TÜRKİYE CUMHURİYET MERKEZ BANKASI

WORKING PAPER NO: 12/33

Nowcasting Turkish GDP Growth

December 2012

Hüseyin Çağrı AKKOYUN
Mahmut GÜNAY

Address:
Central Bank of the Republic of Turkey
Head Office
Research and Monetary Policy Department
İstiklal Caddesi No: 10
Ulus, 06100 Ankara, Turkey

Phone:
+90 312 507 54 02

Facsimile:
+90 312 507 57 33

Nowcasting Turkish GDP Growth

Hüseyin Çağrı Akkoyun and Mahmut Günay[1]

Abstract

In this paper we present backcasts and nowcasts for quarter on quarter Gross Domestic Product (GDP) growth for Turkish economy. GDP growth is one of the most important economic indicators since GDP figures provide comprehensive information regarding the economic activity. GDP data are published with considerable delay, so early estimates of GDP growth may be valuable. For this aim, we use an extended version of the Stock and Watson coincident indicator model that can deal with mixed frequency (such as quarterly and monthly variables), ragged ends (some indicators are published before others), and missing data (data may not be available at the beginning of the sample for some variables). As soft data we use PMI, and as hard data we use industrial production, import and export quantity indices. We perform simulated out of sample forecasting exercise by taking the flow of data releases for 2008Q1-2012Q2 into account. Results show that nowcasts obtained with a model including a soft indicator tracks the GDP growth relatively successfully. Also, the model outperforms benchmark AR model.

Key Words: Time Series, Forecasting, Output Growth

Jel Classification: C22, C53, E37

[1] The views expressed are those of the authors and should not be attributed to the Central Bank of Turkey. We thank an anonymous referee for helpful comments and suggestions. cagri.akkoyun@tcmb.gov.tr, mahmut.gunay@tcmb.gov.tr.

1 Introduction

Gross Domestic Product (GDP) is one of the most comprehensive indicators regarding economic activity; hence GDP growth has the potential to provide important information in the policy making process. However, GDP figures are announced with considerable delay. For example, in Turkey GDP data for the last and first quarters of the year are published approximately 90 days after the end of the quarter and second and third quarter figures are released around 70 days after the end of the quarter. Accordingly, early estimates of the GDP growth may be valuable in assessing the state of the economy. In this study, we use a small scale dynamic factor model that can deal with issues such as *missing data* (data may not be available at the beginning of the sample for some indicators), *ragged ends* (some data may be announced earlier than others), and *mixed frequencies* (for example, GDP is quarterly while industrial production is monthly). We generate backcasts and nowcasts[2] for Turkish GDP growth for the period 2008Q1-2012Q2 by using some combination of hard indicators such as industrial production, export and import quantity indices and soft indicators such as PMI and PMI new orders.[3] We find that inclusion of soft indicators significantly improves the nowcasting performance of the model.

Hard data, such as industrial production or imports, carry a lot of information about GDP growth. However, hard data are also announced with considerable delay. For instance, industrial production and import quantity indices for a given month are announced around 40 days after the end of that month. In contrast, soft indicators such as confidence indicators for a given month become available earlier than many hard indicators. Moreover, surveys may carry information regarding future (such as new orders) That is why, effective use of surveys may help to learn about the state of the economy in a timely manner. An example of soft data that can be used in this context is Purchasing Managers Index (PMI). Koenig (2002) notes that PMIs are timely due to the fact that they become available on the first business day of the following month for a given month, and they are subject to small revisions. These properties make the PMIs appealing candidates for nowcasting, consequently PMIs are commonly used in the nowcasting models. For instance, Lahiri and Monokroussos (2011) study the role of PMIs in nowcasting US GDP growth using a large scale factor model. Lombardi and Maier (2011) use a dynamic factor model and

[2] The term backcast corresponds to the estimation of GDP growth of a quarter by using the information after the end of that quarter. Similarly, nowcast corresponds to the estimation of GDP growth of a quarter by using the information set within that quarter. See Section 2 for detailed explanation.

[3] The Purchasing Managers' Index^{TM} (PMI^{TM}) is a composite index based on five of the individual indexes with the following weights: New Orders : 0.3, Output : 0.25, Employment : 0.2, Suppliers' Delivery Times: 0.15, Stock of Items Purchased :0.1. A value above (below) 50 signals expansion (contraction).

a simple model including PMI. They find that both models outperform a benchmark AR model and dynamic factor model tends to outperform PMI model.

The methodology used in our paper is named as *small scale dynamic one factor* model in literature. This approach can be considered as an extension of Stock and Watson (1991) single index model. In their seminal work, Stock and Watson (1991) present an alternative measure of economic activity to *The Index of Coincident Economic Indicators* calculated by simply weighting the key economic indicators and it is announced by the U.S. Department of Commerce. Their model depends on the idea that movements in many macroeconomic variables have a common unobserved component reflecting the state of the economy apart from idiosyncratic component. They use Kalman filter to obtain the common component of four key variables: industrial production, real personal income, real manufacturing and trade sales, and employment. This model is dynamic since common component depends on its lagged values through the transition equation of the Kalman filter. They find that common component is highly correlated with *The Index of Coincident Economic Indicators*. Although their methodology provides some econometric foundations to the notion of economic indicator, it cannot use variables with different frequencies. This prevents use of quarterly variables such as GDP in calculation of economic indicator. Mariano and Murasawa (2003) use a clever trick and extend the model of Stock and Watson (1991) so as to incorporate both quarterly and monthly data. Their new economic indicator captures the NBER recession dates. Also, the levels of their new economic indicator is economically interpretable. Camacho and Perez-Quiros (2010) follow the methodology in Mariano and Murasawa (2003) to overcome mixed frequency and ragged end problems. They produce nowcasts for Euro Area GDP growth by using soft data [4] as well as hard data [5]. Their nowcasts are compatible with the nowcasts of professional forecasters. Moreover, nowcasts can be updated after each data announcement providing a quantitative assessment to contribution of arriving information on the nowcasts.

All three models above fall in the category of small scale models since they use small number of economic variables in their model. There are also models using many variables and these models are named as large scale models. One important paper in this category is Giannone et al. (2008). They generate GDP nowcasts for U.S. economy by using around 200 macroeconomic indicators. Their methodology combines the principle component analysis and Kalman filter. First, they obtain two common factors

[4] Belgium Overall Business Indicator, Euro-Area Economic Sentiment Indicator, Germany IFO Business Climate Index, PMI Manufacturing and PMI Services.

[5] Euro-Area Industrial Production Index, Euro-Area Total Retail Sales Volume, Industrial New Orders Index, Euro-Area Employment and Euro-Area GDP.

from macroeconomic indicators. In the second step, they smooth these common factors by using Kalman filter. Afterwards, they use smoothed common factors as explanatory variables in the simple OLS to produce nowcasts for GDP. Alvarez et al. (2012) compare the nowcasting performance of small and large scale models.[6] By using both simulated data and U.S. data, they conclude that small scale models outperform large scale models.

In our paper, we follow the methodology proposed by Mariano and Murasawa (2003) to deal with mixed frequency and ragged ends. As in Camacho and Perez-Quiros (2010) and Giannone et al. (2008), we use soft data, for our case PMI and PMI New Orders. We evaluate our models by their pseudo out of sample performance for the post 2008 period. We find that using PMIs in the models improves the nowcasting relative to a simple AR model and also relative to a base model that only includes hard data. Moreover, during the global financial crisis, model provides accurate early signals about the magnitude of decline and increase in GDP.

The flow of the paper is as follows. Section 2 defines terms and illustrates the input in estimation process. Section 3 presents the model which is modified version of Kalman filter. Section 4 gives information about the selection of variables. Section 5 demonstrates the results and Section 6 concludes.

2 Illustration

First of all, we introduce the terminology used in the paper. We can use a general name and denote the estimates for the GDP growth of a quarter as forecasts. On the other hand, considering the fact that we can produce estimates for the GDP growth of a given quarter at different times of the year such as before the quarter, during the quarter and after the quarter, it may be helpful to differentiate the terms used for these estimates. Therefore, we follow the literature and use different names for estimates produced at different times with respect to the reference quarter. As an example suppose that one produces estimates for the first three quarters of 2012 in the middle of May 2012 (Table 2.1). GDP figures are published with considerable delay so that in the middle of the second quarter, GDP data are not available for the first quarter. Since the first quarter already ended (but GDP data are not announced yet), we name estimates for that quarter as *back*cast, the estimate for the second quarter made in the second quarter is named as *now*cast and estimates for all the quarters after the second quarter are named as *fore*casts. As this is a recursive process, in the middle of the third quarter, estimates for the second quarter GDP growth will

[6] Boivin and Ng (2006) and Camacho and Peres-Quiros (2010) discuss the possible flaws in large scale models.

be named as backcast, while estimates for the third quarter growth will be named as nowcast.

Table 2.1. Terminology for the Estimates of a Quarter's GDP and Data Availability

	Jan-12	Feb-12	Mar-12	Apr-12	May-12	Jun-12	Jul-12	Aug-12	Sep-12
Available GDP data in the middle of the given month	2011Q3	2011Q3	2011Q3	2011Q4	2011Q4	2011Q4	2012Q1	2012Q1	2012Q2
Terminology	BACKCAST for 2011Q4			BACKCAST for 2012Q1			BACKCAST for 2012Q2		-
	NOWCAST for 2012Q1			NOWCAST for 2012Q2			NOWCAST for 2012Q3		
	FORECAST for 2012Q2+			FORECAST for 2012Q3+			FORECAST for 2012Q4+		

We get backcast, nowcast and forecast for each quarter but we do not report our findings for forecasts and in fact we concentrate on nowcasting performance of our models. Satisfactory performance for nowcasting GDP growth requires effective use of coincident or leading indicators such as survey data. Surveys are timely and have the potential to provide information about future. For instance, surveys ask questions such as the new orders or expectation of new orders which may have clues about the future path of the economic activity. Although surveys have these nice properties, there will be issue of ragged ends resulting from the fact that they are published before hard data. In other words, at the end of the sample, survey data for a month may be available while hard data may not be available. For example, in the middle of January, information set contains November figures for industrial production, and import and export quantity indices but December data for the PMI (Table 2.2).

Table 2.2 Data Availability in the Middle of a Month

	Jan-12	Feb-12	Mar-12	Apr-12	May-12	Jun-12	Jul-12	Aug-12	Sep-12
GDP	2011Q3	2011Q3	2011Q3	2011Q4	2011Q4	2011Q4	2012Q1	2012Q1	2012Q2
IP	Nov-11	Dec-11	Jan-12	Feb-12	Mar-12	Apr-12	May-12	Jun-12	Jul-12
QM	Nov-11	Dec-11	Jan-12	Feb-12	Mar-12	Apr-12	May-12	Jun-12	Jul-12
QX	Nov-11	Dec-11	Jan-12	Feb-12	Mar-12	Apr-12	May-12	Jun-12	Jul-12
PMI	Dec-11	Jan-12	Feb-12	Mar-12	Apr-12	May-12	Jun-12	Jul-12	Aug-12

Another issue that arises from the use of PMI in our case is the "missing data" problem. PMI data start from May 2005 while hard data goes back to much earlier time. Considering the fact that GDP data start from 1998, missing data for PMI indicates that we will not be able to utilize a lot of data points if we work with a balanced panel data set starting from May 2005. Final issue that we need to tackle is the "mixed frequency" of our data set. GDP is a quarterly variable while other indicators are monthly. There are alternative approaches to dealing with different frequencies of the data. The simplest solution would be working with quarterly averages of the monthly variables with a sample size determined by the shortest data. For example, in January 2012 we may use a quarterly data set starting from the second quarter of 2005 and ending in 2011Q4 by taking averages of monthly variables. However, this may

cause loss of information such as not utilizing pre-May 2005 period. Moreover, December 2011 data for industrial production and quantity indices are not available in January 2012. We need to work with two data points for that quarter for each hard indicator. In summary, task of extracting information content of the available data comes with many challenges. As a matter of fact, we utilize a technique that allows us to work with missing data, ragged ends and mixed frequency to get backcasts/nowcasts.

Table 2.3 shows the structure of our data set that is for the middle of February 2012. We feed the system with this data set that contains both monthly and quarterly variables, including missing data at the beginning of the sample and ragged ends caused by different announcement dates at the end of the sample. In the table, N/A means that data at that point are not observable (either there are no data as in PMI or we use a quarterly variable as in GDP), while "+" means data are available in the information set at the time of estimations and TBA indicates that we will learn the data in the future. GDP is coded in a way that quarterly growth rates appear in the last month of the quarter. Hence note that we coded the last month of the second quarter of 2012 with TBA while entries for April and May are N/A.

Table 2.3. Data Structure

Date	GDP	Industiral Production	Import Quantity Index	Export Quantity Index	PMI New Orders
1998.01	N/A	+	+	+	N/A
1998.02	N/A	+	+	+	N/A
1998.03	N/A	+	+	+	N/A
1998.04	N/A	+	+	+	N/A
.....					
2005.02	N/A	+	+	+	N/A
2005.03	+	+	+	+	N/A
2005.04	N/A	+	+	+	N/A
2005.05	N/A	+	+	+	+
2005.06	+	+	+	+	+
2005.07	N/A	+	+	+	+
.....					
2011.09	+	+	+	+	+
2011.10	N/A	+	+	+	+
2011.11	N/A	+	+	+	+
2011.12	TBA	+	+	+	+
2012.01	N/A	TBA	TBA	TBA	+
2012.02	N/A	TBA	TBA	TBA	TBA
2012.03	TBA	TBA	TBA	TBA	TBA
2012.04	N/A	TBA	TBA	TBA	TBA
2012.05	N/A	TBA	TBA	TBA	TBA
2012.06	TBA	TBA	TBA	TBA	TBA

Notes: + means we observe data, N/A means data are not observale, TBA means data will be observed.

3 Methodology

In this section, we demonstrate the state space representation of our model based on the works of Mariano and Murasawa (2003), and Camacho and Quiros (2010). First, we discuss how the model deals with the non-linearity stemming from mixed frequencies. Afterwards we explain the Kalman filter process in the presence of missing observations and ragged ends.

3.1 Model with Mixed Frequencies

Let $Y_{1,t}$ is the quarterly series observed at t=1,4,7,... and $\mathbf{Y}_{2,t}$ is the vector of n monthly indicators observed at t=1,2,3,... and assume logs of all series are integrated of order 1. We can represent the $Y_{1,t}$ as the simple average of the latent monthly series $\mathrm{X}_{1,t}$ within the quarter:

$$Y_{1,t} = \frac{1}{3}(X_{1,t} + X_{1,t-1} + X_{1,t-2}) \tag{1}$$

However, this definition leads to a non-linearity in state space representation since we work with log-differences. Mariano and Murasawa (2003) suggest using geometric mean as an approximation to arithmetic mean to deal with non-linearity:

$$\ln Y_{1,t} \approx \frac{1}{3}(\ln X_{1,t} + \ln X_{1,t-1} + \ln X_{1,t-2}) \tag{2}$$

Let us denote the quarterly growth of $Y_{1,t}$ with $y_{1,t} = \ln Y_{1,t} - \ln Y_{1,t-3}$ and monthly growth rates of $X_{1,t}$ with $x_{1,t} = \ln X_{1,t} - \ln X_{1,t-1}$. Then we can rewrite the $y_{1,t}$ as follows:

$$y_{1,t} = \frac{1}{3}(\ln X_{1,t} + \ln X_{1,t-1} + \ln X_{1,t-2}) - \frac{1}{3}(\ln X_{1,t-3} + \ln X_{1,t-4} + \ln X_{1,t-5}) \tag{3}$$

$$\begin{aligned} y_{1,t} \quad = \quad & \frac{1}{3}(\ln X_{1,t} - \ln X_{1,t-1}) + \frac{2}{3}(\ln X_{1,t-1} - \ln X_{1,t-2}) + (\ln X_{1,t-2} - \ln X_{1,t-3})... \\ & + \frac{2}{3}(\ln X_{1,t-3} - \ln X_{1,t-4}) + \frac{1}{3}(\ln X_{1,t-4} - \ln X_{1,t-5}) \end{aligned} \tag{4}$$

$$y_{1,t} = \frac{1}{3}x_{1,t} + \frac{2}{3}x_{1,t-1} + x_{1,t-2} + \frac{2}{3}x_{1,t-3} + \frac{1}{3}x_{1,t-4} \tag{5}$$

The last equation enables representing the growth rate of quarterly variable as a linear combination of the growth rate of latent monthly series. So, we can work with linear state space model. However, if

one considers using arithmetic mean instead of geometric mean, then its model should include non-linear terms that complicates the estimation For example, Aruoba et al. (2009) include polynomial trends of order two in their state space representation to capture the effects that cannot be identified by linear terms.

In state representation, we use standardized quarterly growth rate of GDP,$y_{1,t}$, and standardized monthly growth rates of indicators, $y_{2,t}$. We assume that all variables are combination of common factor h_t and idiosyncratic component $i_{j,t}$ $j \in \{1, ..., n+1\}$. Then the dynamic model can be represented as:

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \end{bmatrix} = \begin{bmatrix} \phi_1(\frac{1}{3}h_t + \frac{2}{3}h_{t-1} + h_{t-2} + \frac{2}{3}h_{t-3} + \frac{1}{3}h_{t-4}) \\ \phi_2 h_t \end{bmatrix} + \begin{bmatrix} \frac{1}{3}i_{1,t} + \frac{2}{3}i_{1,t-1} + i_{1,t-2} + \frac{2}{3}i_{1,t-3} + \frac{1}{3}i_{1,t-4} \\ i_{2,t} \end{bmatrix} \tag{6}$$

We also assume that common factor h_t and idiosyncratic component $i_{j,t}$ $j \in \{1, ..., n+1\}$ follow AR(2) processes. Now the system can be represented in the standard state space form:

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \end{bmatrix} = Hs_t + v_t \ , \ v_t \sim N(0, R) \tag{7}$$

$$s_t = Fs_{t-1} + w_t \ , \ w_t \sim N(0, Q) \tag{8}$$

where the first equation is measurement equation and the second equation is the transition equation, s_t is the state vector, v_t and w_t are the i.i.d. errors. H is $(n+1) \times (2n+10)$ matrix relating the observables with common and idiosyncratic components. F is $(2n+10) \times (2n+10)$ matrix that achieves the transition between states.[7]

3.2 Kalman Filter with Missing Observations

We must deal with the missing observations before estimation processes. For this purpose, we make random draws from the normal distribution N(0,1) and replace these draws with missing observations Since we do not want random draws to effect our estimation process, we redefine the objects in the measurement equation of the Kalman filter where ρ_t is the random draw at time t:

$$y^*_{i,t} \triangleq \begin{Bmatrix} y_{i,t} & \text{if } y_{i,t} \text{ is observable} \\ \rho_t & \text{otherwise} \end{Bmatrix} \tag{9}$$

[7] Appendix A contains a simple example for two monthly indicators.

$$H^*_{i,t} \triangleq \begin{Bmatrix} H_i & \text{if } y_{i,t} \text{ is observable} \\ \mathbf{0}_{1x(2n+10)} & \text{otherwise} \end{Bmatrix} \tag{10}$$

$$v^*_{i,t} \triangleq \begin{Bmatrix} 0 & \text{if } y_{i,t} \text{ is observable} \\ \rho_t & \text{otherwise} \end{Bmatrix} \tag{11}$$

$$R^*_{i,t} \triangleq \begin{Bmatrix} 0 & \text{if } y_{i,t} \text{ is observable} \\ 1 & \text{otherwise} \end{Bmatrix} \tag{12}$$

The above system prevents random draws from effecting the sum of log-likelihood function that will be discussed below.

Now, at the beginning assume that we know the model parameters' vector Π that is composed of the parameters in matrices H and F, and standard deviation σ_w. Kalman filter[8] needs initial values for state s_t and its covariance matrix P_t for starting estimation process. So, we assign zero vector for $s_{0|0}$ and identity matrix for $P_{0|0}$. In the first step, Kalman filter uses the information of the previous state to predict the present state. Let $x_{t|\tau}$ be the estimate of variable x_t at time τ. Then the prediction equations for state s_t and its covariance matrix P_t at time t by using the information at time $t-1$ are :

$$s_{t|t-1} = F s_{t-1|t-1} \tag{13}$$

$$P_{t|t-1} = F P_{t-1|t-1} F' + Q \tag{14}$$

Now the prediction error is $\varepsilon_{t|t-1} = y^*_t - H^*_t h_{t|t-1}$ and the associated covariance matrix $\sigma_{t|t-1} = H^*_t P_{t|t-1} H^{*\prime}_t + R^*_t$. The associated log-likelihood of the prediction error:

$$\lambda_t = -\frac{1}{2}\ln(2\pi|\sigma_{t|t-1}|) - \frac{1}{2}\varepsilon'_{t|t-1}(\sigma_{t|t-1})^{-1}\varepsilon_{t|t-1} \tag{15}$$

After estimation we update s_t and P_t by updating equations:

$$s_{t|t} = s_{t|t-1} + K^*_t \varepsilon_{t|t-1} \tag{16}$$

$$P_{t|t} = P_{t|t-1} - K^*_t H^*_t P_{t|t-1} \tag{17}$$

[8] Hamilton (1994) includes detail information and derivations for Kalman filter.

where $K_t^* = P_{t|t-1} H_t^{*\prime} (\sigma_{t|t-1})^{-1}$ is the Kalman gain.

In the second step, we find the model parameters Π that minimizes the sum of log-likelihood function[9] i.e. $\sum_{t \in T} \lambda_t$. If $y_{i,t}$ is missing at time t, then its contribution to log-likelihood λ_t will be function of ρ_t which is independent of model parameters So, random draws do not effect the estimation of parameter vector Π and Π is only effected by the observable monthly indicators at period t and we update our state, i.e. $s_{t|t} \neq s_{t|t-1}$. If we observe no variables $y_{i,t}$ $i \in \{1, ..., n+1\}$at time t, then $H_t^* = 0$ implying that $K_t^* = 0$. So, we cannot adjust the state with measurement equation i.e. $s_{t|t} = s_{t|t-1}$. Also in such a case our estimate for parameter vector Π does not change. Afterwards, we can obtain backcasts and nowcasts by using the optimal Π and the initial state $s_{0|0} = 0$.[10]

4 Data and Selection of Indicators

Potential number of series that can be used in nowcasting GDP growth is very large. In their seminal paper, Stock and Watson (1991) use industrial production, total personal income less transfer payments, total sales and employment. The rationale for using these indicators is that GDP can be calculated with three approaches, namely production approach, income approach and expenditure approach. They choose key indicators for each of these approaches and employment as an additional variable.

Unfortunately, not all of these variables, like sales or personal income, are available for the Turkish economy for using in the nowcasting context. Therefore, we choose a different set of indicators as the base model which includes industrial production, import quantity and export quantity indices[11]. Analysis of the contribution of production approach components to the GDP shows that contributions of industry and some of the components are very similar (Figure B.1.a and Figure B.1.b). Therefore, even though share of industry is around 25 percent of GDP (in real terms), it is an important indicator due to its link with other sectors in the economy. Imports are informative about demand conditions like consumption and investment expenditures. Although correlation of exports with GDP is weaker than correlations of imports and industrial production we use exports in the base model to better capture the developments in global economy (albeit with a shorter sample). We provide quarterly percentage change of these indicators and GDP in the Figure B.2.a-c and we provide the correlations in Table B.1.

[9] We use swarm algorithm in MATLAB for parameter estimation. The codes are avaliable upon request.

[10] Also, we can take forecasts by simply adding missing observations for all variables $y_{i,t}$ $i \in \{1, ..., n+1\}$ at the end of the sample.

[11] We use import and export quantity indices excluding gold for post 2003 period.

We can utilize soft data to improve nowcasting performance since soft data are published earlier than hard data. Moreover, these data may provide signals about future as they may provide information about perceptions, confidence and expectations. There are many candidates to be used as soft data such as PMI, Business Tendency Survey (BTS) indicators like expectation of new orders over the next three months; consumer confidence or capacity utilization. Structure of BTS changed substantially in the 2007 as the sample size increased and some questions changed. We worked by combining old and new surveys for some questions to use BTS from 1998 and alternatively we used new BTS so that data covers post 2007. However, we concluded that it could be better to report findings with BTS in the future when we have more data for the new BTS. PMI starts earlier than new BTS and it is commonly used in nowcasting models for other economies hence as the soft indicator we work with PMI. Set of indicators that can be used in this framework is not limited with survey data. In this respect, we experimented with financial indicators, global variables, fiscal variables and labor market indicators. In this paper we estimate a parsimonious model that includes four variables. In addition to the PMI (total index), we also analyze PMI New Orders sub-index. Table B.2 shows the correlation of PMI and PMI New Orders with GDP growth. Variables used in the analysis are described in Table 4.1.

We use data until mid August 2012 and we check for stationarity of the data using ADF.

Table 4.1: Variables

Series	Frequency	Sample	Source	Publication Delay[12]	Data Transformation
Gross Domestic Product (GDP)	Quarterly	1998.Q1-2012.Q2	TURKSTAT	70-90 days	SA[13], Quarterly Growth
Industrial Production (IP)	Monthly	1998.M01-2012.M06	TURKSTAT	39 days	SA, Monthly Growth
Import Quantity Index (QM)	Monthly	1998.M01-2012.M06	TURKSTAT	41 days	SA, Monthly Growth
Export Quantity Index (QX)	Monthly	2003.M01-2012.M06	TURKSTAT	41 days	SA, Monthly Growth
PMI Indices	Monthly	2005.M05-2012.M07	MARKIT	1 day	SA, Level

[12] approximate delays

[13] SA: Seasonally and working day adjusted, also where necessary outlier correction is also done. Seasonal adjustment is done with Tramo Seats using Demetra.

5 Results

5.1 Simulated Real-Time Analysis

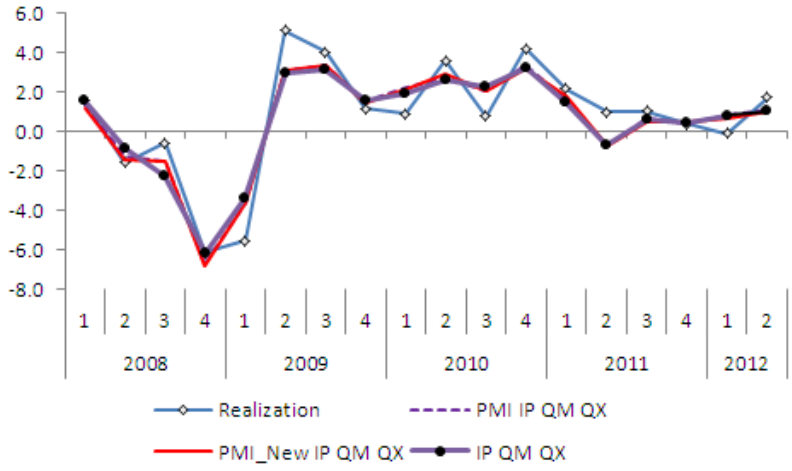
If a model is planned to be used for nowcasting in the future, it is essential to understand the past performance of the model in an out of sample evaluation context. For example, consider that we report our nowcasts from a model for the second quarter of 2012. Then a natural question is how the performance of the nowcasts has been for, say, the last two years. It might be the case that model cannot even capture the sign of the GDP growth in the last two years so that reported nowcast for 2012Q2 may not be very reliable. As a matter of fact, it is important to test the performance of the model visually with a graph showing the nowcasts obtained for the past quarters and realizations. It is also important to test the performance formally by comparing the model with some benchmark to see whether the model or new variables improve the performance, say with respect to a simple Auto Regressive model. Once it is shown that nowcasts were reasonable in the past, nowcasts of the model might be used with more confidence in the future. We choose 2008Q1-2012Q2 period as the evaluation sample. During the financial crisis many series moved together, so we want to see the performance of the model before the global financial crisis. PMI starts from May 2005 so it is not desirable to go too far in the pre-crisis period, hence we start our evaluation sample from 2008. Final data point for the GDP, when we write the paper, is 2012Q2.

Out of sample evaluation exercise requires replicating the data structure for each of the past quarters that we do our evaluation. Suppose in August 2012, we want to analyze the performance of the model in the middle of 2008Q1. For that aim, we estimate the factor using information that would have been available in the middle of 2008Q1 such as industrial production for the December 2007 and PMI January 2008. Then, using the factor, factor loadings and AR parameters, we get nowcast of the model for 2008Q1. Recursively iterating this process by extending information set at each step we can get the nowcasts for each quarter in the evaluation period. After getting nowcasts, we calculate Root Mean Squared Errors (RMSE) of the nowcasting errors to compare the performance of different models.

Ideally, we should work with *real time data* as suggested by Croushore and Stark (2001). Using real time data enables the researcher to replicate exactly what was available in a given time to a forecaster. However, we do not have a real time data set; therefore we do a simulated real-time analysis where we use revised data. In a simulated real time analysis one uses final data (revised data) but when doing nowcasting analysis for a quarter, information availability is taken into account (such as using information only up to mid 2008Q1 while nowcasting 2008Q1).

We start by presenting the results for backcasts obtained from three models. As stated in the Section 2, we can label each estimate after the end of the reference quarter as backcast. In the Figure 5.1 we only present the backcasts obtained 45 days after the end of the reference quarter. Note that, all hard data used in the paper become available approximately 40 days after a reference quarter. As a result, backcasts presented in Figure 5.1 use full information for the reference quarter for hard data used in the paper. We observe that adding a soft indicator does not change the performance too much. Accordingly, we conclude that contribution of soft data may be very limited for backcasting .

Figure 5.1. Backcasting

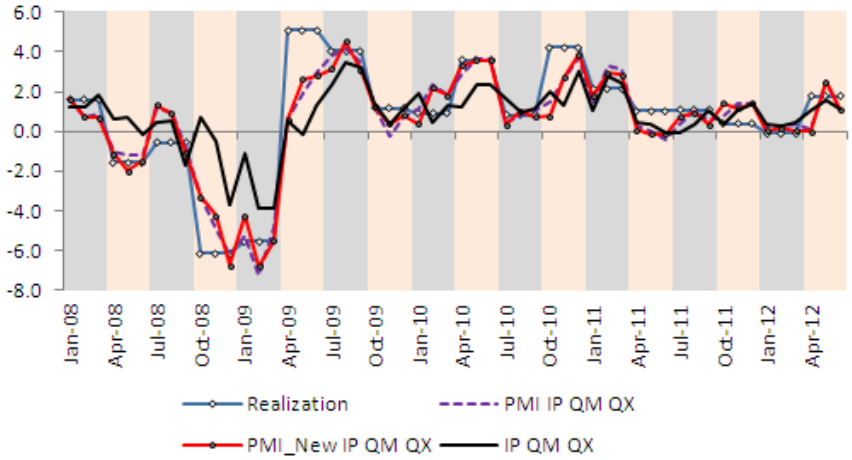


An issue regarding the presentation of nowcasting results is the detail of the analysis of nowcasts. We can produce nowcasts for any day in a given quarter by using what data would be available at that day or we can only report the nowcasts obtained by simulating the available information in the middle of each quarter. Yet, in the paper we take a middle ground and evaluate our models by producing nowcasts with available information in the middle of each of the three months of a quarter. Table 2.2 and Table 4.1 describe in detail the how the information set is constructed for each nowcasting round.

Visual inspection of nowcasts presented in Figure 5.2 shows that using soft data substantially improves nowcasting performance, especially in the periods with abrupt changes in GDP growth. It can be stated that during the global financial crisis our model produces timely and accurate nowcasts for GDP growth. For instance, in the middle of October 2008 our model suggests that GDP might decline by more than 3 percent, in November 2008 nowcasts suggest that decline in GDP might reach to 5 percent, and as of December 2008 model suggests that contraction might be more than 6 percent. Indeed, decline in

GDP was around 6.2 percent in the last quarter of 2008. We observe that model tracks growth relatively successfully for most of the quarters.

Figure 5.2. Nowcasting



Notes: Quarters are shaded by changing colors. Example: For the January 2008; February 2008 and March 2008 grey shaded area shows the nowcasts for 2008Q1 while April to June 2008 points which are shaded by a different color than grey show the nowcasts for 2008Q2

Our method enables us to evaluate the information content of the arriving information. We can check the nowcasting errors after each data release to see whether a specific indicator contributes more to decline in nowcasting errors. It might be the case that an indicator is useful for nowcasting in the first half of the quarter, while second half figures for that variable may be relatively more informative for forecasting next quarter's GDP. For example, if we are using a forward looking financial variable, its contribution to nowcasting might be higher at the beginning of the quarter while towards the end of the quarter that variable may improve forecasting. In this respect, we present RMSE relative to AR for nowcasts calculated at the middle of each month of a quarter for different models in Table.5.1. It is observed from the table that as new information arrives errors decline and we get the best performance in the third month.

Table 5.1: RMSE w.r.t. AR for the Nowcasts

Model	1st Month	2nd Month	3rd Month
IP QM QX	0.83	0.71	0.40
IP QM QX PMI	0.53	0.42	0.28
IP QM QX PMI_NEW	0.57	0.37	0.28

After analyzing the backcasting and nowcasting performance separately, we now compare backcasts and nowcasts since in both cases RMSE for the benchmark AR models are close to each other. Backcasts produced 45 days after the end of a reference quarter and nowcasts produced in the middle of the reference quarter for 2008Q1-2012Q2 are compared in Table 5.2 relative to their AR benchmarks. It can be seen from the table that backcast and nowcast errors are close to each. This is striking considering the fact that information set of nowcasts is relatively limited. We conclude that inclusion of soft data makes nowcast performance compatible with backcast performance and our model produces timely estimates of GDP growth in the evaluation sample.

Table 5.2: Prediction Accuracy (RMSE w.r.t. AR)

Model	Backcast	Nowcast
IP QM QX	0.39	0.71
IP QM QX PMI	0.36	0.42
IP QM QX PMI_NEW	0.35	0.37

Finally, we report factor loadings which show the relation of each variable with the common factor. Factor loadings[14] for the model with PMI New Orders as of 2012Q2 and t-values are presented in Table 5.3. Loadings are positive and statistically significant which implies that indicators are pro-cyclical with the GDP.

Table 5.3: Factor Loadings

GDP	IP	QM	QX	PMI_NEW
0.35	0.55	0.53	0.21	0.51
(5.68)	(5.72)	(5.68)	(3.96)	(3.11)

6 Conclusion

This study generates backcasts and nowcasts for GDP growth based on the idea of single-index dynamic factor model first proposed by Stock and Watson (1991). Our model considers the recent modifications for handling missing observations, ragged ends and mixed frequency. We present the results of a model where we use industrial production, import and export quantity indices as hard data and PMI as soft data.

[14] loadings correspond to γ_i for the system in Appendix.A.

Results show that inclusion of PMI to a model with hard indicators substantially improves nowcasting performance. The improvement becomes particularly evident during the recent global crisis. We evaluate nowcasts on a monthly basis. We observe that as new information arrives, nowcasting error declines. We also observe that although information set is limited for nowcasts of a reference quarter, nowcasting and backcasting performances are compatible.

There are many other survey indicators and financial variables that can be used in the context of our method. In the future, more data will be available for some survey data (such as new Business Tendency Survey) that have a short sample. Therefore, active and continuous monitoring of the performance of soft indicators is a promising area for further research.

References

[1] Alvarez, R., Camacho, M., and Pérez-Quirós, G. (2012). "Finite sample performance of small versus large scale dynamic factor models". CEPR Discussion Papers, 8867.

[2] Aruoba, B., Diebold, F., and Scotti, C. (2009). "Real-Time Measurement of Business Conditions". Journal of Business and Economic Statistics, 27: 417-427.

[3] Boivin, J., and Ng, S. (2006). "Are more data always better for factor analysis?" Journal of Econometrics, Elsevier, vol. 132(1), pages 169-194, May.

[4] Camacho, M., and Perez Quiros, G. (2010). "Introducing the Euro-STING: Short Term INdicator of Euro Area Growth". Journal of Applied Econometrics, 25: 663–694.

[5] Croushore, D., and Stark, T. (2001). "A real time data set for macroeconomists". Journal of Econometrics 105: 111-130.

[6] Giannone, D., Reichlin, L., and Small, D. (2008). "Nowcasting: The real-time informational content of macroeconomic data". Journal of Monetary Economics 55: 665-676.

[7] Hamilton, J. (1994)." State-space models. Handbook of Econometrics", Volume 4. Edited by R.Engle and D. McFadden, North-Holland.

[8] Koenig, E.F., (2002). "Using the Purchasing Managers' Index to assess the economy's strength and the likely direction of monetary policy". Economic and Financial Policy Review, Federal Reserve Bank of Dallas.

[9] Lahiri, K., and Monokroussos, G., (2011). "Nowcasting US GDP: The role of ISM Business Surveys". Discussion Papers 11-01, University at Albany, SUNY, Department of Economics.

[10] Lombardi, M.J., and Maier, P. (2011). "Forecasting economic growth in the euro area during the great moderation and the great recession" Working Paper Series 1379, ECB.

[11] Mariano, R., and Murasawa, Y. (2003). "A new coincident index of business cycles based on monthly and quarterly series" Journal of Applied Econometrics 18: 427-443.

[12] Stock, J., and Watson, M. (1991). "A probability model of the coincident economic indicators" In Kajal Lahiri and Geoffrey Moore editors, Leading economic indicators, new approaches and forecasting records. Cambridge University Press, Cambridge.

7 Appendix

7.1 Appendix A.

The simple example for two monthly indicator case (n=2) has the state space representation:

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \\ y_{3,t} \end{bmatrix} = \begin{bmatrix} \frac{1}{3}\gamma_1 & \frac{2}{3}\gamma_1 & \gamma_1 & \frac{2}{3}\gamma_1 & \frac{1}{3}\gamma_1 & \frac{1}{3} & \frac{2}{3} & 1 & \frac{2}{3} & \frac{1}{3} & 0 & 0 & 0 & 0 \\ \gamma_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ \gamma 3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} h_t \\ h_{t-1} \\ h_{t-2} \\ h_{t-3} \\ h_{t-4} \\ i_{1,t} \\ i_{1,t-1} \\ i_{1,t-2} \\ i_{1,t-3} \\ i_{1,t-4} \\ i_{2,t} \\ i_{2,t-1} \\ i_{3,t} \\ i_{3,t-1} \end{bmatrix} + \begin{bmatrix} v_{1,t} \\ v_{2,t} \\ v_{3,t} \end{bmatrix} \tag{18}$$

$$
\begin{bmatrix} h_t \\ h_{t-1} \\ h_{t-2} \\ h_{t-3} \\ h_{t-4} \\ i_{1,t} \\ i_{1,t-1} \\ i_{1,t-2} \\ i_{1,t-3} \\ i_{1,t-4} \\ i_{2,t} \\ i_{2,t-1} \\ i_{3,t} \\ i_{3,t-1} \end{bmatrix}
=
\begin{bmatrix}
\phi_1 & \phi_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \theta_1 & \theta_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \theta_3 & \theta_4 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \theta_5 & \theta_6 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0
\end{bmatrix}
\begin{bmatrix} h_{t-1} \\ h_{t-2} \\ h_{t-3} \\ h_{t-4} \\ h_{t-5} \\ i_{1,t-1} \\ i_{1,t-2} \\ i_{1,t-3} \\ i_{1,t-4} \\ i_{1,t-5} \\ i_{2,t-1} \\ i_{2,t-2} \\ i_{3,t-1} \\ i_{3,t-2} \end{bmatrix}
+
\begin{bmatrix} w_{1,t} \\ 0 \\ 0 \\ 0 \\ 0 \\ w_{2,t} \\ 0 \\ 0 \\ 0 \\ 0 \\ w_{3,t} \\ 0 \\ w_{4,t} \\ 0 \end{bmatrix}
\tag{19}
$$

The estimated parameter in the Kalman filter process are $\gamma_1, \gamma_2, \gamma_3, \phi_1, \phi_2, \theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6, \sigma_1, \sigma_2, \sigma_3$.

7.2 Appendix B

Figure B.1.a. Share in the GDP

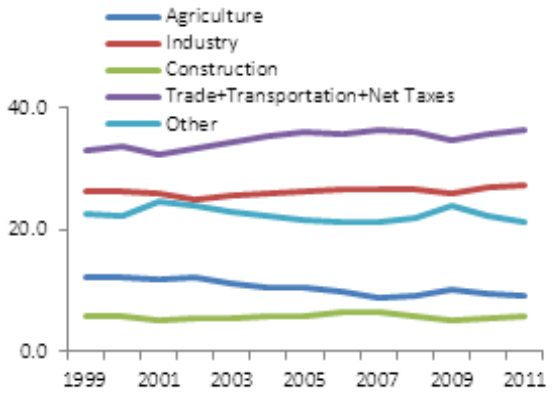


Figure B.1.b. Contribution to GDP Growth

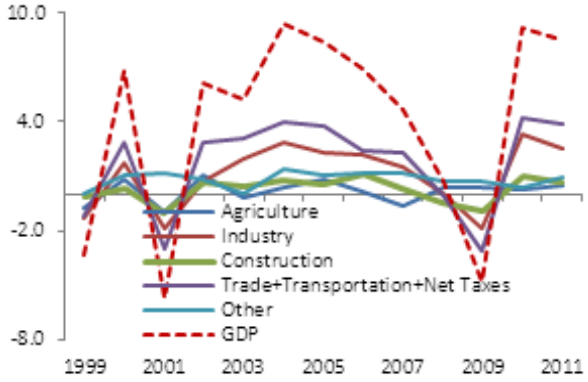


Figure B.2.a. Industrial Production and GDP

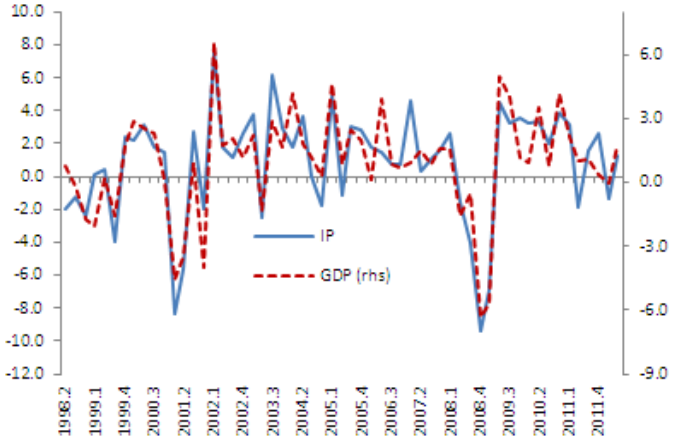


Figure B.2.b. Import Quantity Index and GDP

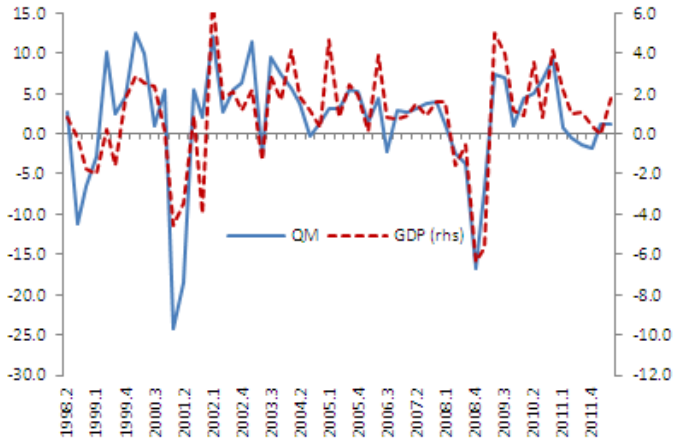


Figure B.2.c. Export Quantity Index and GDP

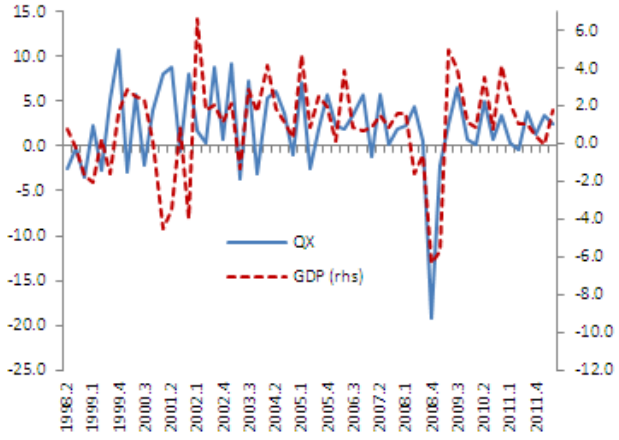


Table B.1.: Correlations of Quarterly Growth in the 1998-2012 and 2003-2012 Periods

1998Q2-2012Q2				
	QM	QX	IP	GDP
QM	1.00	0.20	0.79	0.75
QX	0.20	1.00	0.32	0.30
IP	0.79	0.32	1.00	0.87
GDP	0.75	0.30	0.87	1.00
2003Q2-2012Q2				
	QM	QX	IP	GDP
QM	1.00	0.67	0.82	0.84
QX	0.67	1.00	0.66	0.69
IP	0.82	0.66	1.00	0.85
GDP	0.84	0.69	0.85	1.00

Table B.2.: Correlation of PMI and GDP Growth (2005Q3-2012Q2)

	PMI	PMINEW	GDP
PMI	1.00	0.99	0.86
PMINEW	0.99	1.00	0.90
GDP	0.86	0.90	1.00

Central Bank of the Republic of Turkey
Recent Working Papers
The complete list of Working Paper series can be found at Bank's website
(http://www.tcmb.gov.tr).

Rezerv Opsiyonu Mekanizması ve Optimal Rezerv Opsiyonu Katsayılarının Hesaplanması
(Doruk Küçüksaraç, Özgür Özel Çalışma Tebliği No. 12/32, Kasım 2012)

Finansal Krizlerin Belirleyicileri Olarak Hızlı Kredi Genişlemeleri ve Cari İşlemler Açığı
(Aytül Ganioğlu Çalışma Tebliği No. 12/31, Kasım 2012)

On the Sources and Consequences of Oil Price Shocks: The Role of Storage
(Deren Ünalmış, İbrahim Ünalmış, Derya Filiz Ünsal Working Paper No. 12/30, November 2012)

Mitigating Turkey's Trilemma Tradeoffs
(Yasin Akçelik, Orcan Çörtük, İbrahim M. Turhan Working Paper No. 12/29, October 2012)

Capital Regulation, Monetary Policyand Financial Stability
(Pierre-Richard Agénor, Koray Alper, L. Pereira da Silva Working Paper No. 12/28, October 2012)

Determinants of Precautionary Savings: Elasticity of Intertemporal Substitution vs. Risk Aversion
(Arif Oduncu Working Paper No. 12/27, August 2012)

An Analysis of Intraday Patterns and Liquidity on the Istanbul Stock Exchange
(Bülent Köksal Working Paper No. 12/26, August 2012)

Short Run Import Dynamics in Turkey
(Altan Aldan, İhsan Bozok, Mahmut Günay Working Paper No. 12/25, August 2012)

Required Reserves as a Credit Policy Tool
(Yasin Mimir, Enes Sunel, Temel Taşkın Working Paper No. 12/24, August 2012)

Are Swap and Bond Markets Alternatives to Each Other in Turkey?
(Murat Duran, Doruk Küçüksaraç Working Paper No. 12/23, August 2012)

Health Expenditures Risk, Purchase of Private Health Insurance, and Precautionary Saving in Turkey
(Evren Ceritoğlu Working Paper No. 12/22, August 2012)

Türk İmalat Sanayi Sektörel Reel Efektif Döviz Kuru Endeksleri Üzerine Bir Değerlendirme
(Hülya Saygılı, Gökhan Yılmaz Çalışma Tebliği No. 12/21, Haziran 2012)

Home Production and the Optimal Rate of Unemployment Insurance
(Temel Taşkın Working Paper No. 12/20, June 2012)

Türkiye İçin Bir Reel İktisadi Faaliyet Göstergesi
(S. Borağan Aruoba, Çağrı Sarıkaya Çalışma Tebliği No. 12/19, Haziran 2012)

Using Google Search Index to Nowcast Unemployment Rate: Evidence from Turkey
(Meltem Gülenay Chadwick, Gönül Şengül Working Paper No. 12/18, June 2012)

Küresel Kriz Sonrası Para Politikası
(A. Hakan Kara Çalışma Tebliği No. 12/17, Haziran 2012)