Empir Econ (2019) 56:367–395
https://doi.org/10.1007/s00181-017-1357-8

Forecasting Turkish real GDP growth in a data-rich environment

Bahar Şen Doğan[1] · **Murat Midiliç**[2]

Received: 12 May 2016 / Accepted: 8 August 2017 / Published online: 8 January 2018

Abstract This study generates nowcasts and forecasts for the growth rate of the gross domestic product in Turkey using 204 daily financial series with mixed data sampling (MIDAS) framework. The daily financial series include commodity prices, equity indices, exchange rates, and global and domestic corporate risk series. Forecasting exercises are also carried out with the daily factors extracted from separate financial data classes and from the whole dataset. The findings of the study suggest that MIDAS regression models and forecast combinations provide advantage in exploiting information from daily financial data compared to the models using simple aggregation schemes. In addition, incorporating daily financial data into the analysis improves the forecasts substantially. These results indicate that both the information content of the financial data and the flexible data-driven weighting scheme of MIDAS regressions play an essential role in forecasting the future state of the Turkish economy.

Keywords Real GDP growth · Forecasting · MIDAS

JEL Classification C22 · C53 · G10

1 Introduction

Gross domestic product (GDP) is one of the most important indicators which summarize economic activity. However, it is unavailable at frequencies higher than the

✉ Bahar Şen Doğan
e199134@metu.edu.tr

Murat Midiliç
murat.midilic@ugent.be

1 Economics Department, Middle East Technical University, Ankara, Turkey

2 Department of Financial Economics, Ghent University, Ghent, Belgium

quarterly level and released with a significant lag. Considering the importance of timely assessment of current and future economic developments, forecasting GDP is an important task for policymakers. To forecast GDP data, analysts usually track timelier and high-frequency indicators, including monthly, weekly or daily data. In the literature, leading property of financial market variables about the future state of the economy has been pointed out (Stock and Watson 2003). In this study, we focus on the quarter-on-quarter Turkish GDP growth between 2000-Q2 and 2016-Q2 and generate nowcasts and forecasts for Turkish quarterly GDP growth by employing high-frequency financial data with the mixed data sampling (MIDAS) models. We evaluate forward-looking nature of financial variables for the Turkish economy over the period starting from the second quarter of 2010 until the second quarter of 2016.

Regarding the use of high-frequency data in forecasting GDP, mapping movements in these indicators into real GDP growth is not an easy task because there is mixed frequency in the data. For example, GDP data are sampled quarterly, employment and inflation data are sampled monthly, and asset price data are sampled daily. The simplest approach is to aggregate the high-frequency data to obtain a balanced data set at the same low frequency. Explicitly, a forecaster may time aggregate three monthly samples of employment data into a single observation for each quarterly sample of GDP data by taking an average of the monthly data. However, it is simple; in general, temporal aggregation entails a loss of information (Marcellino 1999). Moreover, it also changes the data generating mechanism, which may lead to considerable difference between the dynamics of the aggregate and high- or mixed-frequency model. This indicates that key econometric features can be spuriously modified when employing the aggregated data (Marcellino 1999). In this paper, we follow MIDAS approach developed by Ghysels et al. (2004). MIDAS models are one of alternative approaches to model the mixed-frequency data to avoid the mentioned problems related to temporal aggregation. They have parsimonious specifications based on distributed lag polynomials, which flexibly deal with data sampled at different frequencies and provide a direct forecast of the low-frequency variable (Ghysels et al. 2004; Clements and Galvão 2008).

The advantages of using MIDAS regressions in terms of improving quarterly macro-forecasts with monthly data, or improving quarterly and monthly macroeconomic predictions with a small set of daily financial data have been investigated in the literature (Kuzin et al. 2011; Armesto et al. 2009; Clements and Galvão 2008, 2009; Galvão 2013; Tay 2007; Schumacher and Breitung 2008; Ghysels and Wright 2009; Hamilton 2008; Monteforte and Moretti 2013). However, these studies limit the number of variables in the regression and do not incorporate large set of data into their analysis. Apart from the previous works, (Andreou et al. 2013) use large cross section of around one thousand financial time series to forecast quarterly US GDP growth rate. They extract a small set of daily factors from these large dataset and apply forecast combination methods. Their application shows that MIDAS model forecasts are able to outperform naive models substantially.

Most of the existing studies employing MIDAS regression models focus on the US or the euro area economy to forecast quarterly series (Tay 2007; Clements and Galvão 2009), but few on the developing economies. In this study, we take Turkey as our laboratory. The reason is that it is a small open economy and exhibits similar macroeconomic dynamics to other emerging economies. Assessing the ability of financial market variables to anticipate Turkish output growth may provide valuable information about the forward-looking nature of financial variables in peer countries.

The literature related to forecasting Turkish GDP is quite limited. In one of the rare studies, Akkoyun and Günay (2012) employ small-scale dynamic factor model. They forecast quarterly GDP growth with monthly data by following Kalman filter approach. They find that some combinations of hard indicators such as industrial production, export and import quantity indices and soft indicators such as Purchasing Managers Index (PMI) and PMI new orders have predictive power for the output growth. Sacakli-Sacildi (2015) tests forecasting performance of Bayesian vector autoregression (BVAR) methodology for Turkish GDP growth rate. However, to our best knowledge, MIDAS regression modeling has not been used for forecasting Turkish GDP growth and this paper is the first attempt in exploring forecasting ability of financial data using MIDAS regression models for Turkish output growth. In this study, we follow the forecasting strategy introduced by Andreou et al. (2013) and generate nowcasts and forecasts of Turkish GDP growth using 204 daily financial series. We analyze whether financial data provide information about the future state of the economy and improve forecasting ability.

Our results suggest that there are statistically significant gains in using daily financial data and MIDAS models. MIDAS models with daily series and quarterly factors perform better than the benchmark model and other linear models at every horizon in consideration. The robustness of the results is proven by different predictive ability tests. Another implication of the results is that daily factors extracted from commodities, forex and domestic corporate risk series give the best forecasts at certain horizons, and in others, their results are competitive with other financial data classes. The results also reveal that combination of small set of daily factors received from separate financial data classes provides substantial forecasting gains especially in the medium run. This implies that the small set of series can be employed in following the Turkish economy.

The article is organized as follows. Section 2 describes the MIDAS models and the forecasting methodology. Section 3 describes the dataset, and Sect. 4 presents the empirical results. Finally, Sect. 5 concludes.

2 Forecasting methodology

This section describes the MIDAS models and forecast combination methods that are used in the study. The aim of the section is to give the intuition and details of MIDAS models and how they are implemented in macroeconomic forecasting. The model is compared and contrasted with other methods in the literature to give a clear picture of its flexibility as well as its limitations. To improve forecast performance, the study uses forecast combination methods. The second part of the section explains the forecast combination methods and how they are used.

2.1 MIDAS models

In order to understand innovations brought by MIDAS to single equation estimation of mixed-frequency data, first consider a distributed lag (DL) model which is typically employed in the literature to describe the distribution over time of the lagged effects of a change in the explanatory variable. In this linear regression methods, once all the high-frequency values are aggregated to the corresponding low-frequency values.

$$Y_t = \beta_0 + +\beta(L)X_t + \varepsilon_t, \tag{1}$$

where $\beta(L)$ is some finite or infinite lag polynomial operator, Y_t is the low-frequency variable, X_t is the high-frequency variable and ε_t is the error term. DL regression involves temporally aggregated series, based, for example, on equal weights of daily data. In this setup, reliable estimations of the low-frequency variable necessitates the correct form of aggregation of the high-frequency variable into low frequency. Therefore, a crucial drawback of the model is that the imposed aggregation schemes might lead to severe aggregation bias.

A potential solution to the aggregation bias is to quit the aggregation and project Y_t on, for instance, every observation of the high-frequency variable for each lag of the low frequency. This approach might work well when the fixed number of high-frequency periods in one single low-frequency period is small (e.g., it is 3 such as in the case of estimation of a quarterly variable with monthly data). However, for instance, in a quarterly low-frequency and daily high-frequency setting, even only one lag of the quarterly data is used, there are more than 60 parameters to estimate corresponding each day in the lagged quarter. Thus, parameter proliferation is another potential problem with the single equation models.

Novelty of the MIDAS models lies in its solution to handling high- frequency data. In its simplest form, the novelty can be described as aggregating data parsimoniously and allowing data to decide on the aggregation weights. MIDAS does these by using polynomial distributed lag functions to weigh the high-frequency data. Consider the basic MIDAS model:[1]

$$Y_t = \beta_0 + \beta_1 \sum_{j=0}^{(q_X-1)} \sum_{i=0}^{(m-1)} \omega(\theta, i + j \times m) X_{(m-i,t-j)} + \varepsilon_{(t)}, \tag{2}$$

where $\omega(\theta, i + j \times m)$ is the polynomial distributed lag function that depends on hyperparameter θ; q_X is the number of low-frequency lags; m is the fixed number of high-frequency periods in one single low-frequency period (e.g., 3 months in a quarter); (i, t) is the ith lagged high-frequency period in the low-frequency period t; and j is low-frequency lag.

Specification of weights is central to the MIDAS models since it gives the parsimony of the model over models such as bridge equations and autoregressive distributed lag (ADL). There is a number of polynomial functions suggested in the literature some of which are beta function, exponential Almon lag function, step function and Almon

[1] For a detailed overview of MIDAS models, please see Ghysels et al. (2004), Wohlrabe (2009), and Foroni and Marcellino (2013).

distributed lag polynomial. This study is using Almon distributed lag polynomial[2] in which the weights can be written as

$$\omega(\theta, i) = \sum_{p=0}^{n_p} \theta_p i^p, \tag{3}$$

and assumes that the weight of the ith lag can be calculated with underlying $(n_p + 1)$ hyperparameters $\theta = (\theta_0, \ldots, \theta_p)$. Now, assume that one wants to estimate quarterly data with daily data and a quarter consists of 66 working days. Instead of assigning weights separately to each lag and estimating 66 coefficients, with two or three hyperparameters, the Almon lag polynomial solves the parameter proliferation problem. Furthermore, weights are data driven which prevents aggregation bias. The parameters of the MIDAS regression model are estimated by nonlinear least squares.

Ghysels et al. (2004) show that MIDAS is always more efficient than simple data aggregation. MIDAS gives a reduced form method which does not require complexity of other alternative methods in the literature such as Kalman filters. The performance of MIDAS models against state-space models has been discussed by Andreou et al. (2011) and Bai et al. (2013). These studies show that in simple frameworks Kalman filters are found to be more efficient. However, comparisons with more complex problems imply that MIDAS models are less "prone" to specification errors and are easier to implement with a few parameters. This is why in a data-rich environment, such as in the empirical study of Andreou et al. (2013) with over 900 daily financial series, MIDAS can be preferred as the forecasting model.

2.2 Forecasting with MIDAS

This study deals with the problem of forecasting GDP growth in a data-rich environment and follows the methodology sketched in detail by Andreou et al. (2013). For h-step-ahead forecasts considering the persistence of growth in an economy, MIDAS Eq. 2 can be written as follows:

$$Y_{t+h} = \beta_0^h + \sum_{k=0}^{(q_Y-1)} \rho_k^h Y_{(t-k)} + \beta_1^h \sum_{j=0}^{(q_X-1)} \sum_{i=0}^{(m-1)} \omega(\theta^h, i + j \times m) X_{(m-i,t-j)} + \varepsilon_{(t+h)}^h, \tag{4}$$

where β_0^h is the constant coefficient, β_1^h is the common slope, $\omega(\theta^h, i + j \times m)$ is the weight for the lagged high-frequency values, q_y is the number autoregressive lags, ρ_k^h are the corresponding coefficients to these lags and $\varepsilon_{(t+h)}^h$ is the error term. Equation 4 is autoregressive distributed lag MIDAS (ADL-MIDAS) model.

As it can be seen in Eq. 4, MIDAS forecasts depend on the forecast horizon; thus, it has to be re-estimated for each horizon of interest and it gives direct forecasts for each of these horizons. This property of MIDAS makes it more robust in comparison with

[2] Almon distributed lag polynomial has delivered the best results for the Turkish economy. Results with other polynomial functions are available upon request.

methods that use iterated forecasting since this forecasting method might cumulate possible model misspecification errors and cause inferior forecasts (Marcellino et al. 2006; Andreou et al. 2011).

As a deviation from the basic MIDAS model, factors extracted by principal component analysis (PCA) are included in the forecast regressions. As noted by Stock and Watson (2002) and Andreou et al. (2013), quarterly factors (i) reduce the dimensionality of variables at hand and (ii) might improve the performance of macroeconomic forecasts. PCA is chosen as the method of factor extraction following Andreou et al. (2013). Marcellino and Schumacher (2010) focus on three factor estimators. They find no clear indication of dramatic differences between the nowcast accuracy of the three-factor models over time.[3]

Two types of factors are included in the regressions. The first type of factors is extracted from quarterly macroeconomic series and included in the regressions as the following:

$$\begin{aligned} Y_{t+h} = \beta_0^h &+ \sum_{k=0}^{(q_Y-1)} \rho_k^h Y_{(t-k)} + \sum_{k=0}^{(q_F-1)} \alpha_k^h F_{(t-k)} \\ &+ \beta_1^h \sum_{j=0}^{(q_X-1)} \sum_{i=0}^{(m-1)} \omega(\theta^h, i + j \times m) X_{(m-i,t-j)} + \varepsilon_{(t+h)}^h, \end{aligned} \tag{5}$$

where F_{t-k} represents quarterly factors and α_k^h are the corresponding coefficients. Models that include quarterly factors are called factor autoregressive distributed lag MIDAS (FADL-MIDAS) models throughout the study.

The second type of factors is extracted from daily financial series. Estimations are carried out with factors both from all series and from separate financial data classes. These factors are high- frequency variables and are inserted as $X_{(m-j,t)}$ in Eqs. 4 and 5.

In the empirical part of the study, ADL-MIDAS and FADL-MIDAS are compared with their non-MIDAS counterparts (i.e., ADL and FADL) in order to measure the added value of the MIDAS method. Let X_t^Q be the aggregation of the high-frequency data for quarter t, then the counterpart models can be given as:

$$Y_{t+h} = \beta_0^h + \sum_{k=0}^{(q_Y-1)} \rho_k^h Y_{(t-k)} + \sum_{j=0}^{(q_X-1)} \beta_j X_{(t-j)}^Q + \varepsilon_{(t+h)}^h, \tag{6}$$

$$Y_{t+h} = \beta_0^h + \sum_{k=0}^{(q_Y-1)} \rho_k^h Y_{(t-k)} + \sum_{k=0}^{(q_F-1)} \alpha_k^h F_{(t-k)} + \sum_{j=0}^{(q_X-1)} \beta_j X_{(t-j)}^Q + \varepsilon_{(t+h)}^h. \tag{7}$$

Equations 6 and 7 are labeled as the ADL and factor ADL (FADL) models, respectively.

[3] The first factor estimator builds upon the one-sided nonparametric dynamic PCA. Second, the expectation-maximization (EM) algorithm is combined with the factor estimator-based static PCA as introduced by Stock and Watson (2002). The third one is the two-step parametric state-space factor estimator.

Within-quarter forecasts are made by using the MIDAS with leads model. MIDAS with leads model was first used by Clements and Galvão (2008) and Kuzin et al. (2011) in a monthly–quarterly mixed data context. To illustrate the method within the daily–quarterly context, assume that we want to obtain 1-month-ahead forecast for a quarterly variable. In that case, we are 2 months into the quarter of interest and have financial data of 44 working days. Let J_x denote the number of months that passed in the quarter (i.e., for the current example, $J_x = 2$). Then, the ADL-MIDAS with leads can be written as:

$$
\begin{aligned}
Y_{t+h} = \beta_0^h + \sum_{k=0}^{(q_Y-1)} \rho_k^h Y_{(t-k)} + \beta_1^h \Bigg[& \sum_{i=(3-J_x)\times m/3}^{(m-1)} \omega(\theta^h, i-m) X_{(m-i,t+1)} \\
& + \sum_{j=0}^{(q_X-1)} \sum_{i=0}^{(m-1)} \omega(\theta^h, i+j\times m) X_{(m-i,t-j)} \Bigg] + \varepsilon_{(t+h)}^h,
\end{aligned}
\tag{8}
$$

where the first term in the brackets represent the leads and the second term represents the lags (i.e., information that belongs to the previous quarters).

In the forecasting literature, nowcasting[4] refers to within period of forecasts and updating them when new information becomes available. Andreou et al. (2011) note two main differences between MIDAS with leads and nowcasting. First, while nowcasting gives an update for the current quarter, MIDAS with leads can provide direct forecasts not only for the current quarter but also for longer horizons. Second difference is related to how models that do nowcasting (e.g., state-space models) and MIDAS with leads handle arrival of new information. Issues such as announcement dates, missing data and ragged end are taken into account in the state-space models. Furthermore, effects of macroeconomic announcements can also be analyzed with Kalman filters. However, daily financial data are available on "daily" basis and do not require updates. This study does not deal with the noted problems and assume that all news effect is already reflected by the daily data.

2.3 Forecast combinations

As a part of the forecasting procedure, individual forecasts are combined by using forecast combination methods suggested in the literature. Combining individual forecasts by a pre-specified weighting rule has shown to improve forecast accuracy (Bates and Granger 1969). This conclusion is reached by the assumption that individual forecasts are unbiased. The intuition behind the forecast gain by combination is that one forecast might include information which is not considered by the others and a better forecast can be generated by combining different information sets. In addition to that,

[4] Another term used in the forecasting literature is *backcasting* which refers to making forecasts for a certain period using data that become available after that certain period ends. This is usually the case for quarterly GDP growth forecasts with monthly data where GDP growth is announced with approximately 3 months lag, while monthly data become available with one or 2 months lags.

as pointed out by Hendry and Clements (2004) and Stock and Watson (2004), forecast combinations can deal with model instability and structural breaks.[5]

Assume that R is the number of h-step-ahead forecasts made for Y_t at time t. Then, the combination of forecasts, $\hat{Y}^{c_R}_{(t+h)}$, can be written as

$$\hat{Y}^{c_R}_{(t+h)} = \sum_{r=1}^{R} w^c_{r,t} \hat{Y}_{(r,t+h)}, \tag{9}$$

where the superscript c_R denotes the combination of R forecasts, $\hat{Y}_{(r,t+h)}$ is the forecast from model r and $w^c_{r,t}$ is the combination weight for the forecast from model r. There are methods suggested in the literature to decide on the weights of individual forecasts. It can be as simple as assigning equal weights to each model (i.e., $w^c_{r,t} = 1/R$). This study employs discounted mean squared forecast error (DMSFE)[6] forecast combination method. With this method, forecast weights can be calculated as follows:

$$w^c_{r,t} = \frac{(\lambda_{r,t})^{\kappa}}{\sum_{r=1}^{R} (\lambda^{-1}_{r,t})^{\kappa}}, \quad \lambda_{r,t} = \sum_{\tau=T_0}^{t-h} \delta^{t-h-\tau} (Y_{t-h} - \hat{Y}_{r,\tau+h})^2, \tag{10}$$

where $\delta \in [0, 1]$ is the discount factor and κ is 1 or 2. The method takes into account the historical performance of a model, and as δ gets higher, the comparative importance of recent forecast accuracy gets higher. When $\delta = 1$, the DMSFE forecast combination method boils down to MSFE forecast combination method. In this study, δ is taken as 0.9 and κ as 2.[7]

One of the issues in the forecast combination literature is the selection of forecasts to combine (Timmermann 2006; Aiolfi and Timmermann 2006; Elliott et al. 2015). In the literature, there is no consensus on how to proceed. Therefore, we first combine forecasts of all series and daily factors. Then, forecasts from different financial data classes are combined to see their performances. Finally, based on their root-mean-squared forecast error (RMSFE) performances, best series for each horizons are selected and combination of these forecasts is analyzed.

The forecasting stage of the empirical research includes several steps. At the first step, for every daily financial series or factor and for each model (e.g., ADL-MIDAS and FADL-MIDAS), forecasts are calculated for different lag lengths. Afterward, for each high- frequency variable, the best-performing lag specification is selected based on Akaike information criterion (AIC) and RMSFE performances. At the last step, forecasts are combined by using the DMSFE method.

[5] For an overview of the forecast combination literature, please see Timmermann (2006).

[6] We have received best results from this combination method.

[7] The initial regression shows that this combination of the parameters delivers the best performance in our case.

2.4 Comparing predictive accuracy

Diebold and Mariano (1995) (DM) suggest equal predictability test in the case of pairwise model comparison of non-nested models. Predictive accuracy of forecasts is compared with squared forecast errors in the DM test. Since we have a small number of forecasts, small sample adjustment of the DM test by Harvey et al. (1997) (HLN) is employed. It is specified as:

$$\mathrm{HLN} = \left(\frac{T + 1 - 2h + h(h-1)/T}{T}\right)^{-1/2} \mathrm{DM}, \tag{11}$$

where T is the number of forecasts, h is the forecast horizon and DM is the original DM statistic. HLN follows a Student t distribution with $T-1$ degrees of freedom (Harvey et al. 1997).

As it can be observed in the models that are used in the forecasting GDP growth, some of the models are nested in others. One issue noted about the DM test in the literature is that when comparing two nested models, the critical values are not asymptotically normally distributed and shown to be undersized (McCracken 2007). Clark and West (2007) (CW) develop a test statistic which can be used to compare out-of-sample forecast accuracy of two nested models. The null hypothesis of the test is that nesting model and the nested model have equal predictive accuracy and the alternative hypothesis is that the nesting model performs better than the nested model. Based on their simulation studies, Clark and West (2007) define two critical values for their CW statistics. If the test statistic is bigger that 1.282, the p value of the test statistic is between 0.05 and 0.10, and if the test statistic is bigger than 1.645, the p value is between 0.01 and 0.05.

3 Data

In this study, since we employ the data belonging to the last two decades of the Turkish economy, as a first step, we try to summarize developments in this period. After suffering important setbacks during 1990s, starting with 1994 domestic crisis and continuing with Asian and Russian crisis of late 1990s, in 2001 Turkey faced the most severe economic crisis in its history in which its GDP shrunk by 5.7%. From then up to the onset of the global crisis, Turkey experienced a strong and robust growth by supportive external conditions, the improved policy frameworks and growth-enhancing reforms. In addition, Turkey, as many emerging economies, also used the decade to implement structural reforms, strengthen policy frameworks which resulted in remarkable fiscal consolidation, improved macrofinancial stability and flexible exchange rate regimes (Gros and Selçuki 2013). The global crisis severely hit the Turkish economy (shrunk by 4.8% in 2009). During 2010 and the first half of 2011, the recovery from the global financial crisis peaked. The growth decelerated in 2012 and by end 2013 with the result of macroprudential measures and weaker external demand (especially weak growth in the EU). The FED took steps toward gradually finalizing the quantitative expansion programs during 2014 and announced that it could initiate an interest rate hike in

2015. This period of heightened global uncertainty and weak external demand due to geopolitical developments put downside risks for the Turkish economy and lead to slow pace of growth in the economy during 2014 and 2015.

The data of this period for the Turkish economy are retrieved from *Datastream* using several databases, and the most recently available data have been used for the regressions at the time of writing. The study focuses on the quarter-on-quarter seasonally adjusted Turkish GDP growth between 2000-Q2 and 2016-Q2. The GDP data series is from the TURKSTAT database. The length of the series is 65, and last 25 data points of the series (i.e., GDP growth data from 2010-Q2 till 2015-Q1) are selected to evaluate out-of-sample performance of MIDAS models. The initial window size is; thus, 40 (i.e., initial estimation sample is from 2000-Q2 till 2010-Q1) and forecasts are made with a recursive window.

Quarterly factors are extracted from six quarterly macroeconomic series.[8] Like the GDP data, these quarterly series are seasonally adjusted. The series are available since 1989-Q2. The total import and export series are transformed as first difference of their logarithms and remaining quarterly series are used in their original form, which gives quarter-on-quarter growth rates. The factors are extracted using the PCA with a window size of 40. Only first lag of the first factor, which explains 50% total variation in quarterly series, is used in the FADL-MIDAS and FADL regressions.

Daily financial series which run from January 3, 2000, to June 30, 2016, consist of 3 asset classes and a corporate risk class. There are 41 commodity, 35 equity and 12 foreign exchange series in the asset classes and 10 corporate risk series with a length of 4301. The commodity prices are mostly from the London Metal Exchange (LME) and the Goldman Sachs Commodity Index (GSCI). We use several forward prices, such as 3-month or 15-month forward, in order to exploit the forward-looking information in these prices. The equity series include sectoral indices, such as banking or industrials, which summarize developments in a certain sector. In addition to the sectoral information, we also use the Borsa Istanbul (BIST) 100 series to get information on the daily performances of major companies in Turkey. The United States Dollar (USD) and Euro are the two major currencies in the trade relation of Turkey, so we expect that the information content of these two currencies is the most useful for the forecasting GDP growth. However, other currency pairs are also added to the list to see whether they can provide improvements beyond our expectations. Finally, an additional *foreign corporate risk* class is used in order to account for the regional and global risk on GDP growth. There are 106 series in this class. The motivation of including the class stems from the possible linkages within trade, investment and growth for a small open economy. High international risk might negatively influence trade and investment flows and thereby the domestic GDP growth.

The whole list of quarterly series and daily financial series used here can be found in Şen Doğan and Midiliç (2016).

The daily data series are transformed to obtain return series in percentage terms as follows:

$$x_t = \log\left(\frac{X_{t-1}}{X_t}\right) \times 100; \tag{12}$$

[8] The selection of the series is solely based on the data availability.

where x_t denotes log daily return at time t. The series are then tested for stationarity.

Five daily factors are extracted for all of the 204 series and for each financial data class separately. The window size used in factor extraction is 40. These factors explain 55–18%, 46–16%, 30–10%, 19–2%, 22–5×10^{-6}% and 26–0.03% of total variation in all series, commodities, equities, foreign exchange series, domestic corporate risk and foreign corporate risk series, respectively.

Daily series and all quarterly series except GDP growth are winsorized at the 1% level in order to avoid any outlier effect.

4 Empirical results

We analyze two types of models: models with financial data (ADL and ADL-MIDAS) and models with financial data and macrofactors (FADL and FADL-MIDAS). We evaluate forecasting performance of these models with respect to benchmark autoregressive model with lag one (AR(1)) using RMSFE in order to determine the contribution of financial data to forecasting GDP growth. Additionally, we try to find out to what extent employing quarterly factors and MIDAS models in forecasting GDP growth of Turkey provide forecasting gains.

The selection of the benchmark model in model comparisons relies on the results of Marcellino (2008), who show that the linear models prove to strong benchmarks for out-of-sample forecasting if they are correctly specified. For the Turkish GDP growth, a linear time-series model is specified based on the AIC and the Bayesian information criterion (BIC). According to the information criteria given in Table 1, the minimum values for both the AIC and the BIC are received form the AR(1) model with a constant; therefore, this model is selected as the benchmark model.

The forecast results are reported in three subgroups as 1-month- ahead, 2-month-ahead, 3-month-ahead nowcasts and forecasts, and horizons beyond 2-quarter-ahead, namely, 3-quarter- and 4-quarter- ahead forecasts. Time left in certain quarter in monthly terms is considered in labeling the forecasts. In order to illustrate the distinction between these forecast horizons, assume that we want to make forecasts for 2013-Q1 and 2013-Q2. The forecast for the first quarter of 2013 done at the beginning of the first quarter (i.e., only use data from the previous periods and not from

Table 1 Lag selection

	AIC	BIC
AR(1)	105.04	107.21
AR(2)	106.69	111.04
AR(3)	108.06	114.58
AR(4)	109.00	117.70
AR(1), constant	**99.98**	**104.33**
AR(2), constant	101.87	108.40
AR(3), constant	103.87	112.57
AR(4), constant	101.95	112.83

GDP growth lag selection
Numbers marked in bold indicate that these values are the best ones in their respective horizons

2013-Q1) is labeled as 3-month-ahead nowcast, the one done at the end of the first month is labeled as 2-month-ahead nowcast, and the one done at the end of the second month is labeled as 1-month-ahead nowcast. Forecasts done for 2013-Q2 in 2013-Q1 are labeled as forecasts, and again the time left in 2013-Q1 in monthwise labels the forecasts (e.g., forecast done for 2013-Q2 at the beginning of 2013-Q1 is labeled as 3-month-ahead forecast).

Firstly, we evaluate contribution of using financial data in forecasting real GDP growth. Top panels of Table 2 include forecast combinations with 204 daily series, while combinations with 5 daily factors are given at the bottom panels. The results with macrofactors (FADL and FADL-MIDAS) are given after the results without macrofactors (ADL and ADL-MIDAS). The table presents the RMSFE of the benchmark model and RMSFE of the combined forecasts of the two types of models with respect to the RMSFE of the benchmark. If the ratio is less than one, it is interpreted as an improvement of the underlying forecast upon the benchmark. The results reveal that all models with financial data perform better than the benchmark model for all forecast horizons. To be explicit, in case of 1-month-ahead nowcast, for the daily series, the ADL and ADL-MIDAS improve upon the AR by 2 and 17%, respectively. In Table 3, predictive ability of MIDAS models with respect to the AR(1) is compared by using the HLN and CW tests. The results are similar and show that financial data provide statistically significant forecasting gains over the benchmark.[9]

Secondly, we compare forecasting performance of ADL-MIDAS and FADL-MIDAS models over the corresponding ADL and FADL models so that we aim to assess predictive ability of MIDAS models for quarterly real GDP growth. RMSFE performances of these models are given in Table 2. An initial comparison shows that ADL-MIDAS models perform better than ADL models in all cases. Similarly, performance of FADL-MIDAS models is better than FADL models. The statistical significance of the differences between the models is tested by HLN and CW tests.[10] Corresponding test statistics are given in Table 4. In these tests, ADL models are compared with ADL-MIDAS models and FADL models with FADL-MIDAS models. According to the HLN test results, MIDAS models do not consistently beat the non-MIDAS models. On the other hand, according to the CW test statistics, the ADL-MIDAS and FADL-MIDAS models perform statistically better except for one case (i.e., FADL vs. FADL-MIDAS at 1-month-ahead forecast).

Thirdly, we evaluate the performance of quarterly macroeconomic factors in forecasting GDP growth. For almost all MIDAS and non-MIDAS models, quarterly factors provide forecasting gains over the corresponding ADL and ADL-MIDAS models for both daily series and daily factors combinations and for each forecast horizon (Table 2).

[9] We carry out the HLN and CW tests with all financial data classes and comparison dimensions (e.g., ADL vs. ADL-MIDAS models); however, only for the whole sample test results are reported here for space considerations. Other test results are available upon request.

[10] In addition to HLN and CW tests, robustness of the results is checked with forecast encompassing and mean squared adjusted test statistics (Clark and McCracken 2001, 2010; McCracken 2007) when the bigger model includes additional variables. The details of the tests are given in Appendix of the working paper version of this study (Şen Doğan and Midiliç 2016). We have significant results similar to the ones in the working paper version. The test results for the forecast results reported in this version are available upon request.

Table 2 RMSFE comparisons for MIDAS and non-MIDAS models

Type of series	Forecast horizon							
	Nowcast			Forecast				
	3-month	2-month	1-month	3-month	2-month	1-month	3-quarter	4-quarter
AR(1)	1.08	1.08	1.08	1.10	1.10	1.10	1.12	1.33
Daily series								
ADL	0.93	0.92	0.98	0.96	0.91	0.88	0.94	0.75
ADL-MIDAS	0.80	0.80	0.83	0.91	0.83	0.87	**0.78**	0.69
FADL	**0.78**	0.90	0.95	0.92	0.87	0.88	0.92	0.74
FADL-MIDAS	**0.78**	**0.76**	0.81	**0.87**	**0.80**	**0.84**	**0.78**	0.68
Daily factors								
ADL	0.96	0.98	0.93	0.93	0.91	0.92	0.93	0.65
ADL-MIDAS	0.89	0.85	0.82	0.90	0.82	0.88	0.92	0.59
FADL	0.95	0.96	0.90	0.92	0.89	0.92	0.90	0.63
FADL-MIDAS	0.89	0.84	**0.80**	**0.87**	0.81	0.86	0.88	**0.55**

Forecast combination results of ADL-MIDAS and FADL-MIDAS models for all daily series and factors. The first line gives the RMSFE of the AR(1) model with constant. Other cells include the relative RMSFE of combined forecasts to the AR(1) model

Numbers marked in bold indicate that these values are the best ones in their respective horizons

Table 3 HLN and CW test statistics for AR(1) versus MIDAS models

AR(1) versus	Forecast horizon							
	Nowcast			Forecast				
	3-month	2-month	1-month	3-month	2-month	1-month	3-quarter	4-quarter
Daily series								
HLN								
ADL-MIDAS	1.52	1.42	1.34	1.15	0.94	0.89	1.07	0.97
FADL-MIDAS	1.47	1.36	1.22	1.14	0.90	0.87	1.06	0.97
CW								
ADL-MIDAS	1.73**	1.50*	1.43*	1.71**	1.56*	1.56*	1.38*	1.45*
FADL-MIDAS	1.62*	1.42*	1.34*	1.74**	1.61*	1.58*	1.42*	1.48*
Daily factors								
HLN								
ADL-MIDAS	1.42	1.87*	1.96*	1.48	0.88	0.88	0.75	1.54
FADL-MIDAS	1.15	1.47	1.75*	1.21	0.84	0.87	0.70	1.37
CW								
ADL-MIDAS	2.10**	2.19**	2.34**	2.16**	1.52*	1.58*	1.41*	1.89**
FADL-MIDAS	1.86**	1.85**	1.96**	2.05**	1.53*	1.66*	1.42*	1.82**

Harvey–Leybourne–Newbold (HLN) and Clark–West (CW) test statistics for comparison of AR(1) benchmark model and MIDAS model forecast combination performances with all daily series and factors. For HLN, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, and for CW, *$0.05 < p < 0.10$, ** $0.01 < p < 0.05$

Table 4 HLN and CW test statistics non-MIDAS versus MIDAS models

ADL and FADL versus	Forecast horizon							
	Nowcast			Forecast				
	3-month	2-month	1-month	3-month	2-month	1-month	3-quarter	4-quarter
Daily series								
HLN								
ADL-MIDAS	1.25	1.16	1.35	1.22	1.74*	0.31	1.03	1.27
FADL-MIDAS	1.50	1.33	1.47	1.23	1.84*	0.77	1.14	1.15
CW								
ADL-MIDAS	1.48*	1.30*	1.48*	1.71**	2.11**	0.93	1.49*	1.40*
FADL-MIDAS	1.89**	1.47*	1.66**	1.78**	2.20**	1.25	1.70**	1.40*
Daily factors								
HLN								
ADL-MIDAS	1.20	1.75*	1.93*	0.59	1.12	0.82	0.15	0.66
FADL-MIDAS	1.07	1.85*	1.96*	0.86	1.25	0.95	1.46	0.64
CW								
ADL-MIDAS	1.92**	1.99**	2.42**	1.52**	1.84**	1.59*	1.73**	1.72**
FADL-MIDAS	1.86**	2.19**	3.13**	1.55*	2.06**	1.79**	1.57*	1.58*

Harvey–Leybourne–Newbold (HLN) and Clark–West (CW) test statistics for comparison of non-MIDAS and MIDAS model forecast combination performances with all daily series and factors. For HLN, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, and for CW, *$0.05 < p < 0.10$, ** $0.01 < p < 0.05$

Test statistics for the significance of better performance of the FADL-MIDAS models against ADL-MIDAS models are given in Table 5. Regarding the observation on Table 2, the test statistics imply that the quarterly factors significantly improve the MIDAS forecasts only with the daily series and only for the 2-month-ahead forecasts.

Fourthly, we investigate forecasting performance of each financial data class. In Table 6, RMSFE of each financial data class with respect to the benchmark model is reported, while the results corresponding to the linear models are tabulated in Table 7. FADL-MIDAS models with the daily factors show the best performance for most of the horizons. For example, for 3-month-ahead nowcasts, FADL-MIDAS models with forex factors improve upon the benchmark by 19%. For 3-month-ahead forecasts, the combinations of the factors extracted from separate financial data classes provide 17% improvement over the benchmark. The improvement goes up to 38% for the 4-quarter-ahead forecasts with the forecast combination of the equity factors.

Lastly, we try to identify whether some small group of series become prominent among 204 daily series in terms of forecasting gains in order to search for evidence for the argument that the results can be replicated by a small number of daily series. Forecast combination performances of the best 5 and 10 series for each horizon are given in Table 8. The table shows that forecasts do not improve in comparison with the results reported before. HLN and CW test statistics comparing the forecasts of best 5 series are given in Table 9, and the test statistics show that the FADL-MIDAS model forecasts statistically perform better than the combination of the best-performing series.

Tables 10 and 11 report the first 5 best series for each horizon based on their RMSFE performances. The table shows that, within the daily financial series, some of the equity series, especially FTSE Oil & Gas index, are very useful in forecasting the Turkish GDP growth with both ADL-MIDAS and FADL-MIDAS models. At the longer horizons, the foreign corporate bond indices dominate the table. This might be a period specific phenomenon, but the predictive ability of both equity and foreign corporate bond series should be further analyzed.

It can be observed on the tables that as the forecast horizon increases, the performance of the MIDAS models gets better compared to the benchmark. This is due to the deterioration of the performance of the benchmark model. The RMSFE of the AR(1) goes up to 1.33 at the 4-quarter-ahead forecasts from 1.08 at the 1-quarter- ahead forecasts. Therefore, the MIDAS models produce arguable more stable forecasts within horizons than the benchmark model.

Overall, the RMSFE comparisons show that better results are received with the combination of a small set of daily factors extracted from separate financial data classes for the 1-month-ahead nowcasts, and 2-month-, 3-month- and 4-quarter-ahead forecasts. The result implies that the dimensions of the financial data can be reduced and the Turkish economy can be tracked by using a small set of daily factors for some horizons.

The results suggest that the framework used in the study works significantly better than naive models. Our findings indicate that employing daily financial data improves forecasting performance of two types of model. Besides, with financial data one does not have to wait for updates of the dataset. MIDAS regression models provide forecasts gains compared to the models using simple aggregation schemes, and in most of the cases, the gain is substantial. Consistent with the findings of previous studies (Stock

Table 5 HLN and CW test statistics for ADL-MIDAS versus FADL-MIDAS models

ADL-MIDAS versus FADL-MIDAS	Forecast horizon							
	Nowcast			Forecast				
	3-month	2-month	1-month	3-month	2-month	1-month	3-quarter	4-quarter
Daily series								
HLN	0.33	0.92	0.35	0.67	0.58	0.67	0.28	0.54
CW	0.65	1.37*	0.82	1.15	1.17	1.15	0.84	0.88
Daily factors								
HLN	−0.00	0.34	0.68	0.59	0.31	0.36	0.56	0.68
CW	0.46	0.88	1.28	1.07	0.79	1.03	0.96	1.21

Harvey–Leybourne–Newbold (HLN) and Clark–West (CW) test statistics for comparison of ADL-MIDAS and FADL-MIDAS model forecast combination performances with all daily series and factors. For HLN, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, and for CW, *$0.05 < p < 0.10$, ** $0.01 < p < 0.05$

Table 6 RMSFE comparisons for MIDAS models, separate financial data classes

Type of series	Forecast horizon							
	Nowcast			Forecast				
	3-month	2-month	1-month	3-month	2-month	1-month	3-quarter	4-quarter
AR(1)	1.08	1.08	1.08	1.10	1.10	1.10	1.12	1.33
Daily series								
Equities								
ADL-MIDAS	**0.81**	0.84	0.88	0.91	0.83	0.91	0.90	0.76
FADL-MIDAS	**0.81**	0.81	0.86	0.87	0.82	0.87	0.89	0.74
Commodities								
ADL-MIDAS	0.86	0.93	1.03	0.91	0.86	0.85	0.93	0.77
FADL-MIDAS	0.86	0.92	1.01	0.89	0.86	0.85	0.91	0.76
Forex								
ADL-MIDAS	0.97	1.03	1.05	1.04	0.92	0.93	0.90	0.75
FADL-MIDAS	0.93	0.99	1.05	1.02	0.90	0.91	0.89	0.73
Corporate risk (Turkey)								
ADL-MIDAS	1.03	0.89	0.97	0.91	0.94	1.00	1.00	0.81
FADL-MIDAS	1.00	**0.77**	0.96	0.89	0.88	0.96	0.99	0.80
Corporate risk (foreign)								
ADL-MIDAS	0.93	0.89	0.90	0.96	0.88	0.96	0.75	0.70
FADL-MIDAS	0.92	0.87	0.88	0.92	0.85	0.93	**0.74**	0.68

Table 6 continued

Type of series	Forecast horizon							
	Nowcast			Forecast				
	3-month	2-month	1-month	3-month	2-month	1-month	3-quarter	4-quarter
Daily factors								
Separate Asset Classes								
ADL-MIDAS	0.84	0.84	0.84	0.86	0.81	0.85	0.88	0.69
FADL-MIDAS	0.82	0.81	0.82	**0.83**	0.78	0.81	0.81	0.67
Equities								
ADL-MIDAS	0.93	0.87	0.86	0.87	0.81	0.87	0.85	0.66
FADL-MIDAS	0.92	0.86	0.85	0.86	**0.76**	0.89	0.81	**0.62**
Commodities								
ADL-MIDAS	0.86	0.86	0.81	0.89	0.93	0.86	0.86	0.68
FADL-MIDAS	0.83	0.80	**0.76**	0.86	0.90	0.84	0.83	0.68
Forex								
ADL-MIDAS	**0.81**	0.85	0.90	0.88	0.78	0.91	0.93	0.75
FADL-MIDAS	**0.81**	0.83	0.88	0.89	0.81	0.89	0.78	0.72
Corporate risk								
ADL-MIDAS	0.92	0.93	0.88	0.95	0.91	0.89	0.92	0.72
FADL-MIDAS	0.87	0.92	0.88	0.91	0.85	**0.79**	0.88	0.74
Corporate risk (foreign)								
ADL-MIDAS	0.84	0.86	0.90	0.90	0.82	0.86	0.97	0.77
FADL-MIDAS	0.86	0.82	0.86	0.87	0.79	0.82	0.86	0.73

Forecast combination results of ADL-MIDAS and FADL-MIDAS models for separate financial data classes and factors derived from these classes. The first line gives the RMSFE of the AR(1) model with constant. Other cells include the relative RMSFE of combined forecasts to the AR(1) model

Numbers marked in bold indicate that these values are the best ones in their respective horizons

Table 7 RMSFE comparisons for non-MIDAS models, separate financial data classes

Type of series	Forecast horizon							
	Nowcast			Forecast				
	3-month	2-month	1-month	3-month	2-month	1-month	3-quarter	4-quarter
AR(1)	1.08	1.08	1.08	1.10	1.10	1.10	1.12	1.33
Daily series								
Equities								
ADL	0.88	0.94	0.98	0.95	0.87	0.89	0.95	0.79
FADL	**0.86**	0.91	0.94	**0.90**	**0.86**	0.86	0.93	0.77
Commodities								
ADL	0.94	0.95	0.98	0.98	0.89	**0.84**	0.95	0.80
FADL	0.92	0.92	0.95	0.96	0.88	0.86	0.93	0.77
Forex								
ADL	0.96	0.96	1.01	0.98	0.95	0.96	0.94	0.78
FADL	0.93	0.93	0.99	0.94	0.90	0.93	0.92	0.76
Corporate risk (Turkey)								
ADL	0.96	1.00	1.00	0.95	0.93	0.96	0.96	0.78
FADL	0.92	0.97	0.96	**0.90**	0.88	0.93	0.94	0.75
Corporate risk (foreign)								
ADL	0.96	0.94	1.00	0.97	0.94	0.93	0.93	0.75
FADL	0.93	0.92	0.97	0.93	0.91	0.90	0.91	0.73

Table 7 continued

Type of series	Forecast horizon							
	Nowcast			Forecast				
	3-month	2-month	1-month	3-month	2-month	1-month	3-quarter	4-quarter
Daily factors								
Separate asset classes								
ADL	0.95	0.95	0.95	0.96	0.93	0.92	0.92	0.77
FADL	0.94	0.92	0.90	0.94	0.90	0.90	0.89	0.75
Equities								
ADL	0.99	0.94	0.93	0.95	0.96	0.97	0.91	0.74
FADL	0.97	**0.90**	0.91	0.94	0.95	0.95	**0.87**	**0.71**
Commodities								
ADL	0.98	0.98	0.96	0.94	0.94	0.95	0.89	0.77
FADL	0.93	0.92	**0.87**	0.93	0.91	0.93	**0.87**	0.76
Forex								
ADL	0.95	0.93	0.98	0.96	0.97	0.95	0.94	0.81
FADL	0.93	0.92	0.96	0.94	0.93	0.93	0.91	0.79
Corporate risk								
ADL	0.97	0.97	0.97	0.98	0.94	0.96	0.96	0.80
FADL	0.97	0.96	0.93	0.96	0.90	0.92	0.93	0.78
Corporate risk (foreign)								
ADL	0.91	0.97	0.94	0.95	0.91	0.87	0.94	0.77
FADL	0.91	0.93	0.90	0.93	0.88	0.86	0.91	0.74

Forecast combination results of ADL and FADL models for separate financial data classes and factors derived from these classes. The first line gives the RMSFE of the AR(1) model with no constant. Other cells include the relative RMSFE of combined forecasts to the AR(1) model
Numbers marked in bold indicate that these values are the best ones in their respective horizons

Table 8 RMSFE comparisons for MIDAS models with best series

	Forecast horizon							
	Nowcast			Forecast				
	3-month	2-month	1-month	3-month	2-month	1-month	3-quarter	4-quarter
AR(1)	1.08	1.08	1.08	1.10	1.10	1.10	1.12	1.33
Best series combinations								
Daily series								
Best 5 Series (ADL-MIDAS)	0.84	0.85	0.77	0.91	0.83	1.02	0.90	0.76
Best 5 Series (FADL-MIDAS)	0.84	0.84	**0.73**	0.97	0.82	0.85	0.91	0.76
Best 10 Series (ADL-MIDAS)	0.85	0.87	0.79	0.94	0.88	0.92	0.89	0.72
Best 10 Series (FADL-MIDAS)	0.84	0.83	**0.73**	0.91	0.84	**0.81**	0.90	0.71
Daily factors								
Best 5 Factors (ADL-MIDAS)	0.99	0.89	0.74	**0.69**	0.73	0.85	**0.82**	0.65
Best 5 Factors (FADL-MIDAS)	0.88	**0.79**	0.74	0.88	0.89	0.92	0.86	0.70
Best 10 Factors (ADL-MIDAS)	0.92	0.87	0.78	0.71	**0.71**	0.84	0.84	0.66
Best 10 Factors (FADL-MIDAS)	**0.82**	0.82	0.74	0.80	0.87	0.82	0.84	**0.64**

Forecasts combination results of FADL-MIDAS models the for the best 5 and 10 daily series. The first line gives the RMSFE of the AR(1) model with constant. Other cells include the relative RMSFE of combined forecasts to the AR(1) model

Numbers marked in bold indicate that these values are the best ones in their respective horizons

Table 9 HLN and CW test statistics for best series versus FADL-MIDAS models

FADL-MIDAS versus best series	Forecast horizon							
	Nowcast			Forecast				
	3-month	2-month	1-month	3-month	2-month	1-month	3-quarter	4-quarter
Daily series								
HLN	−0.39	−1.12	0.99	−0.50	−0.94	−2.38**	−1.53	−0.75
CW	0.62	0.03	2.43**	1.43*	1.33*	−1.40	−0.28	0.93
Daily factors								
HLN	0.72	−0.12	0.35	−0.62	−0.30	−5.21***	−0.28	−1.56
CW	2.68**	2.02**	2.24**	1.28	1.26	−1.02	1.46	0.64

Harvey–Leybourne–Newbold (HLN) and Clark–West (CW) test statistics for comparison of forecast combination of best 5 series and FADL-MIDAS model forecast combination performances. The statistics are calculated using the small sample adjustment for the test. For HLN, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, and for CW, *$0.05 < p < 0.10$, ** $0.01 < p < 0.05$

Table 10 The best series ranked based on their RMSFE performances for each MIDAS model and horizon

Best-performing series 1 3-month	2-month	1-month
Nowcast		
ADL-MIDAS		
FTSE TURKEY OIL & GAS PROD $ - PRICE INDEX	FTSE TURKEY OIL & GAS PROD $ - PRICE INDEX	CGBI WGBI WORLD AA 10–20Y($) - TOT RETURN IND
BIST NATIONAL 100 - PRICE INDEX	FTSE W TURKEY NON-LIFE INSUR $ - PRICE INDEX	CGBI WGBI WORLD NON-EURO 5+y ($) - TOT RETURN IND
FTSE W TURKEY FORESTRY & PAP $ - PRICE INDEX	CBOE SPX VOLATILITY VIX (NEW) - PRICE INDEX	FTSE TURKEY OIL & GAS PROD $ - PRICE INDEX
FTSE W TURKEY GENERAL INDS $ - PRICE INDEX	FTSE W TURKEY ELTRO/ELEC EQ $ - PRICE INDEX	ML MOVE 1M BOND VOLATILITY INDEX - PRICE INDEX
CBOE SPX VOLATILITY VIX (NEW) - PRICE INDEX	BIST NATIONAL 100 - PRICE INDEX	CGBI WGBI WORLD AA 10+Y ($) - TOT RETURN IND
Forecast		
FTSE TURKEY OIL & GAS PROD $ - PRICE INDEX	FTSE W TURKEY HEALTH CARE $ - PRICE INDEX	LME-Aluminum Alloy Cash U$/MT
BOFA ML U$ GLB HY IDX ($) - TOT RETURN IND	FTSE W TURKEY GENERAL INDS $ - PRICE INDEX	CGBI EMUSDGBI-C 5–10 YEARS INDEX - TOT RETURN IND
BARCLAYS EM MID.EAST TURKEY	FTSE TURKEY OIL & GAS PROD $ - PRICE INDEX	CGBI EMUSDGBI 1–5 YEARS INDEX - TOT RETURN IND
BARCLAYS EM TURKEY FIXED RATE	CGBI EMUSDGBI 1–5 YEARS INDEX - TOT RETURN IND	BOFA ML U$ GLB HY IDX ($) - TOT RETURN IND
BARCLAYS EM TURKEY INTL ISSUE	FTSE W TURKEY FORESTRY & PAP $ - PRICE INDEX	CGBI EMUSDGBI 5–10 YEARS INDEX - TOT RETURN IND

Table 10 continued

Best-performing series 1 3-month	2-month	1-month
Nowcast		
FADL-MIDAS		
FTSE TURKEY OIL & GAS PROD $ - PRICE INDEX	FTSE TURKEY OIL & GAS PROD $ - PRICE INDEX	CGBI EMUSDGBI INV.GRADE INDEX - TOT RETURN IND
FTSE W TURKEY FORESTRY & PAP $ - PRICE INDEX	JPM EMBI GLOBAL TURKEY - TOT RETURN IND	FTSE TURKEY OIL & GAS PROD $ - PRICE INDEX
BIST NATIONAL 100 - PRICE INDEX	JPM EMBI GLB.DIVERS TURKEY - TOT RETURN IND	CGBI WGBI WORLD NON-EURO 5+y ($) - TOT RETURN IND
Cocoa-ICCO Daily Price US$/MT	FTSE W TURKEY NON-LIFE INSUR $ - PRICE INDEX	FTSE W TURKEY NON-LIFE INSUR $ - PRICE INDEX
FTSE TURKEY FINANCIALS $ - PRICE INDEX	BARCLAYS EM TURKEY INTL ISSUE	CGBI WGBI WORLD AA 10–20Y($) - TOT RETURN IND
Forecast		
FTSE TURKEY CON & MAT $ - PRICE INDEX	FTSE W TURKEY HEALTH CARE $ - PRICE INDEX	FTSE W TURKEY HEALTH CARE $ - PRICE INDEX
FTSE TURKEY TELECOM $ - PRICE INDEX	JPM ELMI+ TURKEY ($) - TOT RETURN IND	FTSE TURKEY BEVERAGES $ - PRICE INDEX
FTSE W TURKEY CHEMICALS $ - PRICE INDEX	FTSE TURKEY OIL & GAS PROD $ - PRICE INDEX	LME-Aluminum Alloy Cash U$/MT
FTSE TURKEY FINANCIALS $ - PRICE INDEX	FTSE TURKEY BEVERAGES $ - PRICE INDEX	S&P GSCI INDL Metals 1MTH Fwd Cap PI - PRICE INDEX
BOFA ML U$ GLB HY IDX ($) - TOT RETURN IND	CGBI EMUSDGBI 1–5 YEARS INDEX - TOT RETURN IND	S&P GSCI INDL Metals 3 MTH Fwd Cap PI - PRICE INDEX

Table 11 The best series ranked based on their RMSFE performances for each MIDAS model and horizon

Best-performing series 2	
Forecast	
3-quarter	4-quarter
ADL-MIDAS	
CGBI WGBI WORLD AA 10–15 Year($) - TOT RETURN IND	CGBI WGBI WORLD NON-EURO 3–7y ($) - TOT RETURN IND
CGBI WGBI WORLD AA 1–10 Year($) - TOT RETURN IND	CGBI WGBI WORLD NON-EURO 3–5y ($) - TOT RETURN IND
CGBI WGBI WORLD AA ($) - TOT RETURN IND	CGBI WGBI WORLD NON-EURO 5–7y ($) - TOT RETURN IND
CGBI WGBI WORLD AA 1–3Y ($) - TOT RETURN IND	CGBI WGBI WORLD NON-EURO 1–3y ($) - TOT RETURN IND
CGBI WGBI WORLD AA 10–20Y($) - TOT RETURN IND	CGBI WGBI WORLD NON-EURO 7–10y ($) - TOT RETURN IND
FADL-MIDAS	
CGBI WGBI WORLD AA 10–15Y($) - TOT RETURN IND	CGBI WGBI WORLD NON-EURO 3–7y ($) - TOT RETURN IND
CGBI WGBI WORLD AA 10–20Y($) - TOT RETURN IND	CGBI WGBI WORLD NON-EURO 3–5y ($) - TOT RETURN IND
CGBI WGBI WORLD AA 1–10Y($) - TOT RETURN IND	CGBI WGBI WORLD NON-EURO 5–7y ($) - TOT RETURN IND
CGBI WGBI WORLD AA ($) - TOT RETURN IND	CGBI WGBI WORLD NON-EURO 7–10y ($) - TOT RETURN IND
CGBI WGBI WORLD AA 1–3Y ($) - TOT RETURN IND	CGBI WGBI WORLD NON-EURO 1–3y ($) - TOT RETURN IND

and Watson 2003; Andreou et al. 2013), quarterly factors provide forecasting gains over the benchmark models for most of the forecast horizons and financial data classes. However, the models considered here do not make use of any monthly data that were shown to forecast Turkish GDP data quite well in the past. By adding monthly variables in Eq. 5 and updating the forecasts after monthly data releases, the performances of the models can further be improved.

Finally, we may compare our results with one of the limited studies about forecasting Turkish GDP growth, which is Akkoyun and Günay (2012). However, the direct comparison is not possible since we employ different estimation and forecasting period. Thus, we only mention about advantage of our framework over their models. Akkoyun and Günay (2012) employ small-scale dynamic factor model and forecast quarterly GDP growth with monthly data by following Kalman filter approach. They find that some combinations of hard indicators such as industrial production, export and import quantity indices and soft indicators such as PMI and PMI new orders have predictive power for the output growth. In last 2 years, correlation among GDP growth and the monthly data decreased (see Box 4.1 in the third inflation report of 2015 for details (CBRT 2015)) which leads to decline in the performance of these models. Instead, we extract information from large set of daily financial data and make use of historical performance of the data in generating forecasts. Therefore, we can say that the methodology applied in this study is more robust to movements in individual series and deterioration of explanatory power in time.

5 Conclusions

In this study, our aim is to incorporate the information in daily financial data about the future state of the Turkish economy. To this end, the mixed data sampling (MIDAS) models are used for the first time in generating nowcasts and forecasts for quarterly Turkish GDP growth in the analysis. We follow the approach introduced by Andreou et al. (2013) and employ 204 daily financial data consisting of 4 major financial data classes: commodity price, equity, foreign exchange and corporate risk.

Our results suggest that MIDAS regression models and forecast combination methods provide considerable advantage in exploiting information from daily financial data compared to the models using simple aggregation schemes when forecasting the quarterly GDP growth. Additionally, incorporating daily financial data into the analysis improves our forecasts substantially. These indicate that both the information content of the financial data and flexible data-driven weighting scheme of MIDAS regressions play a significant role for forecasting real GDP growth. Besides, consistent with the findings of previous studies (Stock and Watson 2003; Andreou et al. 2013), quarterly factors lead to forecasting gains over the benchmark models for most of the forecast horizons and financial data classes. Additionally, the analysis of best-performing series reveals the usefulness of daily factors from separate financial data classes in forecasting the Turkish GDP growth. This implies that daily factor series might be a good indicator for the Turkish GDP growth and be used as a daily index to track economic activity in. Further research might be carried out by incorporating monthly hard and soft data into the analysis since previous studies prove their significance in improving the nowcasting performance.

Data and computer code availability

The datasets of this paper (1. code and programs, 2. data, 3. detailed readme files) are collected in electronic supplementary material of this article.

Acknowledgements Authors thank the editor and two anonymous referees of the journal, Koen Inghelbrecht and Gert Peersman, for valuable comments and suggestions.

References

Aiolfi M, Timmermann A (2006) Persistence in forecasting performance and conditional combination strategies. J Econ 135(1):31–53
Akkoyun HC, Günay M (2012) Nowcasting Turkish GDP growth. CBRT Working Paper (No: 12/33)
Andreou E, Ghysels E, Kourtellos A (2011) Forecasting with mixed-frequency data. In: Oxford handbook of economic forecasting, pp 225–245
Andreou E, Ghysels E, Kourtellos A (2013) Should macroeconomic forecasters use daily financial data and how? J Bus Econ Stat 31(2):240–251
Armesto MT, Hernandez-Murillo R, Owyang MT, Piger J (2009) Measuring the information content of the beige book: a mixed data sampling approach. J Money Credit Bank 41(1):35–55
Bai J, Ghysels E, Wright JH (2013) State space models and MIDAS regressions. Econ Rev 32(7):779–813
Bates JM, Granger CW (1969) The combination of forecasts. Or 20(4):451–468
CBRT (2015) Inflation report-3
Clark TE, McCracken MW (2001) Tests of equal forecast accuracy and encompassing for nested models. J Econ 105(1):85–110
Clark TE, McCracken MW (2010) Testing for unconditional predictive ability. Federal Reserve Bank of St. Louis Working Paper (No: 2010-031A)
Clark TE, West KD (2007) Approximately normal tests for equal predictive accuracy in nested models. J Econ 138(1):291–311
Clements MP, Galvão AB (2008) Macroeconomic forecasting with mixed-frequency data: forecasting output growth in the United States. J Bus Econ Stat 26(4):546–554
Clements MP, Galvão AB (2009) Forecasting US output growth using leading indicators: an appraisal using MIDAS models. J Appl Econ 24(7):1187–1206
Diebold FX, Mariano RS (1995) Comparing predictive accuracy. J Bus Econ Stat 13(3):253–63
Elliott G, Gargano A, Timmermann A (2015) Complete subset regressions with large-dimensional sets of predictors. J Econ Dyn Control 54:86–110
Foroni C, Marcellino M (2013) A survey of econometric methods for mixed-frequency data. Norges Bank Working Paper Series (No: 2013-06)
Galvão AB (2013) Changes in predictive ability with mixed frequency data. Int J Forecast 29(3):395–410
Ghysels E, Wright JH (2009) Forecasting professional forecasters. J Bus Econ Stat 27(4):504–516
Ghysels E, Santa-Clara P, Valkanov R (2004) The MIDAS touch: mixed data sampling regression models. CIRANO Working Papers (No: 2014s-20)
Gros D, Selçuki C (2013) The changing structure of turkey's trade and industrial competitiveness: implications for the EU. Centre for European Policy Studies (CEPS) Working Paper Series (No: 03)
Hamilton JD (2008) Daily monetary policy shocks and new home sales. J Monet Econ 55(7):1171–1190
Harvey D, Leybourne S, Newbold P (1997) Testing the equality of prediction mean squared errors. Int J Forecast 13(2):281–291
Hendry DF, Clements MP (2004) Pooling of forecasts. Econ J 7(1):1–31
Kuzin V, Marcellino M, Schumacher C (2011) MIDAS vs. mixed-frequency VAR: nowcasting GDP in the euro area. Int J Forecast 27(2):529–542

Marcellino M (1999) Some consequences of temporal aggregation in empirical analysis. J Bus Econ Stat 17(1):129–136

Marcellino M (2008) A linear benchmark for forecasting GDP growth and inflation? J Forecast 27(4):305–340

Marcellino M, Schumacher C (2010) Factor MIDAS for nowcasting and forecasting with ragged-edge data: a model comparison for German GDP. Oxford Bull Econ Stat 72(4):518–550

Marcellino M, Stock JH, Watson MW (2006) A comparison of direct and iterated multistep AR methods for forecasting macroeconomic time series. J Econ 135(1):499–526

McCracken MW (2007) Asymptotics for out of sample tests of Granger causality. J Econ 140(2):719–752

Monteforte L, Moretti G (2013) Real-time forecasts of inflation: the role of financial variables. J Forecast 32(1):51–61

Sacakli-Sacildi I (2015) Do BVAR models forecast Turkish GDP better than UVAR models? Br J Econ Manag Trade 7(4):259–268

Schumacher C, Breitung J (2008) Real-time forecasting of German GDP based on a large factor model with monthly and quarterly data. Int J Forecast 24(3):386–398

Şen Doğan B, Midiliç M (2016) Forecasting Turkish real GDP growth in a data rich environment. CBRT Working Paper (No: 16/11)

Stock JH, Watson MW (2002) Forecasting using principal components from a large number of predictors. J Am Stat Assoc 97(460):1167–1179

Stock JH, Watson MW (2003) Forecasting output and inflation: the role of asset prices. J Econ Lit 41:788–829

Stock JH, Watson MW (2004) Combination forecasts of output growth in a seven-country data set. J Forecast 23(6):405–430

Tay AS (2007) Financial variables as predictors of real output growth. Singapore Management University, School of Economics Working Papers (No:7-2007)

Timmermann A (2006) Forecast combinations. Handb Econ Forecast 1:135–196

Wohlrabe K (2009) Forecasting with mixed-frequency time series models. PhD thesis, Ludwig-Maximilians-Universität München