE for $h_m = -2, -1, 1$. For the remaining forecast horizons the MSE appear to be almost the same over the different values of K. This conclusion can also be drawn out in the case of one GDP lag in the equation (Table A.4).

The use of more lags did not improve the forecast performance in terms of lower MSE values nor the automatic lag selection. Forecast evaluation test were applied to each model across the factor lags yielding no conclusive results about which number of lags performs better. As a conclusion, the best choice is to keep $K \leq 3$.

3.5 Influence of the number of factors

In this section the results of the comparison among the number of factors are shown. Table 3.1 presents some of the relative MSEs of MIDAS projections that were estimated with $r = 1$ and $r = 2$. As can be seen in this table, when a second factor is present, some combinations of factors show worse forecast performance than the case with only one factor. Furthermore, this increase in the MSE is more evident in the predictions made for the backcast horizon. For the prediction horizons that comprehend the forecast for one and two quarters ahead, the use of one or two factors seems to be irrelevant. Lower values MSE values are prominent when $r = 1$.

The forecast evaluation tests MDM-NW and GW were applied here to check if the use of one or two factors indeed made a difference in the forecast accuracy of the models. The outcome of the tests did not lead to conclusive results. There are few comparisons where there was a significative difference, always in favor of the model with one factor. To summarize, it is clear that the use of more than one factor does not generate significant improvements in the forecast performance. For simplicity and parsimony, from now on the results are focused on the models with only one factor.

In addition, it is important to note that the MSEs in the equations with $r = 2$ with no autoregressive GDP terms in the equation and SW-EM factors, showed very high values. Focusing more in this result, it appeared that only one or two forecasts in the evaluation period presented very high errors, while the other points have fairly close errors. The estimated parameters of the windows that produced those forecast do not present an awkward behavior.

3.6 Influence of the set of variables used in the factor estimation

The comparisons of the sets of variables used estimate the factors are presented in Table 3.2. As it is seen in this table there is no pattern that shows improvements in the performance of the predictions when the set ALL is used to estimate the factors or if the CL set is used. For example, the results for the MIDAS-U with SW-EM factors show lower MSEs when they are calculated with the CL set compared with the ones calculated with the ALL set for the forecast horizon $h_m = 5$. Nonetheless, for the prediction horizon $h_m = 6$ the the opposite happens. Also depending on the lags included in the monthly factors there are mixed results even in the same prediction horizon. Forecast evaluation test were applied here as well, finding no significative results except for the MIDAS-U projection with KAL-SS factors where the CL set set seem to perform better than the ALL set for prediction horizons $h_m = 2$ y $h_m = 3$ with significance level 10% for both MDM-NW test and GW. In the end, there are no conclusive results to say that one set of variables performs better than the other.

Projection	Factors	r	backcast before release		nowcast current quarter			forecast 1 quarter			forecast 2 quarters		
		$\backslash h_m$	-2	-1	1	2	3	4	5	6	7	8	9
No lags in GDP													
MIDAS-U0	SW-EM	1	1.57	1.41	3.98	1.99	1.95	2.41	1.83	1.79	1.95	1.59	1.55
		2	3.82	2.68	3.84	1.69	1.80	2.37	1.54	1.45	2.17	1.47	1.29
	KAL-SS	1	1.43	1.32	1.32	0.78	0.77	0.65	0.59	0.62	0.69	0.62	0.66
		2	1.51	1.41	1.30	0.77	0.75	0.64	0.59	0.61	0.69	0.61	0.65
MIDAS-U	SW-EM	1	1.53	2.44	2.53	1.37	3.03	1.23	0.89	2.34	1.17	0.83	1.92
		2	9.76	28.10	4.53	3.94[b]	10.89	1.29	0.70	2.59	2.65	0.79	1.63
	KAL-SS	1	1.45	1.69	0.94	0.60	0.56	0.86	0.79	0.93	1.02	0.80	0.96
		2	6.49	2.11	1.36	2.18	0.65	1.06	1.42	1.02	0.96	0.83	0.97
MIDAS-basic	SW-EM	1	1.20	1.89	1.28	1.48	2.88	0.81	1.19	2.26	0.87	1.06	1.93
		2	8.51	34.34	2.62[b]	2.83[b]	12.79	1.21	1.14	2.35	1.79	1.39	1.79
	KAL-SS	1	1.11	1.31	0.79	0.54	0.54	0.81	0.80	0.93	0.99	0.85	1.01
		2	6.24	2.17	1.47	2.33	0.67	1.08	1.31	0.89	1.05	0.77	1.13
One lag in GDP													
MIDAS-U0	SW-EM	1	0.84	0.92	1.38	0.68	0.67	0.82	0.62	0.60	0.71	0.57	0.55
		2	1.05	1.40	1.87	0.66	0.56	0.74	0.57	0.53	0.53	0.47	0.49
	KAL-SS	1	0.79	0.85	0.85	0.52	0.51	0.47	0.39	0.42	0.42	0.35	0.36
		2	0.80	0.86	0.85	0.52	0.51	0.47	0.39	0.42	0.42	0.35	0.36
MIDAS-U	SW-EM	1	0.87	1.59	0.93	0.65	1.40	0.67	0.56	1.05	0.69	0.60	1.36
		2	1.20	2.58	3.14[b]	2.07	2.09	1.04	0.49	2.40	1.66[b]	0.54	1.55
	KAL-SS	1	0.76	0.97	0.79	0.51	0.52	0.62	0.48	0.57	0.58	0.45	0.53
		2	1.11	1.09	0.97	0.76	0.52	0.59	0.46	0.66	0.55	0.51	0.55
MIDAS-basic	SW-EM	1	0.88	1.58	0.75	0.67	1.33	0.60	0.67	0.93	0.63	0.68	1.03
		2	1.20	2.20	1.87	1.35	1.99	0.96	0.74	2.64	1.08	0.82	1.87
	KAL-SS	1	0.77	0.98	0.79	0.52	0.53	0.58	0.48	0.57	0.55	0.45	0.53
		2	0.89	1.04	0.82	0.57	0.51	0.48	0.45	0.56	0.63	0.52	0.54

[b] MDM-NW significant 10%

[d] GW significant 10%

Table 3.1: Comparison of the Relative MSEs for $r = 1$ and $r = 2$ with $K = 2$ and the variable set ALL

3.7 Comparison of MIDAS projections

In this section the results for the different types of MIDAS projections are discussed. The MSE values of MIDAS projections are compared in Table 3.3 classified for factor estimation method and set of variables used in the estimation of factors. For the MSEs of the MIDAS projections when there are no GDP lags in the equation, there is no clarity about which projection has lower MSE values. However, for the forecast horizons $h_m = 4\dots,9$ the projection MIDAS-U0 with KAL-SS factors has lower MSE values compared to MIDAS-U and MIDAS basic. When one lag of GDP is present in the equation the MSEs are relatively close. In general, there is no projection that has lower MSE values across all the prediction horizons. The forecast evaluation statistics do not show conclusive results about which projection is the best. Recalling what was exposed in section 3.5, is worth to note in Tables A.2 and A.4 the MIDAS-basic projection do not exploit the parsimonious structure of the exponential almon when more monthly factor lags are added. Therefore, there are no big differences in using few or many factor

Projection	Factors		backcast before release		nowcast current quarter			forecast 1 quarter			forecast 2 quarters		
		$\backslash h_m$	-2	-1	1	2	3	4	5	6	7	8	9
					No lags in GDP								
MIDAS-U0	SW-EM	ALL	1.57	1.41	3.98	1.99	1.95	2.41	1.83	1.79	1.95	1.59	1.55
		CL	1.26	1.07	4.40	2.11	2.02	2.25	1.71	1.62	1.84	1.52	1.45
	KAL-SS	ALL	1.43	1.32	1.32	0.78	0.77	0.65	0.59	0.62	0.69	0.62	0.66
		CL	1.14	0.95	0.86	0.49^{bd}	0.48^{b}	0.72	0.63	0.64	0.86	0.76	0.78
MIDAS-U	SW-EM	ALL	1.53	2.44	2.53	1.37	3.03	1.23	0.89	2.34	1.17	0.83	1.92
		CL	1.41	1.64	2.75	1.56	2.53	1.03	0.81	1.90	1.34	0.93	1.56
	KAL-SS	ALL	1.45	1.69	0.94	0.60	0.56	0.86	0.79	0.93	1.02	0.80	0.96
		CL	1.57	1.60	1.28	0.64	0.45	1.18	0.97	0.88	1.26	1.01	0.99
MIDAS-basic	SW-EM	ALL	1.20	1.89	1.28	1.48	2.88	0.81	1.19	2.26	0.87	1.06	1.93
		CL	0.99	1.03	1.37	1.52	2.48	0.83	1.10	1.90	1.09	1.18	1.56
	KAL-SS	ALL	1.11	1.31	0.79	0.54	0.54	0.81	0.80	0.93	0.99	0.85	1.01
		CL	1.06	1.00	0.91	0.50	0.46	1.07	0.90	0.87	1.19	0.97	1.00
					One lag in GDP								
MIDAS-U0	SW-EM	ALL	0.84	0.92	1.38	0.68	0.67	0.82	0.62	0.60	0.71	0.57	0.55
		CL	0.84	0.93	1.61	0.78	0.74	0.79	0.65	0.58	0.91	0.75	0.69
	KAL-SS	ALL	0.79	0.85	0.85	0.52	0.51	0.47	0.39	0.42	0.42	0.35	0.36
		CL	0.85	0.94	0.78	0.47	0.48	0.54	0.46	0.48	0.48	0.41	0.44
MIDAS-U	SW-EM	ALL	0.87	1.59	0.93	0.65	1.40	0.67	0.56	1.05	0.69	0.60	1.36
		CL	0.83	0.83	1.13	0.78	1.18	0.55	0.57	1.42	0.92	0.71	1.06
	KAL-SS	ALL	0.76	0.97	0.79	0.51	0.52	0.62	0.48	0.57	0.58	0.45	0.53
		CL	0.91	0.94	0.98	0.49	0.46	0.86	0.61	0.59	0.71	0.58	0.59
MIDAS-basic	SW-EM	ALL	0.88	1.58	0.75	0.67	1.33	0.60	0.67	0.93	0.63	0.68	1.03
		CL	0.74	0.79	0.90	0.82	1.17	0.64	0.81	1.26	0.84	0.85	0.83
	KAL-SS	ALL	0.77	0.98	0.79	0.52	0.53	0.58	0.48	0.57	0.55	0.45	0.53
		CL	0.85	0.88	0.85	0.48	0.49	0.69	0.53	0.55	0.66	0.55	0.59

[b] MDM-NW significant 10%
[d] GW significant 10%

Table 3.2: Comparison of the Relative MSEs for the variable sets ALL and CL with $r = 1$ and $K = 2$

lags when the MIDAS-basic projection is estimated. When $K = 2$, MIDAS-basic and MIDAS-U have the same number of parameters, being the first one just a more complex parameterization. If they have more or less an equal performance, like in this case, is better to use MIDAS-U.

The forecast evaluation statistics do not show conclusive results about which projection is the best. The only significative results are for comparisons for MIDAS projections that use factors estimated with the SW-EM procedure. In conclusion, there is no clear cut improvement for using one specific projection. These results coincide with the results of Marcellino and Schumacher (2007).

			backcast		nowcast			forecast			forecast		
set	Factors	Projection	before release		current quarter			1 quarter			2 quarters		
		$\backslash h_m$	-2	-1	1	2	3	4	5	6	7	8	9
		No lags in GDP											
ALL	SW-EM	MIDAS-U0	1.57	1.41^{a}	3.98	1.99	1.95^{be}	2.41	1.83	1.79	1.95	1.59	1.55
		MIDAS-U	1.53	2.44	2.53	1.37	3.03	1.23	0.89	2.34	1.17	0.83	1.92
		MIDAS-basic	1.20	1.89	1.28^{gh}	1.48	2.88	0.81	1.19	2.26	0.87	1.06	1.93
	KAL-SS	MIDAS-U0	1.43	1.32	1.32	0.78	0.77	0.65	0.59	0.62	0.69	0.62	0.66
		MIDAS-U	1.45	1.69	0.94	0.60	0.56	0.86	0.79	0.93	1.02	0.80	0.96
		MIDAS-basic	1.11	1.31^{b}	0.79	0.54	0.54	0.81	0.80	0.93	0.99	0.85	1.01
CL	SW-EM	MIDAS-U0	1.26	1.07	4.40	2.11	2.02^{e}	2.25	1.71	1.62^{e}	1.84	1.52	1.45
		MIDAS-U	1.41	1.64	2.75	1.56	2.53	1.03	0.81	1.90	1.34	0.93^{e}	1.56
		MIDAS-basic	0.99	1.03^{b}	1.37^{gh}	1.52	2.48	0.83	1.10	1.90	1.09	1.18^{ad}	1.56
	KAL-SS	MIDAS-U0	1.14	0.95	0.86	0.49	0.48	0.72	0.63	0.64	0.86	0.76	0.78
		MIDAS-U	1.57	1.60	1.28	0.64	0.45	1.18	0.97	0.88	1.26	1.01	0.99
		MIDAS-basic	1.06	1.00	0.91	0.50	0.46	1.07	0.90	0.87	1.19	0.97	1.00
		One lag in GDP											
ALL	SW-EM	MIDAS-U0	0.84	0.92	1.38	0.68	0.67	0.82	0.62	0.60^{befg}	0.71	0.57	0.55
		MIDAS-U	0.87	1.59	0.93	0.65	1.40	0.67	0.56	1.05	0.69	0.60	1.36
		MIDAS-basic	0.88	1.58	0.75^{gh}	0.67	1.33	0.60	0.67^{b}	0.93	0.63	0.68	1.03
	KAL-SS	MIDAS-U0	0.79	0.85	0.85	0.52	0.51	0.47	0.39	0.42	0.42	0.35	0.36
		MIDAS-U	0.76	0.97	0.79	0.51	0.52	0.62	0.48	0.57	0.58	0.45	0.53
		MIDAS-basic	0.77	0.98	0.79	0.52	0.53	0.58	0.48	0.57	0.55	0.45	0.53
CL	SW-EM	MIDAS-U0	0.84	0.93	1.61	0.78	0.74^{e}	0.79	0.65^{b}	0.58^{eg}	0.91	0.75	0.69
		MIDAS-U	0.83	0.83	1.13	0.78	1.18	0.55	0.57	1.42	0.92	0.71	1.06
		MIDAS-basic	0.74	0.79	0.90^{gh}	0.82	1.17	0.64	0.81	1.26	0.84	0.85	0.83
	KAL-SS	MIDAS-U0	0.85	0.94	0.78	0.47	0.48	0.54	0.46	0.48	0.48	0.41	0.44
		MIDAS-U	0.91	0.94	0.98	0.49	0.46	0.86	0.61	0.59	0.71	0.58	0.59
		MIDAS-basic	0.85	0.88	0.85	0.48	0.49	0.69	0.53	0.55	0.66	0.55	0.59

MIDAS-U vs. MIDAS-basic: [a]MDM-NW significant 5%, [b]MDM-NW significant 10%, [d]GW significant 10%
MIDAS-U vs. MIDAS-U0: [e]MDM-NW significant 10%, [f]GW significant 10%
MIDAS-basic vs. MIDAS-U0: [g]MDM-NW significant 10%, [h]GW significant 10%

Table 3.3: Comparison of the Relative MSEs for the MIDAS-U and MIDAS-basic with $r = 1$ and variables set ALL

3.8 Comparisons of the factor estimation methods

The results in Table 3.4 show the comparisons of the two factor estimation method used in this study. The results are classified by MIDAS projection and set of variables used to estimate the factors. These comparisons of the MSEs show in general a lower forecast error of the KAL-SS. Nonetheless, the statistical test only show significant comparisons when the MSEs present a big difference. Also, the test show significance almost all the time at 10%. The rejections are concentrated mostly when there are no lags of GDP on the equation. In this case, the SW-EM factor estimation presents high MSEs, specially for the MIDAS-U projection and the predictions made in the horizons that comprehend the nowcast and the forecast one quarter ahead.

For the case when there is one GDP lag in the equation the MSEs from the projections with KAL-SS factors seem to be a bit better than the SW-EM factors. For this part there are few comparisons that show significant differences in the factor estimation method. In general, the KAL-SS have lower MSEs but the forecast evaluation tests are not conclusive to claim that the forecast accuracy is improved using one or the other.

Projection	Set	Factors	backcast before release		nowcast current quarter			forecast 1 quarter			forecast 2 quarters		
		$\backslash h_m$	-2	-1	1	2	3	4	5	6	7	8	9
No lags in GDP													
MIDAS-U0	ALL	SW-EM	1.57	1.41	3.98	1.99	1.95	2.41	1.83	1.79	1.95	1.59	1.55
		KAL-SS	1.43^{a}	1.32	1.32^{bd}	0.78^{bd}	0.77^{bd}	0.65^{bd}	0.59^{b}	0.62^{b}	0.69	0.62	0.66
	CL	SW-EM	1.26	1.07	4.40	2.11	2.02	2.25	1.71	1.62	1.84	1.52	1.45
		KAL-SS	1.14	0.95	0.86^{ad}	0.49^{ad}	0.48^{ad}	0.72^{b}	0.63	0.64	0.86	0.76	0.78^{d}
MIDAS-U	ALL	SW-EM	1.53	2.44	2.53	1.37	3.03	1.23	0.89	2.34	1.17	0.83	1.92
		KAL-SS	1.45	1.69	0.94^{bd}	0.60^{ad}	0.56^{b}	0.86^{b}	0.79	0.93	1.02	0.80	0.96^{d}
	CL	SW-EM	1.41	1.64	2.75	1.56	2.53	1.03	0.81	1.90	1.34	0.93	1.56
		KAL-SS	1.57	1.60	1.28^{ad}	0.64^{b}	0.45^{bd}	1.18	0.97	0.88^{d}	1.26	1.01	0.99^{d}
MIDAS-basic	ALL	SW-EM	1.20	1.89	1.28	1.48	2.88	0.81	1.19	2.26	0.87	1.06	1.93
		KAL-SS	1.11	1.31	0.79	0.54^{b}	0.54^{b}	0.81	0.80^{d}	0.93	0.99	0.85^{d}	1.01^{d}
	CL	SW-EM	0.99	1.03	1.37	1.52	2.48	0.83	1.10	1.90	1.09	1.18	1.56
		KAL-SS	1.06	1.00	0.91	0.50^{b}	0.46^{b}	1.07	0.90^{d}	0.87^{d}	1.19	0.97	1.00^{d}
One lag in GDP													
MIDAS-U0	ALL	SW-EM	0.84	0.92	1.38	0.68	0.67	0.82	0.62	0.60	0.71	0.57	0.55
		KAL-SS	0.79^{ad}	0.85^{b}	0.85^{b}	0.52	0.51	0.47	0.39	0.42	0.42	0.35	0.36
	CL	SW-EM	0.84	0.93	1.61	0.78	0.74	0.79	0.65	0.58	0.91	0.75	0.69
		KAL-SS	0.85	0.94	0.78^{b}	0.47	0.48	0.54	0.46	0.48	0.48	0.41	0.44
MIDAS-U	ALL	SW-EM	0.87	1.59	0.93	0.65	1.40	0.67	0.56	1.05	0.69	0.60	1.36
		KAL-SS	0.76	0.97	0.79	0.51	0.52	0.62	0.48	0.57^{d}	0.58	0.45	0.53
	CL	SW-EM	0.83	0.83	1.13	0.78	1.18	0.55	0.57	1.42	0.92	0.71	1.06
		KAL-SS	0.91	0.94	0.98	0.49	0.46	0.86	0.61	0.59	0.71	0.58	0.59
MIDAS-basic	ALL	SW-EM	0.88	1.58	0.75	0.67	1.33	0.60	0.67	0.93	0.63	0.68	1.03
		KAL-SS	0.77	0.98	0.79	0.52	0.53	0.58	0.48	0.57^{d}	0.55	0.45	0.53
	CL	SW-EM	0.74	0.79	0.90	0.82	1.17	0.64	0.81	1.26	0.84	0.85	0.83
		KAL-SS	0.85	0.88	0.85	0.48	0.49	0.69	0.53	0.55	0.66	0.55	0.59

[a] MDM-NW significant 5%
[b] MDM-NW significant 10%
[d] GW significant 10%

Table 3.4: comparison of the Relative MSEs for the MIDAS-U and MIDAS-basic with $r = 1$ and variables set ALL. Here the MIDAS-U0 refers to the MIDAS-U with $K = 0$. The other MIDAS projections shown are with $K = 2$

3.9 Influence of the lags in the equation

As it has been seen in the tables 3.1 to 3.4, the inclusion of one GDP lag into the equation lowers the relative MSEs of the models. However, the forecast evaluation tests do not yield to conclusive results here as well, giving the idea that either way the accuracy is not improved indeed.

Chapter 4

Conclusions

There are two main problems in GDP forecasting: low frequency and delays in the publication of data. The methods presented in this research intend to forecast GDP growth taking advantage of data that are available more frequently and contains up to date information in a promptly way. This data also has problems of availability that require the methods to be able deal with it.

The procedures applied here follow the approach used in the recent factor-forecast literature where there are two steps, the first one for forecast estimation and the second one for the generation of the forecast using a model that gets the factors as inputs. For the factors estimation part, the methods used included the estimation of missing values at the end of the sample due to the issues with publication lags mentioned above. For the forecast part the model used was the Factor-MIDAS approach of Marcellino and Schumacher (2007) which handles mixed frequency data.

To compare the performance of the MIDAS projections, the back-, now- and forecasts were contrasted to the ones produced by an AR(1) benchmark model. In addition, forecast evaluation tests were performed to see if these differences were statistically significant. The comparisons of the forecast accuracy of the MIDAS projections, done by MSE point comparisons, against the benchmark model showed mixed results when there are no lags of the GDP in the equation. The backcast estimations, which hold the most updated data previous to the release of the GDP do not show better performance than the benchmark for none of the MIDAS projections. MIDAS unrestricted and MIDAS with exponential distributed lag tend to have low mean squared errors for the nowcast and forecast prediction horizons, when they are combined with factors estimated by the State-Space approach with lags less or equal to three. Conversely, including one lag of the GDP seem to lower the forecast errors compared to the benchmark. Nonetheless, the outcomes of the forecast evaluation tests do not show conclusive evidence to assert that the MIDAS projections perform better than the benchmark.

The comparisons were extended to check the influence of the MIDAS projections, factor estimation methods, number of factors, variables included in the factor estimation and factor lags in the equation. The forecast evaluation tests were applied sequentially to all the models to test if the characteristics presented above have an effect on the forecast performance.

Comparing the models with one and two factors show that it is preferable to work with one factor, since the use of two factors rarely yields better results. Moreover, for the predictions of the backcast and nowcast the use of two factors estimated with the Stock and Watson EM algorithm show high forecast errors sporadically in the rolling windows. The forecast evaluation tests applied to contrast models with one and two factors do not show many significant differences.

Concerning the use of all monthly series in the estimation of the factors compared with the use of only the series from the Coincident and Leading index do not show clear results about which approach is better. Neither the mean square error comparisons nor the forecast evaluation tests evidence a clear pattern where the projections and factor estimation methods, in combination with some set of variables, performs better along the different prediction horizons.

Comparing the factor estimation methods with ragged edge data, the two methods do not seem to have an impact on the backcast performance. For the nowcast and the forecast horizons the State-space factor estimation seem to work a bit better specially in the case where there are no GDP autoregressive terms in the equations, due the large errors presented when the Stock and Watson factor estimation method is used. Forecast evaluation test were applied here as well, only contrasting the factor estimation method and leaving the rest fixed. However, these tests are not conclusive either, making difficult to state that one method is better than the other, since only few contrasts yield significant results.

The MIDAS projections results indicate that both methods perform equally when the number of factor lags is not higher than three. When the factor lags are increased the MIDAS unrestricted projection forecast errors increase. MIDAS unrestricted and MIDAS with exponential distributed lag show lower errors on average than the benchmark for the nowcast and forecast prediction horizons when combined with factors estimated with the State-Space approach and the lags of the factors are less than three. Nevertheless, forecast evaluation tests do not evidence significative differences over the two methods.

The inclusion of one lag of GDP on the equation suggests that it might be the case that the MIDAS projections with autoregressive terms of GDP improve the forecasts. However, the high variability in the forecast errors makes it difficult to test whether there is an improvement of this method indeed or wheter the low values of the relative MSEs for some MIDAS projections against the benchmark are just luck.

To conclude, there is no certainty about the improvement of the MIDAS projections in the forecast performance of the GDP when compared to a simple AR(1) model. There are only slight indications that show they might help to get better forecasts but still more research is needed to try to get a better model. Moreover, the more parsimonious and simple the model is, the better the results are in terms of lowering the magnitude of the forecast errors.

As a path for further research, it will be desirable to use other optimization methods such as genetic algorithms to try to improve the optimization when the MIDAS with exponential almon is applied. Also, it would be useful to apply the White's reality check (2000) to see if there is evidence of some combination of factors and projection methods that really improve the forecasts. In terms of data, the next step is to use an updated version of the data in which the GDP series vintages are spliced in order to have more information to test the models.

Bibliography

[1] Altissimo, F., Cristadoro, R., Forni, M., Lippi, M., Veronese, G. (2006), EuroCOIN: A Real Time Coincident Indicator of the Euro Area Business Cycle, CEPR Working Paper 5633.

[2] Boivin, J. and Ng, S. (2006). Are More Data Always Better For Factor Analysis?, *Journal of Econometrics*, v132, 169-194.

[3] Castro, Carlos A. (2003a). Sistema de modelos multivariados para la proyección del Producto Interno Bruto. *Archivos de Economía*, Departamento Nacional de Planeación, Bogotá. No. 232.

[4] Castro, Carlos A. (2003b). Yet another lagging, coincident and leading index for the Colombian economy. *Archivos de Economía*, Departamento Nacional de Planeación, Bogotá. No. 233.

[5] Diebold, F., Mariano, R. (1995). Comparing predictive accuracy. *Journal of Business and Economic Statistics*, 13: 253-265

[6] Doz, C., Giannone, D., Reichlin, L. (2006), A quasi maximum likelihood approach for large approximate dynamic factor models, ECB Working Paper 674.

[7] Ghysels, E., Santa-Clara P., and Valkanov R. (2004), The MIDAS touch: Mixed data sampling regression models, CIRANO Working Papers 2004s-20, CIRANO.

[8] Ghysels, E., Sinko A., and Valkanov R. (2007), MIDAS Regressions: Further results and new directions, *Econometric Reviews*,26(1):53-90.

[9] Gómez, J., Uribe, J.D., Vargas, H. (2002), The Implementation of Inflation Targeting in Colombia, *Borradores de Economía, No. 202*, Banco de la República, Bogotá.

, E., Sinko A., and Valkanov R. (2007), MIDAS Regressions: Further results and new directions, *Econometric Reviews*,26(1):53-90.

[10] Forni, M., Hallin, m., Lippi, M., Reichlin,L. (2005), The Generalized Dynamic Factor Model: One-sided estimation and forecasting, *Journal of the American Statistical Association*, 100, 830-840.

[11] Harvey, A.C. (1989).*Forecasting, Structural Time Series Models, and the Kalman filter*, Cambridge: Cambridge University Press.

[12] Harvey, D., Leybourne,S., Newbold, P. (1997), Testing the equality of prediction mean squared errors. *International Journal of Forecasting* 13 , 281-291

[13] Marcellino, M., Schumacher,C. (2007). Factor-MIDAS for Now- and Forecasting with Ragged-Edge Data: A Model Comparison for German GDP, Economics Working Papers ECO2008/16, European University Institute.

[14] Newey, W. and West, K., (1987), A simple positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix . *Econometrica*, 55:703-708

[15] Newey, W. and West, K., (1994), Automatic Lag Selection in Covariance Matrix Estimation . *Review of Economic Studies*, 61(4):631-53

[16] Nieto, F.H. (1998), Ex-post and Ex-ante Prediction of Unobserved Economic Time Series: A Case Study, *Journal of Forecasting*, 17, 35-58.

[17] Nieto, F.H., Melo, L.F. (2001). About a coincident index for the state of the economy, *Borradores de Economía, No. 194*, Banco de la República, Bogotá, Colombia.

[18] Nieto, F.H., Melo, L.F., Ramos, M.A. (2003), A Leading Index For The Colombian Economic Activity. *Borradores de Economía, No. 243*, Banco de la República, Bogotá, Colombia.

[19] Nieto, F.H. (2007), Ex Post and Ex Ante Prediction of Unobserved Multivariate Time Series: A Structural-Model Based Approach *Journal of Forecasting*, Vol. 26, 53-76

[20] Sargent, T.J. and Sims, C.A. (1977). Business Cycle Modeling without Pretending to Have Too Much A-priori Economic Theory, in *New Methods in Business Cycle Research* ed. C. Sims et. al., Minneapolis: Federal Reserve Bank of Minneapolis.

[21] Stock, J.H. and Watson, M.W. (2002). Macroeconomic forecasting using diffusion indexes. *Journal of Business and Economic Statistics*, 20, 2.

Appendix A

Code	Name
	Expectations on Production
F1	Current economic conditions
F2	Production activity compared to the previous month
F3	Stocks at the end of the month
F4	Received orders compared with previous month
F5	Volume of orders compared with previous month
F6	Number of orders
F7	Intalled capacity, given the current situation of demand
F8	Production expectations for the next 3 months
F9	Price expectations for the next 3 months
F10	Expectations on the economy during the next 6 months
F11	Actual installed capacity given the number of orders
	Economic Activity
C2	Car Sales
C1	Demand of Energy
C3	Entries of foreing passengers (Air Transport)
C4	International departures of passengers (Air Transport)
C5	Cofee Production
C7	Oil Production
C9	Flight Load
C10	Domestic passengers (Air Transport)
C11	Livestock sacrifice
V22	Value of coffee crops
I2	Comercial employment index
I3	Industrial Production index
I4	Sales index excluding combustibles
	Prices and Wages
P1	International oil price
P2	International cofee price
IU1	Producer price index USA
IU2	Consumer price index
IPP	Producer price index
I5	Index of the real wage in manufacturing industries
	Monetary and financial sector
V2	Real Money Supply (M1)
V9	Private Real Checking Accounts
V10	Public Real Checking Accounts
VU3	Net International Reserves
V1	Real monetary base
V3	Real M2
V4	Real M3 plus bonds
V5	Currency in circulation in real terms
V6	Real loan portafolio of the financial system
V8	Total Real Checking Accounts
V11	Total value of deposits in real terms
V12	Portafolio in real terms
V13	Total value savings accounts in real terms
V14	Total values of certificate of deposits accounts in real terms
V15	Total value of available deposits in real terms
V16	Total values of fiduciary deposits in real terms
V17	Values of Certificates in real terms
V18	Total Bonds in real terms
	Trade
C6	Cofee Exports in real terms
VU1	Current Account Deficit
VU2	Total Exports in real terms (FOB)
V21	Total Imports in real terms (CIF)
V23	Imports of consumption goods in real terms
V24	Imports of intermediate goods in real terms
V25	Imports of capital goods in real terms
	Asset Prices
I1	Terms of trade
I6	Real exchange rate index
T1	Real Interest rates of 90-day deposits for banks and corporations
	Construction Sector
C8	Approved building area

Table A.1: List of the 60 Colombian macroeconomic used in the forecast exercise divided by economic sector.

Variable	Levels			First differences		
	ADF	KPSS	ERS	ADF	KPSS	ERS
C1	-0.32	1.59	112.11	-32.14	0.15	1.08
C10	-2.33	0.93	43.62	-17.63	0.08	1.56
C11	-1.7	0.52	3.82	-21.69	0.23	1.6
C2	-2.2	0.3	6.12	-2.52	0.22	265.06
C3	-0.59	1.75	258.11	-14.12	0.09	1.08
C4	-1.02	1.59	167.53	-17.94	0.16	1.56
C5	-2.22	0.21	18.97	-15.25	0.07	0.48
C6	-1.91	0.65	8.11	-20.78	0.28	0.3
C7	-1.5	0.47	26.54	-20.6	0.21	0.79
C8	-1.03	0.34	18.25	-23.35	0.09	0.59
C9	-2.75	0.52	17.64	-21.3	0.16	0.35
F1	-1.66	0.51	4.31	-6.35	0.08	0.58
F10	-2.67	0.44	1.18	-15.43	0.15	0.42
F11	-1.76	0.66	4.13	-5.55	0.08	1.16
F2	-1.83	0.49	18.05	-22.94	0.07	0.66
F3	-3.35	0.32	1.71	-10.38	0.03	2.57
F4	-1.82	0.47	12.3	-22.95	0.08	0.67
F5	-2.52	0.42	2.2	-17.87	0.11	0.71
F6	-2.72	0.4	0.49	-7.55	0.04	0.55
F7	-1.9	0.68	4.49	-5.32	0.14	0.84
F8	-2.46	0.4	1.74	-11.16	0.06	0.31
F9	-6.23	0.27	0.59	-11.24	0.18	0.48
I1	0.27	1.68	25.02	-12.1	0.31	1.96
I2	-1.71	1.3	12.23	-1.7	0.43	83.17
I3	-0.93	1.05	20.37	-16.7	0.15	0.73
I4	-0.65	1.05	20.16	-24.71	0.23	0.68
I5	-2.15	1.75	717.29	-15.04	0.5	1.9
I6	-1.29	0.91	10.19	-9.76	0.12	0.44
I7	-0.14	1.42	142.04	-3.35	0.16	2.19
IPP	-4.43	1.76	3060.98	-2.21	1.61	11.61
IU1	1.79	1.55	142.58	-2.72	0.64	18.7
IU2	-1.87	1.72	5410.45	-1.58	1.93	305.32
P1	0.53	1.45	25.1	-3.04	0.34	28.39
P2	-1.53	0.21	5.6	-3.26	0.09	1.9
T1	-0.3	1.71	668.16	-18.88	0.18	0.62
V1	-1.61	1.69	547.29	-20.84	0.27	0.66
V10	-0.59	1.63	416.47	-1.87	1.07	39.95
V11	-4.14	1.63	1930.26	-2.04	0.98	28.83
V12	-1.28	1.72	1143.92	-3.39	0.64	14.45
V13	-3.09	1.48	440.36	-2.7	0.68	3.2
V14	-1.84	1.7	359.46	-19.06	0.12	0.32
V15	-1.62	1.67	623.26	-39.97	0.06	0.78
V16	0.97	0.71	30.72	-2.4	0.82	5.64
V17	-2.81	1.09	228.17	-2.55	0.81	2.66
V18	-1.06	1.77	917.6	-17.71	0.2	0.41
V2	-2.03	1.72	273.12	-19.97	0.22	1.95
V21	-0.74	1.72	113.51	-15.07	0.06	0.48
V22	-2.93	1.68	633.33	-18.31	0.38	0.37
V23	-2.28	1.76	588.9	-20.96	0.19	1.97
V24	-2.1	1.66	152.11	-20.34	0.13	1.39
V25	-0.72	1.68	687.13	-2.05	1.42	38.14
V3	-0.48	1.66	420.15	-1.77	1.24	48.71
V4	-1.9	1.79	3832.39	-19.82	0.23	0.37
V5	-3.77	0.31	0.03	-1.41	0.32	14.97
V6	-0.97	1.74	577.18	-17.45	0.19	0.46
V8	-0.66	1.75	495.23	-16.11	0.15	0.43
V9	-1.79	0.29	11.36	-26.78	0.12	0.43
VU1	0.93	1.59	127.45	-22.03	0.29	0.3
VU2	-0.19	1.52	284.09	-4.29	0.2	1.93
VU3	-2.28	0.42	2.1	-6.37	0.04	0.88

ADF critical values 1%:-3.46 5%: -2.87 10%:-2.57

KPSS critical values 1%:0.73 5%: 0.46 10%:0.34

ERS critical values 1%:1.91 5%: 3.17 10%:4.333

Table A.2: Unit root tests for the macroeconomic series described in Table A.5.

			backcast before release		nowcast current quarter			forecast 1 quarter			forecast 2 quarters		
		h_m	-2	-1	1	2	3	4	5	6	7	8	9
		K											
SW-EM	ALL	0	1.57	1.41	3.98	1.99	1.95	2.41	1.83	1.79	1.95	1.59	1.55
		1	1.31	1.52	1.64	2.35	2.49	1.33	1.96	2.01	1.42	1.63	1.68
		2	1.53	2.44	2.53	1.37	3.03	1.23	0.89	2.34	1.17	0.83	1.92
		3	1.75	2.55	2.07	0.93	1.78	1.08	0.62^b	1.97	1.17	0.92	2.64
		4	1.81	2.51	2.61	1.68	1.25	1.30	0.95	1.53	1.33	1.09	3.11
		5	1.49	2.00	2.05	1.22	1.27	1.27	0.93	1.24	1.37	1.03	2.77
		6	1.58	2.05	2.38	1.43	1.29	1.43	1.12	1.22	1.58	1.14	2.84
		7	1.58	1.84	1.88	1.24	1.57	1.14	0.98	1.32	1.67	1.22	3.10
		8	1.50	1.70	1.83	1.14	1.69	1.48	1.06	1.17	1.65	1.28	2.69
		9	1.47	1.70	1.99	1.42	1.94	1.49	1.11	1.19	1.77	1.55	3.27
		10	1.53	1.79	2.29	1.83	2.21	1.38	1.02	1.19	1.81	1.61	3.28
		11	1.94	2.09	2.46	1.99	2.03	1.32	1.09	1.26	2.16	1.62	3.31
		12	2.20	2.37	2.64	2.29	2.12	1.36	1.09	1.17	2.45	1.73	3.59
		BIC_6	1.61	2.44	2.31	1.52	1.99	1.56	1.05	1.35	1.46	1.14	2.48
		BIC_{12}	1.49	2.24	2.21	1.73	2.19	1.53	1.17	1.42	1.57	1.34	2.57
	CL	0	1.26	1.07	4.40	2.11	2.02	2.25	1.71	1.62	1.84	1.52	1.45
		1	1.33	1.37	2.17	2.56	2.43	1.11	1.91	1.82	1.54	1.57	1.50
		2	1.41	1.64	2.75	1.56	2.53	1.03	0.81	1.90	1.34	0.93	1.56
		3	1.41	1.67	2.09	1.09	1.62	0.93	0.54^b	1.41	1.24	0.81	1.67
		4	1.43	1.60	2.16	1.41	1.48	1.02	0.75	1.93	1.28	0.91	2.34
		5	1.39	1.46	1.92	1.28	1.61	1.07	0.76	1.68	1.31	1.10	3.09
		6	1.53	1.77	1.89	1.18	1.68	1.10	0.77	1.84	1.38	1.17	3.20
		7	1.61	1.62	2.19	1.26	1.59	1.13	0.76	1.59	1.42	1.19	3.24
		8	1.93	2.27	2.29	1.27	2.19	1.07	0.69	1.43	1.47	1.31	3.02
		9	1.93	2.12	2.30	1.27	2.00	1.16	0.72	1.39	1.50	1.39	3.03
		10	1.77	1.90	2.65	1.48	1.56	1.27	0.78	1.32	1.47	1.35	2.75
		11	1.84	1.83	2.86	1.59	1.69	1.37	0.86	1.17	1.47	1.47	2.72
		12	2.15	2.15	3.02	1.68	1.39	1.38	0.78	1.28	1.43	1.54	2.91
		BIC_6	1.52	1.76	2.12	1.64	2.10	1.04	0.81	1.90	1.22	0.99	3.08
		BIC_{12}	2.03	2.48	2.74	1.65	2.52	1.04	0.80	1.86	1.21	1.18	2.97
KAL-SS	ALL	0	1.43	1.32	1.32	0.78	0.77	0.65	0.59^b	0.62^b	0.69	0.62^b	0.66
		1	1.22	1.19	0.87	0.58	0.53^b	0.77	0.70	0.75	1.01	0.82	0.94
		2	1.45	1.69	0.94	0.60^b	0.56^b	0.86	0.79	0.93	1.02	0.80	0.96
		3	1.67	1.93	0.80	0.54^b	0.54^b	0.80	0.74	0.89	0.98	0.77	0.94
		4	1.76	2.01	1.20	0.78	0.75	1.07	0.98	1.11	1.13	0.92	1.06
		5	1.43	1.56	1.12	0.75	0.82	1.13	1.04	1.17	1.20	1.03	1.13
		6	1.53	1.76	1.22	0.88	0.90	1.33	1.26	1.31	1.29	1.18	1.21
		7	1.50	1.51	1.19	0.90	0.99	1.29	1.28	1.37	1.39	1.27	1.34
		8	1.49	1.51	1.23	0.90	0.97	1.37	1.36	1.41	1.38	1.27	1.41
		9	1.43	1.43	1.31	0.99	1.14	1.38	1.34	1.48	1.42	1.35	1.57
		10	1.50	1.63	1.40	1.10	1.20	1.39	1.32	1.47	1.41	1.39	1.64
		11	1.90	2.07	1.60	1.36	1.38	1.35	1.44	1.65	1.54	1.51	1.69
		12	1.99	2.30	1.66	1.42	1.52	1.36	1.46	1.78	1.77	1.65	1.85
		BIC_6	1.49	1.77	1.12	0.85	0.90	1.25	1.12	1.22	1.14	1.04	1.12
		BIC_{12}	1.50	1.69	1.16	1.08	1.23	1.18	1.29	1.36	1.22	1.04	1.16
	CL	0	1.14	0.95	0.86	0.49^b	0.48^b	0.72	0.63	0.64	0.86	0.76	0.78
		1	1.40	1.31	1.05	0.42^a	0.41^a	1.05	0.71	0.72	1.20	0.89	0.92
		2	1.57	1.60	1.28	0.64	0.45^b	1.18	0.97	0.88	1.26	1.01	0.99
		3	1.56	1.65	1.08	0.58	0.44^b	1.13	0.97	0.86	1.18	0.96	0.96
		4	1.64	1.74	1.45	0.86	0.70	1.31	1.23	1.08	1.33	1.03	1.06
		5	1.63	1.66	1.55	0.88	0.79	1.39	1.29	1.17	1.35	1.06	1.06
		6	1.92	1.95	1.45	0.88	0.78	1.38	1.27	1.13	1.38	1.10	1.10
		7	1.92	1.94	1.60	1.03	0.99	1.46	1.37	1.24	1.47	1.12	1.10
		8	2.18	2.23	1.72	1.09	1.01	1.45	1.35	1.23	1.52	1.13	1.08
		9	2.12	2.02	1.74	1.08	0.97	1.63	1.45	1.24	1.52	1.16	1.17
		10	2.18	2.04	1.77	1.19	1.05	1.69	1.53	1.28	1.48	1.19	1.17
		11	2.01	1.84	1.78	1.16	1.00	1.79	1.66	1.36	1.51	1.26	1.19
		12	2.29	2.15	2.03	1.35	1.15	2.12	1.55	1.33	1.58	1.40	1.30
		BIC_6	1.91	1.91	1.54	0.88	0.78	1.42	1.31	1.16	1.29	1.00	1.01
		BIC_{12}	1.96	2.00	1.75	0.97	0.84	1.65	1.46	1.19	1.31	1.04	1.07

[a] MDM-NW significant 5%
[b] MDM-NW significant 10%
[d] GW significant 10%

Table A.3: Relative MSEs for the MIDAS-U projection with no lags in the GDP and $r = 1$

			backcast		nowcast			forecast			forecast		
			before release		current quarter			1 quarter			2 quarters		
		h_m	-2	-1	1	2	3	4	5	6	7	8	9
		K											
		2	1.20	1.89	1.28	1.48	2.88	0.81	1.19	2.26	0.87	1.06	1.93
		3	1.37	1.88	1.02	1.04	1.68	0.67	0.87	1.46	0.93	1.10	1.53
		4	1.39	1.80	1.07	0.90	1.25	0.93	1.01	1.22	1.13	1.27	1.59
		5	1.39	1.78	1.03	0.77	0.94	0.79	0.78	0.89	0.85	0.98	1.20
		6	1.58	2.05	1.10	0.79	0.94	0.86	0.88	1.00	0.90	1.06	1.41
		7	1.63	1.94	1.23	0.86	0.95	1.00	1.06	1.22	0.89	1.05	1.39
		8	1.55	1.79	1.25	0.82	0.85	0.91	0.92	1.02	0.96	1.07	1.41
	ALL	9	1.65	1.93	1.41	0.95	0.98	0.93	0.94	1.06	0.95	1.06	1.40
		10	1.69	2.02	1.51	1.01	1.04	0.94	0.97	1.10	0.86	1.04	1.38
		11	1.68	1.92	1.50	1.02	1.03	0.95	0.97	1.10	0.86	1.04	1.38
		12	1.82	2.12	1.59	1.13	1.14	0.95	0.97	1.09	0.86	1.03	1.37
		BIC_6	1.55	1.95	1.09	0.78	0.90	0.81	0.83	0.92	0.94	1.10	1.88
		BIC_{12}	1.79	2.07	1.66	1.11	1.13	0.94	1.00	1.13	0.92	1.07	1.83
		2	0.99	1.03	1.37	1.52	2.48	0.83	1.10	1.90	1.09	1.18	1.56
SW-EM		3	1.10	1.18	1.11	1.09	1.71	0.76	0.87	1.38	1.06	1.10	1.37
		4	1.24	1.39	1.17	0.98	1.45	0.95	1.09	1.49	1.18	1.14	1.39
		5	1.28	1.48	1.13	0.79	1.20	0.94	0.97	1.28	1.07	1.10	1.40
		6	1.41	1.72	1.13	0.73	1.15	0.95	1.00	1.46	0.98	1.05	1.42
		7	1.57	1.79	1.36	0.92	1.26	0.99	1.05	1.45	0.97	1.06	1.35
		8	1.52	1.71	1.39	0.88	1.16	0.98	1.03	1.39	0.97	1.06	1.38
	CL	9	1.60	1.80	1.53	1.00	1.31	0.98	1.03	1.43	0.97	1.05	1.36
		10	1.60	1.87	1.61	1.05	1.36	0.99	1.03	1.40	0.95	1.03	1.31
		11	1.65	1.91	1.65	1.09	1.36	0.99	1.04	1.40	1.06	1.04	1.31
		12	1.73	2.04	1.74	1.17	1.45	0.99	1.04	1.38	1.00	1.04	1.29
		BIC_6	1.41	1.73	1.13	0.73	1.15	0.92	0.99	1.46	1.05	1.08	1.52
		BIC_{12}	1.71	1.99	1.69	1.13	1.42	1.00	1.05	1.54	1.03	0.99	1.42
		2	1.11	1.31	0.79	0.54^b	0.54^b	0.81	0.80	0.93	0.99	0.85	1.01
		3	1.27	1.53	0.81	0.55^b	0.54^b	0.79	0.80	0.93	0.99	0.86	1.02
		4	1.33	1.56	1.01	0.66	0.67	1.05	1.02	1.14	1.08	0.94	1.10
		5	1.36	1.60	1.10	0.71	0.77	0.96	0.96	1.15	0.98	0.87	1.03
		6	1.52	1.81	1.21	0.83	0.94	1.03	1.03	1.23	0.98	0.87	1.03
		7	1.60	1.85	1.36	0.95	1.07	1.11	1.16	1.38	0.98	0.87	1.03
		8	1.56	1.79	1.44	0.99	1.11	1.06	1.09	1.30	0.99	0.87	1.03
	ALL	9	1.66	1.92	1.61	1.16	1.30	1.11	1.15	1.37	0.98	0.87	1.03
		10	1.68	1.94	1.69	1.21	1.36	1.14	1.20	1.42	0.98	0.86	1.03
		11	1.69	1.94	1.70	1.26	1.42	1.13	1.19	1.41	0.98	0.86	1.03
		12	1.80	2.05	1.79	1.39	1.56	1.15	1.21	1.44	0.98	0.86	1.02
		BIC_6	1.52	1.89	1.22	0.83	0.92	1.02	1.01	1.13	1.03	0.90	1.07
		BIC_{12}	1.76	2.00	1.83	1.37	1.54	1.09	1.20	1.37	1.02	0.88	1.06
		2	1.06	1.00	0.91	0.50^b	0.46^b	1.07	0.90	0.87	1.19	0.97	1.00
KAL-SS		3	1.19	1.17	0.96	0.54^b	0.50^b	1.04	0.91	0.90	1.17	0.96	1.01
		4	1.30	1.29	1.18	0.65	0.61	1.20	1.05	1.03	1.16	0.96	1.00
		5	1.38	1.38	1.29	0.73	0.70	1.21	1.07	1.07	1.17	0.96	1.01
		6	1.56	1.59	1.40	0.81	0.80	1.22	1.08	1.10	1.16	0.96	1.01
		7	1.67	1.71	1.56	0.94	0.94	1.25	1.15	1.17	1.15	0.95	1.00
		8	1.66	1.73	1.68	1.02	1.03	1.27	1.15	1.18	1.15	0.95	1.00
	CL	9	1.69	1.75	1.84	1.13	1.14	1.29	1.17	1.21	1.15	0.94	1.00
		10	1.70	1.77	1.89	1.19	1.21	1.29	1.18	1.23	1.16	0.93	0.99
		11	1.72	1.81	1.91	1.25	1.29	1.29	1.19	1.23	1.14	0.94	1.00
		12	1.81	1.89	1.97	1.34	1.37	1.31	1.20	1.25	1.14	0.94	1.00
		BIC_6	1.55	1.59	1.40	0.82	0.80	1.18	1.04	1.05	1.16	0.95	0.99
		BIC_{12}	1.73	1.83	1.92	1.31	1.34	1.27	1.19	1.22	1.14	0.93	0.99

[a] MDM-NW significant 5%
[b] MDM-NW significant 10%
[d] GW significant 10%

Table A.4: Relative MSEs for the MIDAS-Basic projection with no lags in the GDP and $r = 1$

			backcast before release		nowcast current quarter			forecast 1 quarter			forecast 2 quarters		
		h_m	-2	-1	1	2	3	4	5	6	7	8	9
		K											
SW-EM	ALL	0	0.84	0.92	1.38	0.68^b	0.67^b	0.82	0.62	0.60^b	0.71	0.57^b	0.55
		1	0.88	0.99	1.02	0.94	0.97	0.77	0.78	0.75	0.83	0.75	0.74
		2	0.87	1.59	0.93	0.65	1.40	0.67	0.56^b	1.05	0.69	0.60^b	1.36
		3	0.83	1.51	0.87	0.52^b	0.93	0.62	0.51^a	1.08	0.67	0.62	1.68
		4	0.92	1.60	1.08	0.77	1.05	0.73	0.52^b	1.03	0.88	0.71	2.23
		5	1.09	1.75	1.26	1.03	1.09	0.65	0.61	1.12	0.89	0.63^b	1.59
		6	1.05	1.72	1.44	0.98	1.09	0.68	0.73	1.20	1.12	0.68^b	1.66
		7	1.33	2.00	1.65	1.01	1.24	0.73	0.78	1.29	1.18	0.74	1.77
		8	1.13	1.65	1.58	1.10	1.36	1.17	1.27	1.53	0.90	0.68^b	1.65
		9	1.15	1.57	1.57	1.33	1.65	1.08	1.15	1.41	1.00	0.76	2.03
		10	1.62	2.08	2.14	1.89	2.43	1.05	1.30	1.77	1.02	0.82	1.85
		11	1.95	2.09	2.25	2.18	2.29	1.79	1.99	2.38	1.12	0.78	1.73
		12	2.94	3.32	3.14	3.46	3.34	1.87	2.35	2.68	1.40	1.32	1.98
		BIC_6	1.14	1.84	1.64	0.94	1.32	0.75	0.60^b	0.88	0.74	0.59^b	1.38
		BIC_{12}	1.14	1.84	1.64	0.94	1.32	0.75	0.60	0.92	0.68	0.64	1.39
	CL	0	0.84	0.93	1.61	0.78	0.74	0.79	0.65	0.58^b	0.91	0.75	0.69
		1	0.77	0.82	1.19	1.18	1.07	0.70	0.88	0.76	0.98	0.74	0.69
		2	0.83	0.83	1.13	0.78	1.18	0.55	0.57^b	1.42	0.92	0.71	1.06
		3	0.88	1.00	1.08	0.76	0.88	0.49^b	0.43^b	1.28	0.90	0.66	1.09
		4	0.89	1.03	1.15	0.89	0.95	0.55	0.50^b	1.63	0.89	0.87	2.50
		5	0.90	1.02	1.03	0.83	0.78	0.48^b	0.49^b	1.52	0.85	0.96	2.60
		6	0.91	1.01	1.07	0.82	0.91	0.51^b	0.55	1.73	0.95	1.16	2.91
		7	1.01	1.03	1.01	0.70	0.85	0.70	0.59	1.65	1.06	1.23	2.96
		8	1.17	1.39	1.18	0.88	1.50	0.67	0.70	1.34	1.35	1.44	3.31
		9	1.29	1.45	1.25	1.04	1.55	0.90	0.84	1.39	1.63	1.65	3.34
		10	1.87	1.93	1.73	1.38	1.27	1.07	1.01	1.49	1.66	1.57	3.08
		11	2.04	1.96	2.32	1.48	1.16	1.22	1.19	2.14	2.49	1.81	3.32
		12	3.16	3.19	2.53	2.33	1.30	1.44	1.42	1.97	2.24	1.46	3.27
		BIC_6	0.84	0.99	1.28	0.73	0.89	0.68	0.47^b	1.47	0.75	0.75	2.09
		BIC_{12}	0.84	0.99	1.28	0.73	0.89	0.74	0.54	1.46	1.01	1.01	2.28
KAL-SS	ALL	0	0.79	0.85	0.85	0.52^b	0.51^b	0.47^b	0.39^b	0.42^b	0.42^b	0.35^a	0.36
		1	0.82	0.90	0.81	0.50^b	0.49^b	0.60	0.49^b	0.54^b	0.66	0.54	0.59
		2	0.76	0.97	0.79	0.51^b	0.52^b	0.62	0.48^b	0.57	0.58	0.45^b	0.53
		3	0.72	0.97	0.79	0.52^b	0.54^b	0.58	0.43^b	0.51^b	0.55	0.42^b	0.49
		4	0.84	1.10	0.87	0.59	0.59	0.70	0.53^b	0.62	0.78	0.59	0.67
		5	1.04	1.22	1.01	0.79	0.72	0.70	0.57	0.59	0.82	0.61	0.62
		6	1.02	1.12	0.95	0.75	0.68	0.79	0.71	0.70	0.99	0.81	0.79
		7	1.29	1.29	1.17	0.80	0.79	0.89	0.91	0.96	1.10	0.89	0.97
		8	1.13	1.10	1.22	0.74	0.70	0.96	1.15	1.05	0.85	0.58^b	0.76
		9	1.14	1.02	1.41	0.86	0.79	1.06	1.19	1.15	0.94	0.66	0.87
		10	1.62	1.66	1.91	1.11	1.14	0.99	1.18	1.05	0.91	0.63	0.92
		11	1.99	1.96	2.17	1.57	1.62	1.14	1.44	1.24	0.92	0.81	1.08
		12	2.89	3.31	5.69	2.55	3.17	1.37	1.62	1.73	1.36	1.29	1.55
		BIC_6	1.10	1.21	0.99	0.74	0.65	0.62	0.48^b	0.55	0.65	0.65	0.65
		BIC_{12}	1.10	1.21	0.99	0.74	0.65	0.62	0.48^b	0.55	0.56	0.54^b	0.54
	CL	0	0.85	0.94	0.78	0.47^b	0.48^b	0.54	0.46^b	0.48^b	0.48^b	0.41^a	0.44
		1	0.88	0.88	0.90	0.44^b	0.45^b	0.75	0.52^b	0.54^b	0.75	0.58	0.59
		2	0.91	0.94	0.98	0.49^b	0.46^b	0.86	0.61	0.59	0.71	0.58	0.59
		3	0.94	1.01	0.97	0.53^b	0.51^b	0.80	0.63	0.60	0.64	0.54^b	0.57
		4	1.03	1.04	1.20	0.61	0.57	0.88	0.71	0.68	0.76	0.62	0.72
		5	1.05	1.01	1.21	0.63	0.56	0.83	0.74	0.71	0.82	0.61	0.65
		6	1.15	1.06	1.26	0.73	0.60	0.87	0.91	0.84	0.97	0.69	0.76
		7	1.29	1.17	1.23	0.68^b	0.60^b	1.00	1.06	0.95	0.96	0.73	0.80
		8	1.35	1.30	1.22	0.77	0.69	1.02	1.16	1.06	1.24	0.87	0.90
		9	1.74	1.60	1.17	1.01	0.96	1.35	1.33	1.15	1.42	1.13	1.20
		10	2.24	2.04	1.53	1.40	1.41	1.33	1.58	1.29	1.58	1.24	1.21
		11	2.50	2.19	1.61	1.23	1.16	1.78	2.34	1.55	1.41	1.46	1.22
		12	3.22	3.25	1.95	1.78	1.86	2.60	3.23	2.43	1.72	1.52	1.22
		BIC_6	1.00	0.99	1.16	0.53^b	0.48^b	0.87	0.60	0.61	0.59	0.48^b	0.55
		BIC_{12}	1.00	0.99	1.16	0.53^b	0.48^b	0.91	0.66	0.65	0.74	0.83	0.72

[a] MDM-NW significant 5%
[b] MDM-NW significant 10%
[d] GW significant 10%

Table A.5: Relative MSEs for the MIDAS-U projection with one lag in the GDP and $r = 1$

			backcast before release		nowcast current quarter			forecast 1 quarter			forecast 2 quarters		
		h_m	-2	-1	1	2	3	4	5	6	7	8	9
		K											
SW-EM	ALL	2	0.88	1.58	0.75	0.67	1.33	0.60^b	0.67^b	0.93	0.63	0.68	1.03
		3	0.86	1.25	0.78	0.60^b	0.82	0.59^b	0.59^b	0.78	0.70	0.72	0.82
		4	0.88	1.25	0.82	0.64	0.83	0.56^b	0.69^b	0.84	0.70	0.73	0.83
		5	0.85	1.23	0.81	0.62	0.76	0.50^b	0.53^b	0.68	0.69	0.73	0.82
		6	0.85	1.24	0.76	0.59^b	0.74	0.54^b	0.58^b	0.74	0.68	0.73	0.82
		7	0.89	1.27	0.80	0.63	0.77	0.56^b	0.60^b	0.78	0.68	0.81	0.88
		8	0.78	0.97	0.82	0.59	0.66	0.54^b	0.57^b	0.73	0.71	0.61	0.58
		9	0.80	1.02	0.77	0.59^b	0.67	0.52^b	0.54^b	0.70	0.81	0.75	0.70
		10	0.83	1.12	0.79	0.61	0.71	0.51^b	0.53^b	0.69	0.87	0.82	0.78
		11	0.78	0.94	0.82	0.60	0.68	0.52^b	0.56^b	0.73	0.87	0.81	0.77
		12	0.77	0.97	0.76	0.59^b	0.67	0.51^b	0.55^b	0.71	0.53^b	0.82	0.78
		BIC_6	0.89	1.60	0.78	0.61	1.15	0.52^b	0.56^b	0.79	0.61^b	0.68	1.03
		BIC_{12}	0.80	1.00	0.82	0.60	0.68	0.52^b	0.57^b	0.79	0.96	1.13	1.42
	CL	2	0.74	0.79	0.90	0.82	1.17	0.64	0.81	1.26	0.84	0.85	0.83
		3	0.77	0.80	0.90	0.76	1.00	0.62	0.85	1.26	0.90	0.83	0.81
		4	0.81	0.85	0.93	0.80	0.98	0.65	0.87	1.13	0.89	0.82	0.79
		5	0.78	0.86	0.88	0.71	0.92	0.62	0.81	1.06	0.89	0.82	0.79
		6	0.81	0.92	0.77	0.59	0.86	0.67	0.87	1.20	0.92	0.86	0.82
		7	0.86	0.93	0.88	0.70	0.93	0.65	0.87	1.24	0.83	0.66	0.58
		8	0.84	0.87	0.84	0.65	0.83	0.65	0.84	1.17	0.96	0.74^d	0.71
		9	0.86	0.92	0.82	0.71	0.96	0.64	0.85	1.23	1.11	0.87	0.85
		10	0.88	0.98	0.85	0.76	1.04	0.64	0.81	1.21	1.10	0.90	0.85
		11	0.84	0.94	0.82	0.71	0.95	0.66	0.86	1.26	1.11	0.91	0.86
		12	0.92	1.04	0.88	0.80	1.06	0.66	0.86	1.26	1.08	0.88^d	0.84
		BIC_6	0.79	0.90	0.82	0.61	0.88	0.63	0.87	1.28	0.86	0.84	0.83
		BIC_{12}	0.95	1.05	0.94	0.84	1.10	0.65	0.88	1.44	1.33	1.33	1.26
KAL-SS	ALL	2	0.77	0.98	0.79	0.52^b	0.53^b	0.58	0.48^b	0.57	0.55	0.45^b	0.53
		3	0.76	0.98	0.84	0.55^b	0.56	0.57	0.47^b	0.54	0.57	0.47^b	0.53
		4	0.81	0.98	0.87	0.59	0.59	0.58	0.51^b	0.60	0.57	0.48^b	0.53
		5	0.78	0.96	0.86	0.58	0.60	0.56	0.49^b	0.55	0.55	0.47^b	0.52
		6	0.77	0.95	0.83	0.57^b	0.58	0.58	0.49^b	0.56	0.55	0.47^b	0.53
		7	0.82	1.01	0.87	0.61	0.64	0.59	0.50^b	0.57	0.54	0.60	0.66
		8	0.75	0.88	0.89	0.60	0.62	0.57	0.49^b	0.55	0.69	0.71	0.78
		9	0.78	0.94	0.86	0.61	0.64	0.57	0.48^b	0.55	0.80	0.87	0.97
		10	0.78	0.97	0.86	0.62	0.68	0.56	0.48^b	0.55	0.85	0.92	1.04
		11	0.77	0.89	0.87	0.62	0.65	0.57	0.48^b	0.55	0.84	0.89	0.99
		12	0.74	0.88	0.86	0.59	0.63	0.56	0.48^b	0.54	0.85	0.89	0.98
		BIC_6	0.79	1.01	0.85	0.58	0.58	0.58	0.52^b	0.58	0.54	0.45^b	0.53
		BIC_{12}	0.79	1.10	0.90	0.62	0.64	0.58	0.52^b	0.60	0.89	1.00	1.15
	CL	2	0.85	0.88	0.85	0.48^b	0.49^b	0.69	0.53^b	0.55	0.66	0.55	0.59
		3	0.86	0.92	0.87	0.50^b	0.51^b	0.65	0.54^b	0.56	0.65	0.55^b	0.58
		4	0.89	0.94	0.94	0.53	0.54	0.70	0.56	0.58	0.64	0.54^b	0.58
		5	0.86	0.91	0.90	0.52^b	0.52	0.68	0.55	0.58	0.63	0.54^b	0.57
		6	0.89	0.94	0.90	0.52^b	0.52	0.71	0.56	0.59	0.63	0.54^b	0.57
		7	0.93	0.98	0.93	0.55	0.56	0.69	0.56	0.59	0.64	0.54^b	0.73
		8	0.92	0.97	0.93	0.57	0.57	0.69	0.56	0.59	0.95	0.95	0.97
		9	0.94	1.02	0.93	0.64	0.62	0.69	0.56	0.59	1.09	1.11	1.16
		10	0.94	1.00	1.02	0.67	0.64	0.69	0.56	0.61	1.04	1.05	1.10
		11	0.92	1.00	0.99	0.67	0.63	0.69	0.57	0.59	1.02	1.02	1.07
		12	1.00	1.07	1.07	0.77	0.74	0.69	0.56	0.61	0.98	0.91	0.96
		BIC_6	0.92	0.93	0.89	0.53	0.52	0.69	0.55	0.56	0.65	0.55	0.59
		BIC_{12}	1.05	1.06	1.06	0.77	0.73	0.69	0.56	0.57	1.04	1.09	1.16

[a]MDM-NW significant 5%
[b]MDM-NW significant 10%
[d]GW significant 10%

Table A.6: Relative MSEs for the MIDAS-Basic projection with one lag in the GDP and $r = 1$