International Journal of Forecasting 32 (2016) 1369–1384

Contents lists available at ScienceDirect

International Journal of Forecasting

journal homepage: www.elsevier.com/locate/ijforecast

Nowcasting Turkish GDP and news decomposition

Michele Modugno [a], Barış Soybilgen [b,*], Ege Yazgan [b]

[a] Board of Governors of the Federal Reserve System, Washington DC, United States
[b] Istanbul Bilgi University, Turkey

ARTICLE INFO

Keywords:
Dynamic factor model
Nowcasting
Gross domestic product

ABSTRACT

Real gross domestic product (GDP) data in Turkey are released with a very long delay relative those of to other economies, between 10 and 13 weeks after the end of the reference quarter. This means that policy makers, the media, and market practitioners have to infer the current state of the economy by examining data that are more timely and are released at higher frequencies than the GDP. This paper proposes an econometric model that allows us to read through these more current and higher-frequency data automatically, and translate them into nowcasts for the Turkish real GDP. Our model outperforms the nowcasts produced by the Central Bank of Turkey, the International Monetary Fund, and the Organisation for Economic Co-operation and Development. Moreover, our model allows us to quantify the importance of each variable in our dataset for nowcasting Turkish real GDP. In line with the findings for other economies, we find that real variables play the most important role; however, contrary to the findings for other economies, we find that financial variables are as important as surveys.

Published by Elsevier B.V. on behalf of International Institute of Forecasters.

1. Introduction

Policy makers and market participants need to infer the current state of the economy in order to inform their policy decisions and investment strategies. One of the main indicators that they look at is the real gross domestic product (GDP), which shows the overall health of the economy. However, it is usually released with a delay relative to the reference quarter. In particular, for the Turkish economy, GDP is released between 10 and 13 weeks after the end of the reference quarter. This very long delay in the release of the Turkish GDP relative to developed economies is because no early or advance estimates are produced. For example, early estimates of the US and euro-area GDPs are released four and six weeks after the end of the reference quarter, respectively. Therefore, the usual practice in Turkey is to infer the current state of the economy by analyzing data that are released in a more timely manner than GDP, at a higher frequency. For example, two real variables, the industrial production index and the unemployment rate, are released six and 10 weeks after the end of the reference month, respectively, making them more timely than GDP. Moreover, surveys such as the capacity utilization rate, the consumer confidence index, and the real sector confidence index are released a few days before the end of the reference month, while financial data such as the real effective exchange rate are released a few days after the reference month. Inferring the state of the economy by interpreting numerous variables that are characterized by differences in definitions, frequencies, and lags is a difficult task. However, we can overcome this challenge using an econometric framework that translates all sorts

* Corresponding author.
E-mail addresses: michele.modugno@frb.gov (M. Modugno), baris.soybilgen@bilgi.edu.tr (B. Soybilgen), ege.yazgan@bilgi.edu.tr (E. Yazgan).

http://dx.doi.org/10.1016/j.ijforecast.2016.07.001
0169-2070/Published by Elsevier B.V. on behalf of International Institute of Forecasters.

of data into a nowcast of GDP, summarizing scattered information into a unique index of the overall health of the Turkish economy.

Following the seminal paper of Giannone, Reichlin, and Small (2008), this study nowcasts GDP using a dynamic factor model (DFM). DFMs are natural tools for nowcasting variables such as the GDP because they allow us to exploit the more timely variables in order to predict those released with longer delays, by capturing the comovement among a potentially large set of variables. Indeed, these models have been applied successfully to the nowcasting of GDP for various different countries: Belgium by de Antonio Liedo (2014), Brazil by Bragoli, Metelli, and Modugno (2015), China by Giannone, Agrippino, and Modugno (2013) and Yiu and Chow (2010), the Czech Republic by Arnostova, Havrlant, Rùžièka, and Tóth (2011), France by Barhoumi, Darné, and Ferrara (2010), Indonesia by Luciani, Pundit, Ramayandi, and Veronese (2015), Ireland by D'Agostino, McQuinn, and O'Brien (2013), Mexico by Caruso (2015), the Netherlands by de Winter (2011), New Zealand by Matheson (2010), Norway by Aastveit and Trovik (2012), and the BRIC (Brazil, Russia, India and China) countries and Mexico by Dahlhaus, Guénette, and Vasishtha (2015).[1] Moreover, the same framework has also been applied to the nowcasting of variables other than real GDP; see D'Agostino, Modugno, and Osbat (2015) for euro area trade variables and Modugno (2013) for US inflation, among others.

One crucial point when using DFMs for nowcasting is the choice of an estimation methodology that suits the needs of the task at hand: dealing with a dataset that is characterized by different frequencies, time spans, and delays. We follow the procedure proposed by Bańbura and Modugno (2014); i.e., a modified version of the expectation–maximization (EM) algorithm for maximum likelihood estimation.[2] This procedure has two important advantages over competing procedures for estimating a DFM. First, the methodology of Bańbura and Modugno (2014) can easily address mixed-frequency datasets with an arbitrary pattern of data availability, exploiting their information content fully for both the parameter estimations and the signal extraction. Second, maximum likelihood estimation is more efficient for small samples. Turkey is a young, newly industrialized economy, so institutions have only recently started collecting economic data. As a result, our dataset is short and contains data that cover different time periods.

We nowcast both seasonally adjusted (SA) quarter-on-quarter (QoQ) GDP growth rates and non-SA (NSA) year-on-year (YoY) GDP growth rates between 2008:Q1 and 2013:Q4 with a medium-scale mixed frequency dataset that includes 15 variables. The dataset is constructed using publicly available time series that are followed by the media, economists, and financial sector practitioners. We perform two out-of-sample exercises. In the first exercise, we nowcast the SA QoQ GDP and NSA YoY GDP. We abstract from data revisions, but require our dataset to replicate the data availability as it was at the time when the forecast would have been generated. This is a "pseudo real-time" exercise, and we show that the GDP nowcasts obtained with our model in this context outperform those obtained with univariate and "partial" models.[3] In the second exercise, we nowcast NSA YoY GDP growth rates and compare our predictions with those produced by (1) the International Monetary Fund (IMF), available in the World Economic Outlook Database (WEOD); (2) the Organisation for Economic Co-operation and Development (OECD), available in the Economic Outlook (EO), and (3) the Central Bank of Turkey (CBRT), collected in the survey of expectations (SoE). This comparison shows that our DFM outperforms the professional forecasters. In order to have a fair comparison with the nowcasts produced by these institutions, we construct a "partial" real-time dataset to account for the effects of data revisions on the nowcast accuracy. The dataset is a "partial" real-time one because we have vintages for the sample that we analyze for only eight of the variables included in our dataset. The other series are included following the "pseudo" real-time approach.[4]

Only two studies in the literature have nowcasted Turkish GDP growth rates. Akkoyun and Günay (2012) used a dynamic one-factor model with a small-scale dataset, including three real variables and one survey, to nowcast SA QoQ GDP growth rates between 2008:Q1 and 2012:Q2, and compare their model with an autoregressive (AR) model. They find that their model outperforms the benchmark model. We use a more comprehensive dataset than Akkoyun and Günay (2012), which allows us to understand which variables are informative for nowcasting Turkish GDP. We find that financial variables such as the REER and financial accounts, which are monitored closely by both policy makers and the market practitioners, are as important as surveys such as the CCI, the RSCI, and the CUR for nowcasting SA QoQ GDP growth rates. None of these financial and survey variables are considered by Akkoyun and Günay (2012).[5] Ermişoğlu, Akçelik, and Oduncu (2013) nowcast QoQ GDP growth rates for the period between 2011:Q1 and 2012:Q4 by forming bridge-type single equations that include the purchasing manufacturing index (PMI) and credit data. To form bridge equations, monthly variables are converted into quarterly variables. Then, a regression or series of regressions that includes the GDP as a dependent variable and other

[1] See Bańbura, Giannone, Modugno, and Reichlin (2013) for a survey of the literature on nowcasting.

[2] The EM approach for maximum likelihood estimation in the case of small-scale DFMs was first proposed by Shumway and Stoffer (1982) and Watson and Engle (1983). Later, Doz, Giannone, and Reichlin (2012) proved that maximum likelihood estimation is also feasible for large-scale DFMs, and Bańbura and Modugno (2014) modified the EM algorithm to account for arbitrary patterns of missing data and the serial correlation of the idiosyncratic component. Jungbacker and Koopman (2015) and Jungbacker, Koopman, and Van der Wel (2011) further showed how the computational efficiency of the methodology could be improved.

[3] See Bańbura et al. (2013) for a definition of "partial" models.

[4] We do not do the same with SA real GDP because it has been being published for only a short time, and real time vintages are not available.

[5] Our model does not include the purchasing managers index (PMI) used by Akkoyun and Günay (2012) because the PMI is provided by a private company for a fee, and, as a result, is not available to the larger public.

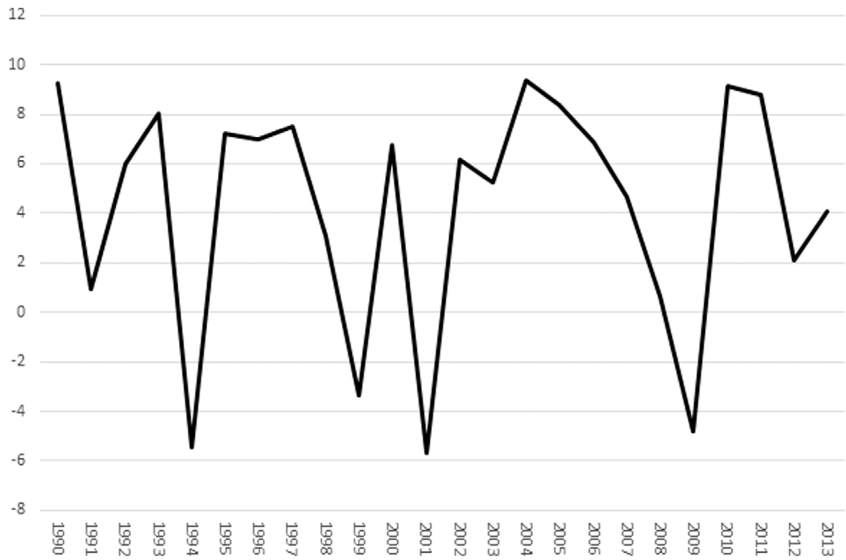


Fig. 1. Annual GDP growth rates, 1990–2013.

economic indicators as explanatory variables is formed. However, as was noted by Bańbura et al. (2013), "partial" models, such as bridge equations, do not allow one to interpret how new data releases revise the nowcast of the target variable, for two reasons. First, single-equation models do not produce forecasts of the input variables, and therefore, it is not possible to isolate the surprise component, i.e., the difference between the realization of the input variable and what the model predicted. As Bańbura and Modugno (2010) showed, it is only the surprise component, or "news", that revises the nowcast of the target variable when new information arrives. Second, single-equation models have to be parameterized to account for the varying availability of the input data and the forecasting horizon for the target variable. This parametrization makes it even more difficult to interpret why the new release of an input variable would revise the nowcast of the target variable. The methodology adopted in this paper can address these issues in a comprehensive unifying framework. Our study computes the "news" and finds that, in line with the findings for other economies, real variables play the most important role in nowcasting GDP. However, contrary the findings for other economies, we find that financial variables play a crucial role in nowcasting SA QoQ GDP growth rates, as important as the role played by surveys.

The remainder of this paper is as follows. Section 2 explains the dataset. Section 3 presents the methodology. Section 4 shows nowcast results of SA QoQ and NSA YoY GDP growth rates. Section 5 presents nowcast results of annual GDP growth rates. Section 6 analyzes the effect of news, and Section 7 concludes.

2. The dataset

We construct a medium-scale dataset for nowcasting Turkish GDP. The literature shows that large-scale DFMs do not necessarily lead to better forecasting performances than smaller DFMs (e.g., Alvarez, Camacho, & Pérez-Quirós, 2012; Boivin & Ng, 2006), and Bańbura and Modugno (2014) found the forecasting accuracies of medium-scale DFMs to be higher than those of large-scale DFMs for euro-area GDP.

Fig. 1 displays the evolution of the annual real GDP growth rates in Turkey between 1990 and 2013. The figure highlights the volatile nature of Turkish growth. The hard landings in 1994, 1999 and 2001, together with the 2008 global crisis and the rapid recovery afterward, depict rather typical behavior for the Turkish economy. It is easy to conclude from the figure that extreme highs and lows are the rule rather than the exception.[6] In terms of international comparisons, the volatility of Turkish growth is one of the highest among its emerging market peers, as was shown by Akat and Yazgan (2013), among others.

Some authors have argued that the high volatility of Turkish GDP growth is related mainly to the intensity of capital inflows (e.g., Akat & Yazgan, 2013; Özatay, 2015),[7] with this heavy reliance on capital flows to attain high growth being attributed in turn to the inadequate level of domestic savings (i.e., if domestic savings are scarce, rapid growth can only be attained with a greater access to foreign capital). However, the volatile nature of capital flows means that this is only a short-term gain, leading to boom–bust cycles in the long run.[8]

6 Clearly, this rollercoaster nature of Turkish growth makes the already difficult task of forecasting even more complicated than usual.

7 See Acemoglu and Ucer (2015) for a recent history of Turkish growth, with an emphasis on the role of structural factors.

8 The view that the Turkish economy is prone to boom–bust cycles originating from capital flows is shared widely among market professionals, some academicians, and even policy makers in Turkey and various international organizations, such as the IMF and World Bank (e.g., IMF, 2012; World Bank and the Ministry of Development of Turkey, 2011). However, it should be emphasized that the postulated causality from capital flows to growth may actually run in the reverse direction or in both directions, in the sense that both variables are affected by a set of possibly unobserved variables lurking in the background, such as increased productivity, better fundamentals, and/or macroeconomic policy stances accompanying these features.

Taking the Turkish economic structure into account, we choose 14 publicly available economic indicators for nowcasting GDP, and categorize them as real variables, survey variables, and financial variables. The real variables used in this study are the industrial production index (IPI), automobile production, the import volume index, the export volume index, the Ercan Türkan Consumer Index, the non-agricultural unemployment rate, and the total employment excluding agriculture. The IPI, which is highly correlated with GDP, is often used as a proxy for GDP when monthly output data are needed for analysis (e.g., Bildirici, Alp, & Kayikçi, 2011; Civcir & Akçağlayan, 2010; Dedeoğlu & Kaya, 2014). Similarly, practitioners predict the YoY GDP growth rate in Turkey simply by taking the YoY quarterly IPI and applying a certain amount of judgment. Automobile production is one of the most important production sectors in Turkey. It is released earlier than the IPI, and is an important determinant of the IPI. Turkey's economy is driven generally by domestic demand; thus, when its economy expands, imports are also expected to increase. In this sense, imports are good predictors of both private consumption and investment expenditure. Another good predictor of private consumption expenditure is the Ercan Türkan Consumer Index, which is based on credit and debit card data. Although Turkey's economy is driven mainly by domestic demand, policy makers curb the domestic demand when the current account deficit reaches unsustainable levels (for instance, in 2012 and 2013). In these periods, growth has relied mainly on exports. Finally, we also include variables related to labor force statistics. We expect the GDP to slow down when the unemployment rate is increasing or total employment is stagnating.

Our survey variables include all publicly available surveys that begin before 2008: the capacity utilization rate (CUR), the Turkish Statistical Agency's (Turkstat) consumer confidence index (CCI), the CNBC-e's CCI, and the real sector confidence index (RSCI).[9] Although financial variables are usually not informative for predicting GDP (see Bańbura et al., 2013), we select three important time series: financial accounts, the real effective exchange rate (REER), and TRLibor 3 Months. As has been explained, Turkey suffers from both a low savings rate and a high current account deficit. Therefore, economists are following developments in financial accounts closely. Another important piece of financial data that is followed closely by both the public and market participants is the exchange rate. We use the real effective exchange rate as a proxy for the exchange rate. One of the reasons for this choice is that the CBRT used the REER as forward guidance for monetary policy in 2013. Finally, we include the three-month interest rate, because it affects domestic demand.

We nowcast both SA QoQ and NSA YoY GDP growth rates. The NSA YoY GDP growth rate is still observed very closely, because market participants in Turkey have only very recently started to use seasonally adjusted data, unlike their counterparts in developed countries.

[9] The RSCI is designed to show short-term tendencies in the manufacturing sector (see CBRT, 2016).

We use NSA data to nowcast NSA YoY GDP growth rates and SA data, when available, to nowcast SA QoQ GDP growth rates. If SA data are not available, we seasonally adjust those variables using Tramo-Seats.[10] SA data are usually announced together with NSA data.

Another aspect of our dataset that should be discussed is that the SA CUR and the SA RSCI begin in 2007:M01, whereas their NSA counterparts begin in 1998:M01. The CUR and the RSCI were revised in 2007 to meet the requirements of the Joint Harmonized European Union Programme of Business and Consumer Surveys. Since then, the CBRT has only announced the CUR and the RSCI, starting from 2007:M01. We nowcast SA QoQ GDP growth rates using official SA figures, and follow the regular market practice of combining new and old surveys for nowcasting NSA YoY GDP growth rates. First, we calculate the monthly percentage differences of the old series. Then, we use the first available data in the new survey and the growth rates of the old survey to calculate new index values for the period in which the new survey lacks data.

In order to have stationary variables, we compute yearly differences of NSA data and monthly differences of SA data. A log transformation is also applied whenever necessary. A list of variables, details about seasonal adjustment procedures, and applied transformations are provided in Appendix A.

3. The methodology

We use a DFM to produce nowcasts of both SA QoQ GDP and NSA YoY GDP growth rates. Adopting a DFM allows us to obtain a parsimonious representation of macroeconomic data, because a small number of dynamic factors is enough to drive most of the co-movement among macroeconomic data series.[11] Our DFM has the following representation:

$$x_{i,t} = \Lambda_i f_t + \epsilon_{i,t}; \quad i = 1, .., n, \tag{1}$$

$$\epsilon_{i,t} = \alpha_i \epsilon_{i,t-1} + \upsilon_{i,t}; \quad \upsilon_{i,t} \sim i.i.d.\ \mathcal{N}(0, \sigma_i^2), \tag{2}$$

where $x_{i,t}$ are our n monthly observations, standardized to a zero mean and unit variance; Λ is an $n \times r$ vector containing factor loadings for monthly variables; $\epsilon_{i,t}$ are the idiosyncratic components of monthly variables that we model as an autoregressive process of order one (AR(1)), as shown in Eq. (2); and f_t is an $r \times 1$ vector of unobserved common factors.[12] f_t is modeled as a stationary vector autoregression process:

$$f_t = \varphi(L) f_{t-1} + \eta_t; \quad \eta_t \sim i.i.d.\ \mathcal{N}(0, R), \tag{3}$$

where $\varphi(L)$ is an $r \times r$ lag polynomial matrix and η_t is an $r \times 1$ vector of innovations.

[10] Official institutions in Turkey also use Tramo-Seats to seasonally adjust data series (e.g., Turkstat, 2013). We follow a simple approach and seasonally adjust all series using the automatic procedure of Tramo-Seats Rev. 941, setting RSA = 4.

[11] See Giannone, Reichlin, and Sala (2005) and Sargent and Sims (1977).

[12] Bańbura and Modugno (2014) argue that modeling the idiosyncratic component as an AR(1) process may improve the forecasts.

We incorporate quarterly variables into the model by constructing a partially observed monthly counterpart for each of them in which the value of the quarterly variable is assigned to the third month of the respective quarter. We assume that the "unobserved monthly" QoQ or YoY growth rate of our quarterly variables (x_t^{UM}) admits the same factor model representation as the monthly real variables:

$$x_{i,t}^{UM} = \Lambda_i^Q f_t + e_{i,t}^Q, \tag{4}$$

$$e_{i,t}^Q = \rho_i^Q e_{i,t-1}^Q + v_{i,t}^Q; \quad v_{i,t}^Q \sim i.i.d.\ \mathcal{N}(0, \sigma_i^2). \tag{5}$$

We link the "unobserved monthly" QoQ growth rate with its quarterly QoQ growth rate counterpart (x^{QQ}) by constructing a partially observed monthly series and using the approximation of Mariano and Murasawa (2003), imposing restrictions on the factor loadings[13]:

$$x_{i,t}^{QQ} = x_{i,t}^{UM} + 2x_{i,t-1}^{UM} + 3x_{i,t-2}^{UM} + 2x_{i,t-3}^{UM} + x_{i,t-4}^{UM}, \tag{6}$$

and we link the "unobserved monthly" YoY growth rate with its quarterly YoY growth rate counterpart ($x_{i,t}^{QY}$) by following Giannone et al. (2013):

$$\begin{aligned} x_{i,t}^{QY} &= X_{i,t}^Q - X_{i,t-12}^Q \\ &= (1 - L^{12})X_{i,t}^Q \\ &\approx (1 - L^{12})(1 + L + L^2)X_{i,t}^M \\ &= (1 + L + L^2)x_{i,t}^{UM} \\ &= x_{i,t}^{UM} + x_{i,t-1}^{UM} + x_{i,t-2}^{UM}; \end{aligned} \tag{7}$$

where X^Q and X^M indicate the log-levels of the quarterly and "fictional" monthly counterpart data, respectively.

DFMs can be estimated in a couple of different ways. Because Turkey is a developing economy where institutions have only recently begun to collect macroeconomic data, our dataset includes series of various different sample lengths and frequencies, and has a small sample size. We therefore adopt the estimation techniques proposed by Bańbura and Modugno (2014), who showed how to modify the expectations maximization algorithm for the estimation of factor models when the data are characterized by an arbitrary pattern of availability. Moreover, this estimation technique is a natural choice because maximum likelihood estimators are more efficient in small samples.

The number of factors in Eq. (1) is selected by using Bai and Ng's (2002) information criteria (BG), modified to take into account the fact that the parameters are estimated through maximum likelihood as per Doz et al. (2012). The number of lags in Eq. (3) is chosen using the Akaike information criteria (AIC).[14]

[13] See Bańbura and Modugno (2014) on the way in which to impose restriction on the factor loadings. Camacho and Perez-Quiros (2010) argue that the aggregation proposed by Mariano and Murasawa (2003) may not be appropriate for emerging market economies. However, models that follow this strategy produce very accurate nowcasts, as was shown by Bragoli et al. (2015) for Brazil, Giannone et al. (2013) and Yiu and Chow (2010) for China, Luciani et al. (2015) for Indonesia, and Caruso (2015) for Mexico, among others.

[14] See Appendix B for details.

4. Nowcasting SA QoQ and NSA YoY GDP growth rates

This section uses the final figures as of November 2014 and ignores historical data revisions. However, we replicate the historical data availability when estimating the model, using only the data that were available at that specific time when producing our nowcasts.

Turkish real GDP data are typically released with a two-quarter delay from the beginning of the reference period. We produce our nowcasts once per month, at the time when labor force statistics are released, i.e., around the 15th day of each month. Because the publication delay is greater than one quarter, we also need to "backcast" the previous quarter's GDP in the months for which the previous quarter's data have not yet been announced. Therefore, we nowcast the first quarter GDP in months corresponding to the first quarter of the year, and both nowcast the second quarter GDP and backcast the first quarter GDP in the months corresponding to the second quarter, because the data on the first quarter GDP are still not released. We continue in the same manner in the third quarter, and nowcast and backcast the third and second quarters, but stop backcasting the first quarter because the data are already available.

As has been indicated, our estimation of the DFM each month uses all of the information that was available at that time. Because of the different publication lags of different variables, the length (or the amount of missing data) of the variables used in the estimation varies from month to month. Although the structure of the monthly variables at each month of estimation is rather stable, GDP has a variable structure. The publication lags are shown more explicitly in Table 1. For example, if we assume that the DFM is estimated in April 2008, the dataset has four months of missing data for GDP, so that our estimation uses past values of GDP, but with four months of missing data. However, the dataset has five and six months of missing data for GDP in May 2008 and June 2008, respectively, so we have fewer lagged values of GDP to use for the estimation. When the first quarter GDP data are released at the end of June, the number of months of missing GDP data in the dataset reduces to four again. As has been mentioned, the number of months of missing data for monthly variables is more stable. For example, the dataset always has two months of missing data for the IPI. The number of months of missing data for a variable at the end of the sample is equal to the publication lag of that variable from the beginning of the reference period. Publication lags for all variables are listed in Appendix A.

4.1. Out-of-sample forecast performance evaluation

We estimate our models recursively using data starting in January 1998, given that Turkish national accounts are available since 1998:Q1. We evaluate the nowcast accuracy of the proposed models on the sample from 2008:Q1 to 2013:Q4, by calculating root mean square forecast errors (RMSFEs). The performance of the DFM is compared with those of an autoregressive model with lags chosen by the AIC, the sample mean of the GDP growth rate, and bridge equations including all variables in the DFM. For

Table 1
The structure of missing and available data in the dataset for GDP and the IPI between April and July 2008.

Variable: GDP							
	15.1.08	15.2.08	15.3.08	15.4.08	15.5.08	15.6.08	15.7.08
15.4.08	NA	NA	NA	NA			
15.5.08	NA	NA	NA	NA	NA		
15.6.08	NA	NA	NA	NA	NA	NA	
15.7.08	A	A	A	NA	NA	NA	NA
Variable: IPI							
	15.1.08	15.2.08	15.3.08	15.4.08	15.5.08	15.6.08	15.7.08
15.4.08	A	A	NA	NA			
15.5.08	A	A	A	NA	NA		
15.6.08	A	A	A	A	NA	NA	
15.7.08	A	A	A	A	A	NA	NA

Note: The table presents an example for the calculation of publication lags for variables. The upper part shows the state of the dataset for GDP, and the lower part the state of the dataset for the IPI. "NA" refers to non-available data, and "A" to available data. The dates in the first column are the dates on which the models are run. The other dates refer to positions in the dataset.

Table 2
RMSFEs for SA QoQ and NSA YoY GDP growth rates, 2008:Q1–2013:Q4.

	SA QoQ GDP				NSA YoY GDP			
	DFM	BE	AR	Mean	DFM	BE	AR	Mean
Q(0)M01	2.29	3.07	3.02	3.81	2.76	7.83	6.40	7.05
Q(0)M02	1.99	3.21	3.02	3.81	2.56	6.89	6.40	7.05
Q(0)M03	1.62	2.93	2.87	3.32	2.09	4.35	5.70	5.89
Q(1)M01	1.60	1.93	2.74	2.91	1.76	3.30	4.11	4.32
Q(1)M02	1.52	1.87	2.74	2.91	1.77	3.14	4.11	4.32
Average	1.80	2.60	2.87	3.35	2.19	5.10	5.34	5.72

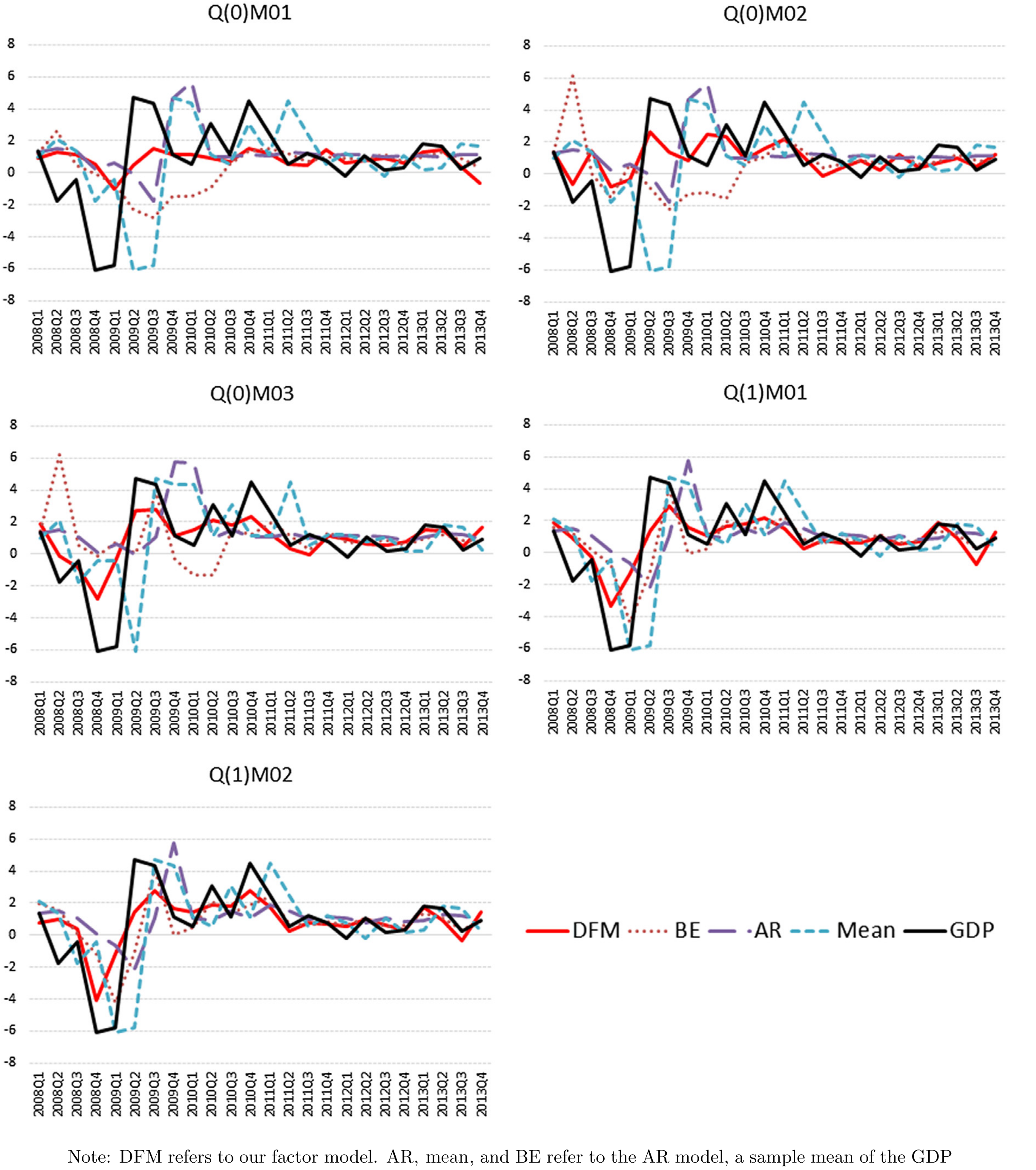
Note: The table reports the RMSFEs of DFMs and benchmark models for SA QoQ and NSA YoY GDP growth rates. DFM refers to our factor model. AR, Mean, and BE refer to the AR model, a sample mean of GDP growth rate, and the bridge equation model, respectively. Average is the average RMSFEs of all nowcast horizons. Q(1)M03 results for SA QoQ GDP growth rates: DFM (1.66), BE (1.81), AR (3.02), Mean (2.34). Q(1)M03 results for NSA YoY GDP growth rates: DFM (2.17), BE (2.43), AR (5.14), Mean (5.19).

bridge equations, we estimate 14 quarterly autoregressive distributed lag (ARDL) models for each predictor. Monthly predictors are transformed into quarterly variables by averaging, and an autoregressive model is used to fill in the missing data when such missing data at the end of the sample period prevent a monthly variable from being transformed into a quarterly variable. The lag structure in ARDL models is determined by the AIC, then the predictions obtained from all ARDL models are combined by simple averaging.

We also use White's (2000) reality check to determine whether the DFM outperforms competing models significantly. White (2000) developed a test of superior predictive ability for multiple models. The null hypothesis of the reality check is that, of the competing/alternative models, none has a superior predictive ability to the benchmark model. In this exercise, we first take the DFM as the benchmark and accept all other models as alternative models, to show that the DFM is not inferior to the other competing models. Then, we choose the rival model with the lowest RMSFE for each nowcast horizon as the benchmark model and use the DFM as an alternative model, in order to show that the DFM has a superior forecasting ability to the best rival model at each nowcast horizon.

Table 2 presents RMSFEs for successive nowcast horizons from Q(0)M01 to Q(1)M02, and average RMSFEs of all of the nowcast horizons for the SA QoQ and NSA YoY GDP growth rates. Figs. 2 and 3 also show actual and forecasted GDP growth rates for all nowcasting horizons. DFM refers to our factor model. AR and Mean refer to the univariate benchmark models, and BE refers to the bridge equations. The current and next quarters with respect to the reference quarter are represented by Q(0) and Q(1), respectively, while the first, second, and third months of a quarter are denoted as M01, M02, and M03, respectively. For each reference quarter, we produce forecasts for the reference quarter from Q(0)M01 to Q(1)M02. We follow this method to produce five predictions, three nowcasts (in Q(0)M01–Q(0)M03) and two backcasts (in Q(1)M01–Q(1)M02), for each reference quarter.[15] Table 3

[15] The GDP data for the fourth and first quarters are usually announced around the end of March and June, respectively. Those for the second and third quarters are usually announced near the 10th of September and December, respectively (since 2013, GDP data for the first quarter have also begun to be announced near the 10th of June). Our fixed estimation time of the middle of each month means that this procedure results in the first and fourth quarter GDP data having six-month delays from the start of the reference quarter, and the third and second quarter GDP data having five-month delays. We could also compute one additional prediction for the first quarter and fourth quarter GDP in Q(1)M03, but we ignore these predictions in order to obtain an equal number of nowcasts for each reference quarter. However, the results for 6th nowcasts are shown in the notes section of Table 2.

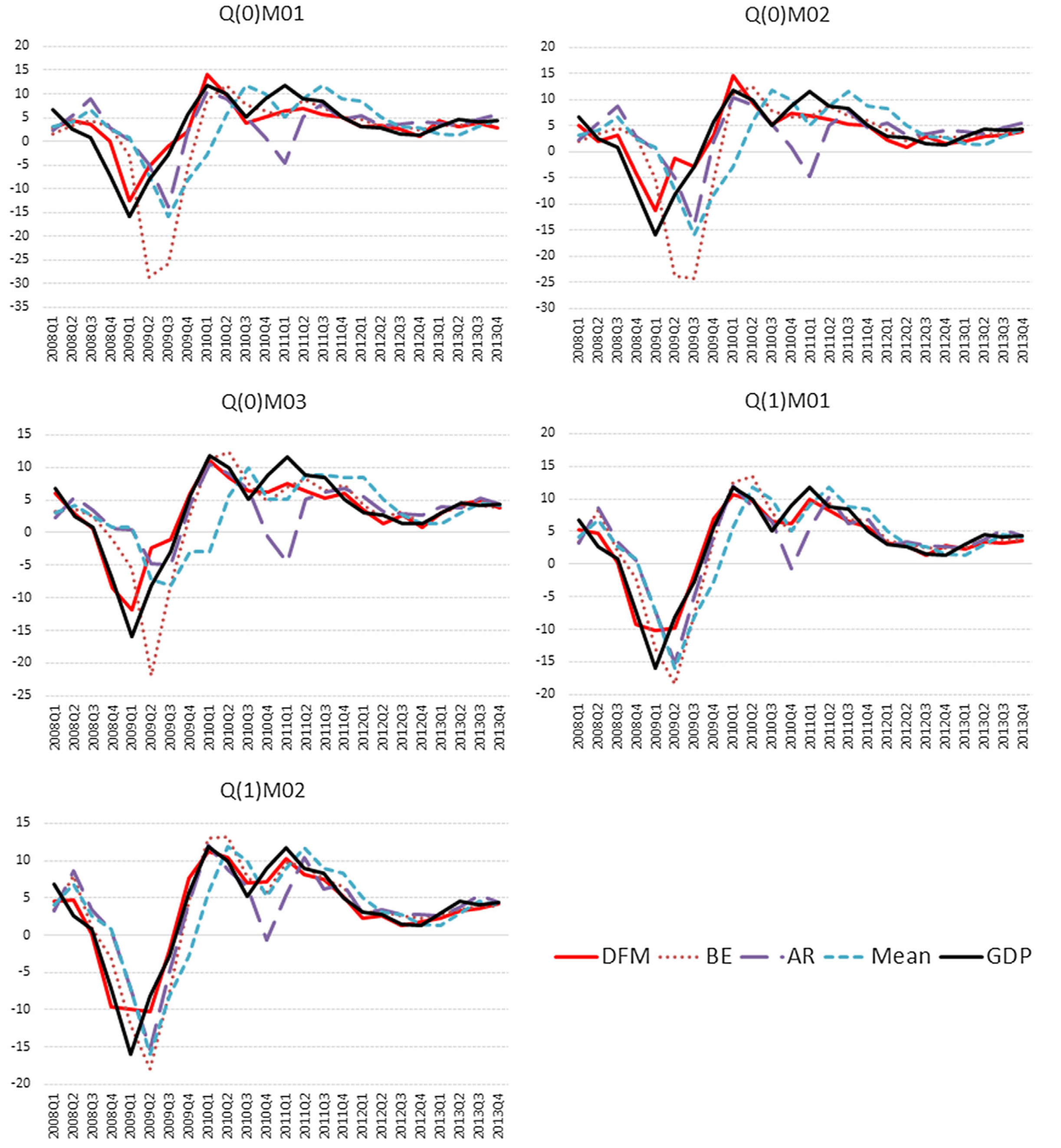


Note: DFM refers to our factor model. AR, mean, and BE refer to the AR model, a sample mean of the GDP growth rate, and the bridge equation model, respectively.

Fig. 2. Actual and forecasted SA QoQ GDP growth rates for successive nowcast horizons, 2008:Q1–2013:Q4.

shows *p*-values for the reality check tests. P-Inf tests whether the DFM is inferior to the competing models and P-Sup tests whether the DFM has a superior predictive ability to the best rival model for each nowcast horizon.

Table 2 shows that the factor models perform better than the other benchmark models at all horizons. For both SA QoQ and NSA YoY GDP growth rates, the bridge equations have the highest forecasting accuracies among the benchmark models. However, the average RMSFE of the DFM for SA QoQ GDP growth rates is 30.8% lower than that of the bridge equations. The average RMSFE of the DFM for the NSA YoY GDP growth rate is 57.0% lower than that of the bridge equation model. Table 3 shows that the DFM is not worse than any alternative model. Furthermore, the DFM outperforms the best competing model at each nowcast horizon for the NSA YoY GDP growth rates,



Note: DFM refers to our factor model. AR, mean, and BE refer to the AR model, a sample mean of the GDP growth rate, and the bridge equation model, respectively.

Fig. 3. Actual and forecasted NSA YoY GDP growth rates for successive nowcast horizons, 2008:Q1–2013:Q4.

whereas for SA QoQ GDP growth rates the DFM beats the best rival model only in the early nowcast horizons. Finally, Figs. 2 and 3 show that the crisis period (2009) and the recovery period (2010) afterward are captured best by DFMs, with the other competing benchmarks performing very poorly in these volatile periods.

The difference in forecasting power between bridge equations and DFMs is especially large for the early nowcast horizons, where many predictors lack data for the reference quarter. The poor performances of bridge equations, especially for early nowcast horizons, clearly show that the joint multivariate modeling strategy in DFM is beneficial for forecasting, as was shown by Angelini, Banbura, and Rünstler (2010); Angelini, Camba-Mendez, Giannone, Reichlin, and Rünstler (2011) and Bańbura et al. (2013). Finally, the RMSFEs of DFMs shrink with each

Table 3
White's reality check test *p*-values for SA QoQ and NSA YoY GDP growth rates, 2008:Q1–2013:Q4.

	SA QoQ GDP		NSA YoY GDP	
	P-Inf	P-Sup	P-Inf	P-Sup
Q(0)M01	0.83	0.07	0.99	0.00
Q(0)M02	0.93	0.00	0.99	0.00
Q(0)M03	0.95	0.00	0.97	0.00
Q(1)M01	0.75	0.21	0.96	0.00
Q(1)M02	0.76	0.20	0.95	0.00

Note: The table reports *p*-values for White's reality check test both for SA QoQ and NSA YoY GDP growth rates. The P-Inf columns show *p*-values when the DFM is the benchmark and other models are the alternative. The P-Sup columns show *p*-values when the DFM is the alternative and the benchmark is the best model according to RMSFE for each horizon.

successive forecasting horizon. In line with the literature (e.g., Bańbura & Modugno, 2014; Bańbura & Rünstler, 2011; Giannone et al., 2008), this shows that timely monthly data increase the forecasting accuracy of DFMs.

5. Nowcasting NSA YoY GDP

As Table 2 shows, the DFM easily beats the simple forecast models that are chosen as benchmarks. Although the use of more sophisticated forecast models or alternative nowcast models as benchmarks would seem to provide a more appropriate evaluation, an analysis that is more interesting from a policy perspective can be provided by comparing the predictions of DFM with those of institutions and experts. However, it is difficult to find public SA QoQ or NSA YoY GDP forecasts for Turkey.[16]

We use forecasts from the IMF's WEOD, the OECD's EO, and the CBRT's SoE. The CBRT's SoE collects the expectations of decision makers in the financial and real sectors on various macroeconomic and financial variables every month. There are two questions about the GDP: "current year annual GDP growth" and "next year annual GDP growth". We use the former for comparisons with our nowcasts. We also use forecasts from the IMF's WEOD and the OECD's EO, which are both updated twice per year.

We make a series of changes to the design of the previous nowcasting exercise in order to enable us to compare our results with annual predictions. Each month of the reference year, we predict NSA YoY GDP growth rates for each quarter of the reference year and for the quarters of the previous year for which the GDP growth rates are not yet available. Next, we use actual and predicted growth rates to calculate the annual real GDP figure (in levels) of the reference year, as well as the annual GDP figure of the previous year if the GDP data of the previous year were not yet available at the time of forecasting. Finally, we compute the annual growth rate of the reference year by using the annual calculated real GDP figure of the reference year and the annual calculated or actual (depending on the period) real GDP figure of the previous year. This method gives us 12 predictions for each reference year.

Table 4
The availability of quarterly GDP data for each month.

	Previous year	Reference year			
	Q4	Q1	Q2	Q3	Q4
M01	NA	NA	NA	NA	NA
M02	NA	NA	NA	NA	NA
M03	NA	NA	NA	NA	NA
M04	A	NA	NA	NA	NA
M05	A	NA	NA	NA	NA
M06	A	NA[a]	NA	NA	NA
M07	A	A	NA	NA	NA
M08	A	A	NA	NA	NA
M09	A	A	A	NA	NA
M10	A	A	A	NA	NA
M11	A	A	A	NA	NA
M12	A	A	A	A	NA

Note: The table shows the availability of GDP data for each month in the reference year. M01, …, M12 denote the first, …, twelfth month of a reference year. "NA" refers to unavailable data and "A" to available data.
[a] Since 2013, GDP data for the first quarter have been being announced around the 10th of June instead of at the end of June.

As was outlined earlier, the exercise in this section includes backcasting, nowcasting, and forecasting NSA YoY GDP growth rates in order to obtain nowcasts of annual GDP growth rates.[17] Table 4 explains this process more clearly, and shows which quarters' GDP data are available in each month of the reference year. For example, in March 2008 we backcast 2007:Q4 GDP, nowcast 2008:Q1 GDP, and forecast GDP data for 2008:Q2, 2008:Q3, and 2008:Q4. We then use these data to compute both 2007 and 2008 annual real GDP figures, thus obtaining a nowcast of the 2008 annual GDP growth rate. GDP data for 2007:Q4 becomes available in the next month, so we backcast 2008:Q1 GDP as previously, as well as nowcasting 2008:Q2 GDP and forecasting the GDP data for 2008:Q3 and 2008:Q4. Then, we estimate only 2008 annual real GDP figures. We obtain the nowcast of the 2008 annual GDP growth rate using actual 2007 real GDP and estimated 2008 real GDP figures. Following this process, we continue to nowcast the 2008 annual GDP growth rate until the end of 2008. In the first month of 2009, we begin to estimate the annual GDP growth rate of 2009. We continue this procedure until the last month of 2013.

Using final revised data might bias the results in favor of the DFM, because institutional and professional forecasts are computed using actual real-time data. Therefore, we also construct a "partial" real-time dataset. We have vintages for GDP, the IPI, the CUR, the import volume index, the export volume index, the Turkstat's CCI, the RSCI, and TRLibor. Unfortunately, we have to use final revised data for the other seven variables in this exercise, due to a lack of vintage data for Turkey. The rest of the exercise's design is the same as the previous one.

Table 5 shows the RMSFEs for successive nowcast horizons and the average RMSFEs of all nowcast horizons for annual GDP growth rates. DFM refers to the model

[16] One notable exception is the CNBC-e analysts' expectations survey, which contains forecasts of GDP growth rates from 2002 to the present. However, the survey has some missing data, due to disturbances in the process of collecting the forecast from the experts. In addition, the survey only includes one prediction for each quarter.

[17] We refer to this exercise as a nowcasting exercise because we predict annual GDP growth rates only during the year of which the growth rate is being predicted.

Table 5
RMSFEs for annual GDP growth rates, 2008–2013.

	DFM	OECD	IMF	SoE
M01	2.18	4.38	4.62	3.90
M02	2.13	4.38	4.62	3.51
M03	2.03	4.38	4.62	3.30
M04	1.46	4.38	4.28	2.88
M05	2.48	4.38	2.50	2.59
M06	2.76	3.04	2.50	2.40
M07	2.08	1.76	2.50	2.24
M08	1.89	1.76	2.50	2.21
M09	0.75	1.76	2.50	2.21
M10	0.62	1.76	1.55	1.81
M11	0.67	1.76	1.55	1.58
M12	0.66	1.20	1.55	1.23
Average	1.64	2.91	2.94	2.49

Note: The table reports RMSFEs of the DFM, the institutions, and the survey for annual GDP growth rates for the "partial" real-time dataset. DFM refers to our factor model. OECD, IMF, and SoE refer to the OECD's EO, the IMF's WEOD, and the CBRT's SoE, respectively. Average shows the average RMSFEs of all nowcast horizons. M01, …, M12 denote the first, …, twelfth month of a reference year.

Table 6
White's reality check test p-values for annual GDP growth rates, 2008–2013.

	P-Inf	P-Sup
M01	0.98	0.00
M02	0.98	0.00
M03	0.99	0.00
M04	1.00	0.00
M05	0.78	0.42
M06	0.38	0.81
M07	0.27	0.85
M08	0.40	0.73
M09	1.00	0.00
M10	1.00	0.00
M11	1.00	0.00
M12	1.00	0.00

Note: The table reports p-values for White's reality check test both for annual growth rates. The P-Inf columns show p-values when the DFM is the benchmark and professional forecasters are the alternative. The P-Sup columns show p-values when the DFM is the alternative and the benchmark is the best professional forecaster according to RMSFE for each horizon. M01, …, M12 denote the first, …, twelfth month of a reference year.

presented here. OECD, IMF and SoE refer to the OECD's EO, the IMF's WEOD, and the CBRT's SoE, respectively. Table 5 shows that, on average, the DFM performs better than the IMF, OECD, and SoE. As expected, SoE has the best forecast accuracy after the DFM. However, the average RMSFE of the DFM is still 34.1% lower than that of SoE. Table 6 also shows that the DFM is not worse than any professional forecaster, and outperforms the best professional forecaster at most nowcast horizons.

We analyze the results further by presenting actual and forecasted annual GDP growth rates in Fig. 4. Fig. 4 clearly shows that the biggest differences between forecasts of the DFM and institutional forecasts occur during the crisis and in the subsequent recovery period. Turkey's economy contracted sharply in 2009 as a result of the global economic crisis, then enjoyed a very rapid recovery in 2010 and 2011. Institutional forecasts are conservative in these volatile periods, especially at the beginning of such periods. The DFM incorporates new information quickly and efficiently, even though there are a few significant misses in periods around the middle of the year. These results are highly interesting because Ang, Bekaert, and Wei (2007) claimed that forecasts from surveys are superior to model-based forecasts, and the literature shows that nowcasting models perform well, especially for short forecasting horizons (Bańbura et al., 2013).

6. The effects of news and model re-estimation

The previous out-of-sample forecasting evaluation updated GDP nowcasts each month based on the new data releases. Because our DFM produces forecasts for all of the variables used in the dataset, instead of only for GDP, only the news or "unexpected" component from the newly released data should be used to revise nowcasts of GDP. In other words, the change between two consecutive nowcasts of GDP can be the result of both news from all variables, i.e., the unpredicted component of the dataset, and model re-estimation, i.e., the change in the model parameters as a result of newly released data.

We denote two consecutive datasets collected one month apart[18] as Ω_{v+1} and Ω_v, and newly released data that is included in Ω_{v+1} but not in Ω_v as x.[19] Defining the nowcast of quarterly GDP x_t^Q as an orthogonal projection of itself on the available dataset, the nowcast can be shown as follows:

$$\mathbb{E}[x_t^Q \mid \Omega_{v+1}] = \mathbb{E}[x_t^Q \mid \Omega_v] - \mathbb{E}[x_t^Q \mid I_{v+1}] \quad (8)$$

where $\mathbb{E}[x_t^Q \mid \Omega_{v+1}]$ and $\mathbb{E}[x_t^Q \mid \Omega_v]$ are the new and old nowcasts respectively, and $\mathbb{E}[x_t^Q \mid I_{v+1}]$ is the revision between the two consecutive nowcasts. I_{v+1} denotes news, which is the unexpected part of the release with respect to the model, and is shown as $I_{v+1} = x - \mathbb{E}[x \mid \Omega_v]$. Eq. (8) shows that the nowcast of GDP between successive months changes only if the values of newly released variables and the DFM's predictions of those variables based on Ω_v differ. In addition, the effect of parameter re-estimation on nowcasts with each expansion of the dataset is also taken into account.[20]

We determine which types of news on nowcast revisions are most important in a tractable and compact fashion by computing the relative impacts of news and model re-estimation. For a given horizon, the relative impact metric shows how much each news or model re-estimation revises the nowcast as a percentage of the sum of the absolute contributions of news and model re-estimation. Let us define the contribution of news or a model re-estimation for the reference quarter r at the nowcast horizon h as $C_{i,r,h}$. We can then calculate the relative impact of news or a model re-estimation for the reference quarter r at the nowcast horizon h as follows:

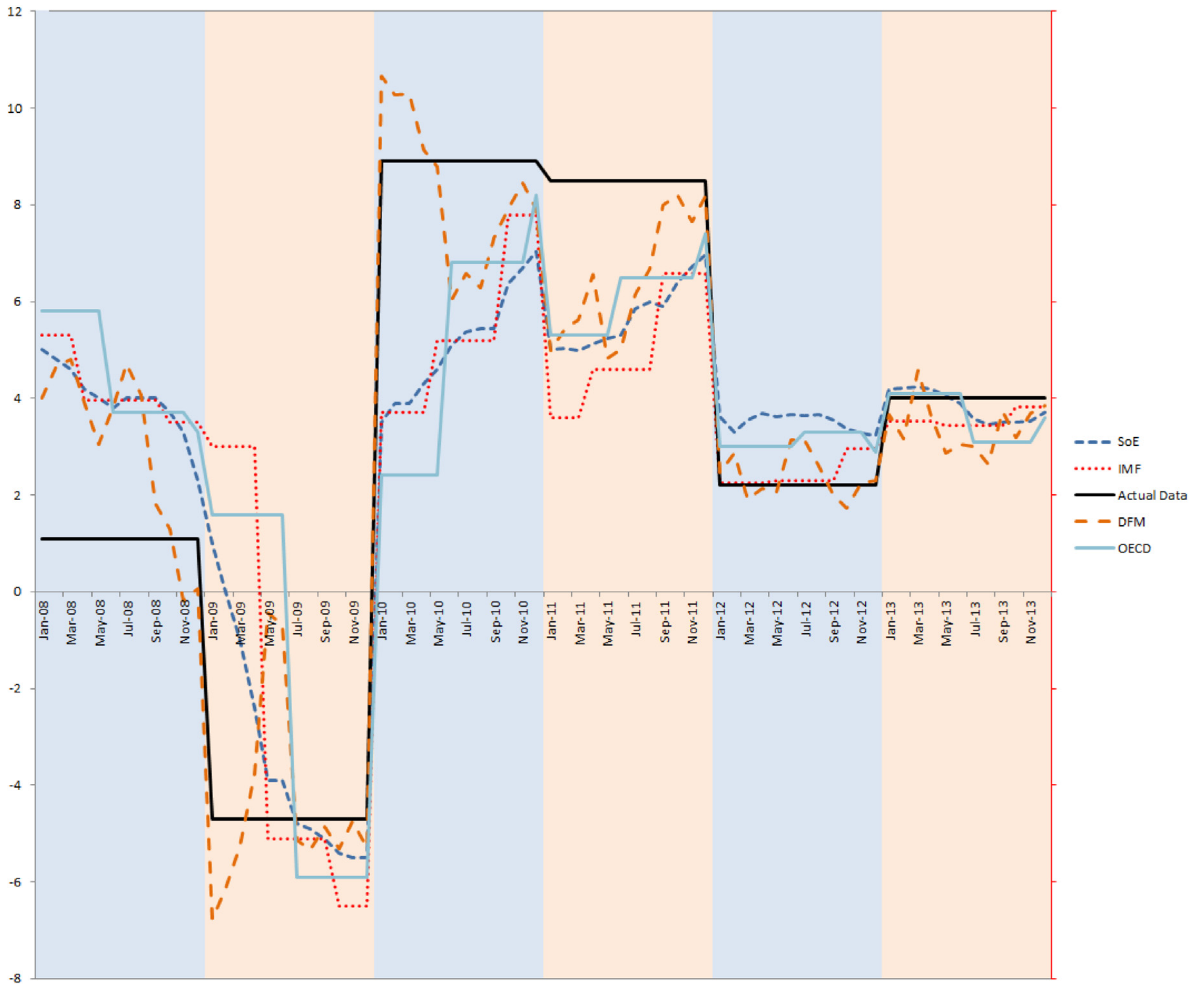
$$\text{Relative Impact}_{i,r,h} = \frac{|C_{i,r,h}|}{\sum_{i=1}^{S} |C_{i,r,h}|}, \quad (9)$$

where $i = 1, \ldots, S$ is an index that contains all news and model re-estimation.

[18] In principle, any frequency can be used, but we use a one-month frame in this study.

[19] We abstract from data revisions here.

[20] See Bańbura and Modugno (2010) for a more detailed explanation.



Note: DFM refers to our factor model. OECD, IMF, and SoE refer to the OECD's EO, the IMF's WEOD, and the CBRT's SoE, respectively.

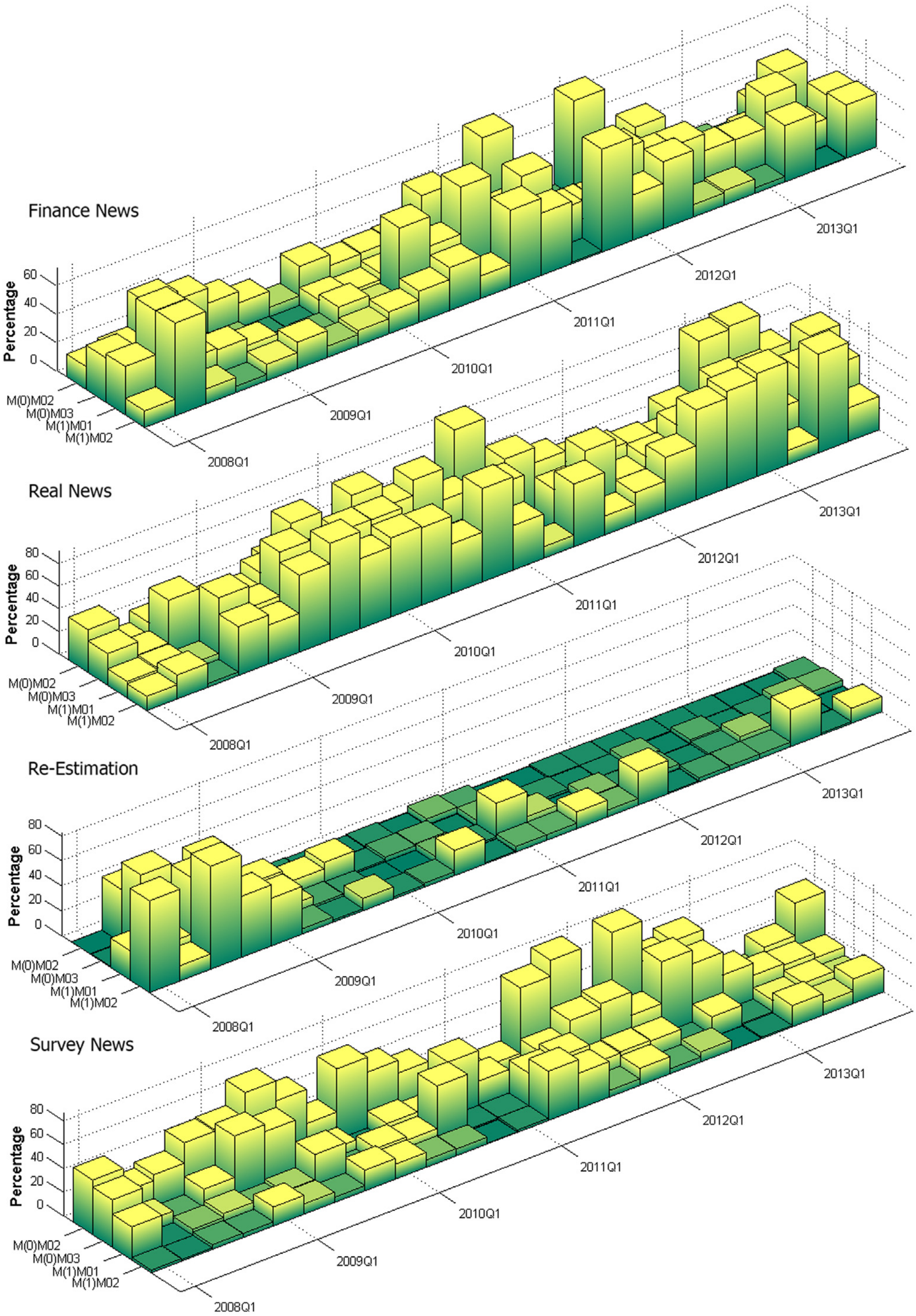
Fig. 4. Actual and forecasted annual GDP growth rates, 2008–2013.

Figs. 5 and 6 present the relative impact of news and model re-estimation on SA QoQ and NSA YoY GDP growth rates, calculated for the four consecutive nowcasts from Q(0)M02 to Q(1)M02 for each reference quarter in our evaluation period of 2008:Q1 to 2013:Q4.[21] For the sake of tractability, we group the variables into three categories: variables relating to the real economy (Real), those obtained from surveys (Survey), and financial variables (Finance).[22] Re-estimation (ReEst) denotes the impact of model re-estimation. The Y-axis shows nowcast periods. The X-axis shows reference quarters, and the Z-axis shows the relative absolute impact of news or model re-estimation as a percentage of the sum of all absolute impacts. Table 7 summarizes the information contained in Figs. 5 and 6.

The graphs for SA QoQ GDP and NSA YoY GDP clearly show that the real variable group is the most prominent factor in nowcast revisions. After the first period, the relative impact of real news accounts for more than 40% of all contributions. Furthermore, survey news have a significant impact on nowcast revisions in the early period because some real variables have not yet been announced at those times. Financial news have only a small impact in many periods, when nowcasting NSA YoY GDP growth rates, but seems to matter more when nowcasting SA QoQ GDP growth rates. There are quite a few periods in which the relative impact of finance news reaches high levels for SA QoQ GDP nowcasts. Table 7 also shows that the average relative impact of finance news (23.57%) is higher than that of survey news (22.91%) for SA QoQ GDP growth rates. We analyze the impact of financial variables

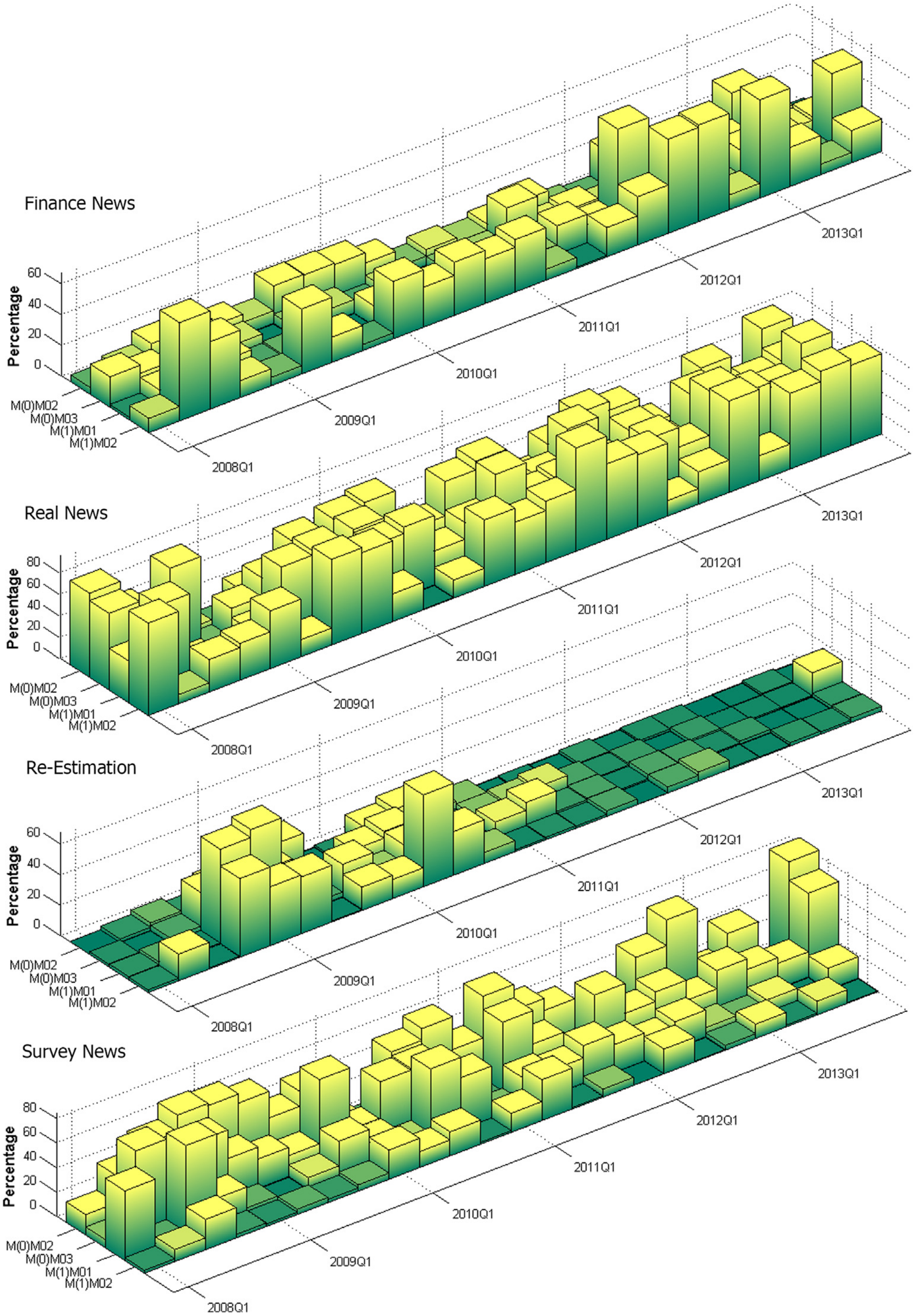
[21] Q(0)M01 is not included because nowcast revisions can only be calculated starting from Q(0)M02.

[22] See Appendix A for the variables' associated groups.



Note: The Z-axis shows percentages. The X-axis shows reference quarters (2008:Q1–2013:Q4). The Y-axis shows nowcast periods.

Fig. 5. Relative impacts of news and model re-estimation for SA QoQ GDP growth rates, 2008Q1–2013Q4.



Note: The Z-axis shows percentages. The X-axis shows reference quarters (2008:Q1–2013:Q4). The Y-axis shows nowcast periods.

Fig. 6. Relative impacts of news and model re-estimation for NSA YoY GDP growth rates, 2008:Q1-2013:Q4.

Table 7
Average relative impacts of news and model re-estimation, 2008:Q1–2013:Q4.

	SA QoQ GDP				NSA YoY GDP			
	Finance	Real	ReEst	Survey	Finance	Real	ReEst	Survey
Q(0)M02	23.34	32.50	3.57	40.59	14.27	37.58	4.75	43.41
Q(0)M03	15.05	48.76	9.33	26.86	11.49	47.24	7.68	33.58
Q(1)M01	31.94	43.57	10.23	14.26	14.83	52.26	11.01	21.90
Q(1)M02	23.93	49.24	16.89	9.94	27.88	50.60	12.40	9.12
Average	23.57	43.52	10.01	22.91	17.12	46.92	8.96	27.00

Note: The table shows the average relative impacts of news and model re-estimation over all reference quarters. "Average" is the average of the relative impact of news and model re-estimation over all reference quarters and nowcast periods.

Table 8
RMSFEs for DFMs with and without financial variables, 2008:Q1–2013:Q4.

	Full dataset		Partial dataset	
	SA	NSA	SA	NSA
Q(0)M01	2.29	2.76	2.46	2.68
Q(0)M02	1.99	2.56	2.47	2.63
Q(0)M03	1.62	2.09	2.29	2.28
Q(1)M01	1.60	1.76	2.09	1.74
Q(1)M02	1.52	1.77	2.11	1.72
Average	1.80	2.19	2.28	2.21

Note: The table reports RMSFEs of DFMs for the full and partial datasets. The partial dataset does not include financial variables. SA refers to SA QoQ GDP growth rates and NSA refers to NSA YoY GDP growth rates. Average shows average RMSFEs for all nowcast horizons. Q(1)M03 results for the full dataset: SA (1.66), NSA (2.17). Q(1)M03 results for the partial dataset: SA (2.48), NSA (2.08).

on nowcasting further by dropping financial variables from the DFM and nowcasting GDP growth rates without financial variables.[23] Table 8 shows that the average RMSFE of DFM with financial variables is 21.1% lower than that without for SA QoQ GDP growth rates, but only 0.9% lower for NSA YoY GDP growth rates. The results of this nowcasting exercise are in line with the outcome of the news decomposition. The results for SA QoQ GDP growth rates are highly interesting, because the literature shows that financial variables are not very helpful for forecasting SA QoQ GDP (e.g., Bańbura et al., 2013), and the effect of finance news on nowcast revisions for SA QoQ GDP is low (e.g., Bańbura & Modugno, 2010). Finally, the effect of re-estimation is very high during the global economic crisis because of the high volatility experienced at that time; however, the impact of model re-estimation becomes negligible after the crisis.

7. Conclusion

This paper nowcasts SA QoQ, NSA YoY, and annual GDP growth rates by using the methodology of Bańbura and Modugno (2014). In addition to handling mixed-frequency datasets with an arbitrary pattern of missing data efficiently, the adopted methodology also offers a comprehensive unifying solution that allows us to compute news. For nowcasting, we adopt a medium-scale dataset consisting of 15 variables and perform sample forecasts recursively between 2008:Q1 and 2013:Q4.

We compare SA QoQ and NSA YoY GDP nowcast results of the DFM with those of the AR model, a sample mean of GDP growth rate, and the bridge equation model. The results show that the DFM outperforms competing benchmark models.

Furthermore, we compare the annual GDP nowcasts from the DFM with those from the IMF, the OECD, and the CBRT's Survey of Expectations. We demonstrate that the DFM can even beat professional forecasters, and find that the biggest difference between institutional forecasts and the forecasts of the DFM exists in volatile periods. The DFM incorporates new information quickly and efficiently, whereas professional forecasters seem to remain conservative in these periods.

Finally, we evaluate the impact of news. We find that survey news have a significant impact on nowcast revisions in earlier forecasting periods, but that the impact diminishes quickly in later periods. Real variables have a strong impact at all times, especially in later periods. We find that the impact of finance news on nowcast revisions is more prominent for SA QoQ GDP growth rates than NSA YoY GDP growth rates. We also analyze the usefulness of financial variables for nowcasting by comparing DFMs with and without financial variables. In contrast to the literature, we find that removing financial variables from the dataset causes the nowcasting accuracy of DFMs to deteriorate at all nowcasting horizons for SA QoQ GDP growth rates. Finally, the effect of model re-estimation is very high during the global economic crisis, but fades quickly after the crisis.

Acknowledgments

The opinions in this paper are those of the authors and do not necessarily reflect the views of the Board of Governors of the Federal Reserve System.

Appendix A. Description of the dataset

See Table A.1.

[23] The lag and factor structure for newly estimated DFMs are recalculated as is shown in Appendix B.

Table A.1
Description of the dataset.

Group	Variables	Publication lags	Transformation		Released by	SA by[b]	Starting date
			Log	Difference[a]			
Real	Industrial production index	2	1	1	Turkstat	Turkstat	2005M01
Survey	Capacity utilization rate	1	0	1	CBRT	CBRT	2007M01, 1998M01[c]
Real	Export volume index	2	1	1	Turkstat	Turkstat	1998M01
Real	Import volume index	2	1	1	Turkstat	Turkstat	1998M01
Real	Ercan Türkan consumer index	2	1	1	Ercan Türkan	TDM	2006M04
Real	Total car production	1	1	1	AMS	Authors	1999M01
Survey	Turkstat consumer confidence index	1	1	1	Turkstat	Authors	2004M01
Survey	CNBC-E consumer confidence index	1	1	1	CNBC-E	Authors	2002M01
Survey	Real sector confidence index	1	1	1	CBRT	CBRT	2007M01, 1998M01[c]
Real	Non-agricultural unemployment rate	3	0	1	Turkstat	Turkstat	2005M01
Real	Total employment excl. agriculture	3	1	1	Turkstat	Turkstat	2005M01
Financial	Real effective Exch. rate by CPI	1	1	1	CBRT	NSA	2003M01
Financial	TRLIBOR 3 months	1	0	1	BAT	NSA	2002M08
Financial	Financial account	2	0	1	CBRT	Authors	1998M01
Real	Real gross domestic product	5, 6[d]	1	1	Turkstat	Turkstat	1998Q1

Notes: The table shows the variables, their associated groups, their publication lags from the start of the reference period, applied transformations, the institutions and people by which the variables are released and seasonally adjusted, and their starting dates. Turkstat refers to the Turkish Statistical Institute. CBRT refers to the Central Bank of Republic of Turkey. AMS refers to the Automotive Manufacturers Association. BAT refers to the Banks Association of Turkey. TDM refers to the Turkish Data Manager.

[a] We use yearly differences for NSA data and monthly differences for SA data.

[b] Only applicable for SA data.

[c] 1998M01 is the starting date for NSA data, 2007M01 is the starting date for SA data.

[d] GDP data for the fourth quarter and GDP data for the first quarter (except the first quarter of 2013) have a maximum six-month announcement lag. Others have maximum a five-month announcement delay. For more information on publication lags, see Section 4.

Appendix B. Factor and lag selection

Bai and Ng (2002) proposed an information criterion for determining the number of factors in an approximate factor model with a balanced dataset. The original information criterion was for principal components. We use a modified information criterion for maximum likelihood estimation. In addition to the notation already defined in Section 3, T and n refer to the numbers of observations and variables in the dataset, respectively. Then, the modified information criterion is as follows:

$$IC(r) = \ln(V(r, f)) + r\ln(g(n, T))/g(n, T) \tag{10}$$

where $g(n, T) = \min(\sqrt{T}, n/\ln(n))$ and $V(r, f) = \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}(x_{it} - \Lambda_i f_t)^2$.

We test up to five factors and try to find the number of factors that minimizes $IC(r)$. As our dataset is highly unbalanced, we reorganize it to make it balanced. We ensure the robustness of our test results by using three different reorganized datasets. First, we use part of the estimation period (January 1998–December 2007) in which all variables have data available. This reorganization yields eight observations for the SA dataset and six for the NSA dataset. Second, we remove some variables with a short data span in order to make the first one longer. We remove three variables from the SA dataset, which yields 32 observations. We remove one observation from the NSA dataset, and our number of observations expands to 21. Third, we use part of the overall period (January 1998–March 2014) in which all variables have data available. This yields 83 observations for the SA dataset and 81 for the NSA dataset. The results for both the SA and NSA datasets are shown in Table B.1. The results for all datasets indicate one factor. Therefore, we use one factor in Eq. (1).

We determine the number of lags in Eq. (3) using the AIC. We estimate one factor (as was shown by BG) for the estimation period, then choose the number of lags that minimizes the AIC. The maximum number of lags tested is five. The AIC chooses three lags for the SA full dataset,

Table B.1
Bai and Ng criterion for determining the number of factors.

	SA dataset			NSA dataset		
	1st	2nd	3rd	1st	2nd	3rd
1	−6.26 (−15.67)	−6.61 (−15.33)	−5.91 (−14.75)	−3.63 (−11.23)	−4.32 (−10.67)	−4.38 (−10.58)
2	−5.93 (−14.98)	−6.30 (−14.60)	−5.62 (−14.27)	−3.27 (−10.26)	−3.99 (−10.05)	−4.07 (−9.96)
3	−5.63 (−14.37)	−6.01 (−14.05)	−5.35 (−13.79)	−2.88 (−9.76)	−3.65 (−9.61)	−3.77 (−9.51)
4	−4.77 (−13.99)	−5.35 (−13.58)	−4.82 (−13.38)	−2.51 (−9.47)	−3.33 (−9.16)	−3.38 (−9.10)
5	−4.44 (−13.12)	−5.83 (−12.71)	−4.38 (−12.81)	−2.26 (−8.52)	−3.04 (−8.55)	−3.17 (−8.49)
T	**8**	**32**	**83**	**6**	**21**	**81**
n	**14** (**11**)	**11** (**8**)	**14** (**11**)	**14** (**11**)	**13** (**10**)	**14** (**11**)

Notes: The table shows Bai and Ng information criterion values for the SA and NSA datasets with and without financial variables. The numbers in parentheses are the results for the dataset without financial variables, and the others are the results for the full dataset. The first column shows the numbers of factors tested in the corresponding rows. T is the number of observations in the datasets, and n denotes the number of variables in the datasets.

two lags for the SA dataset without financial variables, and four lags for both NSA datasets.

References

Aastveit, K. A., & Trovik, T. (2012). Nowcasting Norwegian GDP: The role of asset prices in a small open economy. *Empirical Economics*, *42*(1), 95–119.

Acemoglu, D., & Ucer, M. (2015). *The ups and downs of Turkish growth, 2002–2015: Political dynamics, the European union and the institutional slide*. Working Paper No. 21608, National Bureau of Economic Research.

Akat, A. S., & Yazgan, E. (2013). Observations on Turkey's recent economic performance. *Atlantic Economic Journal*, *41*(1), 1–27.

Akkoyun, H.Ç., & Günay, M. (2012). *Nowcasting Turkish GDP growth*. Working Paper No. 12/33, Central Bank of the Republic of Turkey.

Alvarez, R., Camacho, M., & Pérez-Quirós, G. (2012). *Finite sample performance of small versus large scale dynamic factor models*. CEPR Discussion Papers No. DP8867.

Ang, A., Bekaert, G., & Wei, M. (2007). Do macro variables, asset markets, or surveys forecast inflation better? *Journal of Monetary Economics*, *54*(4), 1163–1212.

Angelini, E., Banbura, M., & Rünstler, G. (2010). Estimating and forecasting the Euro area monthly national accounts from a dynamic factor model. *OECD Journal: Journal of Business Cycle Measurement and Analysis*, *2010*(1).

Angelini, E., Camba-Mendez, G., Giannone, D., Reichlin, L., & Rünstler, G. (2011). Short-term forecasts of Euro area GDP growth. *The Econometrics Journal*, *14*(1), C25–C44.

Arnostova, K., Havrlant, D., Rùžièka, L., & Tóth, P. (2011). Short-term forecasting of czech quarterly GDP using monthly indicators. *Czech Journal of Economics and Finance*, *61*(6), 566–583.

Bai, J., & Ng, S. (2002). Determining the number of factors in approximate factor models. *Econometrica*, *70*(1), 191–221.

Bańbura, M., Giannone, D., Modugno, M., & Reichlin, L. (2013). *Handbook of economic forecasting*: *Vol. 2. Now-casting and the real-time data flow* (pp. 195–237). Elsevier, (Chapter 0).

Bańbura, M., & Modugno, M. (2010). *Maximum likelihood estimation of factor models on datasets with arbitrary pattern of missing data*. Working Paper Series No. 1189, European Central Bank.

Bańbura, M., & Modugno, M. (2014). Maximum likelihood estimation of factor models on datasets with arbitrary pattern of missing data. *Journal of Applied Econometrics*, *29*(1), 133–160.

Bańbura, M., & Rünstler, G. (2011). A look into the factor model black box: Publication lags and the role of hard and soft data in forecasting GDP. *International Journal of Forecasting*, *27*(2), 333–346.

Barhoumi, K., Darné, O., & Ferrara, L. (2010). Are disaggregate data useful for factor analysis in forecasting french GDP? *Journal of Forecasting*, *29*(1–2), 132–144.

Bildirici, M., Alp, E. A., & Kayikçi, F. (2011). Relationship between hot money and economic growth: Tar-cointegration and MS-VAR analysis. *African Journal of Business Management*, *5*(3), 1060–1075.

Boivin, J., & Ng, S. (2006). Are more data always better for factor analysis? *Journal of Econometrics*, *132*(1), 169–194.

Bragoli, D., Metelli, L., & Modugno, M. (2015). The importance of updating: Evidence from a Brazilian nowcasting model. *Journal of Business Cycle Measurement and Analysis*, *2015*(1), 5–22.

Camacho, M., & Perez-Quiros, G. (2010). Introducing the Euro-sting: Short-term indicator of Euro area growth. *Journal of Applied Econometrics*, *25*(4), 663–694.

Caruso, A. (2015). *Nowcasting Mexican GDP*. Working Paper No. 2015-40, ECARES.

CBRT, (2016). *Methodological information on "Business Tendency Survey and Real Sector Confidence Index"*. Mimeo. URL: http://goo.gl/Ymwltv (accessed on 26.07.16).

Civcir, I., & Akçağlayan, A. (2010). Inflation targeting and the exchange rate: Does it matter in Turkey? *Journal of Policy Modeling*, *32*(3), 339–354.

D'Agostino, A., McQuinn, K., & O'Brien, D. (2013). Nowcasting Irish GDP. *OECD Journal: Journal of Business Cycle Measurement and Analysis*, *2012*(2), 21–31.

D'Agostino, A., Modugno, M., & Osbat, C. (2015). *A global trade model for the Euro area*. Finance and Economics Discussion Series 2015-13, Board of Governors of the Federal Reserve System.

Dahlhaus, T., Guénette, J.-D., & Vasishtha, G. (2015). *Nowcasting BRIC+ M in real time*. Working Paper 2015-38, Bank of Canada.

de Antonio Liedo, D. (2014). *Nowcasting Belgium*. Working Paper Research 256, National Bank of Belgium.

Dedeoğlu, D., & Kaya, H. (2014). Pass-through of oil prices to domestic prices: Evidence from an oil-hungry but oil-poor emerging market. *Economic Modelling*, *43*, 67–74.

de Winter, J. (2011). *Forecasting GDP growth in times of crisis: Private sector forecasts versus statistical models*. DNB Working Papers No. 320, Netherlands Central Bank, Research Department.

Doz, C., Giannone, D., & Reichlin, L. (2012). A quasi–maximum likelihood approach for large, approximate dynamic factor models. *The Review of Economics and Statistics*, *94*(4), 1014–1024.

Ermişoğlu, E., Akçelik, Y., & Oduncu, A. (2013). Nowcasting GDP growth with credit data: Evidence from an emerging market economy. *Borsa Istanbul Review*, *13*(4), 93–98.

Giannone, D., Agrippino, S. M., & Modugno, M. (2013). *Nowcasting China Real GDP*. Mimeo.

Giannone, D., Reichlin, L., & Sala, L. (2005). Monetary policy in real time. In M. Gertler, & K. Rogoff (Eds.), *NBER macroeconomic annual* (pp. 161–200). Cambridge: MIT Press.

Giannone, D., Reichlin, L., & Small, D. (2008). Nowcasting: The real-time informational content of macroeconomic data. *Journal of Monetary Economics*, *55*(4), 665–676.

IMF (2012). *Turkey: 2011 article IV consultation. Country Report No. 12/16*. IMF.

Jungbacker, B., & Koopman, S. J. (2015). Likelihood-based dynamic factor analysis for measurement and forecasting. *The Econometrics Journal*, *18*(2), C1–C21.

Jungbacker, B., Koopman, S. J., & Van der Wel, M. (2011). Maximum likelihood estimation for dynamic factor models with missing data. *Journal of Economic Dynamics and Control*, *35*(8), 1358–1368.

Luciani, M., Pundit, M., Ramayandi, A., & Veronese, G. (2015). *Nowcasting Indonesia*. Finance and Economics Discussion Series 2015-100.

Mariano, R. S., & Murasawa, Y. (2003). A new coincident index of business cycles based on monthly and quarterly series. *Journal of Applied Econometrics*, *18*(4), 427–443.

Matheson, T. D. (2010). An analysis of the informational content of New Zealand data releases: the importance of business opinion surveys. *Economic Modelling*, *27*(1), 304–314.

Modugno, M. (2013). Now-casting inflation using high frequency data. *International Journal of Forecasting*, *29*(4), 664–675.

Özatay, F. (2015). Turkey's distressing dance with capital flows. *Emerging Markets Finance and Trade*, 1–15.

Sargent, T.J., & Sims, C.A. (1977). *Business cycle modeling without pretending to have too much a priori economic theory*. Working Papers No. 55, Federal Reserve Bank of Minneapolis.

Shumway, R. H., & Stoffer, D. S. (1982). An approach to time series smoothing and forecasting using the EM algorithm. *Journal of Time Series Analysis*, *3*(4), 253–264.

Turkstat, (2013). *Mevsim ve Takvim Etkilerinden Arındırılmış Verilerin Yorumlanmasıve Üretim Sürecine İlişkin Kamuoyu Duyurusu*. Mimeo.

Watson, M. W., & Engle, R. F. (1983). Alternative algorithms for the estimation of dynamic factor, mimic and varying coefficient regression models. *Journal of Econometrics*, *23*(3), 385–400.

White, H. (2000). A reality check for data snooping. *Econometrica*, *68*(5), 1097–1126.

World Bank and the Ministry of Development of Turkey (2011). *Sustaining high growth: the role of domestic savings. Turkey country economic synthesis Report No: 66301-TR*. The World Bank.

Yiu, M. S., & Chow, K. K. (2010). Nowcasting Chinese GDP: Information content of economic and financial data. *China Economic Journal*, *3*(3), 223–240.

Michele Modugno is an economist at the Federal Reserve Board, in Washington DC.

Barış Soybilgen is a researcher at Bilgi University, in Istanbul.

Ege Yazgan is full professor of Economics at at Bilgi University, in Istanbul.