# Should macroeconomic forecasters use daily financial data and how?*

**Elena Andreou**[†] **Eric Ghysels**[‡] **Andros Kourtellos**[§]

First Draft: May 2009
This Draft: October 1, 2009

**Abstract**

There are hundreds of financial times series available on a daily basis that contain information about the future states of the economy. Can we efficiently use all this daily financial information for improving and/or updating macroeconomic forecasts? The literature on macroeconomic forecasting has not address this question. Instead, it has focused on how the use of a small set of financial series - usually aggregated at a monthly or quarterly frequency. In the paper we introduce two methods for predicting inflation and real activity: (1) methods which rely on combinations of regressions that involve regressors with different sampling frequency, such as quarterly macro series and daily financial series and (2) a small set of daily financial factors extracted from the large cross-section of daily series along with quarterly frequency factors dominated by macroeconomic variables. Both methods have the following important features: (1) they allow us to clearly show the incremental value of daily financial series in terms of forecast improvements, (2) they provide a succinct summary of a huge amount of information in daily financial data, (3) they rely on extremely simple parsimonious regression methods that are easy to implement in practice. The analysis of daily financial factors are of independent interest and other potential applications too as they span equity premium, foreign exchange, fixed income and commodity price information.

*The second author benefited from funding by the Federal Reserve Bank of New York through the Resident Scholar Program. We would like to thank Tobias Adrian, Jennie Bai, Jushan Bai, Frank Diebold, Rob Engle, Ana Galvão, Michael Fleming, Serena Ng, Simon Potter, Lucrezia Reichlin, Jim Stock, Mark W. Watson as well as seminar participants at Columbia University, the Federal Reserve Bank of New York, MEG Conference, Queen Mary University, the University of Pennsylvania and the CIRANO/CIREQF Financial Econometrics Conference, the NBER Summer Institute, and the Deutsch Bundesbank for comments. We also thank Constantinos Kourouyiannis and Michael Sockin for providing excellent research assistance.

[†]Department of Economics, University of Cyprus, P.O. Box 537, CY 1678 Nicosia, Cyprus, e-mail: elena.andreou@ucy.ac.cy.

[‡]Department of Economics, University of North Carolina, Gardner Hall CB 3305, Chapel Hill, NC 27599-3305, USA, and Department of Finance, Kenan-Flagler Business School, e-mail: eghysels@unc.edu.

[§]Department of Economics, University of Cyprus, P.O. Box 537, CY 1678 Nicosia, Cyprus, e-mail: andros@ucy.ac.cy

# 1 Introduction

There are literally hundreds of financial times series available on a daily basis that contain information about the future states of the economy. Can we use them for the purpose of improving and/or updating macroeconomic forecasts? The literature on macroeconomic forecasting has not address this question. Instead, it has focused on how the use of a small set of financial series - usually aggregated at a monthly or quarterly frequency - to improve macroeconomic forecasts. Since macroeconomic data are typically sampled at low frequency, that is quarterly or monthly, the standard approach is to match such data with monthly or quarterly aggregates of financial data to build prediction models for the macro economy. For example, typically one does not use a daily measure of the slope of the term structure, but instead a quarterly or monthly average. Such aggregation often implies loss of information which translates to forecasting losses. Moreover, the low frequency nature of macroeconomic data releases creates another difficulty. For example, GDP growth is announced with considerable delay and is only released every quarter. Monthly series, like inflation, or industrial production are also announced with at least one month delay. Similarly while economic theory suggests that the forward looking behavior of financial asset prices should be considered as good predictors for economic conditions, the empirical evidence based on quarterly and monthly financial data is mixed and not robust (for example see Stock and Watson (1989), Stock and Watson (2002), Forni, Hallin, Lippi, and Reichlin (2000) and Forni, Hallin, Lippi, and Reichlin (2003)).

In the paper we introduce two types of prediction methods for inflation and real activity: (1) methods which rely on combinations of regressions that involve regressors with different sampling frequency, such as quarterly macro factors and daily financial series and (2) a small set of daily financial factors extracted from the large cross-section of daily series along with quarterly frequency factors dominated by macroeconomic variables. Both methods have the following important features: (1) they allow us to clearly show the incremental value of daily financial series in terms of forecast improvements, (2) they provide a succinct summary of a huge amount of information in daily financial data, (3) they rely on extremely simple parsimonious regression methods that are easy to implement in practice. The daily financial factors - what happen to find about 10 of them that summarize the cross-section- are of independent interest as they span equity premium, fixed income, commodity price information.

We employ regression models that involve data sampled at different frequencies, the so called Mi(xed) Da(ta) S(ampling), or MIDAS, regression models. MIDAS was introduced in both a forecasting and regression context in a number of recent papers, including recent work by Ghysels, Santa-Clara, and Valkanov (2002), Ghysels, Santa-Clara, and Valkanov (2006) and Andreou, Ghysels, and Kourtellos (2008). Such regressions have been used in the context of improving quarterly macro forecasts with monthly data (see e.g. Armesto, Hernandez-Murillo, Owyang, and Piger (2008), Clements and Galvão (2008a), Clements and Galvão (2008b), Galvão (2006), Schumacher and Breitung (2008), Tay (2007) ), or improving quarterly and monthly macroeconomic predictions with daily financial data (see e.g. Monteforte and Moretti (2008), Ghysels and Wright (2008), Hamilton (2006), Tay (2006)).

Our results provide some interesting findings for forecasting inflation and economic activity for the period 1999-2008. We find that the best MIDAS models provide substantial forecasting gains relative to the benchmark forecasting models considered in the literature. More importantly, the best MIDAS models improve substantially the forecasts of the best traditional models that use quarterly lags of averaged daily data to forecast CPI inflation, IP and GDP growth. Interestingly, we find that the top percentile of best predictors for $h = 1$ and $h = 4$ comprise commodity prices and interest rate spreads while the best models are generally given by MIDAS ADL with no quarterly factors. In the case of GDP growth we find the best predictors include daily Canadian/US Dollar returns and the daily Aruoba, Diebold and Scotti (ADS) indicator for $h = 1$ and interest rate spreads for $h = 4$. In contrast to the case of forecasting CPI inflation, we find that models with quarterly factors are important for forecasting GDP growth. Finally, while MIDAS models with daily factors cannot outperform the MIDAS models with the best predictor they do provide substantial improvements over the univariate benchmarks (UCSV, AO, and RW) as well as the quarterly factor augmented AR models.

The paper is organized as follows. In section 2 and 3 we describe the MIDAS Regression Models and discuss our data. In section 4 we present our results and section 5 concludes.

## 2 MIDAS Regression Models

We are interested in forecasting at a quarterly horizon, namely one quarter ahead up to four quarters, a series denoted $Y_t^Q$. The time index $t$ will refer to the quarterly frequency of forecast horizons. The typical Augmented Distributed Lag, $ADL(p_Y^Q, q_X^Q)$, model is

$$Y_{t+1}^Q = \mu + \sum_{k=0}^{p_Y^Q - 1} \alpha_k Y_{t-k}^Q + \sum_{k=0}^{p_X^Q - 1} \beta_k X_{t-k}^Q + u_{t+1}, \tag{2.1}$$

as it involves $p_Y^Q$ lags of $Y_t^Q$ and $q_X^Q$ lags of a regressor sampled also at a quarterly frequency. This regression is fairly parsimonious as it only requires $p_Y^Q + q_X^Q + 1$ parameters to be estimated. One such example in the literature is the case where $Y_{t+1}^Q$ represents quarterly inflation and $X_{t-k}^Q$ represents interest rate yields.

In many situations we do have monthly, weekly or daily data available for regressors. We will denote them respectively $X_{j,t}^M$, $X_{j,t}^W$, or $X_{j,t}^D$, with $j$ referring to the $j^{th}$ month, week or day in quarter $t$. An illustrative example is the prediction of (quarterly) inflation with a spot commodity price as regressor. While we can formulate equation (2.1) using past quarterly inflation and past (say quarterly averaged) commodity prices, we could also consider:

$$Y_{t+1}^Q = \mu + \sum_{k=0}^{p_Y^Q - 1} \alpha_k Y_{t-k}^Q + \sum_{j=1}^{N_D} \beta_j X_{j,t}^Q + u_{t+1} \tag{2.2}$$

assuming for simplicity that we take one quarter of daily data, namely $N_D$ lags (number of trading days per quarter - we assume it to be constant for simplicity). Take again the case of lags that amount to one past quarter's information. In such a case we have $(N_D + 1 + 1)$ parameters to estimate. When, say $N_D = 66$, we have a total of 68 parameters. This illustrative parameter proliferation problem example also explains why high frequency data has been avoided.

We develop two strategies to address the use of high frequency financial data for forecasting key macroeconomic variables. One involves the use of MIDAS regressions with a single high frequency regressor and then combines the forecasts they generate using a cross-section of daily financial series. The second involves extracting factors from two large cross-sections that involve quarterly data and daily financial data. The latter approach involves extracting financial factors that span many series within the equities, foreign exchange, fixed income and commodity prices. These daily financial factors can be used for many other application beyond the present forecasting analysis.

One appealing approach is to solve parameter proliferation by using MIDAS regression models. The main idea - which is rather simple and fairly straightforward to implement - is to hyper-parameterize the lag coefficient weights. Hence, one can rewrite equation (2.2) to define $ADL - MIDAS(p_Y^Q, N_D)$

$$Y_{t+1}^Q = \mu + \sum_{k=0}^{p_Y^Q - 1} \alpha_k Y_{t-k}^Q + \beta \sum_{j=1}^{N_D} w_j(\theta^D) X_{j,t}^D + u_{t+1}. \tag{2.3}$$

Following Ghysels, Santa-Clara, and Valkanov (2006) and Ghysels, Sinko, and Valkanov (2006), we employ a two parameter exponential Almon lag polynomial

$$w_j(\theta_1, \theta_2) = \frac{\exp\{\theta_1 j + \theta_2 j^2\}}{\sum_{j=1}^{m} \exp\{\theta_1 j + \theta_2 j^2\}} \tag{2.4}$$

with $\theta = (\theta_1, \theta_2)$. This approach allows us to obtain a linear projection of high frequency data $X_t^D$ onto $Y_t^Q$ with a small set of parameters.

At this point several issues emerge. Some issues are theoretical in nature. For example, to what extend is this tightly parameterized formulation in (2.3) able to approximate the unconstrained (albeit practically infeasible) projection in equation (2.2)? There is also the question whether a more traditional approach involving the Kalman filter would be more suitable.[1] We do not deal directly with these types of questions here, as they have been addressed notably in Bai, Ghysels, and Wright (2009) and Kuzin, Marcellino, and Schumacher (2009). However, some short answers to these questions are as follows: (1) the approximation errors are typically small as the exponential Almon appears quite versatile, and (2) related to the latter - it turns out that a MIDAS regression can be viewed as a reduced form approximation to the linear projection that emerges from a state space model approach. The latter, while clearly optimal as far as linear projections goes, has two main

[1]The mismatch of sampling frequency has been address in the context of state space models by Harvey and Pierse (1984), Harvey (1989), Bernanke, Gertler, and Watson (1997), Zadrozny (1990), Mariano and Murasawa (2003), Mittnik and Zadrozny (2004), Aruoba, Diebold, and Scotti (2009), Ghysels and Wright (2008), Kuzin, Marcellino, and Schumacher (2009), among others.

disadvantages (1) it is more prone to specification errors as a full system of equations for $Y$, $X$, and latent factors is required and (2) as a consequence it typically requires a lot more parameters to achieve the same goal. This is particularly relevant for the cases we cover in this paper. Namely handling a combination of quarterly and daily data leads to large state space system equations prone to mis-specification. MIDAS regressions, in comparison, are frugal in terms of parameters and achieve the same goal. In the remainder of this section we expand on the main theme addressed so far. Namely, we will present several MIDAS regression specifications that cover more general cases.

It is worth pointing out that there is a more subtle relationship between the ADL regression appearing in equation (2.1) and the ADL-MIDAS regression in equation (2.3). Note that the ADL regression involves temporally aggregated series, based for example on equal weights of daily data, i.e

$$X_t^Q \equiv (X_{1,t}^D + X_{2,t}^D + ... + X_{N_D,t}^D)/N_D$$

If we take the case of one $N_D$ days of past daily data in an ADL regression, then implicitly through the aggregation we have picked the weighting scheme $\beta_1/N_D$ for the daily data $X_{.,t}^D$. We will sometimes refer this scheme as a *flat* aggregation scheme. While these weights might be "natural"for purposes of temporal aggregation, it may not be optimal for the purpose of forecasting as more recent data may probably get more weight. In some sense the ADL-MIDAS regression lets the data decide what those weights should be.

The comparison with temporal aggregation prompts us to consider two MIDAS regression models that allow for quarterly lags. First, define the following filtered parameter-driven *quarterly* variable

$$X_t^Q(\theta_X^D) = \sum_{i=1}^{p_X^D} w_i(\theta_X^D) X_{i,t}^D,$$

where $p_X^D \geq N_D$. Notice that the number of daily lags $p_X^D$ is equal or greater to the number of trading days in a quarter, $N_D$. This is important because it allows ADL-MIDAS to model persistent memory structures in a very parsimonious way. Then, we can define the $ADL - MIDAS(p_Y^Q, p_X^Q, p_X^D)$ and $ADL - MIDAS - M(p_Y^Q, p_X^Q, p_X^D)$

$$Y_{t+1}^Q = \mu + \sum_{k=0}^{p_Y^Q-1} \alpha_k Y_{t-k}^Q + \sum_{k=0}^{p_X^Q-1} \beta_k X_{t-k}^Q(\theta_X^D) + u_{t+1} \tag{2.5}$$

$$Y_{t+1}^Q = \mu + \sum_{k=0}^{p_Y^Q-1} \alpha_k Y_{t-k}^Q + \beta \sum_{k=0}^{p_X^Q-1} w_k(\theta_X^Q) X_{t-k}^Q(\theta_X^D) + u_{t+1}. \tag{2.6}$$

While both models apply MIDAS aggregation for the daily lags they differ in the way they treat the quarterly lags. Equation (2.5) does not restrict the coefficients of the quarterly lags. However, equation (2.6) hyper-parameterizes these coefficients by employing a multiplicative MIDAS polynomial.[2] This multiplicative structure has certain advantages: (1) this specification nests the equally weighted aggregation scheme and it is more parsimonious than equation (2.5), and (2) Bai, Ghysels, and Wright (2009) show that the weighting scheme in equation (2.6) corresponds to the structure of a steady state Kalman filter linear projection with mixed sampling frequencies. The downside of the MIDAS specification in equation (2.6) is that it is less parsimonious than the single weighting scheme in equation (2.3) and more restrictive than equation (2.5).

## 2.1 MIDAS Regression Models with Factors

Recently, a large body of recent work has developed factor model techniques that are tailored to exploit a large cross-sectional dimension; see for instance, Bai and Ng (2002), Bai (2003), Forni, Hallin, Lippi, and Reichlin (2000), Forni, Hallin, Lippi, and Reichlin (2005), Stock and Watson (1989), Stock and Watson (2003), among many others. These factors are usually estimated at quarterly frequency using a large cross-section of time-series. Following this literature we investigate first whether we can improve factor model forecasts by augmenting such models with high frequency information, especially daily financial data. Subsequently, we will construct *daily* factors, using the large cross-section of financial series.

We augment the aforementioned MIDAS models with factors, $F_t$, obtained by following

[2] The multiplicative MIDAS scheme was originally suggested for purpose of dealing with intra-daily seasonality in high frequency data, see Chen and Ghysels (2009).

factor model

$$\begin{aligned} X_t &= \Lambda_t F_t + u_t \\ F_t &= \Phi F_{t-1} + \eta_t \\ u_{it} &= a_{it}(L) u_{it-1} + \varepsilon_{it}, \quad i = 1, 2, ..., n. \end{aligned} \tag{2.7}$$

The data used to implement the factor representation will be described in the next section. Suffice it here to say that we use series similar to those used by Stock and Watson (2008). The number of factors are chosen based on the information criteria proposed by Bai and Ng (2002).

We augment the MIDAS regression models from the previous subsection by adding quarterly factors. For instance, equation (2.5) generalizes to the $FADL - MIDAS(p_Y^Q, p_F^Q, p_X^Q, p_X^D)$ model

$$Y_{t+1}^Q = \mu + \sum_{k=0}^{p_Y^Q - 1} \alpha_k Y_{t-k}^Q + \sum_{k=0}^{p_F^Q - 1} \beta_k F_{t-k}^Q + \sum_{k=0}^{p_X^Q - 1} \gamma_k X_{t-k}^Q(\theta_X^D) + u_{t+1} \tag{2.8}$$

Note that we can also formulate a $FADL - MIDAS - M(p_Y^Q, p_F^Q, p_X^Q, p_X^D)$ model, which involves the multiplicative MIDAS weighting scheme, hence generalizing equation (2.6). Notice also that equation (2.8) simplifies to the traditional factor model with additional regressors when the MIDAS features are turned off - i.e. say a flat aggregation scheme is used. When the lagged dependent variable is excluded then we have a projection on daily data, combined with aggregate factors. This brings us to the following benchmark models of $FADL(p_Y^Q, p_F^Q, p_X^Q)$

$$Y_{t+1}^Q = \mu + \sum_{k=0}^{p_Y^Q - 1} \alpha_k Y_{t-k}^Q + \sum_{k=0}^{p_F^Q - 1} \beta_k F_{t-k}^Q + \sum_{k=0}^{p_X^Q - 1} \gamma_k X_{t-k}^Q + u_{t+1} \tag{2.9}$$

and $FAR(p_Y^Q, p_X^Q)$ when the regressor $X^Q$ is not present

$$Y_{t+1}^Q = \mu + \sum_{k=0}^{p_Y^Q - 1} \alpha_k Y_{t-k}^Q + \sum_{k=0}^{p_F^Q - 1} \beta_k F_{t-k}^Q + u_{t+1} \tag{2.10}$$

Finally, we consider model selection in the traditional setting, i.e. with respect to the choice between autoregressive (same frequency) models versus factor models or both combined. We consider, between zero and four quarterly (low frequency) lags, $p_Y$, of $Y_t(\theta_Y^M)$ and between one and four quarterly lags, $q_X$, of $X_t(\theta_X^D)$ and $F_t^Q$. In terms of the daily lags we consider and $p_X^D = 66, 132, 198, 264$ daily lags of $X_t^D$. We estimate the models with fixed lags but we also use AIC to select the number of quarterly and/or daily lags.

## 2.2 Nowcasting and Leads

Giannone, Reichlin, and Small (2008), among others, have formalized the process of updating the nowcast and forecasts as new releases of data become available. This process can be mimicked via MIDAS regression models with *leads*. Say we are one or two months into quarter $t+1$. Namely, we consider the MIDAS models with leads in order to incorporate real-time information available mainly on financial variables. Our objective is to forecast quarterly economic activity and in practice we often have a monthly release of macroeconomic data within the quarter and the equivalent of at least 44 trading days of financial data observed with no measurement error. This means that if we stand on the first day of the last month of the quarter and wish to make a forecast for the current quarter we could use and around 44 leads of daily data for financial markets that trade on weekdays.

Consider the Factor ADL model with MIDAS in equation (2.8), which allows for $J_X^D$ daily leads for the daily predictor. Then we can specify the $FADL-MIDAS(p_Y^Q, p_F^Q, p_X^Q, p_X^D, J_X^D)$ model

$$Y_{t+1}^Q = \mu + \sum_{k=0}^{p_Y^Q-1} \alpha_k Y_{t-k}^Q + \sum_{k=0}^{p_F^Q-1} \beta_k F_{t-k}^Q + \sum_{k=0}^{p_X^Q-1} \gamma_k X_{t-k+J_X^D}^Q(\theta_X^D) + u_{t+1}, \tag{2.11}$$

where

$$X_{t+J_X^D}^Q(\theta_X^D) \quad = \quad \sum_{i=1}^{p_X^D} w_i(\theta_X^D) X_{i+J_X^D,t}^D$$

When the aggregation weights are flat, $\theta_X^D = 0$, then equation (2.11) becomes a simple

$FADL$ forecasting model estimated by LS, in which $X^Q$ is recalculated by flat aggregation.

# 3 Daily and Quarterly Factors and Data

We forecast the US quarterly inflation rate and the growth rate of economic activity using various measures. For inflation we use monthly Consumer Price Index (CPI) and Core inflation (CPILFESL). For economic activity we use monthly Industrial Production (IP) and quarterly Real Gross Domestic Product (RGDP). We use two sample periods of US data for the post great moderation period: 1984:Q1-2008:Q4 (T=100) and 1999:Q1-2008:Q4 (T=40). The two subperiods involve different numbers, N, of daily financial predictors due to daily data availability. We chose to focus on two post samples in the post Great Moderation period because this period appears to mark a structural change in many US macroeconomic variables (Stock and Watson, 2008, van Dijk and Sensier, 2004) and it is also documented that it is relatively difficult to predict key macroeconomic variables vis-a-vis the pre-1985 period and vis-a-vis simple univariate models such as the RW or AO models. The relatively shorter sample enables us to examine the role of a larger cross-section of daily financial predictors in improving macroeconomic forecasts in the last two decades.

For the longer sample starting from the mid 1980s, we estimate our models using the period 1986:Q1-1997:Q1 while forecasts are obtained for the period 1997:Q2-2008:Q4. For the shorter sample the estimation and forecasting windows are given by 1999:Q1-2005Q4 and 2006Q1-2008Q4, respectively. Using a recursive estimation method we provide pseudo out-of-sample forecasts (see for instance, Stock and Watson, 1993) to evaluate the predictive ability of our models for various forecasting horizons $h = 1, 2, 4$. For each model we obtain the absolute MSFE:

$$MSFE(h) = \frac{1}{T_1 - T_2 - h + 1} \sum_{t=T_1}^{T_2-h} (\widehat{Y}_{t+h} - Y_{t+h})^2. \tag{3.12}$$

where the model is estimated for the period $1, ..., T_1$ and the forecasting period is given by $T_1 + h, ..., T_2$.

For conciseness we present the results of the shorter sample. We use two databases of two sampling frequencies of macroeconomic and financial indicators. The first is a quarterly dataset of 89 quarterly series of real output and income, capacity utilization, employment and hours, price indices, money e.t.c., described in detail in the Appendix, to extract the quarterly factors. The second is a daily dataset with large cross-section of 1125 daily series for the recent period of 1999-2008 and a relative smaller cross-section of for four categories of financial assets: (i) The equity category has 216 series of the major international stock market returns indices and Fama-French factors and portfolio returns as well as US stock market volume of indices and option volatilities of market indices. (ii) The foreign exchange rate category comprises the exchange rates of the 74 trading partners in the broad index as well as international currency rates of major indices and effective exchange rate indices. (iii) The fixed income securities involves 598 series such as government treasury bonds rates and yields, certificates of deposits, commercial paper, interest rate futures, TIPS yield, inflation compensation, LIBOR, corporate bonds indices yields and spreads. Finally, (iv) the commodities category has 237 series of US individual commodity prices, commodity indices and futures. Using the daily database we extract daily factors using this large cross-section of 1125 series as well as extract the factors from a smaller subset of 96 daily financial variables described in the Appendix.

The tables in the Appendix refer to the variables names, short description and transformations. The data source for the quarterly and daily series are the FRB and Haver Analytics, a data warehouse that collects the data series from their individual sources (such as the Federal Reserve Board (FRB) to Chicago Board of Trade (CBOT) and others), the Global Financial Database (GFD) and FRB, unless otherwise stated.

Following the methodology of Stock and Watson (2008) we estimate Dynamic Factor models to construct the factors. The series were transformed in order to eliminate trends by first differencing (in many cases after taking logarithms as reported in the Appendix). There are alternative approaches of choosing the number of factors. One approach is to use the information criteria (ICP) proposed by Bai and Ng (2002). In the case of quarterly factors ICP criteria suggest the choice of the first two factors. This result is consistent with the finding of Stock and Watson (2008). Interestingly, although our quarterly factors differ from the Stock and Watson factors by not including the series for which daily information is available, our first two factors highly correlate with those of Stock and Watson. This suggests that the excluded aggregated daily series were not important for the first two factors. In the case of daily factors, however, ICP criteria are not helpful because they always suggest 10 factors and therefore, we study all 10 daily factors one at a time. Table 1 shows the total variation explained by each daily factor as well as the series with the highest explanatory power for each factor. For instance, the first daily factor, which explains 22% of total variation, is explained by 40% by the Reuters/Jeffries CRB Commodity Index.

# 4 Empirical results

Table 2 presents the RMSFEs of the benchmark models for CPI inflation, namely the UCSV and AO models, as well as the RW model, which the literature adopts as the benchmark model for GDP and IP growth. In addition, Table 1 provides the summary results of the relative RMSFEs for the following categories of models: univariate AR models, traditional FAR models, (F)ADL and (F)-ADL-MIDAS models with individual daily predictors with and without quarterly factors, F as well as (F)ADL-DF and (F)-ADL-MIDAS-DF models, which are the same as the latter two categories with daily factors (instead of daily predictors). The main message from Table 2 is that best MIDAS models (with the lowest RMSFE) provide substantial forecasting gains relative to the benchmark forecasting models considered in the literature, such as the UCSV, AO, RW, AR and FAR, during the last ten years of US data. In particular, comparing best MIDAS models with the traditional models that use quarterly lags of averaged daily data to forecast CPI inflation, IP and GDP growth, MIDAS models improve around 45%, 40%, and 35% of the RMSFE of the UCSV, AO and RW benchmark models, respectively, for both 1 and 4 quarters ahead. Moreover, it is for 4 quarters ahead that MIDAS models for forecasting CPI inflation and GDP growth, yield substantial gains vis-a-vis the traditional (F)ADL models that impose a flat aggregation scheme for the daily data. Let us first focus on the results of CPI with no leads presented in Table 2. We find that the best model (with the lowest RMSFE across all models) belongs to the ADL-MIDAS models with relative RMSFE being 40% of the UCSV for both $h = 1$ and 4, as well as forecasting gains over the traditional ADL of 33% and 65% for $h = 1$ and 4, respectively. Similar results are found for real GDP growth forecasts. The model with the lowest RMSFE for predicting GDP growth is the FADL-MIDAS model with relative gains of 37% and 44%, for both $h = 1$ or 4, vis-a-vis the RW models, respectively. It also interesting to compare these forecasting gains of MIDAS based on individual daily predictors with the best models which use daily financial factors. The results in the lowest four rows of Table correspond to forecasting models with such Daily Factors (DF). It is interesting to point out that the corresponding best forecasting models for CPI inflation and GDP growth are

ADL-MIDAS-DF and FADL-MIDAS-DF, respectively, as opposed to the traditional models with a flat aggregation scheme. What is also especially interesting is that although the daily factors can not beat the forecasting performance of the best predictor for $h = 1$ or $h = 4$ for either real GDP growth and CPI inflation they do provide substantial improvements over the univariate benchmarks (UCSV, AO and RW) and over the quarterly FAR models

In Table 3 we look deeper into the details of the exact model specifications and predictors that yield the lowest RMSFEs. We start with the results for forecasting CPI inflation using no leads presented in Panel A which summarize the best three models in terms of lowest RMSFEs. We find that these are MIDAS specifications with no quarterly factors given by the ADL-MIDAS model and more precisely by equations (2.3) for $h = 1$ and (2.5) for $h = 4$ for CPI inflation forecasts. We find that the best daily predictor is Platinum returns at $h = 1$ and the A2/P2/F2 Nonfinancial commercial paper spread (with respect to AA Financial Commercial Paper or Federal Funds) at $h = 4$. What is particularly interesting is that the top percentile of best predictors for $h = 1$ and $h = 4$ involve commodity prices and interest rate spreads. Moreover, while the best predictor for CPI inflation at $h = 1$ is Platinum returns and the best 32 models in terms of lowest RMSFE involve Platinum returns, the second or third best predictors are Reuters Commodity Price Index and the 3Month Tbills based on traditional ADL models, with RMSFE which is almost twice as that of the best ADL-MIDAS models with Platinum returns. Similar results are obtained when forecasting CPI inflation with leads shown in Panel B. Again for both for $h = 1$ and 4 the models with the lowest RMSFE are ADL-MIDAS models with no factors given by equation (2.6). Similarly, among the best predictors we have again commodity prices and the nonfinancial commercial paper spread in addition to the Inflation Compensation for 5 years which features as the best predictor of CPI inflation using leads at $h = 1$.

We now turn to the results for the best models and best predictors for forecasting GDP growth found in Table 3. In Panels A and B we find that the best models for forecasting GDP growth (with no leads) for $h = 1$ are quarterly factor ADL-MIDAS models, FADL-MIDAS, given by equation (2.8) with daily Canadian/US Dollar returns as well as ADL-MIDAS-M models (with no quarterly factors) with the daily Aruoba, Diebold and Scotti (ADS) indicator given by equation (2.6). For $h = 4$ the best predictors of GDP growth are interest rate spreads with the best predictors being the A2/P2/F2 Nonfinancial Commercial Paper spread (minus the AA Financial Commercial Paper) as well as the 1month Eurodollar spread (with respect to the FF) rate. For using daily predictors with leads to forecast GDP growth, we find that mainly short term interest rates (and in particular LIBOR rates of 3months and 1 year) as well as corporate bond spreads of Merill Lynch AA and A ratings, Moody's Baa ratings minus the 10 Year Treasury bond rate are among the best predictors.

-TO BE COMPLETED-

# 5 Conclusion

-TO BE COMPLETED-

**Table 1: Analysis of Daily Factors**

The entries show the total variation explained by each daily factor - derived and discussed in section 3 - as well as the series with the highest explanatory power for each factor. Details of the data series appear in the Appendix, Table A2.

| Daily Factor | Total Variation | Series | Factor Variation |
|---|---|---|---|
| DF1 | 0.22 | Reuters/Jeffries CRB Index | 0.40 |
| DF2 | 0.16 | One Month A2P2F2 Non Fin. CP - One Month AA Fin.CP | 0.35 |
| DF3 | 0.15 | S&P 500 Index | 0.74 |
| DF4 | 0.12 | 10 Year Treasury Bond - Federal Funds Rate | 0.66 |
| DF5 | 0.08 | S&P GSCI Soybeans Index | 0.36 |
| DF6 | 0.07 | AAA - 10 year Treasury Bonds | 0.28 |
| DF7 | 0.07 | Swiss Franc/US$ Exchange Rate | 0.18 |
| DF8 | 0.05 | 6-Month LIBOR | 0.64 |
| DF9 | 0.04 | 6-Month Treasury Bills | 0.26 |
| DF10 | 0.04 | S&P GSCI Coffee Index | 0.27 |

**Table 2: Best Root Mean Square Forecast Error (RMSFE) and relative RSMFE using Daily Financial Predictors and Daily Factors**

Entries in the Table refer to Root Mean Squared Forecast Errors (RMSFE) for the benchmark models of AO, UCSV and RW. The rest of the entries refer to the relative RMSFE vis-a-vis the UCSV, for forecasting CPI inflation and vis-a-vis the RW for forecasting economic activity (GDP and IP). All models are summarized in Table A1. Entries below one imply improvements compared to the benchmark. The estimation period is 1999:Q1-2005:Q4 and the forecasting period is 2006:Q1-2008:Q4.

| | Panel A: No Leads | | | | | | Panel B: Lead of Two Month | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | CPI-all Inflation | | IP Growth | | GDP Growth | | CPI-all Inflation | | IP Growth | | GDP Growth | |
| h | 1 | 4 | 1 | 4 | 1 | 4 | 1 | 4 | 1 | 4 | 1 | 4 |
| UCSV | 4.28 | 1.44 | | | | | 4.28 | 1.44 | | | | |
| AO | 6.79 | 1.85 | | | | | 6.79 | 1.85 | | | | |
| RW | | | 6.1 | 4.26 | 3.35 | 1.69 | | | 6.1 | 4.26 | 3.35 | 1.69 |
| $AR$ | 1.04 | 1.17 | 0.62 | 1.1 | 0.96 | 1.03 | | | | | | |
| $FAR$ | 1.02 | 1.48 | 0.54 | 0.88 | 0.69 | 0.77 | 1.04 | 0.7 | 0.44 | 0.9 | 0.69 | 0.77 |
| $ADL$ | 0.57 | 0.73 | 0.36 | 0.52 | 0.59 | 0.31 | 0.81 | 1.11 | 0.3 | 0.47 | 0.52 | 0.45 |
| $FADL$ | 0.59 | 1.01 | 0.3 | 0.46 | 0.37 | 0.44 | 0.86 | 1.06 | 0.24 | 0.33 | 0.48 | 0.45 |
| $ADL-MIDAS$ | 0.43 | 0.44 | 0.28 | 0.5 | 0.48 | 0.3 | 0.68 | 1.19 | 0.3 | 0.53 | 0.52 | 0.35 |
| $FADL-MIDAS$ | 0.47 | 0.72 | 0.37 | 0.43 | 0.36 | 0.28 | 0.77 | 0.85 | 0.28 | 0.41 | 0.33 | 0.36 |
| $ADL-DF$ | 0.89 | 0.93 | 0.36 | 0.44 | 0.58 | 0.54 | 0.98 | 1.49 | 0.34 | 0.73 | 0.58 | 0.37 |
| $FADL-DF$ | 0.94 | 1.28 | 0.4 | 0.66 | 0.61 | 0.53 | 0.98 | 1.37 | 0.37 | 0.55 | 0.6 | 0.46 |
| $ADL-MIDAS-DF$ | 0.83 | 0.77 | 0.42 | 0.89 | 0.61 | 0.68 | 0.98 | 1.55 | 0.35 | 0.72 | 0.62 | 0.72 |
| $FADL-MIDAS-DF$ | 0.83 | 1.26 | 0.39 | 0.65 | 0.61 | 0.50 | 0.98 | 1.35 | 0.36 | 0.52 | 0.49 | 0.47 |

Table continued on next page ...

Table 2 continued

| | CPI and CPIcore: Best Models | Predictor | Best relative RMSFE | Ranking of Best Predictors: | | Relative RMSFE |
|---|---|---|---|---|---|---|
| | | | Panel C: No leads - CPI | | | |
| h=1 | | | | | | |
| 1 | MIDASq-ADL(4,0,1,1) | Platinum | 0.38 | | Platinum | 0.38 |
| 2 | MIDASq-ADL(4,0,1,3) | Platinum | 0.38 | 43 | DJAIG Comm | 0.62 |
| 3 | MIDASq-ADL(4,0,1,4) | Platinum | 0.38 | 48 | RJ CRB Comm | 0.63 |
| h=4 | | | | | | |
| 1 | MIDASq-ADL(AIC,0,AIC,1) | M1 A2P2F2 NFCP - M1 AA NCP | 0.30 | 1 | M1 A2P2F2 NFCP - M1 AA NFCP | 0.30 |
| 2 | MIDASq-ADL(AIC,0,AIC,1) | M1A2P2F2 NFCP - M1AA Fin CP | 0.40 | 2 | M1A2P2F2 NFCP - M1AA FCP | 0.40 |
| 3 | MIDASq-ADL(AIC,0,AIC,1) | M1A2P2F2 NFCP - FF | 0.53 | 3 | M1A2P2F2 NFCP - FF | 0.53 |
| | | | Panel D: Lead of Two Month - CPI | | | |
| h=1 | | | | | | |
| 1 | ADL(AIC,0,AIC) | SP GSCI TER | 0.53 | 1 | SP GSCI TER | 0.53 |
| 2 | ADL(AIC,0,AIC) | SP GSCI Energy Commod TR | 0.53 | 2 | SP GSCI Energy Commod TR | 0.53 |
| 3 | MIDASm-ADL(AIC,1,0,AIC,1) | BKEVEN5 | 0.54 | 3 | BKEVEN5 | 0.54 |
| h=4 | | | | | | |
| 1 | MIDASq-ADL(4,1,1,4) | Sugar | 0.86 | 1 | Sugar | 0.86 |
| 2 | ADL(4,0,4) | Zinc | 0.99 | 2 | Zinc | 0.99 |
| 3 | ADL(4,1,4) | Sugar | 1.00 | 4 | SP GSCI TR | 1.02 |

Table 2 continued

| | CPI and CPIcore: Best Models | Predictor | Best relative RMSFE | Ranking of Best Predictors: | | Relative RMSFE |
|---|---|---|---|---|---|---|
| | | | Panel E: No leads - CPI Core | | | |
| h=1 | | | | | | |
| 1 | MIDASq-ADL(4,0,1,1) | Platinum | 0.38 | 1 | Platinum | 0.38 |
| 2 | MIDASq-ADL(4,0,1,3) | Platinum | 0.38 | 43 | DJAIG Comm | 0.62 |
| 3 | MIDASq-ADL(4,0,1,4) | Platinum | 0.38 | 48 | RJ CRB Comm | 0.63 |
| h=4 | | | | | | |
| 1 | MIDASq-ADL(AIC,0,AIC,1) | M1 A2P2F2 NFCP - M1 AA NCP | | 1 | M1 A2P2F2 NFCP - M1 AA NFCP | 0.30 |
| 2 | MIDASq-ADL(AIC,0,AIC,1) | M1A2P2F2 NFCP - M1AA Fin CP | 0.40 | 2 | M1A2P2F2 NFCP - M1AA FCP | 0.40 |
| 3 | MIDASq-ADL(AIC,0,AIC,1) | M1A2P2F2 NFCP - FF | 0.53 | 3 | M1A2P2F2 NFCP - FF | 0.53 |
| | | | Panel F: Lead of One Month - CPI Core | | | |
| h=1 | | | | | | |
| 1 | ADL(AIC,0,AIC) | SP GSCI TER | 0.53 | 1 | SP GSCI TER | 0.53 |
| 2 | ADL(AIC,0,AIC) | SP GSCI Energy Commod TR | 0.53 | 2 | SP GSCI Energy Commod TR | 0.53 |
| 3 | MIDASm-ADL(AIC,1,0,AIC,1) | BKEVEN5 | 0.54 | 3 | BKEVEN5 | 0.54 |
| h=4 | | | | | | |
| 1 | MIDASq-ADL(4,1,1,4) | Sugar | 0.86 | 1 | Sugar | 0.86 |
| 2 | ADL(4,0,4) | Zinc | 0.99 | 2 | Zinc | 0.99 |
| 3 | ADL(4,1,4) | Sugar | 1.00 | 4 | SP GSCI TR | 1.02 |

**Table 3: Output growth out-of-sample relative RSMFE Distribution using Daily Financial Predictors**

Entries in the Table refer to Root Mean Squared Forecast Errors (RMSFE) for the RW benchmark model and the rest of the entries to relative RMSFE vis-a-vis the RW, for forecasting GDP and IP growth. All models are summarized in Table A1. Some point estimates of the empirical distribution of the relative RMSFEs are reported given by $BEST$, $WORST$ which refer to the best and worst models in terms of minimum and maximum RMSFEs, respectively, as well as the percentile estimates of the relative RMSFEs of $P0i$, $i = 05, 10, 50, 90, 95, 99$. The estimation period is 1999:Q1-2005:Q4 and the forecasting period is 2006:Q1-2008:Q4.

| | | Panel A: No Leads | | | | | | Panel B: Lead of One Month | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | GDP Growth | | | IP growth | | | GDP Growth | | | IP Growth | | |
| | h | 1 | 2 | 4 | 1 | 2 | 4 | 1 | 2 | 4 | 1 | 2 | 4 |
| UCSV | | 4.28 | 1.44 | | | | | 4.28 | 1.44 | | | | |
| AO | | 6.79 | 1.85 | | | | | 6.79 | 1.85 | | | | |
| $ADL-MIDAS/RW$ | | | | | | | | | | | | | |
| | BEST | 0.43 | 0.44 | 0.48 | 0.3 | 1.31 | 1.74 | 0.68 | 1.19 | 0.52 | 0.35 | 0.3 | 0.53 |
| | P99 | 0.51 | 0.79 | 0.54 | 0.57 | 0.4 | 0.55 | 0.92 | 1.43 | 0.56 | 0.48 | 0.32 | 0.57 |
| | P95 | 0.9 | 0.95 | 0.8 | 0.8 | 0.5 | 0.83 | 1.05 | 1.55 | 0.62 | 0.75 | 0.41 | 0.76 |
| | P90 | 0.97 | 1 | 0.83 | 0.89 | 0.55 | 0.89 | 1.14 | 1.62 | 0.68 | 0.82 | 0.43 | 0.86 |
| | P50 | 1.08 | 1.2 | 0.97 | 1.05 | 0.64 | 1.09 | 1.27 | 1.87 | 0.95 | 1.03 | 0.49 | 1.08 |
| | P10 | 1.15 | 1.68 | 1.05 | 1.19 | 0.75 | 1.18 | 1.42 | 2.33 | 1.08 | 1.23 | 0.61 | 1.2 |
| | P05 | 1.21 | 1.94 | 1.09 | 1.27 | 0.79 | 1.23 | 1.56 | 2.62 | 1.16 | 1.29 | 0.67 | 1.25 |
| | WORST | 1.36 | 2.59 | 1.5 | 1.66 | 0.28 | 0.5 | 1.88 | 3.34 | 1.47 | 1.59 | 1.02 | 1.73 |
| $ADL/RW$ | | | | | | | | | | | | | |
| | BEST | 0.57 | 0.73 | 0.59 | 0.31 | 0.36 | 0.52 | 0.81 | 1.11 | 0.52 | 0.35 | 0.3 | 0.47 |
| | P99 | 0.7 | 0.79 | 0.69 | 0.64 | 0.43 | 0.6 | 0.89 | 1.27 | 0.56 | 0.48 | 0.33 | 0.51 |
| | P95 | 0.92 | 0.97 | 0.8 | 0.77 | 0.5 | 0.85 | 1.01 | 1.52 | 0.62 | 0.75 | 0.4 | 0.71 |
| | P90 | 1 | 1.03 | 0.84 | 0.86 | 0.55 | 0.9 | 1.11 | 1.64 | 0.68 | 0.82 | 0.43 | 0.85 |
| | P50 | 1.09 | 1.38 | 0.98 | 1.03 | 0.64 | 1.1 | 1.27 | 1.98 | 0.95 | 1.03 | 0.49 | 1.1 |
| | P10 | 1.17 | 1.78 | 1.07 | 1.22 | 0.77 | 1.2 | 1.44 | 2.4 | 1.08 | 1.23 | 0.59 | 1.21 |
| | P05 | 1.21 | 1.93 | 1.12 | 1.21 | 0.8 | 1.28 | 1.57 | 2.73 | 1.16 | 1.29 | 0.63 | 1.25 |
| | WORST | 1.41 | 2.55 | 1.45 | 1.87 | 1. 04 | 1.93 | 1.87 | 3.1 | 1.47 | 1.59 | 0.89 | 1.88 |

Table continued on next page ...

Table 3 continued

| | | Panel A: No Leads | | | | | | Panel B: Lead of One Month | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | GDP Growth | | | IP Growth | | | GDP Growth | | | IP Growth | | |
| $FADL/RW$ | | | | | | | | | | | | | |
| | BEST | 0.59 | 1.01 | 0.37 | 0.44 | 0.3 | 0.46 | 0.86 | 1.06 | 0.48 | 0.55 | 0.24 | 0.33 |
| | P99 | 0.73 | 1.04 | 0.58 | 0.5 | 0.4 | 0.56 | 0.96 | 1.54 | 0.51 | 0.5 | 0.31 | 0.48 |
| | P95 | 0.91 | 1.3 | 0.66 | 0.71 | 0.47 | 0.65 | 1.07 | 1.75 | 0.56 | 0.59 | 0.37 | 0.67 |
| | P90 | 0.97 | 1.47 | 0.69 | 0.78 | 0.5 | 0.75 | 1.17 | 1.88 | 0.62 | 0.68 | 0.4 | 0.77 |
| | P50 | 1.12 | 2.08 | 0.8 | 1 | 0.58 | 0.9 | 1.34 | 2.64 | 0.81 | 0.99 | 0.49 | 0.93 |
| | P10 | 1.21 | 2.47 | 0.93 | 1.27 | 0.74 | 1.1 | 1.58 | 3.16 | 0.99 | 1.28 | 0.67 | 1.13 |
| | P05 | 1.23 | 2.62 | 0.99 | 1.35 | 0.79 | 1.24 | 1.66 | 3.29 | 1.08 | 1.41 | 0.72 | 1.22 |
| | WORST | 1.35 | 3.6 | 1.25 | 2.02 | 1.02 | 1.67 | 2.01 | 5.09 | 1.76 | 1.91 | 0.96 | 1.67 |
| $FADL-MIDAS/RW$ | | | | | | | | | | | | | |
| | BEST | 0.47 | 0.72 | 0.36 | 0.28 | 0.39 | 0.65 | 0.77 | 0.85 | 0.33 | 0.36 | 0.28 | 0.41 |
| | P99 | 0.56 | 1.08 | 0.55 | 0.5 | 0.41 | 0.65 | 0.95 | 1.42 | 0.44 | 0.48 | 0.31 | 0.53 |
| | P95 | 0.9 | 1.25 | 0.61 | 0.7 | 0.45 | 0.66 | 1.07 | 1.67 | 0.52 | 0.63 | 0.38 | 0.67 |
| | P90 | 0.97 | 1.34 | 0.65 | 0.82 | 0.46 | 0.71 | 1.2 | 1.76 | 0.57 | 0.72 | 0.41 | 0.75 |
| | P50 | 1.09 | 1.89 | 0.77 | 1.02 | 0.58 | 0.91 | 1.31 | 2.41 | 0.75 | 1.01 | 0.48 | 0.91 |
| | P10 | 1.18 | 2.35 | 0.88 | 1.28 | 0.67 | 1.09 | 1.52 | 3.01 | 0.91 | 1.3 | 0.62 | 1.06 |
| | P05 | 1.24 | 2.46 | 0.94 | 1.34 | 0.7 | 1.12 | 1.63 | 3.23 | 0.99 | 1.42 | 0.68 | 1.13 |
| | WORST | 1.54 | 3.28 | 1.22 | 1.84 | 0.91 | 1.3 | 2.46 | 4.26 | 1.56 | 2.19 | 1 | 1.64 |

Table 3 continued

| | GDP and IP Growth: Best Models | Predictor | Best relative RMSFE | Ranking of Best Predictors: | | Relative RMSFE |
|---|---|---|---|---|---|---|
| | | | Panel C: GDP No leads | | | |
| h=1 | | | | | | |
| 1 | | Platinum | 0.38 | 1 | Platinum | 0.38 |
| 2 | | Platinum | 0.38 | 43 | DJAIG Comm | 0.62 |
| 3 | | Platinum | 0.38 | 48 | RJ CRB Comm | 0.63 |
| h=4 | | | | | | |
| 1 | MIDASq-ADL(AIC,0,AIC,1) | M1 A2P2F2 NFCP - M1 AA NCP | 0.30 | 1 | M1 A2P2F2 NFCP - M1 AA NFCP | 0.30 |
| 2 | MIDASq-ADL(AIC,0,AIC,1) | M1A2P2F2 NFCP - M1AA Fin CP | 0.40 | 2 | M1A2P2F2 NFCP - M1AA FCP | 0.40 |
| 3 | MIDASq-ADL(AIC,0,AIC,1) | M1A2P2F2 NFCP - FF | 0.53 | 3 | M1A2P2F2 NFCP - FF | 0.53 |
| | | | Panel D: GDP Lead of Two Month | | | |
| h=1 | | | | | | |
| 1 | ADL(AIC,0,AIC) | SP GSCI TER | 0.53 | 1 | SP GSCI TER | 0.53 |
| 2 | ADL(AIC,0,AIC) | SP GSCI Energy Commod TR | 0.53 | 2 | SP GSCI Energy Commod TR | 0.53 |
| 3 | MIDASm-ADL(AIC,1,0,AIC,1) | BKEVEN5 | 0.54 | 3 | BKEVEN5 | 0.54 |
| h=4 | | | | | | |
| 1 | MIDASq-ADL(4,1,1,4) | Sugar | 0.86 | 1 | Sugar | 0.86 |
| 2 | ADL(4,0,4) | Zinc | 0.99 | 2 | Zinc | 0.99 |
| 3 | ADL(4,1,4) | Sugar | 1.00 | 4 | SP GSCI TR | 1.02 |

# References


Andreou, E., E. Ghysels, and A. Kourtellos, 2008, Regression Models With Mixed Sampling Frequencies, *Journal of Econometrics*, forthcoming.

Armesto, M.T., R. Hernandez-Murillo, M. Owyang, and J. Piger, 2008, Measuring the Information Content of the Beige Book: A Mixed Data Sampling Approach, *Journal of Money, Credit and Banking (forthcoming)*.

Aruoba, S.B., F.X. Diebold, and C. Scotti, 2009, Real-time measurement of business conditions, *Journal of Business and Economic Statistics* (forthcoming).

Bai, J., 2003, Inferential theory for factor models of large dimensions, *Econometrica* pp. 135–171.

——— , E. Ghysels, and J.H. Wright, 2009, State Space Models and MIDAS Regressions, Discussion Paper, NY Fed, UNC and Johns Hopkins.

Bai, J., and S. Ng, 2002, Determining the number of factors in approximate factor models, *Econometrica* pp. 191–221.

Bernanke, Ben, Mark Gertler, and Mark Watson, 1997, Systematic monetary policy and the effects of oil price shocks, *Brookings Papers on Economic Activity* 1, 91–157.

Chen, X., and E. Ghysels, 2009, News - good or bad - and its impact on predicting future volatility, Discussion Paper, UNC.

Clements, M.P., and A.B. Galvão, 2008a, Forecasting US output growth using Leading Indicators: An appraisal using MIDAS models, *Journal of Applied Econometrics (forthcoming)*.

Clements, M., and A. Galvão, 2008b, Macroeconomic Forecasting with Mixed Frequency Data: Forecasting US output growth, *Journal of Business and Economic Statistics* 26, 546–554.

Forni, M., M. Hallin, M. Lippi, and L. Reichlin, 2000, The generalized dynamic-factor model: Identification and estimation, *Review of Economics and Statistics* 82, 540–554.

——— , 2003, Do financial variables help forecasting inflation and real activity in the euro area?, *Journal of Monetary Economics* 50, 1243–1255.

———, 2005, The generalized dynamic factor model, *Journal of the American Statistical Association* 100, 830–840.

Galvão, A.B., 2006, Changes in Predictive Ability with Mixed Frequency Data, Discussion Paper QUeen Mary.

Ghysels, Eric, Pedro Santa-Clara, and Rossen Valkanov, 2002, The MIDAS touch: Mixed data sampling regression models, Working paper, UNC and UCLA.

———, 2006, Predicting volatility: getting the most out of return data sampled at different frequencies, *Journal of Econometrics* 131, 59–95.

Ghysels, E., A. Sinko, and R. Valkanov, 2006, MIDAS regressions: Further results and new directions, *Econometric Reviews* 26, 53–90.

Ghysels, Eric, and Jonathan Wright, 2008, Forecasting professional forecasters, *Journal of Business and Economic Statistics* (forthcoming).

Giannone, D., L. Reichlin, and D. Small, 2008, Nowcasting: The real-time informational content of macroeconomic data, *Journal of Monetary Economics* 55, 665–676.

Hamilton, J.D., 2006, Daily Monetary Policy Shocks and the Delayed Response of New Home Sales, working paper, UCSD.

Harvey, Andrew, 1989, *Forecasting, Structural Time Series Models and the Kalman Filter* (Cambridge University Press, Cambridge).

Harvey, Andrew C., and Richard G. Pierse, 1984, Estimating missing observations in economic time series, *Journal of the American Statistical Association* 79, 125–131.

Kuzin, V., M. Marcellino, and C. Schumacher, 2009, MIDAS versus mixed-frequency VAR: nowcasting GDP in the euro area, Discussion Paper 07/2009 Deutsche Bundesbank.

Mariano, R.S., and Y. Murasawa, 2003, A new coincident index of business cycles based on monthly and quarterly series, *Journal of Applied Econometrics* 18, 427–443.

Mittnik, S., and P. Zadrozny, 2004, *Forecasting quarterly German GDP at monthly intervals using monthly Ifo business conditions data* (Springer).

Monteforte, Libero, and Gianluca Moretti, 2008, Real time forecasts of inflation: The role of financial variables, Discussion paper Bank of Italy.

Schumacher, C., and J. Breitung, 2008, Real-time forecasting of German GDP based on a large factor model with monthly and quarterly data, *International Journal of Forecasting* 24, 386–398.

Stock, J.H., and M.W. Watson, 1989, New indexes of coincident and leading economic indicators, *NBER macroeconomics annual* pp. 351–394.

——— , 2002, Macroeconomic forecasting using diffusion indexes, *Journal of Business and Economic Statistics* 20, 147–162.

——— , 2003, Forecasting output and inflation: the role of asset prices, *Journal of Economic Literature* pp. 788–829.

——— , 2008, Forecasting in Dynamic Factor Models Subject to Structural Instability, in Jennifer Castle, and Neil Shephard, ed.: *The Methodology and Practice of Econometrics, A Festschrift in Honour of Professor David F. Hendry*. Oxford University Press.

Tay, A., 2006, Financial Variables as Predictors of Real Output Growth, Discussion Paper SMU.

Tay, A.S., 2007, Mixing Frequencies: Stock Returns as a Predictor of Real Output Growth, Discussion Paper SMU.

Zadrozny, P.A., 1990, Forecasting US GNP at monthly intervals with an estimated bivariate time series model, *Federal Reserve Bank of Atlanta Economic Review* 75, 2–15.

## Table A1: Summary of Models

| Name | Eq. # | Equation |
|---|---|---|
| $AR(p_Y^Q)$ | | $Y_{t+1}^Q = \mu + \sum_{k=0}^{p_Y^Q-1} \alpha_k Y_{t-k}^Q + u_{t+1}$ |
| $EL-AR(p_Y^Q)$ | | $Y_{t+1}^Q = \mu + \alpha \sum_{k=0}^{p_Y^Q-1} w_k(\theta_X^Q) Y_{t-k}^Q + u_{t+1}$ |
| $FAR(p_Y^Q, p_X^Q)$ | 2.10 | $Y_{t+1}^Q = \mu + \sum_{k=0}^{p_Y^Q-1} \alpha_k Y_{t-k}^Q + \sum_{k=0}^{p_F^Q-1} \beta_k F_{t-k}^Q + u_{t+1}$ |
| $ADL-MIDAS(p_Y^Q, p_X^Q, p_X^D)$ | 2.5 | $Y_{t+1}^Q = \mu + \sum_{k=0}^{p_Y^Q-1} \alpha_k Y_{t-k}^Q + \sum_{k=0}^{p_X^Q-1} \beta_k X_{t-k}^Q(\theta_X^D) + u_{t+1}$ |
| $ADL(p_Y^Q, p_X^Q)$ | 2.1 | $Y_{t+1}^Q = \mu + \sum_{k=0}^{p_Y^Q-1} \alpha_k Y_{t-k}^Q + \sum_{k=0}^{p_X^Q-1} \beta_k X_{t-k}^Q + u_{t+1}$ |
| $ADL-MIDAS-M(p_Y^Q, p_X^Q, p_X^D)$ | 2.6 | $Y_{t+1}^Q = \mu + \sum_{k=0}^{p_Y^Q-1} \alpha_k Y_{t-k}^Q + \beta \sum_{k=0}^{p_X^Q-1} w_k(\theta_X^Q) X_{t-k}^Q(\theta_X^D) + u_{t+1}$ |
| $EL-ADL(p_Y^Q, p_X^Q)$ | | $Y_{t+1}^Q = \mu + \sum_{k=0}^{p_Y^Q-1} \alpha_k Y_{t-k}^Q + \beta \sum_{k=0}^{p_X^Q-1} w_k(\theta_X^Q) X_{t-k}^Q + u_{t+1}$ |
| $FADL-MIDAS(p_Y^Q, p_F^Q, p_X^Q, p_X^D)$ | 2.8 | $Y_{t+1}^Q = \mu + \sum_{k=0}^{p_Y^Q-1} \alpha_k Y_{t-k}^Q + \sum_{k=0}^{p_F^Q-1} \beta_k F_{t-k}^Q + \sum_{k=0}^{p_X^Q-1} \gamma_k X_{t-k}^Q(\theta_X^D) + u_{t+1}$ |
| $FADL(p_Y^Q, p_F^Q, p_X^Q)$ | 2.9 | $Y_{t+1}^Q = \mu + \sum_{k=0}^{p_Y^Q-1} \alpha_k Y_{t-k}^Q + \sum_{k=0}^{p_F^Q-1} \beta_k F_{t-k}^Q + \sum_{k=0}^{p_X^Q-1} \gamma_k X_{t-k}^Q + u_{t+1}$ |
| $FADL-MIDAS-M(p_Y^Q, p_F^Q, p_X^Q, p_X^D)$ | | $Y_{t+1}^Q = \mu + \sum_{k=0}^{p_Y^Q-1} \alpha_k Y_{t-k}^Q + \sum_{k=0}^{p_F^Q-1} \beta_k F_{t-k}^Q + \gamma \sum_{k=0}^{p_X^Q-1} w_k(\theta_X^Q) X_{t-k}^Q(\theta_X^D) + u_{t+1}$ |
| $FADL-MIDAS(p_Y^Q, p_F^Q, p_X^Q, p_X^D, J_X^D)$ | 2.11 | $Y_{t+1}^Q = \mu + \sum_{k=0}^{p_Y^Q-1} \alpha_k Y_{t-k}^Q + \sum_{k=0}^{p_F^Q-1} \beta_k F_{t-k}^Q + \sum_{k=0}^{p_X^Q-1} \gamma_k X_{t-k+J_X^D}^Q(\theta_X^D) + u_{t+1}$ |

**Table A2: Summary of Data Series**

| Name | Definition |
|---|---|