Probability Forecasts of Macro Aggregates in Turkish Economy

Hüseyin Kaya
Bahcesehir University

M. Ege Yazgan*
Istanbul Bilgi University

November 2, 2012

Abstract

In this paper we provide probability forecasts of key Turkish macro economic variables such as inflation and output growth. Some probabilistic forecasts of different scenarios associated to those variables are also calculated. The probability forecasts take different types of uncertainties regarding future, model and structural breaks into account and are derived from a core vector error correction model of Turkish Economy and its several variants. Model and window averaging are used to address uncertainties arising from estimated models and possible structural breaks. The performance of different models and their combinations are evaluated by using relevant forecast accuracy tests in different pseudo out of sample settings. The results indicate that successful directional forecasts can be obtained for output growth and inflation. Both averaging over models and estimation windows improve the level of accuracy of forecasts. Some successful scenario forecasts associated to output growth are also presented.

JEL Classifications: C32, C53, E17

Key Words: Probability forecasts; Forecast combinations; Forecasting and structural breaks, Turkish economy

*Corresponding author. E-mail:ege.yazgan@bilgi.edu.tr

1 Introduction

Forecasting is so pervasive in almost every area of economics. Given its practical importance in policy and decision making it has received a considerable attention from policy makers, market professional and academics.[1]

As emphasized by Gneiting (2008), among others, forecasts are ought to be probabilistic in nature. However, empirical macro studies forecasts are usually presented in the form of point forecasts and their uncertainty is characterized by forecast confidence intervals if considered. As emphasized by Garratt et al. (2003a) and Garratt et al. (2006), point forecasts are reliable only when the decision problem is linear in constraint and quadratic in the loss function (LQ form). Hence, since many decision problems in economics may not be in LQ form, probability forecasts serve useful. As discussed in Diebold and Lopez (1995) and Gneiting (2008), there are other good reasons to believe that why probability forecasts can be beneficial and, therefore may become increasingly prominent in economic applications. For example, in its inflation report Bank of England has been issuing density forecast of inflation and output growth, called "Fan Charts", since 1997.[2]

Despite their importance the usage of probability forecasts in applied works are not so widespread and point forecasts continue to be dominant form of formulating predictions about the future of economic variables. In this paper we aim to contribute to the applied literature on probability forecasts by providing an application in the context of Turkish Economy.[3] To generate the probability forecasts of Turkish macro variables, following Garratt et al. (2003a), Garratt et al. (2003b), Garratt et al. (2006), Assenmacher-Wesche and Pesaran (2008),[4] the paper adopts a long-run structural modelling approach and develops a vector error correction model for Turkish Economy. We make use some variants of this model for the purpose of model averaging and achieve better forecasting performance in combination with window averaging.

[1] The theoretical and empirical literature on forecasting is voluminous. See Elliott and Timmermann (2008) for a review.

[2] For a detailed discussion and evaluation of Bank of England's probability forecasts see Britton et al. (1998) and Clements (2004).

[3] To the best of our knowledge this paper constitutes the first probability forecast exercise using Turkish macroeconomic data

[4] Fair (1980) proposes a pioneering work for estimating the uncertainty of a forecast from an macroeconometric model by considering future uncertainty, parameter uncertainty, model uncertainty and uncertainty due to the exogenous-variable forecast.

Forecast in economics are usually carried out by using time series models. Model based forecasts are subject to a number of uncertainties that make difficult to predict the future values of the economic variables. In this paper, we attempt to address some of these uncertainties by combining model and window averaging in the context of probability forecasts. Future uncertainty, which refers to the unobserved future shocks on forecasts, is tried to be controlled by using probability forecasts.

Model uncertainty, in general, arises from the fact that no model can be able to capture all features of the data generating process under consideration. Therefore, to allow this type of uncertainty we use forecast combination approach. Since the seminal work of Bates and Granger (1969) combining forecasts of different models, instead of relying on forecasts of individual models, has come to be viewed as an effective way of improving the accuracy of predictions regarding a certain target variable. A significant number of theoretical and empirical studies, e.g. Timmermann (2006) and Stock and Watson (2004) have been able to show the superiority of combined forecasts over single-model based predictions.

Another type of future uncertainty is related to changes in some features of data generation process, namely structural breaks. To reduce the undesirable effects of structural breaks on forecasting performance we introduce windows averaging and estimate models over different estimation windows and pool forecasts from different sample periods as suggested by Pesaran and Timmermann (2007) and Pesaran and Pick (2012). Finally, as in Assenmacher-Wesche and Pesaran (2008) and Pesaran et al. (2009) we also consider both model and window averaging together to obtain a total average of forecasts.[5]

The rest of the paper is organized as follow. Section two describes data and its time series properties. Section three introduces a cointegrating VAR-X model for Turkey. Section four tests for the existence of the long-run relations embedded in the model. Section five explains the derivation of the

[5]The other type of uncertainties that are expected to affect forecast performance are those of parameter, policy and measurement. Parameter uncertainty is concerned with robustness of forecast to the selection (estimation) of parameters for a given model. A method for dealing with parameter uncertainty can be easily incorporated into our framework as we comment below. However, handling with policy and measurement uncertainty (data inadequacies and measurement errors) is not as straightforward as in the case of parameter uncertainty and requires using other modelling approaches than the one used here.

probability forecasts. Section six outlines model uncertainty and structural breaks are taken into account while forecasting with the model. Section seven provides some results of directional forecasts generated by the model and explain their evaluation within a pseudo out of sample framework. Section eight presents the probability forecasts of some events associated to inflation and growth for the recent period of Turkish Economy and confront it with the real outcome. Section nine concludes.

2 Data Description and Time Series Properties

Our model consists of 4 domestic and 3 foreign variables. The domestic variables are the Turkish output (y_t), price level (p_t), interest (r_t) and exchange rate (e_t) which is an average of TL/Euro and TL/US Dollar exchange rates. The first two foreign variables, foreign price (p_t^*) and output (y_t^*) are constructed as the trade weighted average of OECD countries' prices and outputs. Unlike output and price, foreign interest rate (r_t^*) is calculated as the average of, only, Euro (Euribor) and USA interest rates reflecting the fact that financial linkages can be better captured by a small number of countries dominating the Turkish financial market. All variables are expressed in logarithms and the data cover the period of 1982Q1-2009Q4 for the estimation period, but are extended to 2011Q4 for forecast evaluation exercise. They are defined more precisely in Appendix where the data sources are also explained.

Since the modelling procedure, described in the following sections, depends on the time series properties of these variables we employ three different unit root tests to obtain, possibly robust results. We applied several unit root to these series [6] and found that all series contain unit root in their levels. Their first differences, on the other hand, appear to be stationary. The only exception is the inflation ($\triangle p_t$) where only unit root tests taking into account structural breaks indicate the its stationarity. As is well

[6] They include 3 different ADF tests, KPSS test and Phillips-Perron test. The ADF-WS (Park and Fuller, 1995) and the ADF-GLS test (Elliott, 1996) tests are used apart from the standard ADF test. To control structural breaks we used Perron (1989, 1997) unit root tests. To save space we do not report these results, however the are available upon request.

known, Perron (1987) and Perron (1997) assert that a stationary series can be spuriously detected as non-stationary in the presence of breaks.[7]

3 Cointegrating VAR-X Model

The probability forecasts are generated from a Cointegrating VAR-X model of Turkish Economy. The model is developed along the similar lines of cointegrating VAR-X models of UK and Swiss Economy described in Garratt et al. (2003b) and Assenmacher-Wesche and Pesaran (2008). To construct the model we begin by assuming that three following equilibrium relationships hold, in the long-run, for Turkish economy;

$$p_t - p_t^* - e_t = a_{10} + a_{11}t + \varepsilon_{1,t} \tag{1}$$

$$r_t - r_t^* = a_{20} + \varepsilon_{2,t} \tag{2}$$

$$y_t - y_t^* = a_{30} + \varepsilon_{3,t} \tag{3}$$

Where t stands for the linear trend and $\varepsilon_{i,t+1}$, $i = 1, 2, 3$, are stationary errors capturing the deviations from long-run equilibriums implied by corresponding relations. Based on international good market arbitrage (1) represents Purchasing Power Parity (PPP) relationship. Similarly, based on arbitrage between domestic and foreign bond holding (2) defines Interest Rate Parity (IRP) relationship. Finally, based on a stochastic version of the Solow growth model, (3) represents an "Output Gap" (OG) relationship. The empirical validity of this relationship, i.e. the stationarity of Turkish output gap, can provide evidence in favour of the long-run convergence hypothesis between OECD and Turkey's output levels.

These three long-run relationships of the model, (1)-(3), can be compactly written as

$$\boldsymbol{\varepsilon}_t = \boldsymbol{\beta}' \boldsymbol{z}_{t-1} - \boldsymbol{a}_1(t-1) - \boldsymbol{a}_0 \tag{4}$$

[7]See Kaya and Yazgan (2011), among others, for a statistical evidence on the apparent structural break in Turkish inflation rate which occurred on the first month of 2002.

where

$$\boldsymbol{z}_t = (p_t, e_t, r_t, r_t^*, y_t, y_t^*, p_t^*)'$$

$$\boldsymbol{a}_1 = (a_{11}, 0, 0)$$

$$\boldsymbol{a}_0 = (a_{10}, a_{20}, a_{30})$$

$$\boldsymbol{\varepsilon}_t = (\varepsilon_{1,t}, \varepsilon_{2,t}, \varepsilon_{3,t})$$

and

$$\boldsymbol{\beta}' = \begin{pmatrix} 1 & -1 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 \end{pmatrix}$$

The matrix $\boldsymbol{\beta}'$ imposes all the over identifying restrictions that are necessary to correspond to the long-run relationships.

In this setting we partition the variables where $\boldsymbol{y}_t = (p_t, e_t, r_t, y_t)$ is treated as an I(1) vector of endogenous variables and $\boldsymbol{x}_t = (y_t^*, p_t^*, r_t^*)$ is treated as an I(1) vector of weakly exogenous variables, in which, changes in weakly exogenous variables have direct influence on $\boldsymbol{y}_t$, but they are not affected by disequilibria in Turkish economy whose extend is measured by error correction terms This standard small open economy set up seems to be a natural choice in the case of Turkey.

Under the assumption of weakly exogenous variables, parameters can be estimated based on following conditional error correction model (see Pesaran et al. (2000) for example);

$$\boldsymbol{\triangle y_t = a_y - \alpha_y[\beta' z_{t-1} - a_1(t-1)] + \sum_{i=1}^{p-1} \Psi_y \triangle z_{t-i} + \psi_{yx} \triangle x_t + v_{yt}} \quad (5)$$

where $\boldsymbol{v}_{yt}$ is an 4×1 vector of serially uncorrelated shocks, $\boldsymbol{\alpha}_y$ is an 4×3 matrix of error correction coefficients, $\boldsymbol{\Psi}_y$ are 4×7 matrices of short-run coefficients and $\boldsymbol{\psi}_{yx}$ is an 4×1 vector of coefficients that represents impact of effects of changes in exogenous variables on $\triangle \boldsymbol{y}_t$.

To produce forecasts of endogenous variables by the conditional model requires forecasts of exogenous variables. In order to construct the forecasts of exogenous variables we specify the following marginal model.

$$\triangle \boldsymbol{x_t} = \boldsymbol{c_x} + \sum_{\boldsymbol{i=1}}^{\boldsymbol{p-1}} \boldsymbol{\Psi_x} \triangle \boldsymbol{z_{t-i}} + \boldsymbol{v_{xt}} \tag{6}$$

Where Ψ_x's are 7×1 matrices of unknown coefficient and $\boldsymbol{c}_x$ is 3×1 vector of intercepts and $\boldsymbol{v}_{xt}$ is an 3×1 vector of shocks are assumed to be uncorrelated with $\boldsymbol{v}_{yt}$.

Combining (5) and (6), and solving for $\triangle \boldsymbol{z}_t$ we have following VECM:

$$\triangle \boldsymbol{z_t} = \boldsymbol{a} - \boldsymbol{\alpha}[\boldsymbol{\beta}' \boldsymbol{z}_{t-1} - \boldsymbol{a}_1(t-1)] + \sum_{i=1}^{p-1} \boldsymbol{\Gamma}_i \triangle \boldsymbol{z}_{t-i} + \boldsymbol{u}_t \tag{7}$$

where

$\boldsymbol{a} = (\boldsymbol{c_x}, \boldsymbol{a'_y} - \boldsymbol{c_x \psi'_{yx}})'$, $\boldsymbol{\alpha} = (\boldsymbol{0}, \boldsymbol{\alpha'_y})'$, $\boldsymbol{\Gamma_i} = (\boldsymbol{\Psi'_x}, \boldsymbol{\Psi'_y} - \boldsymbol{\Psi'_x \psi'_{yx}})'$, and $\boldsymbol{u_t} = (\boldsymbol{v_{xt}}, \boldsymbol{v'_{yt}} - \boldsymbol{v_{xt} \psi'_{yx}})'$ is the vector of reduced form errors assumed to be iid$(\boldsymbol{0}, \boldsymbol{\Sigma})$ where $\boldsymbol{\Sigma}$ is a positive definite covariance matrix.

4 Testing for Long Run Relations

To provide evidence on the above specified long run relations, we test for the number of cointegration vectors (r) among the four endogenous variables of the model by employing Johansen's framework. Using model selection criteria such as AIC and SBC we confirm that VAR(2) model is appropriate and find evidence, based on the trace statistics displayed in Table 1, at 10 percent significance level for the existence of three cointegration relationships.[8]

Table 1

[8]We conduct test with unrestricted intercept and restricted trend.

Given the evidence on three cointegrating relations, the exact identification our model requires three restrictions on each of the three cointegration vectors, (on each row of $\boldsymbol{\beta}$), which in turn amounts to nine restrictions on $\boldsymbol{\beta}$ in total. However, the long run economic theory suggests the relationships, (1) to (3), in which 14 over identifying restrictions on $\boldsymbol{\beta}$ is required. It is well documented that LR test over-reject over-identifying restrictions in the case, therefore we make use of Bartlett correction factor suggested by Johansen (2000) overcome this problem. When the Bartlett correction is used, we find that over-identifying restrictions cannot be rejected with $\chi^2(15) = 3.69$, and correction factor is equal to 21.42. This provides empirical support for the long run relations presented above in the case of Turkish economy.

5 Derivation Probability Forecasts

As is emphasized above, future uncertainty is among the different forms of uncertainties which we aim to address. To tackle with future uncertainty we compute probability forecasts. Probability forecasts are generated via stochastic simulations of the underlying over identified VECM model in (7). For forecasting purpose we write (7) in its level form:

$$\boldsymbol{z_t} = \sum_{\boldsymbol{j=1}}^{\boldsymbol{s}} \boldsymbol{\Phi_j z_{t-j}} + \boldsymbol{b} + \boldsymbol{u_t}, \quad t = 1, 2, ..., T, \tag{8}$$

where

Let $\hat{\boldsymbol{\Phi}}_j, j = 1, 2, ..., s$, $\hat{\boldsymbol{b}}$ and $\hat{\boldsymbol{\Sigma}}$ refer to the estimators of $\boldsymbol{\Phi}_j, j = 1, 2, ..., s$, $\boldsymbol{b}$ and $\boldsymbol{\Sigma}$ respectively. These estimators are derived using the estimated parameters of Equations (5) and (6) through the following relations; $\boldsymbol{\Phi}_1 = \boldsymbol{I}_m - \boldsymbol{\alpha\beta}' + \boldsymbol{\Gamma}_1$, $\boldsymbol{\Phi}_i = \boldsymbol{\Gamma}_i + \boldsymbol{\Gamma}_{i-1}$, $i = 2, 3, ..., s-1$, $\boldsymbol{\Phi}_s = -\boldsymbol{\Gamma}_{s-1}$ which links the parameters of Equations of (7) and (8). Then *h-step* ahead forecast of $\boldsymbol{z}_{T+h}$, which is denoted by $\hat{\boldsymbol{z}}_{\boldsymbol{T+h}}$ can be obtained iteratively by successively feeding the previous period forecasts in

$$\hat{\boldsymbol{z}}_{\boldsymbol{T+h}} = \sum_{\boldsymbol{j=1}}^{\boldsymbol{s}} \hat{\boldsymbol{\phi}}_{\boldsymbol{j}} \hat{\boldsymbol{z}}_{\boldsymbol{T+h-j}} + \hat{\boldsymbol{b}}, \quad h = 1, 2, \ldots \tag{9}$$

We obtain simulation of values of $\boldsymbol{z}_{T+h}$ by

$$\hat{\boldsymbol{z}}_{\boldsymbol{T+h}}^{(i)} = \sum_{j=1}^{s} \hat{\boldsymbol{\phi}}_j \hat{\boldsymbol{z}}_{\boldsymbol{T+h-j}}^{(i)} + \hat{\boldsymbol{b}} + \boldsymbol{u}_{\boldsymbol{T+h}}^{(i)}, \quad h = 1, 2, ...; i = 1, 2, ..., S \tag{10}$$

where the subscript "(i)" represents the *i-th* replication. The $\boldsymbol{u}_{T+h}^{(i)}$ is drawn by parametric stochastic simulation method in which we simply take h random draws from the a multivariable distribution with zero means and the covariance matrix $\boldsymbol{\Sigma}$. Consider the matrix $\boldsymbol{P}$, where $\boldsymbol{P}^{-1}$ is the lower triangular Choleski decomposition of $\boldsymbol{\Sigma}$ such that $\boldsymbol{\Sigma} = \boldsymbol{P}\boldsymbol{P}'$. To obtain the simulated errors for m variables over h periods we first generate mh draws from the standard normal distribution, denoted by $\boldsymbol{\xi_{T+k}}, k = 1, 2, ..., h$. Since we have vector of standard normal disturbances, $\boldsymbol{\xi_t = P^{-1}u_t}$, then these are used to obtain simulated errors by $\boldsymbol{u_{T+h} = P\xi_{T+h}}$.[9]

The probability of an event $\boldsymbol{\varphi}[\boldsymbol{z}_{T+1}^{s}, \boldsymbol{z}_{T+2}^{s}, ..., \boldsymbol{z}_{T+h}^{s}] < \boldsymbol{a}$, is computed as

$$\boldsymbol{\pi(a, h; \varphi(\cdot))} = \frac{1}{S} \sum_{i=1}^{S} I\{\boldsymbol{a} - \boldsymbol{\varphi}(\boldsymbol{z}_{T+1}^{s}, \boldsymbol{z}_{T+2}^{s}, ..., \boldsymbol{z}_{T+h}^{s})\}$$

where $I(\cdot)$ is the indicator function that takes value 1 if $\boldsymbol{a} - \boldsymbol{\varphi}[\boldsymbol{z}_{T+1}^{s}, \boldsymbol{z}_{T+2}^{s}, ..., \boldsymbol{z}_{T+h}^{s}] < 0$ and zero otherwise.

6 Model Uncertainty and Structural Breaks

As is emphasized above and detailed further below, our approach to forecasting aims to address model uncertainty among other forms of uncertainties. To take model uncertainty into account, we adopt forecast combination approach which relies on the idea that a combination of forecasts from different models can perform better than those of single models constituting the combination. We therefore do not only consider forecasts resulted from our long-run theory consistent specification, but also produce forecasts from different specifications of our cointegrating VAR model to combine them to generate an average forecast. The alternative specifications include VECM

[9]Parameter uncertainty can be captured by using bootstrap procedure in which by simulating S*(in-sample)* values of $\boldsymbol{z}_t$, S set of simulated in-sample values are obtained and then models are estimated S times to obtain the ML estimator. Then for each boostrap replications, R replications of the *h-step* ahead forecast can be computed.

with no cointegration relationship, i.e. VAR on first differences only, VECM with one, two, and three exactly identified cointegration relationships, in additon to our overidentified VECM. Morevover we consider two alternative specifications for the marginal model, namely (6) and its random walk counterpart:

$$\triangle \boldsymbol{x_t} = \boldsymbol{\delta_{x0}} + \boldsymbol{u_{xt}} \tag{11}$$

Hence, a total of 10 models are considered (5 VECM models corresponding to each marginal model). To allow for the effects of model uncertainty on forecast performance, we use weighted averages of the forecasts of these models. The weights are derived from Bayesian Model Averaging procedure using AIC, SBC and equal weighting schemes as outlined by Garratt et al. (2003a) and Garratt et al. (2006).

Standard applications of Bayesian Model Averaging for forecasting purposes implicitly assume that all model under consideration are stable. But in reality some or all macroeconometric models under considerations may be subject to structural break which is regarded as among the main sources of forecast failure in many cases. To take into consideration the possible structural breaks we estimate models over different estimation windows and pool forecasts from different sample periods as suggested by Pesaran and Timmermann (2007) and Pesaran and Pick (2012). Like Assenmacher-Wesche and Pesaran (2008) and Pesaran et al. (2009) we also consider both model and window averaging together as mentioned below.

7 Directional Forecasts and Their Evaluations

The probability forecasts are computed for directional events of interest. We calculate the probability that r_t, r_t^*, e_t rise next period, namely $Pr[\triangle r_t > 0 \mid \Im_{t-1}]$, $Pr[\triangle r_t^* > 0 \mid \Im_{t-1}]$,$Pr[\triangle e_t > 0 \mid \Im_{t-1}]$ and probability that changes in y_t, y_t^*, p_t, p_t^* rise next period i.e $Pr[\triangle^2 y_t > 0 \mid \Im_{t-1}]$, $Pr[\triangle^2 y_t^* > 0 \mid \Im_{t-1}]$, $Pr[\triangle^2 p_t > 0 \mid \Im_{t-1}]$, $Pr[\triangle^2 p_t^* > 0 \mid \Im_{t-1}]$ where $\Im_{t-1}$ denotes the information available at time t.

To evaluate the probability forecasts, we use a statistical approach in which an event was forecast to be realized if its probability forecast is greater than a threshold value 0.5. We use Hit score, Kuipers score (KS) and, Pesaran and Timmermann (1992) statistics (PT) for comparisons of forecasts and realization. Hit score is defined as;

$$Hit\ score = \frac{UU + DD}{(UU + UD + DU + DD)}$$

where "UU", (*upward-upward*) indicates that forecast and realization are in the same upward direction, "DD" (*downward-downward*)indicates that they are in same downward direction, "UD" (*upward-downward*) indicates that forecast is in upward direction and realization is in downward direction and lastly "DU" (*downward-upward*) indicates that forecast is in downward direction and realization is in upward direction.

KS is defined by $H - F$, where H is the proportion of ups that were correctly forecast to occur, and F is the proportion of downs that were incorrectly predicted.

$$H = \frac{UU}{(UU + DU)}, \quad F = \frac{DU}{(UU + DU)}$$

These two proportions are known as the "hit rate" and "false alarm rate", respectively. In the case where the outcome is symmetric, in the sense that we value the ability to forecast ups and downs equally, then the score statistic of zero means no accuracy, whilst high positive and negative values indicate high and low predictive power. KS provides a statical measure of the accuracy of directional forecast however, it does not provide a statistical test. To compensate this shortcoming of KS, we employ PT which provides a formal test. As shown in Granger and Pesaran (2000), PT stat turns out to be equivalent to a test based on KS. PT statistics is defined by

$$PT = \frac{\widehat{P} - \widehat{P^*}}{[\widehat{V}(\widehat{P}) - \widehat{V}(\widehat{P^*})]^{\frac{1}{2}}}$$

where $\widehat{P}$ is the proportion of correctly predicted upward movements, $\widehat{P^*}$ is the estimate of the probability of correctly predicting the event under the null hypothesis that forecasts and realizations are independently distributed, $\widehat{V}(\widehat{P})$ and $\widehat{V}(\widehat{P^*})$ are the consistent estimates of the variances of $\widehat{P}$ and $\widehat{P^*}$, repectively. Under the null hypothesis, PT has standard normal distribution.

To evaluate the forecast accuracy of our models we employ pseudo out of sample methodology and estimate 10 models over different sample periods and compute one-step-ahead probability forecasts. First, we estimate the models using the data up to 2007Q3 and forecast for 2007Q4. Then repeat the process recursively by moving one quarter at time, until having the forecast for 2009Q4. This procedure provides 9 1-step ahead forecasts to be used in the evaluation of the accuracy of forecasts for the period 2007Q4-2009Q4. To deal with structural break we consider estimation windows starting between 1982Q1 to 1996Q4. Our first sample period is 1982Q1-2007Q3 and one to eight quarter ahead forecasts are produced for the period of 2007Q4-2009Q4. The sample period is then reduced one observation, i.e. 1982Q2-2007Q3 and another set of forecasts is produced for the same period. This process is repeated till the last sample period we consider, 1996Q4-2007Q3. Then we obtain 60 different forecasts, produced by estimation of the models over the different windows, for each point in the period of 2007Q4-2009Q4.

First we average forecasts over different model specifications for every sample periods in 2007Q4-2009Q4 by one of the averaging method mentioned above. The average forecasts obtained in this manner are denoted by *AveM*. Then, for each point in the period of 2007Q4-2009Q4, we calculate the simple average of forecasts over different estimation windows for each model. The window average forecasts are denoted by $AveW$.[10] Finally we average window averages over models and denote them by *AveAve*.

Table 2 shows the results of directional forecasts obtained under model averaging ($AveM$).[11] Table 3 depicts $AveW$ directional forecasts of over-identified long-run structural model with marginal model of (11)[12] and the Table 4 reports the results of directional *AveAve* forecasts. High values of UU and DD imply that forecasting ability of the model is high and, high values of UD and DU imply that forecasting ability of model is poor.

Table 2, 3 and 4

As can be followed from the tables the highest hit rates are obtained in *AveM* forecasts for the period of 2007Q4-2009Q4. The change in inflation and interest rates are the most successively predicted variables among all with hit ratios equal to 0.778, i.e. the direction of change is correctly predicted in 7 out of 9 quarters between 2007Q4-2009Q4 The forecast of change in the quarterly growth rates is turned to be correct in 6 out of 9 quarters leading to a hit rate equals to 0.667. The change in exchange rate, on the other

[10]Obviously we average 10 forecasts of different models, in the case of *AveM* and 60 forecasts for each model in the case $AveW$.

[11]Hereafter when not mentioned we report averages calculated using AIC weights, the results with other weighting schemes are available upon requests.

[12]($AveW$ forecasts of other model specifications are available upon request.

hand, remains as the most poorly forecasted variable among the 4 domestic variables considered in the model with a hit rate of 0.444.

Table 5 reports the summary statistics for directional *AveM*, *AveW*, and *AveAve* forecasts. We also report the results by using all 3 weighting schemes used in obtaining model averages. Hit scores are the average of hit rates of the changes in the 7 variables denoted as the total hit rate in the previous tables. They constitute an indicator of the overall forecasting success of the multivariate models. Hit score of without time average is 0.635 and its PT statistics is significant indicating that the null of independency of forecasts and realizations is rejected at 5 percentage level. The other forecasting models' PT statistics fail to be significant and their Kuipers scores are very low.

Table 5

It is not possible to assert that this is an impressive forecasting performance. Only the performance of *AveM* forecasts with AIC weigths can be considered satisfactory considering the properties of the period chosen for forecast evaluation. It should be emphasized that the evaluation period covers the period of 2008 crisis and right afterwards, hence it belongs to an extremely difficult period to forecast. Except the performance of *AveM* forecasts with AIC weigths, which is the only statistically significant one by the PT stat, the hit rates that are around 0.5. This indicates that only half of the direction correctly forecasted, which is not better than one that can do by using a fair coin. The results also indicate that window averaging does not improve, on the contrary it worsens in some cases, the forecasting performance of models. This is not a surprising result given the fact that the crisis period, for which the forecast evaluation is carried out, is not included in the estimation windows. On the other hand averaging over models provide higher forecast accuracy.

In order to evaluate forecasting performance of our models in a relatively "calmer" period, we compute forecasts for 2005Q4 to 2007Q4 by estimation the models by using data up to 2004Q4 and applying the same procedures as outlined above. Table 6 to 9 report the results of directional forecasts in the same order as above.

Table 6, 7, 8, and 9

The highest hit rates are achieved by the procedure which takes into account both model and averaging, i.e *AveAve* forecasts leading to an average hit rate of 0.683. Remarkably window averaging also delivers exactly same performance except the change in foreign inflation for which the i.e *AveAve* forecasts have better results. Both *AveAve* and *AveW* forecast the change in inflation and growth rate accurately with a hit rate of 0.889 for both. Table 9 shows that all PT statistics are significant with the exception of that of equal weight *AveM* forecasts. Overall the results point out that our forecasting accuracy are much better in this period and averaging increases significantly this accuracy.

8 Probability Forecast of Events for Inflation and Growth

The above forecast evaluation exercise revealed that the models and their combinations over different estimation windows forecast the change in inflation and output growth more accurately compared to the change in interest and exchange rate where the latter has the poorest results.

Based on this property of our VECMs, we consider the probability forecast of following seven events for inflation and output growth for the period of 2010Q1-2011Q4 and evaluate their performance by confronting them with their realized values. This time we use four quarter moving average of inflation and output growth, where the inflation and output growth rates are defined as quarter to quarter yearly rates, i.e $100[(p_{t+h} - p_{t+h-4})]$ and $100[(y_{t+h} - y_{t+h-4})]$.

As is known the Turkish economy has been characterized by high levels of inflation and numerous stabilization attempts since the early 1980s. Turkey has succeeded to lower inflation to single-digit level with the implementation of inflation targeting regime (IT) which is started at January 2002.

Over the period of 2002-2005 implicit inflation targeting was implemented. The plan was to reduce inflation to 35% in 2002, 20% in 2003, 12% in 2004 and 8% in 2005.[13] By the help of tight and credible fiscal policy, the outcome turned to be successful and inflation lowered from 68% at the end of 2001 to 7.7% at the end of 2005. After bringing inflation down from historically high levels, formal IT started to be implemented at the beginning of 2006.[14] The inflation target was set as a point target in which end-year targets were

[13]Targeted inflation was formulated as december to december changes in CPI.

[14]For a detailed information about implicit IT implementation see Kara (2008).

5% in 2006, 4% in 2007 and 2008, and 7.5% in 2009. An "uncertainty band" around target is defined by CBRT and it is set as 2% in both direction considering uncontrollable development such as oil prices, international liquidity conditions, taxes etc. (CBRT, 2006). To provide a performance criteria for the monetary policy, Central Bank of Turkey (CBRT) announced a quarterly inflation path consistent with the target and set uncertainty band for the end of each quarter (Ersel and Ozatay, 2008)). Table 10 provides the 4-quarter moving average of inflation paths consistent with the target and their uncertainty bands for the period of 2010Q1-2011Q4 for which we aim to produce forecasts.

Table 10

Table 10 also provides four quarterly moving average of actual, quarter to quarter, yearly rates of inflation and output growth. As can be followed from Table 10, CBRT was quite successful in keeping the actual inflation within the band and Turkish Economy registered a strong growth performance for this period. Having this background information in mind we consider the following events:[15]

- E1: A single event of inflation in which the inflation target of CBRT is met. The attainment of the target is defined as four quarterly moving average of inflation is between four quarterly moving average of uncertainty band provided in Table 10.
- E2: A single event of recession in which quarterly output growth is negative for two consecutive quarters.
- E3: A single event of weak growth in which four quarterly moving average of output growth is less than 4 percent.
- E4: A single event of strong growth in which four quarterly moving average of output growth is higher than 7 percent.

Conditional on the information available at the end of 2009Q4, we calculate the probability forecasts of the above events at $h = 1, 2, 3, ...8$ forecast horizons via stochastic simulations as described above. We report $AveM$,

[15]It is also possible to define some joint events and calculate their associated probabilities.

AveW, and *AveAve* probability forecasts for these events in Table 13, Table 14, and Table 15 respectively.[16]

Table 11, 12, and 13

The emboldened entries indicate the correctly predicted events obtained by using the calculated probabilities reported in the tables.[17] As can be followed tables the events E2 and E3 are successfully predicted across all models. On the other hand, the event E4 can only be predicted correctly with window averaging. However the event E1 can be correctly forecasted only once in the last quarter of 2010, when the CBRT missed its inflation target first and last time in the period under consideration. Overall, while the events associated growth are predicted quite well, our models turn out to be unsuccessful in forecasting the inflation within the band targeted by the CBRT. We observe the same property in the directional forecasts of the change in growth and inflation such that even though the hit rates associated the growth forecasts are equal to 0.889 for all the models, those of inflation are as high as 0.750 only with *AveAve* forecasts.

9 Conclusion

In this paper we presented probability forecasts of directional events related to a number of macro aggregates based on a cointegrating VAR-X model of Turkey. The directional forecasts of output and inflation are found quite successful except the period of the 2008 crisis in which their performance deteriorated to a certain extent. The probability forecasts of certain events associated to output and inflation are also illustrated and confronted with their actual outcomes for a recent of Turkish economy. This exercise provided successful results too, at least for the events defined on output. Our results indicated that averaging over models and estimation windows help to increase the forecast accuracy.

Overall the probability forecasting approach adopted in this paper can be deemed promising to serve as a forecasting tool. Especially, the events

[16]Note also that while the hit rates associated directional forecasts for the change in quarterly growth for this period are as high as for the period of 2005Q1-2007Q4 across all models, the hit rates of the change in inflation is slightly worse.

[17]The decision is based on 0.5 threshold as above.

described in the previous section can be increased in numbers towards several interesting directions depending on the characteristic of period for which they are formed. It seems that to follow this line research by extending the basic model and considering some others for combining their forecasts can be useful in terms of increasing forecast accuracy. In this paper, for the purpose of combining forecasts we only considered simple variants of our over-identified model. Although this approach is in line with the applied forecast combination literature, where combinations from significantly different models are not usually taken into account, including forecasts from different, especially non-linear models may be helpful in improving forecast performance. Another interesting future research area may be to develop similar sectoral models with linkages to the aggregate economy to produce forecasts at sectoral level and combine them for the aggregates.

References

Assenmacher-Wesche, K. and M. H. Pesaran (2008). Forecasting the swiss economy using vecx* models: An exercise in forecast combination across modelsand observation windows. *National Institute Economic Review 203*, 91–108.

Bates, J. and C. Granger (1969). The combination of forecasts. *Operations Research Quarterly 20*, 451–468.

Britton, E., P. Fisher, and J. Whitley (1998). The inflation report projections: Understanding the fan chart. *Bank of England Quarterly Bulletin 38*(1), 30.

CBRT (2006). *The booklet of Inflation Targeting (in Turkish)*. CBRT.

Clements, M. P. (2004). Evaluating the bank of england density forecasts of inflation. *Economic Journal 114*(498), 844–866.

Dees, S., F. di Mauro, M. H. Pesaran, and L. V. Smith (2007). Exploring the international linkages of the euro area: A global var analysis. *Journal of Applied Econometrics 22*, 1–38.

Diebold, F. X. and J. A. Lopez (1995). Forecast evaluation and combination. Research Paper 9525, Federal Reserve Bank of New York.

Elliott, G. and A. Timmermann (2008). Economic forecasting. *Journal of Economic Literature 46*(1), 3–56.

Ersel, H. and F. Ozatay (2008). Fiscal dominance and inflation targeting: Lessons from turkey. *Emerging Markets Finance and Trade 44*(6), 38–51.

Fair, R. C. (1980). Estimating the expected predictive accuracy of econometric models. *International Economic Review 21*(2), 355–78.

Garratt, A., K. C. Lee, M. H. Pesaran, and Y. Shin (2003a). Forecast uncertainties in macroeconomic modeling: An application to the u.k. economy. *Journal of the American Statistical Association 98*(2), 829–838.

Garratt, A., K. C. Lee, M. H. Pesaran, and Y. Shin (2003b). A long run structural macroeconometric model of the uk. *Economic Journal 113*(1), 412–455.

Garratt, A., K. C. Lee, M. H. Pesaran, and Y. Shin (2006). *Global and National Macroeconometric Modelling: A Long-Run Structural Approach.* Oxford University Press.

Gneiting, T. (2008). Editorial: Probabilistic forecasting. *Journal Of The Royal Statistical Society Series A 171*(2), 319–321.

Granger, C. and M. Pesaran (2000). Economic and statistical measures of forecast accuracy. *Journal of Forecasting 19*, 537–560.

Johansen, S. (2000). A small sample correction of the test for cointegrating rank in the vector autoregressive model. *Econometric Theory 16*(eco2000/15), 740–778.

Kara, A. H. (2008). Turkish experience with implicit inflation targeting. *Central Bank Review 8*(1), 1–16.

Kaya, H. and M. E. Yazgan (2011). Has inflation targeting increased predictive power of term structure about future inflation: Further evidence from turkish market. *Applied Financial Economics 21(20)*, 1539–1547.

Perron, P. (1987). The great crash, the oil price shock, and the unit root hypothesis. *Econometrica 57*, 1361–1401.

Perron, P. (1997). Further evidence on breaking trend functions in macroeconomic variables. *Journal of Econometrics, 80, 80*, 355–385.

Pesaran, M., T. Schuermann, and L. Vanessa Smith (2009). Forecasting economic and financial variables with global vars. *International Journal of Forecasting 25*, 642–675.

Pesaran, M. and A. Timmermann (1992). A simple nonparametric test of predictive performance. *Journal of Business and Economic Statistics 10*, 461–465.

Pesaran, M. and A. Timmermann (2007). Selection of estimation window in the presence of breaks. *Journal of Econometrics 137*, 134–61.

Pesaran, M., S. Y., and R. Smith (2000). Structural analysis of vector error correction models with i(1) variables. *Journal of Econometrics 97*, 293–343.

Pesaran, M. H. and A. Pick (2012). Forecast combination across estimation windows. *Journal of Business Economics and Statistics 29*, 307–318.

Stock, J. and M. Watson (2004). Combination forecasts of output growth in a seven country data set. *Journal of Forecasting 23*, 405–430.

Timmermann, A. (2006). Forecast combinations. *In: G. Elliot, C. W. J. Granger, and A. Timmermann (eds), Handbook of Economic Forecasting 1*, 135–196. Elsevier.

Tables

Table 1: Cointegration Test

r	Trace Statistics	P-Value
0	167.948	0.000
1	88.916	0.000
2	41.746	0.063
3	12.297	0.473

Table 2: $AveM$(2007Q4-2009Q4)

Variable	Threshold	UD	DD	DU	UU	Hit rate
p_t	$\triangle^2 p_t > 0$	2	3	0	4	0.778
y_t	$\triangle^2 y_t > 0$	2	6	1	0	0.667
e_t	$\triangle e_t > 0$	1	3	4	1	0.444
r_t	$\triangle r_t > 0$	2	1	0	6	0.778
y_t^*	$\triangle^2 y_t^* > 0$	1	2	2	4	0.667
r_t^*	$\triangle r_t^* > 0$	2	4	1	2	0.667
p_t^*	$\triangle^2 p_t^* > 0$	3	1	2	3	0.444
Total		13	20	10	20	
Hit rate	0.635					

Table 3: *AveW* Overidentifed Model (2007Q4-2009Q4)

Variable	Threshold	UD	DD	DU	UU	Hit rate
p_t	$\triangle^2 p_t > 0$	3	2	0	4	0.667
y_t	$\triangle^2 y_t > 0$	3	5	1	0	0.556
e_t	$\triangle e_t > 0$	0	4	5	0	0.444
r_t	$\triangle r_t > 0$	1	2	3	3	0.556
y_t^*	$\triangle^2 y_t^* > 0$	2	1	1	5	0.667
r_t^*	$\triangle r_t^* > 0$	4	2	1	2	0.444
p_t^*	$\triangle^2 p_t^* > 0$	3	1	3	2	0.333
Total		16	17	14	16	
Hit rate	0.524					

Table 4: *AveAve* (2007Q4-2009Q4)

Variable	Threshold	UD	DD	DU	UU	Hit rate
p_t	$\triangle^2 p_t > 0$	3	2	0	4	0.667
y_t	$\triangle^2 y_t > 0$	3	5	1	0	0.556
e_t	$\triangle e_t > 0$	0	4	5	0	0.444
r_t	$\triangle r_t > 0$	3	0	1	5	0.556
y_t^*	$\triangle^2 y_t^* > 0$	1	2	2	4	0.667
r_t^*	$\triangle r_t^* > 0$	2	4	1	2	0.667
p_t^*	$\triangle^2 p_t^* > 0$	3	1	3	2	0.333
Total		15	18	13	17	
Hit rate	0.556					

Table 5: Forecast Evaluation (Forecast Period: 2007Q4-2009Q4)

	AveM		
	Equal Weight	AIC Weight	SBC Weight
Hit score	0.540	0.635	0.556
Kuipers score	0.103	0.273	0.116
PT test	0.787	2.182*	0.923

	***AveW* (over-identified model)**
Hit score	0.524
Kuipers score	0.048
PT test	0.038

	AveAve		
	Equal Weight	AIC Weight	SBC Weight
Hit score	0.540	0.556	0.524
Kuipers score	0.096	0.112	0.065
PT test	0.748	0.896	0.503

* indicates that statistics significant at 5 percent level

Table 6: *AveM* (2005Q4-2007Q4)

Variable	Threshold	UD	DD	DU	UU	Hit rate
p_t	$\triangle^2 p_t > 0$	2	3	0	4	0.778
y_t	$\triangle^2 y_t > 0$	1	0	1	7	0.778
e_t	$\triangle e_t > 0$	5	0	0	4	0.444
r_t	$\triangle r_t > 0$	0	7	1	1	0.889
y_t^*	$\triangle^2 y_t^* > 0$	1	4	1	3	0.778
r_t^*	$\triangle r_t^* > 0$	0	6	3	0	0.667
p_t^*	$\triangle^2 p_t^* > 0$	4	1	0	4	0.556
Total		13	21	6	23	
Hit rate	0.698					

Table 7: $AveW$ of over-identified model (2005Q4-2007Q4)

Variable	Threshold	UD	DD	DU	UU	Hit rate
p_t	$\triangle^2 p_t > 0$	1	4	0	4	0.889
y_t	$\triangle^2 y_t > 0$	1	0	0	8	0.889
e_t	$\triangle e_t > 0$	5	0	1	3	0.333
r_t	$\triangle r_t > 0$	4	3	1	1	0.444
y_t^*	$\triangle^2 y_t^* > 0$	0	5	1	3	0.889
r_t^*	$\triangle r_t^* > 0$	3	3	0	3	0.667
p_t^*	$\triangle^2 p_t^* > 0$	4	1	2	2	0.333
Total		18	16	5	24	
Hit rate	0.635					

Table 8: $AveAve$ (2005Q4-2007Q4)

Variable	Threshold	UD	DD	DU	UU	Hit rate
p_t	$\triangle^2 p_t > 0$	1	4	0	4	0.889
y_t	$\triangle^2 y_t > 0$	1	0	0	8	0.889
e_t	$\triangle e_t > 0$	5	0	1	3	0.333
r_t	$\triangle r_t > 0$	4	3	1	1	0.444
y_t^*	$\triangle^2 y_t^* > 0$	0	5	1	3	0.889
r_t^*	$\triangle r_t^* > 0$	0	6	3	0	0.667
p_t^*	$\triangle^2 p_t^* > 0$	3	2	0	4	0.667
Total		14	20	6	23	
Hit rate	0.683					

Table 9: Forecast Evaluation (Forecast Period 2005Q4-2007Q4)

	AveM		
	Equal Weight	AIC Weight	SBC Weight
Hit score	0.54	0.698	0.587
Kuipers score	0.167	0.417	0.282
PT test	1.14	3.31*	1.97*
	***AveW* (over-identified model)**		
Hit score	0.635		
Kuipers score	0.333		
PT test	2.522*		
	AveAve		
	Equal Weight	AIC Weight	SBC Weight
Hit score	0.635	0.683	0.635
Kuipers score	0.358	0.391	0.389
PT test	2.635*	3.089*	2.77*

* indicates that statistics significant at 5 percent level

Table 10: Output Growth and Inflation Path

	Uncertainty band (upper limit)	Uncertainty band (lower limit)	Path consistent with the target	Actual inflation	Actual growth
2010Q1	9.825	5.825	7.825	6.510	1.9
2010Q2	9.250	5.250	7.250	7.390	6.4
2010Q3	8.750	4.750	6.750	8.150	8.4
2010Q4	8.500	4.500	6.500	8.580	9.2
2011Q1	8.250	4.250	6.250	7.340	9.2
2011Q2	8.000	4.000	6.000	6.510	8.8
2011Q3	7.750	3.750	5.750	6.000	9.6
2011Q4	7.500	3.500	5.500	6.433	8.6

Note: Inflation targets and their uncertainty bands obtained from CBRT inflation reports. 4-quarter moving average figures are used.

Table 11: The Event Probabilities (*AveM*)

	E1	E2	E3	E4
2010Q1	0.285	**0.000**	0.277	**0.113**
Q2	0.265	**0.071**	**0.083**	0.673
Q3	0.234	**0.078**	**0.141**	0.477
Q4	**0.176**	**0.083**	**0.230**	0.397
2011Q1	0.157	**0.089**	**0.239**	0.414
Q2	0.129	**0.091**	**0.230**	0.397
Q3	0.105	**0.096**	**0.251**	0.388
Q4	0.088	**0.091**	**0.279**	0.382

Note: Emboldened items refer to the correctly predicted events

Table 12: The Event Probabilities (*AveW* over-identifed model)

	E1	E2	E3	E4
2010Q1	0.152	**0.000**	0.110	**0.347**
Q2	0.097	**0.016**	**0.005**	0.906
Q3	0.078	**0.018**	**0.025**	**0.821**
Q4	**0.054**	**0.032**	**0.044**	**0.789**
2011Q1	0.050	**0.048**	**0.063**	**0.731**
Q2	0.053	**0.066**	**0.019**	**0.648**
Q3	0.062	**0.086**	**0.162**	**0.552**
Q4	0.070	**0.102**	**0.223**	0.491

Note: Emboldened items refer to the correctly predicted events.

Table 13: The Event Probabilities (*AveAve*)

	E1	E2	E3	E4
2010Q1	0.301	**0.000**	0.335	**0.108**
Q2	0.242	**0.117**	**0.068**	0.613
Q3	0.208	**0.118**	**0.245**	**0.404**
Q4	**0.160**	**0.142**	**0.371**	0.317
2011Q1	0.146	**0.159**	**0.381**	0.301
Q2	0.116	**0.170**	**0.419**	0.270
Q3	0.089	**0.164**	**0.439**	0.268
Q4	0.075	**0.163**	**0.442**	0.272

Note: Emboldened items refer to the correctly predicted events.

Appendix: Data and Sources

- y_t: The logarithm of domestic output is measured by Real GDP Volume. It is obtained from International Financial Statistics (IFS). The quarterly series for period 1980-1986 were interpolated from yearly figures by using the methodology described in Dees et al. (2007). y_t are seasonally adjusted by using X12 method.

- p_t: The logarithm of domestic price level is measured by CPI. They are obtained from IFS. p_t are seasonally adjusted by using X12 method.

- e_t: The logarithm of exchange rate is measured by the average of the period averages of TL/US Dollar and TL/Euro exchange rates. They are obtained from IFS. For the construction of Euro exchange rate, for the period before 2000, TL/German Mark exchange rate series that are converted to Euro by using the Euro conversion rate.

- r_t: Domestic interest rate is measured by annualized three-month Treasury Bill rate, R_t. It is calculated as $0.25ln\left(1+\frac{R_t}{100}\right)$ and obtained from IFS. A few missing data in the three-month Treasury Bill rates are constructed by using the fitted values of their regression on deposit rates.

- y_t^*: Foreign output is constructed as the trade-weighted average of OECD countries' GDPs. It is calculated as $y_t^* = ln\left(\sum_{i=1}^{n} w_{it}Y_{it}\right)$ where Y_{it} is the real GDP of country i, n is the number of OECD countries and w_{it} is the trade shares of country i in Turkey's total trade with OECD. Trade is calculated as the sum of its imports and exports. The bilateral trade data are gathered from Direction of Trade Statistics. GDP Volume series are obtained from IFS. Y_{it} are seasonally adjusted by using X12 method when necessary.

- p_t^*: Foreign price is constructed as the trade-weighted average of OECD countries' CPIs. $p_t^* = ln\left(\sum_{i=1}^{n} w_{it}P_{it}\right)$ where P_{it} is the CPI of country i. CPI series are obtained from IFS. For the period before German unification, in 1990 Q4, West German CPI is used to obtain a common index. P_{it} are seasonally adjusted by using X12 method when necessary.

- r_t^*: Foreign interest rate is measured by the average annualized three-month USA Treasury Bill and Euribor rates. The period before 2000 of the Euribor series is completed by using three-month German Treasury Bill rate. It is calculated as $0.25ln\left(1+\frac{R_t^*}{100}\right)$. Where R_t^* are the average values of foreign rates. They are obtained from IFS and http:/www.euribor.org.