Banco de México

Documentos de Investigación

Banco de México

Working Papers

N° 2017-17

# Are daily financial data useful for forecasting GDP? Evidence from Mexico

Luis M. Gomez-Zamudio
Banco Central de Chile

Raul Ibarra
Banco de México

September 2017

Documento de Investigación
2017-17

Working Paper
2017-17

# Are daily financial data useful for forecasting GDP? Evidence from Mexico*

Luis M. Gomez-Zamudio†
Banco Central de Chile

Raul Ibarra‡
Banco de México

**Abstract:** This article evaluates the use of financial data sampled at high frequencies to improve short-term forecasts of quarterly GDP for Mexico. In particular, the mixed data sampling (MIDAS) regression model is employed to incorporate both quarterly and daily frequencies while remaining parsimonious. To preserve parsimony, factor analysis and forecast combination techniques are used to summarize the information contained in a dataset containing 392 daily financial series. Our findings suggest that the MIDAS model that incorporates daily financial data lead to improvements for quarterly forecasts of GDP growth over traditional models that either rely only on quarterly macroeconomic data or average daily financial data. Furthermore, we explore the ability of the MIDAS model to provide forecast updates for GDP growth (nowcasting).

**Keywords:** GDP Forecasting, Mixed Frequency Data, Daily Financial Data, Nowcasting.

**JEL Classification:** C22, C53, E37.

**Resumen:** Este artículo evalúa el uso de datos financieros muestreados en altas frecuencias para mejorar los pronósticos de corto plazo del PIB trimestral para México. En particular, se emplea un modelo de regresión con muestreo de datos mixto (MIDAS, por sus siglas en inglés) para incorporar frecuencias tanto trimestrales como diarias mientras permanece parsimonioso. Para preservar parsimonia, se utilizan técnicas de análisis de factores y combinaciones de pronósticos para resumir la información contenida en una base de datos que contiene 392 series financieras diarias. Nuestros resultados sugieren que el modelo MIDAS que incorpora información financiera diaria conduce a mejoras en los pronósticos trimestrales del crecimiento del PIB sobre los modelos tradicionales que se basan únicamente en datos trimestrales macroeconómicos o que promedian datos financieros diarios. Además, exploramos la habilidad del modelo MIDAS de proporcionar actualizaciones de los pronósticos de crecimiento del PIB (nowcasting).

**Palabras Clave:** Pronósticos del PIB, Datos con Frecuencias Mixtas, Datos Financieros Diarios, Nowcasting.

*We thank Nicolás Amoroso, Santiago Bazdresch, Julio A. Carrillo, Yoosoon Chang, Bernardo Guimaraes, Juan R. Hernandez, Jorge Herrera, José Gonzalo Rangel, Abel Rodríguez, five anonymous referees and seminar participants at Banco de México, El Colegio de México and the 2016 Asian Meeting of the Econometric Society for valuable comments. Jose A. Jurado and Andrea Miranda provided excellent research assistance. Support provided by CONACYT is gratefully acknowledged. Luis M. Gómez-Zamudio contributed to this paper when he was working at Banco de México. The views on this paper correspond to the authors and do not necessarily reflect those of Banco de México.

† Banco Central de Chile. Email: lgomezz@bcentral.cl.

‡ Dirección General de Investigación Económica. Email: ribarra@banxico.org.mx.

# 1. Introduction

Forecasting GDP growth is important for policymakers, firms and investors in their decision making process. The global financial crisis of 2008-2009 together with the occurrence of the Great Recession have contributed to the need to reassess the role of financial markets to anticipate the business cycle (Espinoza et al., 2012). Financial variables are frequently associated with expectations of future economic events. For instance, stock prices can be interpreted as expected discounted values of future dividend payments, thus capturing future firms' profitability and future discount rates, which in turn are linked to the future growth rates of the economy. Similarly, interest rates can be interpreted as indicators of the stance of monetary policy, which can have effects on the real economy in the short term (Friedman and Schwartz, 2008). In the same way, commodity prices are associated with production costs and affect future growth, and exchange rate depreciations tend to encourage exports and thus output growth. However, the empirical evidence about the role of monthly or quarterly financial variables to forecast GDP growth is rather mixed or not robust (Stock and Watson, 2003; Forni et al., 2003).

Financial data are potentially useful for making predictions not only because of their forward looking nature, but also because there is a large number of series that are available on a continuous basis with no informational lag, as opposed to real activity data that are published with a significant delay. However, there are two main challenges that must be addressed to exploit this type of data. The first is the fact that financial information is sampled at a much higher frequency than macroeconomic variables (e.g., GDP). These macro variables are typically available on a quarterly basis, whereas many financial variables are sampled on a daily basis. The standard approach to use this information to make forecasts is to average the high frequency financial data in the quarter, i.e., a flat aggregation weighting scheme, to be able to estimate a regression with quarterly data. This method, however, might not be optimal, for instance, if more recent data are more informative. In this case, recent data should receive a higher weight than earlier data. A simple linear regression using each daily value of the predictor variable as an individual regressor would require estimating a large number of parameters, thus leading to high estimation uncertainty. One possible way to overcome this difficulty is to use the mixed data sampling (MIDAS) approach proposed by Ghysels et al. (2007). The MIDAS approach consists of regressions that allow the forecasted variable and the regressors to be sampled at different frequencies, using distributed lag polynomials to achieve parsimony. This family of models has been used in recent literature, such as in Clements and Galvão (2008) and Marcellino and Schumacher (2010), to improve the accuracy of predictions of quarterly GDP with monthly indicators for the US and Germany, respectively. More recently, the specific usage of financial data paired with the MIDAS model to forecast GDP growth in the US has been explored in Andreou et al. (2013). In short, these articles have concluded that the use of mixed frequency data improves forecast accuracy.

A second challenge is how to incorporate all the available information in such a way that the model remains parsimonious. In this regard, some methods are potentially useful to deal with large datasets of financial variables such as factor models and forecast combinations, as well as a wide variety of model parameterization options that considerably reduce the number of estimated coefficients. Factor models are useful to summarize the information content of large datasets with a few common factors (Stock and Watson, 2002a). The idea behind this framework is to extract the common component of a set of variables, filtering out the idiosyncratic variations that are uncorrelated. For instance, Stock and Watson (2002b), use a database containing 215 economic series such as real activity, prices and financial variables to extract a small set of factors. These factors in turn are used to construct forecasts for macroeconomic variables such as output and inflation which outperform alternative univariate and multivariate models. Forecast combinations have been found in empirical studies to improve accuracy over individual forecasts by exploiting information from a set of models rather than relying on a single model (Timmermann, 2006). Stock and Watson (2003) have suggested that, by combining forecasts from poorly performing models based on individual financial variables, the predictive role of financial information is rescued. In this paper, we employ factor models and forecast combinations as complementary approaches. That is, we use forecast combinations of MIDAS models estimated with a single daily financial factor in the spirit of the work by Andreou et al. (2013).

In this paper, we follow the forecasting approach proposed by Andreou et al. (2013) to investigate whether the use of financial variables and a MIDAS regression model lead to improvements in short-term forecasting of the Mexican GDP growth rate. For this purpose, a large set of 392 financial variables was obtained from Bloomberg. These variables can be grouped in the following categories: commodities, equities, corporate risk, foreign exchange and fixed income. The study period is from 1999 to 2013. This dataset will be used as the main information source. The financial variables that we select are frequently monitored by policymakers and practitioners and have been proposed in the literature as good predictors of economic activity. Because of the large number of variables, factor analysis is used to summarize all the information. Using these factors, the MIDAS model is estimated and forecasts are obtained for different specifications at horizons of one and four quarters ahead. The performance of the MIDAS models with financial variables is then compared to traditional factor models that only use quarterly macroeconomic data, which in turn have been successful in the literature to predict GDP growth (Stock and Watson, 2002b). For comparison purposes, we also provide benchmark models including random walk, autoregressive, vector autoregressive and Bayesian vector autoregressive models, as well as forecasts from the Survey of Professional Forecasters.[1] In addition, forecast combinations are carried out to further improve accuracy. We also present the GDP forecasts from a MIDAS regression model using a monthly dataset of macroeconomic variables as in Marcellino and Schumacher (2010). Thus, we are able to assess the role of daily financial variables compared to the approach of using only monthly variables.

This paper contributes to the literature in at least two important ways. First, to the best of our knowledge, this is the first paper applying the MIDAS approach to forecast GDP in a developing economy. In this way, we provide further evidence about the potential benefits

[1] There are alternative methods for using high frequency data to predict quarterly GPD growth, such as bridge models (Baffigi et al. 2004), state space models (Mariano and Murasawa, 2003) and factor models (Giannone et al., 2008). While bridge models and state space models rely on small sets of variables, factor models allow exploiting large datasets by summarizing the information into a few common factors. Our paper is focused exclusively on MIDAS models, although comparisons of forecasts from MIDAS models with some of these methods would clearly be of interest for future research.

of this recent methodology. This forecasting exercise is relevant because the volatility of economic and financial variables in these countries tends to be higher, which affects forecast accuracy. Although this might imply greater noise, it might also have relevant predictive content. In addition, as developing economies present lower levels of development in financial markets, financial variables will not necessarily have the same predictive role as in advanced economies. Second, this is the first paper that investigates whether financial variables have an important role at forecasting GDP growth in Mexico.

Our article examines three main questions about of the forecasting ability of the MIDAS model. First, we investigate whether the MIDAS model that incorporates daily financial data leads to improvements for quarterly forecasts of GDP growth over traditional models that rely only on quarterly macroeconomic data. Second, we would like to find out how the MIDAS model compares against a flat aggregation weighting scheme. Third, we explore the ability of the MIDAS model to provide forecast updates of GDP growth using recent information (nowcasting).[2]

The most important result is that the inclusion of daily financial data and the use of the MIDAS regression model help to improve GDP forecasting in Mexico. In particular, we find that the model with financial data and quarterly macroeconomic data outperforms a model that only employs quarterly macroeconomic variables. Furthermore, we show that the MIDAS model outperforms the flat aggregation scheme in terms of accuracy. The MIDAS model is useful to provide updates of GDP growth, although the forecasts with leads seem to have a similar predictive accuracy compared to the short-run forecasts without leads. Furthermore, in line with existing literature, we find that forecast combinations are effective at improving the predictive ability of a set of models. We conclude that the methodologies described herein are successful at incorporating additional information while preserving parsimony.

[2] Nowcasting refers to the process of updating the forecasts of the current quarter GDP growth as new information becomes available. For instance, if we are one month into the current quarter, that is, at the end of January, April, July or October, we will have one month of daily data to forecast quarterly economic growth.

The rest of the article is organized in the following way. Section 2 introduces the MIDAS regression model, factor analysis and forecast combination. An overview of the dataset is shown in Section 3. Section 4 presents the results. Section 5 concludes the article. Lastly, the Appendix provides a detailed description of the dataset and supplemental results.

## 2. Methodology

### 2.1. The MIDAS Model

Our methodology is based on the MIDAS model and follows closely the forecasting approach of Andreou et al. (2013). To illustrate the MIDAS model, consider two of the variables used in this article, the Mexican quarterly growth of GDP as the dependent variable and the daily return for the Mexican stock price index as the independent variable. GDP growth is sampled quarterly, while the GSCI index is sampled daily.

Now, define $Y_t^Q$ as the quarterly growth of GDP, and $X_{m,t}^D$ as the daily return for the Mexican stock price index, where $Q$ stands for quarterly, $D$ for daily and $m$ is the number of trading days in a quarter. Using this notation, a prediction of the GDP growth $h$ periods into the future with the model proposed by Ghysels et al. (2007) has the following form:

$$Y_{t+h}^{Q,h} = \mu^h + \sum_{j=0}^{p_Y^Q-1} \rho_{j+1}^h Y_{t-j}^Q + \beta^h \sum_{j=0}^{q_X^D-1} \sum_{i=0}^{m-1} w_{i+j*m}^{\theta^h} X_{m-i,t-j}^D + u_{t+h}^h.$$

This model has a constant, the traditional AR terms with $p_Y^Q$ quarterly lags of the dependent variable $Y_t^Q$, and a term that incorporates $q_X^D$ times $m$ daily lags for the independent variable. The term multiplying the daily variable $w_{i+j*m}^{\theta^h}$ deserves special attention. This term is the weighting scheme that will reduce the number of parameters to estimate and lead to a more parsimonious model instead of having to estimate a coefficient for each high frequency lag. The weights are normalized to sum up to unity in order to allow for the identification of $\beta^h$. Note that this model can be used to generate direct (rather than iterated) multiperiods ahead forecasts.

As explained in Ghysels et al. (2007), there are several weighting schemes, which are helpful to reduce the number of parameters to estimate. These include the unrestricted MIDAS, the normalized Beta probability function, the normalized exponential Almon lag polynomial, the Almon lag polynomial and the step functions. Excluding the U-MIDAS and the Almon lag polynomial, those schemes are estimated by nonlinear least squares.

We describe the Beta probability density function and the exponential Almon lag polynomial as they have been successful in the literature for forecasting purposes due to their parsimonious representation and flexible shapes (Andreou et al., 2013). The Normalized Beta probability function has the following form consisting of three parameters,

$$w_i(\theta_1, \theta_2, \theta_3) = \frac{a_i^{\theta_1 - 1}(1-a_i)^{\theta_2 - 1}}{\sum_{i=1}^{N} a_i^{\theta_1 - 1}(1-a_i)^{\theta_2 - 1}} + \theta_3,$$

where $a_i = \frac{(i-1)}{(N-1)}$, with $i = 1,2, \ldots, N$. This scheme can be made more parsimonious by restricting the first parameter to be one and/or the third parameter to be zero.[3] If all of these parameters are unrestricted, this weighting scheme is called Beta Non Zero. $N$ denotes the total number of high frequency lags used in the regression. For $\theta_1$=$\theta_2$=1, $\theta_3$=0, we have equal weights.

The normalized exponential Almon lag polynomial consists of two parameters represented as,

$$w_i(\theta_1, \theta_2) = \frac{\exp(\theta_1 i + \theta_2 i^2)}{\sum_{i=1}^{m} \exp(\theta_1 i + \theta_2 i^2)},$$

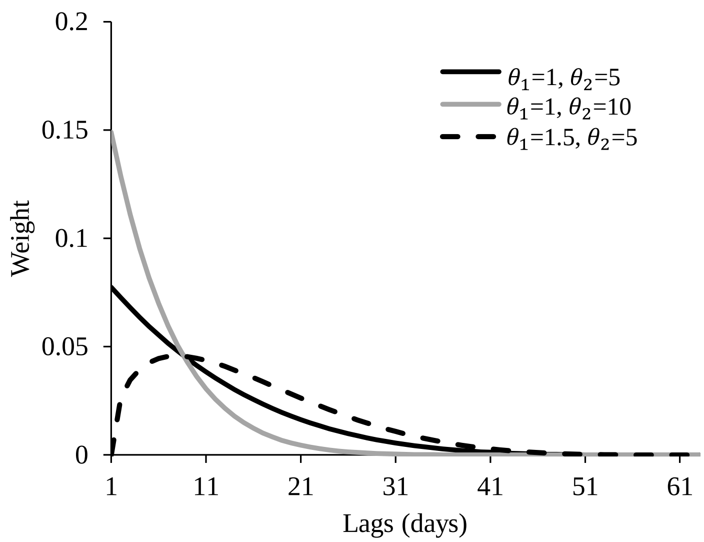
where $i = 1,2, \ldots, N$. As with the previous weighting scheme, the second parameter can be restricted to be zero.

As described in Ghysels et al. (2007), the exponential Almon lag and the Beta probability functions are flexible enough to accommodate various shapes, such as slow-declining, fast-

[3] The beta function described above follows from Ghysels (2015) and approximates the beta function described in Galvão (2013) as $\mathrm{Beta}(\theta_1, \theta_2) = \frac{a_i^{\theta_1 - 1}(1-a_i)^{\theta_2 - 1}\Gamma(\theta_1 + \theta_2)}{\Gamma(\theta_1)\Gamma(\theta_2)}$, where $\Gamma$ is the gamma function.

declining or hump-shaped patterns. A declining shape implies that recent information receives higher weight than earlier information. In contrast, the unrestricted MIDAS and the step-function schemes impose less structure on the function. Those schemes can be conveniently estimated through OLS, but require a larger number of parameters to estimate. We find that in most cases, the beta function performs better in terms of forecasting accuracy. Figure 1 shows various shapes of the Beta function for several values of the parameters, where the third parameter is restricted to 0. As can be seen, the rate of decay is governed by the values of the parameters. As a comparison, the more traditional way of using high frequency data is to make an average, which is called a flat aggregation scheme. In our case, that would mean averaging the GSCI daily index for each quarter, i.e., assigning the same weight to all the lags in a quarter. Although this scheme has been widely used in the literature, it may not be optimal for time series that exhibit memory decay. Thus, the MIDAS regression allows us to choose the optimal shape of the weights.

Figure 1: Beta probability weighting function




Note: The figure plots the weights on the first 63 lags of the beta probability function for different values of the parameters.

## 2.2. Factor Models

Following Stock and Watson (2002a), we use factor models to condense the information of a large number of variables into a few factors. Stock and Watson (2002b) have found that factor models are useful to improve the forecasts of key macroeconomic variables, such as output and inflation. The goal is to obtain a small set of factors that explains an important part of the variation in the entire set of variables. Formally, suppose there is a large set of variables *X* that will be used for forecasting. This set contains *N* variables with *T* observations each. It is possible that *N*>*T*. The goal is to find a set of factors *F* and a set of parameters Λ that best explain *X*.

The factor model can be written as:

$$X_t = \Lambda \mathrm{F}_t + e_t,$$

where $e_t$ are idiosyncratic disturbances with limited cross-sectional and temporal dependence. Another way to look at a factor is to think of it as an unobservable variable that explains an important part of the variation of the observed variables.

To estimate the factors, Stock and Watson (2002a) propose the use of the method of principal components which consists of minimizing the following expression:

$$V(\tilde{F},\tilde{\Lambda}) = (NT)^{-1}\sum_i \sum_t (X_{it} - \tilde{\lambda}_i \tilde{F}_t)^2,$$

subject to the normalization that $\frac{\tilde{F}'\tilde{F}}{T} = I_T$, where $\tilde{F} = (\tilde{F_1}\ \tilde{F_2} \cdots\ \tilde{F_T})'$, $\tilde{\lambda}_i$ is the ith row of $\tilde{\Lambda}$ and $I_T$ is the identity matrix. The estimated factor matrix is $\sqrt{T}$ times the eigenvectors corresponding to the $r$ largest eigenvalues of the $T \times T$ matrix $XX'$. This method produces a set of orthogonal factors that can be ordered according to their contribution to the overall variance of the entire set of variables.

Most of the literature has focused on extracting factors at low frequencies, such as quarterly or monthly data. Following this approach, we will extract factors from a large set of daily financial variables. Once the factors are estimated, they are incorporated into the MIDAS regression as a high frequency variable. For instance, if we use the factor that explains the largest variation of the entire set of financial variables, denoted as $F^1$, as the high frequency regressor, our MIDAS regression model can be written as:

$$Y_{t+h}^{Q,h} = \mu^h + \sum_{j=0}^{p_Y^Q-1} \rho_{j+1}^h Y_{t-j}^Q + \beta^h \sum_{j=0}^{q_X^D-1} \sum_{i=0}^{m-1} w_{i+j*m}^{\theta^h} F_{m-i,t-j}^1 + u_{t+h}^h.$$

In our case, the first factor accounts for 23% of the variability of the 392 daily time series used. The first 5 factors explain 42.7% of underlying variation. Section 4 presents more details about the dataset. To preserve parsimony, we consider forecasting models that include the daily financial factors one at a time, and follow the approach of Andreou et al. (2013) by using forecast combinations of these models that include a single factor.

Following Marcellino et al. (2003), the series are standardized before the factors are obtained, by subtracting their means and dividing by their standard deviations. This is necessary as a wide variety of series are employed and they differ in their units of measurement. In addition, the series are transformed to achieve stationarity, if necessary. Following Stock and Watson (2002a, 2008), the principal components method that we use to estimate the factors is at the same time parsimonious and robust to having temporal instability in the model, as long as the instability is relatively small and idiosyncratic. That is, the estimated factors and forecasts are consistent even in the presence of time variation in the model (Stock and Watson, 2002a). An alternative method proposed by Forni et al. (2000) is to extract the principal components from the frequency domain using spectral methods. However, Boivin and Ng (2005) find that the method of Stock and Watson has smaller forecast errors in simulations as well as in empirical applications. By imposing fewer constraints and having to estimate a smaller number of auxiliary parameters, this approach seems to be less vulnerable to misidentification and produces better forecasts than the method of Forni et al. (2000).

An important issue is the determination of the number of factors to include in the model. For this purpose, we evaluate the marginal contribution of each principal component in explaining the total variation of the series. As a result, we use 3 quarterly macro factors, which explain nearly 76% of the total variation of the 20 macroeconomic series.[4] Similarly, we use 5 daily financial factors, which explain a sufficiently large percentage of the total variation of the 392 financial series (43%).

### 2.3. Forecast Combinations

To employ the information contained in several of the estimated factors without increasing the number of parameters in the model, we use forecast combination methods. By preserving parsimony, we achieve lower parameter uncertainty, thus improving forecasting accuracy. Hence, forecast combinations deal with the problem of model uncertainty by using information from alternative models instead of focusing on a single model. A survey on forecast combination methods can be found in Timmermann (2006).

As a general result in the literature, forecast combinations improve forecast accuracy (Timmermann, 2006). Following Andreou et al. (2013), we present a few combinations that improve the forecasting accuracy of the individual predictions. Formally,

$$\hat{Y}^{Q,h}_{C_M,t+h} = \sum_{i=1}^{M} w^h_{i,t} \hat{Y}^{Q,h}_{i,t+h}.$$

Thus, a forecast combination $\hat{Y}_{C_M,t+h}$ can be interpreted as a weighted average of the $M$ forecasts $\hat{Y}_{i,t+h}$ for the horizon $h$ of $M$ models. Again, an important decision is to select the weighting scheme. For this purpose, we need to think in terms of a loss function. Formally, a combination of $n$ forecasts is preferred to a single forecast if,

$$E\left[\mathcal{L}\left(\hat{Y}^{Q,h}_{i,t+h}, Y_{t+h}\right)\right] > \min_{C(\cdot)} E\left[\mathcal{L}\left(C\left(\hat{Y}^{Q,h}_{1,t+h}, \hat{Y}^{Q,h}_{2,t+h}, \dots, \hat{Y}^{Q,h}_{M,t+h}\right), Y_{t+h}\right)\right],$$

for $i = \{1,2,\dots,M\}$.

[4] We find that the estimated factors from our set of macroeconomic variables are highly related to relevant subsets of key macroeconomic variables such as output and inflation. In particular, the first factor correlates highly with inflation, while the second factor correlates highly with output growth. That is, the estimated factors seem to be informative and interpretable from an economic point of view. Regarding the financial factors, the first factor is highly correlated with equities, and the second factor is correlated with fixed income and commodities.

In the inequality above: $\mathcal{L}$ is a loss function that relates the forecasted and the observed values. Intuitively, the loss function is expected to grow as the forecasted value drifts further from the actual value. $C$ on the other hand, is the combination function that relates the individual forecasts. Thus, we would like to select a function $C$ that minimizes the expected loss, and the forecast combination would be preferred if the expected value of the loss function for that combination is smaller than each of the expected losses for each of the individual forecasts.

Given the previous assumptions, the solution is a linear combination of individual forecasts. To finish this derivation let us denote $\widehat{\boldsymbol{Y}}_{t+h}^{Q,h}$ a vector containing all individual forecasts and $\boldsymbol{w}_{t+h}^{h}$ a vector of parameters. Then, the combination function can be rewritten as $C\left(\widehat{\boldsymbol{Y}}_{t+h}^{Q,h}\,;\boldsymbol{w}_{t+h}^{h}\right)$. The last step requires to define a loss function. Following Andreou et al. (2013), the Mean Squared Forecast Error (MSFE) is used as it has been found to provide the highest improvement in forecasts. Thus, the MSFE weights are selected by analyzing the historical forecasting performance of the model and assigning to each of them a weight inversely proportional to their MSFE.

### 2.4. Forecasting with Leads (Nowcasting)

The MIDAS models have the ability of incorporating recent information to improve the forecasts. To understand this, suppose that current quarter GDP growth needs to be predicted. If we are one month into the current quarter, that is, at the end of January, April, July or October, we will have about 21 trading days (1 month) of daily data to forecast quarterly economic growth. Using the information up to date to forecast the next value of a variable of interest is called nowcasting.

Formally, the MIDAS model is augmented with leads in the following way:

$$Y_{t+h}^{Q,h} = \mu^h + \sum_{j=0}^{p_Y^Q-1} \rho_{j+1}^h Y_{t-j}^Q + \beta^h \left[ \sum_{i=(3-J_X)*\frac{m}{3}}^{m-1} w_{i-m}^{\theta^h} X_{m-i,t+1}^D + \sum_{j=0}^{q_X^D-1} \sum_{i=0}^{m-1} w_{i+j*m}^{\theta^h} X_{m-i,t-j}^D \right] + u_{t+h}^h$$

The new term has two noticeable aspects. First, the subindex *t+1* for the financial variable $X^D$ implies that the forecasting equation includes high frequency information generated during the present quarter. The other important thing to notice is the values of *i* and $J_X$. Let's suppose *m*=63, that means there are 63 trading days in a quarter. $J_X$ denotes the number of months of current quarter information available at the daily frequency. Accordingly, if the first month of the quarter has just finished, there are 21 days of data available, thus, $J_X = 1$ needs to be selected to obtain the appropriate limits of the sum.

As opposed to traditional nowcasting that involves state-space models potentially implying a large number of parameters and measurement equations, the MIDAS approach provides a parsimonious framework to deal with a large number of high frequency predictors. The advantage of using financial data is that they are not subject to revisions as occurs with many real activity variables. Thus, in our model financial data absorb the news into asset prices to provide forecast updates of GDP growth.

### 2.5. Forecast Evaluation

To compare the forecasting ability of alternative models, we use the Diebold and Mariano (1995) test. That is, we test for the null hypothesis that two different models have the same forecasting ability. To that end, we define a quadratic forecast loss function for model $i$ as $g(u_{i,t}) = u_{i,t}^2$. Under the null hypothesis, both models have equal forecasting ability, that is:

$$H_o: g(u_{1,t}) = g(u_{2,t})$$

Diebold and Mariano (1995) first define the difference between the loss functions for two alternative models as $d_t = g(u_{1,t}) - g(u_{2,t})$. Then, they propose the following test statistic:

$$DM = \frac{\bar{d}}{\sqrt{var(\bar{d})}},$$

where $\bar{d}$ is the sample mean of $d_t$ and $\sqrt{var(\bar{d})}$ is defined as $\sqrt{var(\bar{d})} = \frac{\gamma_0 + 2\gamma_1 + \cdots + 2\gamma_q}{H-1}$. H is the number of forecasted periods and $\gamma_j = cov(d_t, d_{t-j})$. The statistic has a t-student distribution with H-1 degrees of freedom. The p-values shown later in the paper are derived from a regression with Huber-White robust errors of $d_t$ on a constant and testing whether the constant is statistically significant.

### 2.6. Alternative Models

To analyze the relative performance of the MIDAS model, we estimate the following alternative models: an autoregressive (AR), a random walk (RW), a vector autoregressive (VAR) and a Bayesian vector autoregressive (BVAR) model. We also compare our results to the Survey of Professional Forecasters. The aforementioned models and survey have been widely used by both central banks and the empirical literature as benchmarks for GDP forecasting (Chauvet and Potter, 2013). The order of the AR model was chosen using the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC), resulting in one autoregressive lag. Both the AR and the RW models contain seasonal dummy variables. For all cases, the dependent variable is the quarterly growth of GDP.

VAR models represent a systematic way to capture the dynamics and comovements of a set of time series without restricting for a specific functional form and have been particularly useful for forecasting purposes since the influential paper by Sims (1980). The VAR model can be written as:

$$Y_t = A_0 + \sum_{i=1}^{p} A_i Y_{t-i} + \varepsilon_t,$$

where $Y_t$ is the vector of variables being forecasted, $A_i$ are the matrices of coefficients to estimate and $\varepsilon_t$ is a vector of residuals. The variables included in the VAR model are the growth rate of GDP, quarterly inflation rate, interest rate and US GDP growth rate.[5] To determine the number of lags $p$ we use the AIC and set the maximum number of lags to four. The model can also contain seasonal dummy variables that are not included in the equation above for simplicity.

A limitation of VAR models is that they often imply a large number of parameters to estimate, resulting in a loss of degrees of freedom, thus leading to inefficient estimates and lower

[5] Herrera-Hernandez (2004) and Capistrán and Lopez-Moctezuma (2010) find that US GDP is useful to improve Mexican GDP forecasts in a VAR framework.

forecasting performance. To deal with this limitation, we estimate a Bayesian VAR (BVAR) model (Litterman, 1986; Doan et al., 1984). The idea is to use an informative prior to shrink the unrestricted VAR model towards a parsimonious naïve benchmark, thus reducing parameter uncertainty and improving forecasting accuracy. Previous studies have found that BVAR models have a good forecasting performance compared to conventional macroeconomic models for different countries and periods, including Litterman (1986), McNees (1986), Artis and Zhang (1990), Bańbura et al. (2010), among others.

A BVAR model requires specifying the mean and standard deviation of the prior distribution of the parameters. In particular, we follow the Minnesota prior, in which each variable follows a random walk around a deterministic component (Litterman, 1986). If the model is specified in first differences, this prior specification shrinks all of the elements of $A_i$ for the previous VAR model toward zero. This implies that each variable depends mainly on its own first lag. In addition, the Minnesota prior incorporates the belief that more recent lags should provide more reliable information than more distant ones and that own lags explain more of the variation of a given variable than lags of other variables in the equation. The prior beliefs are imposed by setting the following moments for the prior distribution of the parameters:

$$E\left[(A_k)_{ij}\right] = 0, \qquad V\left[(A_k)_{ij}\right] = \begin{cases} \dfrac{\lambda^2}{k^{2\tau}}, & j = i \\ \dfrac{\lambda^2\gamma^2}{k^{2\tau}}\dfrac{{\sigma_i}^2}{{\sigma_j}^2}, & \text{otherwise} \end{cases}$$

Thus, the Minnesota prior can be described by three hyperparameters, the overall tightness parameter $\lambda$, the relative cross-lags parameter $\gamma$ and the decay parameter $\tau$. Changes in these parameters imply changes in the variance of the prior distribution. The overall tightness parameter $\lambda$ indicates the tightness of the random walk restriction, or the relative weight of the prior distribution with respect to the information contained in the data. For $\lambda = 0$, the data does not influence the estimates. As $\lambda \rightarrow \infty$, the posterior estimates converge to the OLS estimates. The parameter $\gamma < 1$ indicates the extent to which the lags of other variables are less informative than own lags. The parameter $\tau \geq 0$ captures the extent to which more recent lags contain more information than more distant ones. Thus, the factor $1/k^{2\tau}$ represents the rate at which prior variance decreases with increasing lag length. ${\sigma_i}^2/{\sigma_j}^2$ accounts for the different scale and variability of the series. $\sigma_i$ and $\sigma_j$ are estimated as the standard errors of an univariate AR regression for each variable. Finally, we use a non-informative (diffuse) prior for the deterministic variables. The BVAR model is estimated using Theil's mixed estimation method (Theil and Goldberger, 1961).

The hyperparameters are chosen based on forecasting performance. In particular, we estimate the BVAR model for the combinations resulting from setting the following parameters: $\lambda$={0.1,0.2}, $\gamma$={0.3,0.5}, $\tau$=1, and the number of lags p={1,2,3,4}.[6] From these 16 combinations of hyperparameters, we select the combination that minimizes the RMSFE in a pseudo out-of-sample forecasting exercise.

To provide further evidence of the forecasting accuracy of the MIDAS model, our forecasts are also compared with those of the Survey of Professional Forecasters, which is maintained by Banco de Mexico. Capistrán and Lopez-Moctezuma (2010) find that the forecasts from this survey outperform forecasts from traditional univariate and multivariate time series models. There are about 30 survey participants, including financial, consulting and academic institutions. Capistrán and Lopez-Moctezuma (2010) provide an in depth description of this survey. We use the consensus forecast for the GDP growth rate, defined as the mean across forecasters. For the forecasting period used in this paper, the data are only available at the one quarter ahead horizon.

## 3. Data

We use three databases in our analysis at different sampling frequencies: daily, monthly and quarterly. The daily database is divided into 5 different categories of financial information: commodities (166 series), equities (94 series), foreign exchange (27 series), corporate risk (53 series) and fixed income (52 series). As previously stated, the dependent variable is the Mexican GDP. These daily financial series have been found in the literature to be good predictors of output growth (Andreou et al., 2013). The study period is from 1999Q1 to

---

[6] Those values for the hyperparameters have been used in previous literature (e.g., Dua and Ray, 1995; LeSage, 1999; Canova, 2007).

2013Q4. The initial estimation period is from 1999Q1 to 2009Q4 and the period of forecasting is 2009Q4+h to 2013Q4. Although the sample is relatively small for nonlinear least squares estimation, Bai et al. (2013) provide evidence based on Monte Carlo simulations showing that the forecasting performance of MIDAS regression models may not be affected.

The time series of the Mexican GDP, though not as long as that of developed countries, is available since 1993. Nevertheless, the estimation period is effectively shorter because an important number of financial variables is available from 1999 onwards. Although it might be a short period for forecasting purposes, it allows for the inclusion of useful daily information. Moreover, we use a sample period during which Mexico has followed exclusively a floating exchange regime and exclude the 1995 economic crisis from the estimation period, which could affect our estimations. The adoption of a flexible exchange rate implies lower output volatility as the economy is less affected by external shocks, which affects economic growth (Levy-Yeyati and Sturzenegger, 2003). In addition to the adoption of a flexible exchange rate, other reforms were implemented after the 1995 financial crisis that have promoted the development of financial markets in Mexico, particularly for derivative markets, pension funds, government securities and the banking system (Sidaoui, 2006). The development of financial markets is possibly associated with a more important role of financial variables to predict GDP.

The database constructed is primarily a subset of the time series suggested by Andreou et al. (2013), which has been shown to provide good predictive content for US GDP. Nonetheless, there are a few notable remarks regarding the Mexican data. First, the CETES 28 day rate is included in the fixed income group.[7] It is especially important to include this information since the interest rate is the monetary policy instrument for Mexico. In turn, the 28 day CETES rate mimics the behavior of the interest rate target. Second, the foreign exchange rates are expressed in terms of Mexican pesos. Furthermore, in terms of equity, we use two indexes of the Mexican Stock Market, IPC and INMEX. Finally, some of the financial variables specific for Mexico that could be relevant to forecast GDP, such as corporate bonds

[7] CETES are Mexican Treasury Bills, that is, debt issued by the federal government through the Ministry of Finance and Banco de México (the Mexican Central Bank).

and commodities, are unavailable for the entire study period and thus were not included in our study. For detailed information concerning the series used, please refer to the Appendix. All the financial information was retrieved from Bloomberg.

Following Marcellino et al. (2003), some of the series employed here were transformed because they were nonstationary. For each variable, we tested the hypothesis of unit root by means of an augmented Dickey Fuller (ADF) test with 12 lags. Nonstationary series were transformed to first log-differences. Then, to ensure stationarity the transformed series were tested for unit roots using the ADF test. In general, we transform commodity prices, stock prices and exchange rates into daily returns (i.e., first log differences). Interest rates for US corporate bonds are transformed to first differences. Domestic interest rates are found to be stationary in levels. The forecasting variable, i.e., the GDP growth rate, is not seasonally adjusted. Therefore, regressions are estimated using seasonal dummy variables. However, we find that the results are robust if we use seasonally adjusted data.[8]

Another important set of information included in our regressions is the quarterly macro data. This set comprises 20 macro variables whose high explanatory and predictive power for GDP has been previously documented (Andreou et al., 2013). In particular, this set contains information such as price indexes, international trade variables, inflation rate and economic activity indexes for Mexico and the US. Part of this set of variables is available on a monthly basis. To transform these variables into quarterly data, monthly data are averaged for every quarter. The macroeconomic variables are also transformed if necessary to achieve stationarity, as indicated by an ADF test. In general, real activity variables, prices, and monetary aggregates are transformed into quarterly growth rates.

In addition to the daily set of financial variables and the quarterly set of macroeconomic variables, a dataset of monthly macro data is used as the high frequency data for the MIDAS

[8] Following previous studies on forecasting, including Stock and Watson (2002) and Marcellino et al. (2003), we have not filtered the series using the method by Hodrick and Prescott (1997). As shown by Cogley and Nason (1995), when the HP filter is applied to integrated processes, it can generate business cycle fluctuations even if they are not present in the original series, which would potentially misguide our forecasts.

regression. This set consists of 18 variables, such as price indexes, economic activity indexes for Mexico and US CPI. The same procedure is followed to preserve parsimony, i.e., a set of factors is estimated and different forecasts using each factor are combined to obtain the final forecast. The Mexican data were obtained from the National Institute of Statistics and Geography (INEGI, by its acronym in Spanish) and Banco de México (the Mexican Central Bank). US data were retrieved from the Federal Reserve Bank of St. Louis database (FRED).

## 4. Results

### 4.1. Forecasting Exercise and Model Selection

Before presenting results for the forecasting exercise, a few points that require further clarification will be discussed. First, a recursive window is used for all the model specifications and horizons. For instance, consider the forecast $i$, with $i = 0, 1, .., n - 1$, where $n$ is the number of one-step ahead forecasts. Then, the start date of estimation is fixed at 1999Q1, whereas the end date changes with each forecasted value, which is 2009Q4+i. Thus, the model is estimated each time the window changes and the forecasts are computed one-step ahead. This window grows with each forecasted point as it includes the next observed value. The recursive window is expected to improve the forecasts over a fixed estimation window, as each new estimation includes more recent information.

The second important aspect is to specify whether the exercise is in real time or not. That is, as GDP is subject to revisions (as well as other macroeconomic variables used as regressors), the data actually available at a particular quarter may differ from the final values that will be released by statistical offices. Although it would be of interest to perform a real time forecasting exercise by using the vintages of data that were actually available to the forecasters, real time data for Mexico are unavailable. Thus, we use revised data in our estimations. Notice that our models are still comparable in the forecasting evaluation exercise as all of them use the same information. In addition, this issue is of less relevance in our case as daily financial data are not subject to revisions. Thus, as in Stock and Watson (2003) we follow the view that the best way to evaluate a predictive relationship is to use final data rather than early vintages of GDP.

As discussed in the introduction section, an interesting capability of the MIDAS model is nowcasting, which allows to forecast using current date information. We perform a forecasting excise for GDP growth using information one month farther into the quarter (i.e., 21 trading days), at horizons of one and four quarters ahead. We have also analyzed the forecasts using 2 months of leads of financial data. The conclusions are similar to those reported in this paper.

Since there is a wide variety of specifications available, we use the AIC and BIC to select the number of lags for both the autoregressive terms and the high-frequency terms. In our preferred forecasting framework, we use the information from five factors. In particular, we follow a similar approach as in Andreou et al. (2013) and use a forecast combination from the five models estimated with each of the five factors extracted from the entire set of financial variables. That is, we use both factor models and forecast combinations to deal with the large dataset of financial variables. To determine the number of daily factors, we consider the marginal contribution of each factor to explain the total variation of the series. We find that five factors explain a sufficiently large percentage of the variation.

We use the beta function as it presented in most cases the lowest RMSFE. In addition, the variance of the RMSFE of this weighting scheme is smaller. The tests to identify the best models were implemented using a maximum of 5 lags of the dependent variable and 1 to 6 quarters of information of the independent factor ($q_X^D$). As the number of trading days in a quarter is $m$=63, the maximum number of daily lags is 63×6=378. The selection of the models was done following the AIC and BIC. As explained before, regardless of the high frequency lags specified, the model estimates only 2 parameters for the Beta weighting scheme.

### 4.2. Forecasting Results for Models with Daily Financial Factors

Table 1 presents the RMSFE for different specifications estimated for two different forecasting horizons: 1 quarter ahead ($h = 1$) and 1 year ahead ($h = 4$). Out of the alternative benchmark models, the BVAR and the SPF have the best forecasting performance. The RW model shows the highest RMSFE. The forecasting accuracy of the BVAR model is consistent with previous studies for different countries, including Artis and Zhang (1990) and Bańbura et al. (2010).

The Table also presents the relative RMSFE of the MIDAS model with respect to the benchmark AR model. The optimal number of lags according to the BIC is shown in parenthesis. As can be seen, the RMSFE of the MIDAS model that employs the first factor is outperformed by the benchmark models. A possible explanation is that the benchmark models contain macroeconomic variables that have a good predictive content to forecast GDP which are not contained in the MIDAS model. In the last part of this subsection, we will present an exercise that incorporates macroeconomic variables into the MIDAS model to provide evidence of the forecasting ability of this methodology and the use of high frequency data.

Factor estimation is also applied to each group of financial variables. From this decomposition, 5 factors are extracted, one for each of the 5 groups of financial variables. Table 1 shows the forecasting results with the first factor of each group. We use the Beta weighting scheme and select the number of lags using the Akaike and Bayesian Information Criteria. We only include the first factor in each regression as the variables in each group are highly related among them. Even though this is a parsimonious weighting specification, the predictive power for all variable groups, except for exchange rates, do not seem to improve over the benchmark models. In other words, the uncertainty associated with parameter estimation for these specifications outweighs the additional predictive power incorporated through the individual sets of financial series. The role of the exchange rates to forecast GDP could be explained in part by the status of Mexico as a small open economy. Exchange rate depreciations tend to encourage exports and thus increase output growth.

Table 1: RMSFE comparison for models with no leads

| | Model | h=1 RMSFE | h=1 RMSFE as % of AR | h=4 RMSFE | h=4 RMSFE as % of AR |
|---|---|---|---|---|---|
| Alternative models | AR | 1.1348 | 1.0000 | 1.1136 | 1.0000 |
| | RW | 1.2890 | 1.1359 | 3.2330 | 2.9032 |
| | VAR | 1.1156 | 0.9831 | 1.2112 | 1.0877 |
| | BVAR | 0.9285 | 0.8182 | 0.9899 | 0.8889 |
| | SPF | 0.9688 | 0.8537 | | |
| Factor 1 | Beta (p=2, q=6) | 1.6978 | 1.4961 | 1.8168 | 1.6315 |
| Commodities F1 | Beta (p=1, q=1) | 1.5170 | 1.3368 | 1.3902 | 1.2483 |
| Equities F1 | Beta (p=3, q=5) | 1.4217 | 1.2528 | 1.5319 | 1.3756 |
| Corporate F1 | Beta (p=1, q=2) | 1.4375 | 1.2667 | 1.5536 | 1.3951 |
| FX F1 | Beta (p=1, q=1) | 1.0367 | 0.9135 | 1.0429 | 0.9365 |
| Fixed Income F1 | Beta (p=1, q=5) | 1.9653 | 1.7319 | 1.9792 | 1.7773 |
| Forecast Combinations | | | | | |
| Factors 1 to 5 | Beta Best AIC/BIC | 1.0453 | 0.9211 | 1.1936 | 1.0718 |

Note: The table shows the root mean square forecast error (RMSFE) for $h = 1$ and $h = 4$ step ahead horizons of the GDP for the sample 1999Q1-2013Q4. The study period is from 1999Q1 to 2013Q4. The initial estimation period is from 1999Q1 to 2009Q4 and the period of forecasting is 2009Q4+h to 2013Q4. The RMSFE are also presented as a percentage of the AR. First, the forecasts are estimated for each of the alternative models described in the paper. Second, the table shows the results for the MIDAS model using the first daily factor of the 392 financial variables shown in Appendix A. Then, the forecasts are also estimated using the first factor of each group of financial variables. Finally, a forecast combination based on the first five factors is presented. A recursive window is used for all estimations.

While it is readily apparent from the dataset that corporate risk and fixed income are two groups that focus mainly on the US economy and even though there are some variables such as interest rates for the Mexico, these do not seem to provide sufficient information to predict Mexican GDP growth by themselves. Equities might also present a similar problem.

The last section of the table presents a forecast combination based on the MSFE using five MIDAS specifications, one for each of the first five factors. Each of these models is optimal in the AIC-BIC sense but for different factors. As expected, the combination yields a lower RMSFE. This improvement can be explained by the fact that it considers the information contained in each factor. Thus, although forecasting accuracy does not seem to improve when the individual groups of financial variables are included in the model, when all variables are included together and the factors contain mixed information it is clear that they are successful at improving the forecasting accuracy of the model. These findings are consistent with Stock and Watson (2003), who have found that the predictive power of financial variables is rescued by combining forecasts based on individual variables.

### 4.3. Forecasting Results for Models with Daily Financial Factors and Quarterly Macroeconomic Factors

The goal of the final part of this section is to investigate whether introducing daily financial data into a MIDAS regression framework is useful for forecasting GDP beyond macroeconomic data. We also compare the forecasting accuracy of the MIDAS model with the traditional models that take a simple average of daily financial data, i.e., a flat aggregation scheme.

Table 2 contains a summary of the RMSFE for several models. The model denoted as *FAR (factor autorregresive)* under *quarterly macro data* incorporates the quarterly macro data to the AR model using a factor model as in Stock and Watson (2002b). In particular, we extract 3 quarterly macroeconomic factors from the database of 20 quarterly macroeconomic series described in the data section. As a result, the first 3 factors explain nearly 76% of the overall variation of the 20 quarterly macroeconomic series.[9] These estimated factors augment the benchmark AR model to obtain the FAR models. A second family of MIDAS models is

[9] Ibarra (2012) finds that, for the case of Mexico, the estimated factors from a broad set of macroeconomic variables for the period 1992-2009 are highly related to relevant subsets of key macroeconomic variables such as output and inflation. That is, the estimated factors seem to be informative and interpretable from an economic point of view. Our results are consistent with those findings.

presented as *monthly macro + quarterly macro data*. This family consists of models where the high frequency variables used for the estimation of the model are the same set of monthly macroeconomic variables that were averaged using a flat aggregating scheme. *Flat* is used to denote the family of models that use a flat aggregation scheme for high-frequency data as well as quarterly macro data. In other words, the values for all trading days of the daily financial assets within the quarter were averaged to obtain a single value per quarter.[10] *Combined MIDAS* is used to refer to a combination using the MSFE of 5 MIDAS specifications: one for each of the first 5 factors. Finally, the models denoted as *financial data* incorporate the information contained in the 392 daily financial series.

As before, the RMSFE for different specifications is presented in Table 2. The results show that adding quarterly data to the AR model improves forecasting accuracy in terms of the RMSFE at both horizons. In particular, the factor model that includes quarterly macroeconomic data outperforms the AR, VAR, BVAR and SPF forecasts. That is, quarterly macroeconomic data such as consumption, investment, trade, inflation and foreign macroeconomic variables seem to provide important information to predict future GDP. Similarly, the monthly macro data improve forecast accuracy relative to the AR model at the one quarter ahead horizon. However, the gains in terms of accuracy seem to be lower compared to those obtained from adding quarterly macroeconomic data, possibly due to the estimation uncertainty associated with a larger number of parameters.

The results of adding financial data are of particular interest. As the RMSFE for these specifications show, including these variables within a combined MIDAS model improves forecasting accuracy. Notably, the results also suggest that gains in terms of RMSFE derived from the inclusion of financial data are larger under a MIDAS regression scheme than under a flat aggregation scheme.

[10] We use quarterly averages instead of end of the quarter data to smooth out short-term fluctuations that could potentially misguide our forecasts of GDP growth.

Table 2: RMSFE comparisons of alternative models not seasonally adjusted GDP

| Model | h=1 RMSFE | h=1 RMSFE as % of AR | h=4 RMSFE | h=4 RMSFE as % of AR |
|---|---|---|---|---|
| Traditional Models | | | | |
| AR | 1.1348 | 1 | 1.1136 | 1 |
| RW | 1.2899 | 1.1367 | 3.2331 | 2.9033 |
| VAR | 1.1156 | 0.9831 | 1.2112 | 1.0877 |
| BVAR | 0.9285 | 0.8182 | 0.9899 | 0.8889 |
| SPF | 0.9687 | 0.8537 | | |
| | | | | |
| Quarterly macro data | | | | |
| FAR | 0.7407 | 0.6527 | 0.7308 | 0.6563 |
| | | | | |
| Monthly macro + quarterly macro data | | | | |
| Beta (p=4, q=3) | 0.9671 | 0.8522 | 1.1151 | 1.0014 |
| Combined MIDAS | 0.9827 | 0.8659 | 1.1415 | 1.0250 |
| | | | | |
| Financial data | | | | |
| Flat | 1.8181 | 1.6021 | 1.5537 | 1.3952 |
| Beta (p=2, q=6) | 1.6978 | 1.4961 | 1.8168 | 1.6315 |
| Combined MIDAS | 1.0453 | 0.9211 | 1.1936 | 1.0718 |
| | | | | |
| Financial data + quarterly macro data | | | | |
| Flat | 0.7159 | 0.6309 | 0.6682 | 0.6000 |
| Beta (p=2, q=1) | 0.4709 | 0.4150 | 0.5131 | 0.4608 |
| Combined MIDAS | 0.4614 | 0.4066 | 0.5003 | 0.4492 |

Note: The table shows the root mean square forecast error (RMSFE) for $h = 1$ and $h = 4$ step ahead horizons of the GDP. The study period is from 1999Q1 to 2013Q4. The initial estimation period is from 1999Q1 to 2009Q4 and the period of forecasting is 2009Q4+h to 2013Q4. The RMSFEs are also presented as a percentage of the AR. The 5 MIDAS forecasts estimated from each of the daily factors are combined to obtain the Combined MIDAS. A recursive window is used for all forecasts.

The most important results of this paper are shown in the last part of Table 2, in which we add the financial data to the specifications that include quarterly macro data. The results illustrate that adding daily financial data with a MIDAS regression scheme improves forecasting accuracy over a traditional model that contains only quarterly macro data at both forecast horizons. That is, financial data have an important role to predict GDP, possibly due to their forward looking nature. Another important result is that forecast combinations of MIDAS regression models based on different groups of financial variables improve forecasting accuracy. Thus, the forecasting gains from adding financial data seem to be largely attributed to the flexible weighting scheme in the MIDAS regression approach.

In sum, the combined MIDAS model with financial data has, in general, lower RMSFE than the benchmark models.[11] However, it is also important to notice that the macroeconomic regressors help to improve forecasting accuracy in both models. This is not surprising as they are highly correlated with GDP. Finally, we find that the MIDAS regression approach that incorporates daily financial variables seems to outperform the flat aggregation scheme.[12]

The tests for equal forecasting ability between selected models can be found in Table 3. In particular, the table shows the p-values obtained from a Diebold-Mariano (1995) test as described earlier.[13] The results show that the null hypothesis of equal forecasting accuracy between the benchmark AR model and the AR model augmented with financial data cannot be rejected. Similarly, the null hypothesis of equal forecasting accuracy between the MIDAS model with financial data and the AR model with quarterly data cannot be rejected at the conventional significance levels.

[11] Although the plots of the forecasts versus actual values are not presented in the paper to save space, we find that, in general, the results about the forecasting performance of the MIDAS model are robust over the entire forecasting period. The results are available from the authors upon request.

[12] We have also conducted the forecasting exercise using seasonally adjusted data. The results are presented in Table B.1 of the Appendix. The conclusions are similar to those using not seasonally adjusted data. In particular, we find that at the one quarter ahead horizon, the use of the MIDAS approach and the inclusion of the daily financial variables improve the forecasting accuracy over traditional models that use quarterly macroeconomic data or average daily financial data.

[13] Note that the forecast comparisons are non-nested because the models potentially have different lag structures. For this reason, we use the Diebold and Mariano (1995) test instead of the tests of equal predictive ability designed for nested models.

Table 3: Tests of equal predictive ability

| | h=1 | h=4 |
|---|---|---|
| Model | DM | DM |
| Financial data vs AR | | |
| Combined MIDAS | 0.2198 | 0.2000 |
| Financial data vs Quarterly macro data | | |
| Combined MIDAS | 0.1830 | 0.0920 |
| Financial + macro data vs Quarterly macro data | | |
| Combined MIDAS | 0.0023 | 0.5010 |
| Financial + macro data vs Monthly macro + quarterly macro data | | |
| Combined MIDAS | 0.0108 | 0.0590 |
| MIDAS vs Flat | | |
| Financial data | 0.1178 | 0.5560 |
| Financial + macro data | 0.0000 | 0.0270 |

This table reports p-values of a test for the null hypothesis that the models shown in the left column have equal predictive ability. The comparison is based on a Diebold-Mariano test. The study period is from 1999Q1 to 2013Q4. The initial estimation period is from 1999Q1 to 2009Q4 and the period of forecasting is 2009Q4+h to 2013Q4. A recursive window is used for all forecasts.

Notably, according to the Diebold and Mariano test, the MIDAS model with financial and quarterly data outperforms the model with quarterly macro data. That is, the forecasting gains of adding financial data through a combination of MIDAS regression model over the traditional approach of using only macroeconomic data are statistically significant at the 5% at the one quarter ahead horizons. Another important result is that the MIDAS model that includes financial variables is superior to the MIDAS model that includes monthly variables. Although the MIDAS model with financial data and the flat aggregation scheme have similar predictive ability when they exclude macroeconomic data, the MIDAS model that includes quarterly data outperforms the flat aggregation scheme that includes quarterly data. This result suggests that financial factors need to be used alongside macroeconomic variables to extract their full forecasting potential.

In short, from tables 2 and 3 we find that using quarterly macro data and financial data through a MIDAS regression model improves forecasting ability over traditional models that only include macro data. The results suggest that the inclusion of financial data provides the model with useful information to forecast GDP. It is also possible to observe that, the forecast gains of the MIDAS model over the flat aggregation scheme are significant at the conventional levels. We conclude that MIDAS is superior to a simple flat aggregation scheme. Overall, our results about the role of financial variables and the MIDAS regression model to forecast Mexican GDP are in line with those of Andreou et al. (2013) for the US. However, for h=4 Andreou et al. (2013) find statistically significant differences in predictive power that favor the MIDAS model, whereas our results are more in line with those of Marcellino and Schumacher (2010) and Arnesto et al. (2010). The latter conclude that the forecasting gains of the MIDAS approach over alternative methodologies that employ high frequency information are smaller for long horizons.

### 4.4. MIDAS Forecasts with Leads

Table 4 shows the results for predicting GDP at horizons of one and four quarters ahead using information one month farther into the quarter. In this exercise, we investigate whether the inclusion of recent information is helpful to improve the forecasts of GDP, and whether the MIDAS approach outperforms the traditional flat aggregation scheme. As before, the daily financial variables within a MIDAS approach lead to important gains over the benchmark model, especially when the quarterly macroeconomic data are also included. The RMSFE from the nowcasting exercise are similar to the forecasts shown in Table 4 for most of the specifications. This exercise illustrates the use of the MIDAS approach for nowcasting, as current quarter information is introduced to provide updates of quarterly GDP growth.

Table 4: RMSFE comparisons of alternative models not seasonally adjusted GDP

| | h=1 | | h=4 | |
|---|---|---|---|---|
| Model | RMSFE | RMSFE as % of AR | RMSFE | RMSFE as % of AR |
| Financial Data | | | | |
| Flat | 1.3790 | 1.2150 | 1.3190 | 1.1621 |
| Combined MIDAS | 0.9370 | 0.8256 | 1.0338 | 0.9109 |
| | | | | |
| Financial Data + macro data | | | | |
| Flat | 0.6891 | 0.6072 | 0.7130 | 0.6282 |
| Combined MIDAS | 0.4886 | 0.4305 | 0.5301 | 0.4670 |

Note: The table shows the root mean square forecast error (RMSFE) for $h = 1$ and $h = 4$ step ahead horizons of the GDP. The study period is from 1999Q1 to 2013Q4. The initial estimation period is from 1999Q1 to 2009Q4 and the period of forecasting is 2009Q4+h to 2013Q4. The RMSFEs are also presented as a percentage of the AR. The 5 MIDAS forecasts estimated from each of the daily factors are combined to obtain the Combined MIDAS. A recursive window is used for all forecasts.

Finally, Table 5 shows the Diebold and Mariano test for the nowcasting exercise. From this table, we cannot reject the null hypothesis that nowcasting and forecasting with MIDAS have similar predictive ability. That is, the information contained in the current month does not seem to improve the predictive accuracy for GDP growth. This result for the nowcasting exercise is similar to that of Andreou et al. (2013). A possible explanation is that, due to the forward-looking nature of financial data, the role of financial variables to predict GDP in the near future may be more important than their role to provide forecast updates of current GDP. Importantly, we find that MIDAS with leads is statistically superior in its predictive ability over the flat aggregation scheme with leads, when both models contain quarterly macroeconomic information.

Table 5: Tests of equal predictive ability

| | h=1 | h=4 |
|---|---|---|
| Model | DM | DM |
| Nowcasting vs forecasting | | |
| Flat (financial data) | 0.0810 | 0.2080 |
| Combined MIDAS (financial data) | 0.2117 | 0.0767 |
| Flat (financial+macro data) | 0.4110 | 0.2350 |
| Combined MIDAS (financial+macro data) | 0.1841 | 0.3610 |
| | | |
| Flat vs MIDAS | | |
| Financial data | 0.0760 | 0.2294 |
| Financial+macro data | 0.0001 | 0.0229 |

This table reports p-values of a test for the null hypothesis that the models shown in the left column have equal predictive ability. The comparison is based on a Diebold-Mariano test. Sample 1999Q1-2013Q4. Estimation period: 1999Q1-2009Q4. Forecasting period: 2010Q1-2013Q4. A recursive window is used for all forecasts.

## 5. Conclusion

Following the methodology proposed by Ghysels et al. (2007), we estimate a MIDAS model, which incorporates the information contained in a data rich environment of daily financial variables. Subsequently, this model is used to generate out-of-sample forecasts for the Mexican GDP for horizons of one and four quarters ahead. We find that the use of this methodology and the inclusion of the daily financial variables improve the forecasting accuracy over traditional models that use quarterly macroeconomic data. The MIDAS framework helps to circumvent the problems initially found when dealing with data at different sampling frequencies, while remaining parsimonious. To deal with large datasets of financial variables, we use factor analysis and forecast combinations.

The model comparisons favor the use of the MIDAS approach against the flat aggregation scheme. In addition, we find that a MIDAS model has a better forecasting performance than the AR model augmented with factors based on macro variables. In a nowcasting exercise, the results favor the MIDAS model over the flat aggregation scheme. However, the MIDAS model with leads seems to have a similar predictive accuracy compared to the MIDAS model without leads.

The results presented in this paper have important implications for practitioners, policy makers and researchers. In particular, our results suggest that financial variables contain useful information to predict GDP, possibly due to their forward looking nature. This could be associated with the development of financial markets in Mexico in recent years. From a policy point of view, our findings imply that financial variables should be monitored closely to anticipate the business cycle fluctuations. In terms of modelling, our results point towards the importance of linking the financial and the real sectors of the economy in macroeconomic models. The role of financial variables to predict GDP growth is not only attributed to their forward looking nature, but also to the data dependent weighting scheme used to aggregate the daily series in which recent information can be more informative than lagged information. Our results also suggest that factor analysis is useful to extract the common factors, filtering out the idiosyncratic variations that could affect forecasting performance. Similarly, forecast combinations are useful to pool the forecasts based on individual factors in a parsimonious way, thus improving the efficiency of GDP growth forecasts.

We conclude that this methodology improves forecasts even in an emerging economy that displays higher volatility. In order to improve or extend this work, a group of financial variables that is more directly related to the Mexican economy could be employed. The unavailability of historic data on those useful variables might be a limitation that could be solved by conducting the exercise in a few years. This methodology could also be used to predict other monthly or quarterly macro variables such as unemployment. Finally, as we consider only Mexican data, it would be useful to evaluate whether the MIDAS model that includes daily financial data is also successful at predicting GDP for other developing countries. We leave those extensions for further research.

## Appendix A: Data

Financial Data

| | Type | Description |
|---|---|---|
| 1 | Commodity | RJ CRB |
| 2 | Commodity | SILVER |
| 3 | Commodity | brent oil |
| 4 | Commodity | PL-NYD |
| 5 | Commodity | ZINC |
| 6 | Commodity | XPD-D |
| 7 | Commodity | WHEAT |
| 8 | Commodity | C-US2D |
| 9 | Commodity | SOYB |
| 10 | Commodity | COTTON |
| 11 | Commodity | SUGAR |
| 12 | Commodity | COFFEE |
| 13 | Commodity | COCOA |
| 14 | Commodity | CATTLE |
| 15 | Commodity | HOGS |
| 16 | Commodity | GOLD |
| 17 | Commodity | ALUMINUM |
| 18 | Commodity | WTI OIL |
| 19 | Commodity | LEAD |
| 20 | Commodity | NICKEL |
| 21 | Commodity | TIN |
| 22 | Commodity | ALUM FUT |
| 23 | Commodity | LEAD FWD |
| 24 | Commodity | NICKEL FWD |
| 25 | Commodity | TIN FWD |
| 26 | Commodity | wti first |
| 27 | Commodity | heating oil first |
| 28 | Commodity | gas oil first |
| 29 | Commodity | nat gas first |
| 30 | Commodity | corn first |
| 31 | Commodity | soybean first |
| 32 | Commodity | wheat first |
| 33 | Commodity | rough rice first |
| 34 | Commodity | lumber first |
| 35 | Commodity | sugar first |

| | | |
|---|---|---|
| 36 | Commodity | Gold |
| 37 | Commodity | copper stock |
| 38 | Commodity | nickel stock |
| 39 | Commodity | aluminum stock |
| 40 | Commodity | zink stock |
| 41 | Commodity | lead stock |
| 42 | Commodity | tin stock |
| 43 | Commodity | wti spot Midland |
| 44 | Commodity | european brent crude |
| 45 | Commodity | copper closing plrice |
| 46 | Commodity | aluminum closing Price |
| 47 | Commodity | nickel closing Price |
| 48 | Commodity | zink closing Price |
| 49 | Commodity | tin closing Price |
| 50 | Commodity | copper fwd Price |
| 51 | Commodity | aluminum fwd Price |
| 52 | Commodity | nickel fwd Price |
| 53 | Commodity | zink fwd Price |
| 54 | Commodity | lead fwd Price |
| 55 | Commodity | tin fwd Price |
| 56 | Commodity | fiber all ítems |
| 57 | Commodity | fiber metal |
| 58 | Commodity | fiber textiles |
| 59 | Commodity | fiber oil |
| 60 | Commodity | gsci heating oil |
| 61 | Commodity | gsci crude oil |
| 62 | Commodity | gsci gasolina |
| 63 | Commodity | gsci gasoil |
| 64 | Commodity | gsci gas |
| 65 | Commodity | gsci metals |
| 66 | Commodity | gsci aluminum |
| 67 | Commodity | gsci copper |
| 68 | Commodity | gsci lead |
| 69 | Commodity | gsci nickel |
| 70 | Commodity | gsci zink |
| 71 | Commodity | gsci precious metals |
| 72 | Commodity | gsci gold |
| 73 | Commodity | gsci silver |
| 74 | Commodity | gsci agricultura |

| | | |
|---|---|---|
| 75 | Commodity | gsci wheat |
| 76 | Commodity | gsci soy |
| 77 | Commodity | gsci cotton |
| 78 | Commodity | gsci sugar |
| 79 | Commodity | gsci coffee |
| 80 | Commodity | gsci cocoa |
| 81 | Commodity | gsci energy |
| 82 | Commodity | gsci livestock |
| 83 | Commodity | gsci cattle |
| 84 | Commodity | gsci hogs |
| 85 | Commodity | gsci softs |
| 86 | Commodity | gsci light energy |
| 87 | Commodity | gsci energy metals |
| 88 | Commodity | gsci non livestock |
| 89 | Commodity | gsci grains |
| 90 | Commodity | gsci all wheat |
| 91 | Commodity | gsci all crude |
| 92 | Commodity | gsci biofuel |
| 93 | Commodity | gsci 1 m fwd |
| 94 | Commodity | gsci 3m fwd |
| 95 | Commodity | gsci 2m fwd |
| 96 | Commodity | gsci 4m fwd |
| 97 | Commodity | gsci 5m fwd |
| 98 | Commodity | excess return total |
| 99 | Commodity | excess return crude oil |
| 100 | Commodity | excess return Brent |
| 101 | Commodity | excess return gasolina |
| 102 | Commodity | excess return heating oil |
| 103 | Commodity | excess return gasoil |
| 104 | Commodity | excess return nat gas |
| 105 | Commodity | excess return metals |
| 106 | Commodity | excess return aluminum |
| 107 | Commodity | excess return copper |
| 108 | Commodity | excess return lead |
| 109 | Commodity | excess return zinc |
| 110 | Commodity | excess return precious metals |
| 111 | Commodity | excess return gold |
| 112 | Commodity | excess return silver |
| 113 | Commodity | excess return agri and livestock |

| | | |
|---|---|---|
| 114 | Commodity | excess return soybean |
| 115 | Commodity | excess return corn |
| 116 | Commodity | excess return cotton |
| 117 | Commodity | excess return sugar |
| 118 | Commodity | excess return coffee |
| 119 | Commodity | excess return cocoa |
| 120 | Commodity | excess return livestock |
| 121 | Commodity | excess return hogs |
| 122 | Commodity | excess return non energy |
| 123 | Commodity | excess return light energy |
| 124 | Commodity | excess return ultra energy |
| 125 | Commodity | excess return energy metals |
| 126 | Commodity | excess return petroleum |
| 127 | Commodity | excess return grains |
| 128 | Commodity | excess return all wheat |
| 129 | Commodity | excess return all crude |
| 130 | Commodity | excess return biofuel |
| 131 | Commodity | total return total |
| 132 | Commodity | total return crude oil |
| 133 | Commodity | total return energy |
| 134 | Commodity | total return gasoline |
| 135 | Commodity | total return heating oil |
| 136 | Commodity | total return gasoil |
| 137 | Commodity | total return nat gas |
| 138 | Commodity | total return metals |
| 139 | Commodity | total return aluminum |
| 140 | Commodity | total return copper |
| 141 | Commodity | total return zinc |
| 142 | Commodity | total return precious metals |
| 143 | Commodity | total return gold |
| 144 | Commodity | total return silver |
| 145 | Commodity | total return agri and livestock |
| 146 | Commodity | total return agriculture |
| 147 | Commodity | total return soybean |
| 148 | Commodity | total return corn |
| 149 | Commodity | total return cotton |
| 150 | Commodity | total return sugar |
| 151 | Commodity | total return coffee |
| 152 | Commodity | total return cocoa |

| | | |
|---|---|---|
| 153 | Commodity | total return livestock |
| 154 | Commodity | total return cattle |
| 155 | Commodity | total return hogs |
| 156 | Commodity | total return non energy |
| 157 | Commodity | total return light energy |
| 158 | Commodity | total return ultra energy |
| 159 | Commodity | total return energy metals |
| 160 | Commodity | total return petroleum |
| 161 | Commodity | total return grains |
| 162 | Commodity | total return all crude |
| 163 | Commodity | philadelphia semiconductor |
| 164 | Commodity | corn spot price |
| 165 | Commodity | palladium |
| 166 | Commodity | platinum |
| 167 | Equity | INMEX |
| 168 | Equity | S&P500 |
| 169 | Equity | S&P 500 Industrials Sector Index GICS Level 1 |
| 170 | Equity | Dow Jones Industrial Average - DJI |
| 171 | Equity | NASDAQ |
| 172 | Equity | NASDAQ 100 |
| 173 | Equity | VIX |
| 174 | Equity | IPC |
| 175 | Equity | Dow Jones Industrial Goods and Services Titans 30 Index Euros |
| 176 | Equity | Dow Jones Transportation Average |
| 177 | Equity | Dow Jones Utilities Average |
| 178 | Equity | Dow Jones Composite Average |
| 179 | Equity | Dow Jones Internet Commerce Index |
| 180 | Equity | Dow Jones Internet Composite Index |
| 181 | Equity | S&P 500 Industrials Sector TR Index |
| 182 | Equity | S&P 500 Financials Sector Index GICS Level 1 |
| 183 | Equity | S&P Smallcap 600 Index |
| 184 | Equity | NASDAQ Industrial Index |
| 185 | Equity | Russell 2000 Index |
| 186 | Equity | Value Line Arithmetic |
| 187 | Equity | Value Line Geometric |
| 188 | Equity | Dow Jones Equity REIT Total Return Index |
| 189 | Equity | Dow Jones US Completion Total Stock Market Total Return Index |
| 190 | Equity | CBOE Equity Put/Call Ratio |
| 191 | Equity | CBOE US Put/Call Ratio Composite Intraday |

| | | |
|---|---|---|
| 192 | Equity | PUT/CALL RATIOS COMPOSITE |
| 193 | Equity | CBOE US SPX Put/Call Ratio Intraday |
| 194 | Equity | Put/Call Volume Ratios on SPX Pt/Cl |
| 195 | Equity | CBOE US Index Put/Call Ratio Intraday |
| 196 | Equity | Put/Call Ratios OEX Pt/Cl |
| 197 | Equity | CBOE US OEX Put/Call Ratio Intraday |
| 198 | Equity | Put/Call Ratios RUT Pt/Cl |
| 199 | Equity | US Option Call Volumes on CBOE |
| 200 | Equity | US Option Put Volumes on CBOE |
| 201 | Equity | New York Stock Exchange Advancing Stocks |
| 202 | Equity | New York Stock Exchange Declining Stocks |
| 203 | Equity | NASDAQ Total Volume Composite |
| 204 | Equity | NASDAQ Advancing Stocks Index |
| 205 | Equity | Nasdaq Declining Stocks Index |
| 206 | Equity | FTSE 100 Index |
| 207 | Equity | FTSEurofirst 300 Index |
| 208 | Equity | Dow Jones Islamic Market World Index |
| 209 | Equity | Dow Jones Sustainability World Composite Index |
| 210 | Equity | Dow Jones World Technology Index |
| 211 | Equity | Dow Jones Sustainability World Total Return Index Composite USD |
| 212 | Equity | Dow Jones Sustainability World Total Return Index Composite Euro |
| 213 | Equity | Dow Jones Islamic Market World Developed Index |
| 214 | Equity | Dow Jones Islamic Market World Total Return Index |
| 215 | Equity | Dow Jones World Financials Index |
| 216 | Equity | Dow Jones Sustainability World Index in EUR |
| 217 | Equity | Dow Jones World Consumer Goods Index |
| 218 | Equity | Dow Jones Sustainability World Total Return Index Ex AGTAF Euro |
| 219 | Equity | Dow Jones World Consumer Services Index |
| 220 | Equity | Dow Jones Sustainability World Ex Alcohol Tobacco Gambling Armaments & Firearms |
| 221 | Equity | Dow Jones World Oil & Gas Index |
| 222 | Equity | Dow Jones Islamic Market World Malaysia Index USD |
| 223 | Equity | Dow Jones Islamic Market World Developed Total Return Index |
| 224 | Equity | Dow Jones World Basic Materials Index |
| 225 | Equity | Dow Jones Islamic Market World Excluding US Index |
| 226 | Equity | Dow Jones Sustainability World Excluding Tobacco Index |
| 227 | Equity | Dow Jones World Healthcare Index |

| | | |
|---|---|---|
| 228 | Equity | Dow Jones Islamic Market World Emerging Markets Total Return Index |
| 229 | Equity | Dow Jones Sustainability World Index in CHF |
| 230 | Equity | Dow Jones Sustainability World Total Return Index Ex Gambling USD |
| 231 | Equity | Dow Jones Sustainability World Total Return Index Ex Tobacco USD |
| 232 | Equity | Dow Jones Sustainability World Total Return Index Ex Tobacco Euro |
| 233 | Equity | Dow Jones Islamic Market World Developed Excluding US Total Return Index |
| 234 | Equity | Dow Jones Islamic Market World Excluding US Total Return Index |
| 235 | Equity | Dow Jones Sustainability World ex Australia Index ex Tobacco AUD |
| 236 | Equity | Dow Jones Sustainability World ex Australia Total Return Index ex Tobacco AUD |
| 237 | Equity | Dow Jones World Industrials Index |
| 238 | Equity | Dow Jones Sustainability World Total Return Index in CHF |
| 239 | Equity | Dow Jones Sustainability World Total Return Index Ex AGTAF USD |
| 240 | Equity | Dow Jones Sustainability World Excluding Gambling Index |
| 241 | Equity | Dow Jones Sustainability World Total Return Index Ex Gambling Euro |
| 242 | Equity | Dow Jones World Excluding US Technology Index |
| 243 | Equity | Dow Jones World Excluding US Utilities Index |
| 244 | Equity | Dow Jones World Developed - Ex. U.S. Index |
| 245 | Equity | Dow Jones World Ex Asia Basic Materials Index |
| 246 | Equity | Dow Jones World - Ex Asia/Pacific Consumer Services Index |
| 247 | Equity | Dow Jones World - Ex Asia/Pacific Oil & Gas Index |
| 248 | Equity | Dow Jones World - Ex Asia/Pacific Financials Index |
| 249 | Equity | Dow Jones World Ex Asia Healthcare Index |
| 250 | Equity | Dow Jones World - Ex Asia/Pacific Industrials Index |
| 251 | Equity | Dow Jones World - Ex Asia/Pacific Consumer Goods Index |
| 252 | Equity | Dow Jones World Ex Asia Technology Index |
| 253 | Equity | Dow Jones World Ex Asia Telecommunications Index |
| 254 | Equity | Dow Jones World Ex Asia Utilities Index |
| 255 | Equity | EURO STOXX Index |
| 256 | Equity | Merrill Lynch Option Volatility Estimate MOVE Index |
| 257 | Equity | Merrill Lynch Option Volatility Estimate 3-Month |

| | | |
|---|---|---|
| 258 | Equity | Merrill Lynch Swaption Option Volatility Estimate 3 Month |
| 259 | Equity | Merrill Lynch Swaption Option Volatility Estimate 6 Month |
| 260 | Equity | Morgan Stanley Cyclical Index |
| 261 | Corporate | 1MLIBOR |
| 262 | Corporate | 3MLIBOR |
| 263 | Corporate | 6MLIBOR |
| 264 | Corporate | 1YLIBOR |
| 265 | Corporate | Fed Funds |
| 266 | Corporate | Federal Reserve 30 Day A2 P2 Nonfinancial Commercial Paper Interest Rate |
| 267 | Corporate | Federal Reserve 30 Day A Commercial Paper |
| 268 | Corporate | Federal Reserve 30 Day AA Financial Commercial Paper Interest Rate |
| 269 | Corporate | Federal Reserve US Treasury Note Constant Maturity Not Averaged 3 Month |
| 270 | Corporate | Moody's Bond Indices Corporate AAA |
| 271 | Corporate | Moody's Bond Indices Corporate BAA |
| 272 | Corporate | US Generic Govt 10 Year Yield |
| 273 | Corporate | BlackRock Corporate Bond Fund (London) |
| 274 | Corporate | BlackRock Corporate Bond Fund (London) |
| 275 | Corporate | BlackRock Total Return Fund (Trade Reporting Facility LLC) |
| 276 | Corporate | BlackRock Global Funds - US Dollar High Yield Bond Fund (Luxembourg) |
| 277 | Corporate | BlackRock Global Funds - US Dollar High Yield Bond Fund (Luxembourg) |
| 278 | Corporate | BlackRock Global Funds - Global High Yield Bond Fund (Luxembourg) |
| 279 | Corporate | Merrill Lynch 10-year U.S. Treasury Futures Total Return |
| 280 | Corporate | Merrill Lynch 5-year U.S. Treasury Futures Excess Return |
| 281 | Corporate | Merrill Lynch 2-year U.S. Treasury Futures Total Return |
| 282 | Corporate | Merrill Lynch 30-year U.S. Treasury Futures Total Return |
| 283 | Corporate | Merrill Lynch 5-year U.S. Treasury Futures Total Return |
| 284 | Corporate | Merrill Lynch 10-year U.S. Treasury Futures Excess Return |
| 285 | Corporate | Merrill Lynch 2-year U.S. Treasury Futures Excess Return |
| 286 | Corporate | Merrill Lynch 30-year U.S. Treasury Futures Excess Return |
| 287 | Corporate | ICE LIBOR USD 1 Week |
| 288 | Corporate | Federal Reserve Commercial Paper Financial Discount Basis 1 Day |
| 289 | Corporate | Federal Reserve 7 Day AA Financial Commercial Paper Interest Rate |

| | | |
|---|---|---|
| 290 | Corporate | Federal Reserve Commercial Paper Financial Yield Basis 1 Day |
| 291 | Corporate | Federal Reserve Commercial Paper Non-Financial Yield Basis 1 Day |
| 292 | Corporate | Federal Reserve 60 Day AA Financial Commercial Paper Interest Rate |
| 293 | Corporate | Federal Reserve Overnight AA Financial Commercial Paper Interest Rate |
| 294 | Corporate | Federal Reserve Commercial Paper Non-Financial Discount Basis 1 Day |
| 295 | Corporate | Federal Reserve 7 Day AA Nonfinancial Commercial Paper Interest Rate |
| 296 | Corporate | Federal Reserve 15 Day A2 P2 Nonfinancial Commercial Paper Interest Rate |
| 297 | Corporate | Federal Reserve 15 Day AA Nonfinancial Commercial Paper Interest Rate |
| 298 | Corporate | Federal Reserve 7 Day A2 P2 Nonfinancial Commercial Paper Interest Rate |
| 299 | Corporate | Federal Reserve Overnight A2 P2 Nonfinancial Commercial Paper Interest Rate |
| 300 | Corporate | Mtge Current Cpns Fnma 30 Year |
| 301 | Corporate | Fannie Mae Commitment Rates 30 Year Fixed Rate 30 Day |
| 302 | Corporate | Mtge Current Cpns FNMA 30 Year Spread |
| 303 | Corporate | Fannie Mae Commitment Rates 30 Year Fixed Rate 60 Day |
| 304 | Corporate | Fannie Mae Commitment Rates 15 Year Fixed Rate 30 Day |
| 305 | Corporate | Fannie Mae Commitment Rates 30 Year Fixed Rate 10 Day |
| 306 | Corporate | Fannie Mae Commitment Rates 20 Year Fixed Rate 30 Day |
| 307 | Corporate | Fannie Mae Commitment Rates 30 Year Fixed Rate 90 Day |
| 308 | Corporate | Fannie Mae Commitment Rates 10 Year Fixed Rate 30 Day |
| 309 | Corporate | DBIQ MBS TBA: FNMA: 30 Years FNCL |
| 310 | Corporate | Fannie Mae Commitment Rates 15 Year Biweekly Fixed Rate 30 Day |
| 311 | Corporate | Citigroup Mortgage 30 Year FNMA Sector Local Currency |
| 312 | Corporate | Fannie Mae Commitment Rates 30 Year Biweekly Fixed Rate 10 Day |
| 313 | Corporate | Fannie Mae Commitment Rates 30 Year Biweekly Fixed Rate 30 Day |
| 314 | FX | USA cross rate |
| 315 | FX | suiza cross rate |
| 316 | FX | Fwd USD |

| | | |
|---|---|---|
| 317 | FX | Canada |
| 318 | FX | Japan |
| 319 | FX | Costa Rica |
| 320 | FX | Dominican Republic |
| 321 | FX | Iceland |
| 322 | FX | Israel |
| 323 | FX | Kazakhstan |
| 324 | FX | United Arab Emirates |
| 325 | FX | Australia |
| 326 | FX | Switzerland |
| 327 | FX | Chile |
| 328 | FX | Euro |
| 329 | FX | Great Britain |
| 330 | FX | Guatemala |
| 331 | FX | New Zealand |
| 332 | FX | USA |
| 333 | FX | Peru |
| 334 | FX | Paraguay |
| 335 | FX | Romania |
| 336 | FX | Saudi Arabia |
| 337 | FX | Slovakia |
| 338 | FX | Thailand |
| 339 | FX | Turkey |
| 340 | FX | South Africa |
| 341 | Fixed Income | 3MTB |
| 342 | Fixed Income | 6MTB |
| 343 | Fixed Income | 6 M Treasury Bill |
| 344 | Fixed Income | Federal Reserve US Treasury Note Constant Maturity Not Averaged 1 Year |
| 345 | Fixed Income | Federal Reserve US Treasury Note Constant Maturity Not Averaged 2 Year |
| 346 | Fixed Income | Federal Reserve US Treasury Note Constant Maturity Not Averaged 3 Year |
| 347 | Fixed Income | Federal Reserve US Treasury Note Constant Maturity Not Averaged 5 Year |
| 348 | Fixed Income | Federal Reserve US Treasury Note Constant Maturity Not Averaged 10 Year |
| 349 | Fixed Income | Federal Reserve US Treasury Note Constant Maturity Not Averaged 20 Year |

| | | |
|---|---|---|
| 350 | Fixed Income | US Treasury Yield Curve Rate T Note Constant Maturity 2 Year |
| 351 | Fixed Income | US Treasury Yield Curve Rate T Note Constant Maturity 3 Year |
| 352 | Fixed Income | US Treasury Yield Curve Rate T Note Constant Maturity 3 Month |
| 353 | Fixed Income | US Treasury Yield Curve Rate T Note Constant Maturity 7 Year |
| 354 | Fixed Income | US Treasury Yield Curve Rate T Note Constant Maturity 20 Year |
| 355 | Fixed Income | US Treasury Yield Curve Rate T Note Constant Maturity 6 Month |
| 356 | Fixed Income | Federal Reserve US Treasury Note Constant Maturity Not Averaged 7 Year |
| 357 | Fixed Income | US Treasury Yield Curve Rate T Note Constant Maturity Composite Over 10 Year |
| 358 | Fixed Income | Federal Reserve US Treasury Note Constant Maturity Not Averaged 6 Month |
| 359 | Fixed Income | Federal Reserve US Treasury Note Constant Maturity Not Averaged Composite 30+ Ye |
| 360 | Fixed Income | US Generic Govt 30 Year Yield |
| 361 | Fixed Income | USD US Treasury Bonds/Notes (FMC 82) Zero Coupon Yield 10 Year |
| 362 | Fixed Income | USD Treasury Actives (IYC 25) Zero Coupon Yield 10 Year |
| 363 | Fixed Income | USD Treasury Actives (IYC 25) Zero Coupon Yield 3 Month |
| 364 | Fixed Income | USD Treasury Actives (IYC 25) Zero Coupon Yield 1 Year |
| 365 | Fixed Income | USD US Treasury Bonds/Notes (FMC 82) Zero Coupon Yield |
| 366 | Fixed Income | USD US Treasury Bonds/Notes (FMC 82) Zero Coupon Yield 1 Year |
| 367 | Fixed Income | USD US Treasury Bonds/Notes (FMC 82) Zero Coupon Yield 30 Year |
| 368 | Fixed Income | USD US Treasury Bonds/Notes (FMC 82) Zero Coupon Yield 6 Month |
| 369 | Fixed Income | USD US Treasury Bonds/Notes (FMC 82) Zero Coupon Yield 5 Year |
| 370 | Fixed Income | USD US Treasury Bonds/Notes (FMC 82) Zero Coupon Yield 7 Year |
| 371 | Fixed Income | USD Treasury Actives (IYC 25) Zero Coupon Yield 30 Year |
| 372 | Fixed Income | USD Treasury Actives (IYC 25) Zero Coupon Yield 5 Year |
| 373 | Fixed Income | USD Treasury Actives (IYC 25) Zero Coupon Yield 2 Year |
| 374 | Fixed Income | USD US Treasury Bonds/Notes (FMC 82) Zero Coupon Yield 2 Year |
| 375 | Fixed Income | USD Treasury Actives (IYC 25) Zero Coupon Yield 3 Year |
| 376 | Fixed Income | USD US Treasury Bonds/Notes (FMC 82) Zero Coupon Yield 9 Year |

| | | |
|---|---|---|
| 377 | Fixed Income | USD US Treasury Bonds/Notes (FMC 82) Zero Coupon Yield 20 Year |
| 378 | Fixed Income | USD Treasury Actives (IYC 25) Zero Coupon Yield 6 Month |
| 379 | Fixed Income | USD Treasury Actives (IYC 25) Zero Coupon Yield 20 Year |
| 380 | Fixed Income | USD Treasury Actives (IYC 25) Zero Coupon Yield 4 Year |
| 381 | Fixed Income | USD US Treasury Bonds/Notes (FMC 82) Zero Coupon Yield 6 Year |
| 382 | Fixed Income | USD Treasury Actives (IYC 25) Zero Coupon Yield 7 Year |
| 383 | Fixed Income | USD US Treasury Bonds/Notes (FMC 82) Zero Coupon Yield 4 Year |
| 384 | Fixed Income | USD US Treasury Bonds/Notes (FMC 82) Zero Coupon Yield 3 Year |
| 385 | Fixed Income | USD Treasury Actives (IYC 25) Zero Coupon Yield 15 Year |
| 386 | Fixed Income | USD Treasury Actives (IYC 25) Zero Coupon Yield 8 Year |
| 387 | Fixed Income | USD Treasury Actives (IYC 25) Zero Coupon Yield 6 Year |
| 388 | Fixed Income | USD Treasury Actives (IYC 25) Zero Coupon Yield 9 Year |
| 389 | Fixed Income | USD US Treasury Bonds/Notes (FMC 82) Zero Coupon Yield 8 Year |
| 390 | Fixed Income | USD US Treasury Bonds/Notes (FMC 82) Zero Coupon Yield 15 Year |
| 391 | Fixed Income | Cetes 28 |
| 392 | Fixed Income | Cetes 91 |

Monthly Macro data

| | Description |
|---|---|
| 1 | Industrial production |
| 2 | IMSS permanent workers |
| 3 | Net investment |
| 4 | Non-oil exports |
| 5 | Oil exports |
| 6 | CPI: food, drinks and tobacco |
| 7 | CPI: housing |
| 8 | CPI: clothing, shoes and accesories |
| 9 | CPI: furniture and domestic appliances |
| 10 | CPI: health |
| 11 | CPI: transport |
| 12 | CPI: education and recreation |
| 13 | CPI: other services |

| | |
|---|---|
| 14 | Industrial activity |
| 15 | M1 |
| 16 | M2 |
| 17 | IGAE |
| 18 | USA CPI |

Quarterly Macro data

| | Description |
|---|---|
| 1 | Private consumption |
| 2 | Government consumption |
| 3 | Imports |
| 4 | Net investment |
| 5 | Non-oil exports |
| 6 | Oil exports |
| 7 | CPI: food, drinks and tobacco |
| 8 | CPI: housing |
| 9 | CPI: clothing, shoes and accesories |
| 10 | CPI: furniture and domestic appliances |
| 11 | CPI: health |
| 12 | CPI: transport |
| 13 | CPI: education and recreation |
| 14 | CPI: other services |
| 15 | Industrial activity |
| 16 | M1 |
| 17 | M2 |
| 18 | IGAE |
| 19 | USA GDP |
| 20 | USA CPI |

## Appendix B: Supplemental Results

Table B.1.: RMSFE comparisons of alternative models using seasonally adjusted GDP

| | h=1 | | h=4 | |
|---|---|---|---|---|
| Model | RMSFE | RMSFE as % of AR | RMSFE | RMSFE as % of AR |
| Alternative Models | | | | |
| AR | 0.7162 | 1 | 0.6030 | 1 |
| RW | 0.7869 | 1.0987 | 0.9555 | 1.5845 |
| VAR | 0.7246 | 1.0117 | 0.6810 | 1.1294 |
| BVAR | 0.6303 | 0.8801 | 0.6187 | 1.0261 |
| SPF | 1.0043 | 1.4022 | | |
| | | | | |
| Quarterly macro data | | | | |
| FAR | 0.6691 | 0.9342 | 0.6253 | 1.0368 |
| | | | | |
| Financial data | | | | |
| Flat | 0.8392 | 1.1718 | 0.6359 | 1.0545 |
| Beta (p=1, q=3) | 0.8434 | 1.1775 | 0.8690 | 1.4409 |
| Combined MIDAS | 0.6772 | 0.9455 | 0.7317 | 1.2133 |
| | | | | |
| Financial Data + quarterly macro data | | | | |
| Flat | 0.6294 | 0.8788 | 0.5921 | 0.9818 |
| Beta (p=1, q=3) | 0.6768 | 0.9450 | 0.7185 | 1.1915 |
| Combined MIDAS | 0.6047 | 0.8442 | 0.6096 | 1.0109 |

Note: The table shows the root mean square forecast error (RMSFE) for h=1 and h=4 step ahead horizons of the GDP. The study period is from 1999Q1 to 2013Q4. The initial estimation period is from 1999Q1 to 2009Q4 and the period of forecasting is 2009Q4+h to 2013Q4. The RMSFEs are also presented as a percentage of the AR. The 5 MIDAS forecasts estimated from each one of the daily factors are combined to obtain the Combined MIDAS. A recursive window is used for all forecasts.

**References**


1) Andreou, E., Ghysels, E. and Kourtellos, A. (2013). Should Macroeconomic Forecasters Use Daily Financial Data and How?, *Journal of Business and Economic Statistics*, 31(2), 240-251.

2) Arnesto, M. T., Engerman, K. M. and Owyang, M. T. (2010). Forecasting with Mixed Frequencies, *Federal Reserve Bank of St. Louis Review*, 92(6), 521-536.

3) Artis, M. J., and Zhang, W. (1990). BVAR Forecasts for the G-7. *International Journal of Forecasting*, 6(3), 349-362.

4) Baffigi, A., Roberto G., and Giuseppe P. (2004). Bridge Models to Forecast the Euro Area GDP. *International Journal of Forecasting*, 20(3) 447-460.

5) Bai, J., Gysels, E. and Wright, J. H. (2013). State Space Models and MIDAS Regressions. *Econometric Reviews*, 32(7), 779-813.

6) Bańbura, M., Giannone, D., and Reichlin, L. (2010). Large Bayesian vector auto regressions. *Journal of Applied Econometrics*, 25(1), 71-92.

7) Boivin, J., and Ng, S. (2005). Understanding and Comparing Factor-Based Forecasts. *International Journal of Central Banking*, 1(3).

8) Canova, F. (2007). *Methods for Applied Macroeconomic Research* (Vol. 13). Princeton University Press.

9) Capistrán, C., and López-Moctezuma, G. (2010). Las expectativas macroeconómicas de los especialistas: Una evaluación de pronósticos de corto plazo en México. *El Trimestre Económico,* 77(306 (2)), 275-312.

10) Chauvet, M., and Potter, S. (2013), Forecasting Output, in *Handbook of Economic Forecasting* (vol. 2), eds. G. Elliott, and A. Timmermann, Amsterdam, Netherlands: Elsevier.

11) Clements, M. P. and Galvão, A. B. (2008). Macroeconomic Forecasting With Mixed-Frequency Data. *Journal of Business & Economic Statistics*, 26(4), 546-554.

12) Cogley, T., and Nason, J. M. (1995). Effects of the Hodrick-Prescott Filter on Trend and Difference Stationary Time series Implications for Business Cycle Research. *Journal of Economic Dynamics and Control*, 19(1), 253-278.

13) Diebold, F. X., and Mariano, R. S. (1995). Comparing Predictive Accuracy. *Journal of Business and Economic Statistics*, 13(3), 253–265.

14) Doan, T., Litterman, R., and Sims, C. (1984). Forecasting and Conditional Projection using Realistic Prior Distributions. *Econometric Reviews*, 3(1), 1-100.

15) Dua, P., and Ray, S. C. (1995). A BVAR model for the Connecticut economy. *Journal of Forecasting*, 14(3), 167-180.

16) Espinoza, R., Fornari, F., and Lombardi, M. J. (2012). The Role of Financial Variables in Predicting Economic Activity. *Journal of Forecasting*, 31(1), 15-46.

17) Forni, M., Hallin, M., Lippi, M., and Reichlin, L. (2000). The Generalized Dynamic-Factor Model: Identification and Estimation. *Review of Economics and Statistics*, 82(4), 540-554.

18) Forni, M., Hallin, M., Lippi, M., and Reichlin, L. (2003). Do Financial Variables Help Forecasting Inflation and Real Activity in the Euro Area? *Journal of Monetary Economics*, 50(6), 1243-1255.

19) Friedman, M., and Schwartz, A. J. (2008). *A Monetary History of the United States*, 1867-1960. Princeton University Press.

20) Galvão, A. B. (2013). Changes in Predictive Ability with Mixed Frequency Data. *International Journal of Forecasting*, 29(3), 395-410.

21) Ghysels, E. (2015). Matlab Toolbox for Mixed Sampling Frequency Data Analysis using MIDAS Regression Models. Unpublished Manuscript.

22) Ghysels, E., Sinko, A., and Valkanov, R. (2007). MIDAS Regressions: Further Results and New Directions. *Econometric Reviews*, 26(1), 53-90.

23) Giannone, D., Reichlin, L., and Small, D. (2008). Nowcasting: The Real-Time Informational Content of Macroeconomic Data. *Journal of Monetary Economics,* 55(4), 665-676.

24) Forni, M., Hallin, M., Lippi, M., and Reichlin, L. (2003). Do Financial Variables Help Forecasting Inflation and Real Activity in the Euro Area? *Journal of Monetary Economics*, 50(6), 1243-1255.

25) Herrera-Hernandez, J. (2004). Business Cycles in Mexico and the United States: Do They Share Common Movements? *Journal of Applied Economics*, 7(2), 303-323.

26) Hodrick, R. J., and Prescott, E. C. (1997). Postwar US Business Cycles: An Empirical Investigation. *Journal of Money, Credit, and Banking*, 29(1), 1-16.

27) Ibarra, R. (2012). Do Disaggregated CPI Data Improve the Accuracy of Inflation Forecasts? *Economic Modelling*, 29(4), 1305-1313.

28) LeSage, J. P. (1999). Applied Econometrics using MATLAB. Manuscript, Dept. of Economics, University of Toronto, 154-159.

29) Levy-Yeyati, E., and Sturzenegger, F. (2003). To Float or to Fix: Evidence on the Impact of Exchange Rate Regimes on Growth. The American Economic Review, 93(4), 1173-1193.

30) Litterman, R. B. (1986). Forecasting with Bayesian Vector Autoregressions—Five Years of Experience. *Journal of Business & Economic Statistics*, 4(1), 25-38.

31) Marcellino, M. and Schumacher, C. (2010). Factor MIDAS for Nowcasting and Forecasting with Ragged-Edge Data: A Model Comparison for German GDP, *Oxford Bulletin of Economics and Statistics*, 72(4), 518-550.

32) Marcellino, M., Stock, J. H., and Watson, M. W. (2003). Macroeconomic Forecasting in the Euro Area: Country Specific versus Area-Wide Information. *European Economic Review*, 47(1), 1-18.

33) Mariano, R. S., and Murasawa, Y. (2003). A New Coincident Index of Business Cycles based on Monthly and Quarterly Series. *Journal of Applied Econometrics*, 18(4), 427-443.

34) McNees, S. K. (1986). Forecasting Accuracy of Alternative Techniques: A Comparison of US Macroeconomic Forecasts. *Journal of Business & Economic Statistics*, 4(1), 5-15.

35) Sidaoui, J. J. (2006). The Mexican Financial System: Reforms and Evolution 1995-2005. *BIS Papers*, 28, 277-93.

36) Sims, C. A. (1980). Macroeconomics and Reality. *Econometrica: Journal of the Econometric Society*, 48(1), 1-48.

37) Stock, J. H., and Watson, M. W. (2002a). Forecasting Using Principal Components From a Large Number of Predictors, *Journal of the American Statistical Association*, 97(460), 1167–1179.

38) Stock, J. H., and Watson, M. W. (2002b). Macroeconomic forecasting using diffusion indexes. *Journal of Business & Economic Statistics*, 20(2), 147-162.

39) Stock, J. H., and Watson, M. W. (2003). Forecasting Output and Inflation: The Role of Asset Prices. *Journal of Economic Literature*, 41(3), 788-829.

40) Stock, J. H., and Watson, M. W. (2008). Forecasting in Dynamic Factor Models subject to Structural Instability, In *The Methodology and Practice of Econometrics, A Festschrift in Honour of Professor David F. Hendry*, eds. J. Castle and N. Shephard, Oxford: Oxford University Press.

41) Theil, H., and Goldberger, A. S. (1961). On Pure and Mixed Statistical Estimation in Economics. *International Economic Review*, 2(1), 65-78.

42) Timmermann, A. (2006). Forecast Combinations, in *Handbook of Economic Forecasting* (Vol. 1), eds. G. Elliott, C. Granger, and A. Timmermann, Holland, pp. 136–196. Amsterdam: Elsevier-North.