CERGE-EI FORECASTING MODEL

1 Methodology

The CERGE-EI forecasting procedure is based on a variety of econometric methods. The objective is to obtain quarterly predictions of the core macro variables such as the real Gross Domestic Product (GDP), inflation, and the unemployment rate. The final forecast is a composite (pooled) term which includes the projections obtained from the following methods:
1) VAR in Levels
2) VAR in Differences
3) Exponential Smoothing
4) State Space Models
5) Bayesian Vector Autoregression (BVAR)
6) Diffusion Index
7) Discounted Least Squares (DLS)
8) Forecast on forecast
9) Univariate - ARIMA, Rolling regression

We employ EViews 5.1 to perform the econometric analysis and forecasting. The program offers powerful estimation techniques and advanced methods for working with data, and, therefore, serves as an efficient tool to implement the above-listed forecasting methods. We maintain a database of macroeconomic series for every country for which we produce the forecasts. Initially, we accumulate the information in Excel data files. Afterwards, the data is imported into the corresponding EViews workfiles. Moreover, for each country we have a separate program file. Even though we keep the standardized structure of the computer code for various countries, the program still may comprise country-specific features such as the set of exogenous variables, data length, and VAR model specification.

At present, we compute the macroeconomic forecast for 13 transition countries: Czech Republic, Slovakia, Slovenia, Croatia, Romania, Russia, Bulgaria, Poland, Hungary, Lithuania, Latvia, Estonia, and Ukraine. We describe the methods in some detail below.

1.1 VAR (in levels and differences, Holt-Winters exponential smoothing)

Vector autoregression (VAR) is commonly used for forecasting systems of interrelated time series and for analyzing the dynamic impact of random disturbances on a system of variables.

This class of models does not require strong assumptions and can be applied when the economic theory is not rich enough to provide a dynamic specification that identifies the relationships among the variables of interest. The VAR approach avoids the need for structural modeling by treating every endogenous variable in the system as a function of the lagged values of all of the endogenous variables in the system. For a more detailed description of this approach see Hamilton (1994), Chapters 11, 19, and 20.

Our baseline VAR specification includes this set of endogenous variables: real GDP growth rate, consumer price inflation, M2 growth rate, change in the unemployment rate, and the percentage change in the exchange rate. The set of exogenous variables varies from country to country and may incude the Euro area real GDP growth rate, EURO/USD exchange rate, oil prices (for example, for Russia) as well as "event-based" dummy variables. We perform the necessary transformation of the data - differencing and/or logarithmic transformation. The choice of ILO vs. registered unemployment is determined by the availability and length of ILO unemployment series. We use the registered unemployment if a long ILO series is unavailable. The major problem we face when developing the model is over-parametrization, that the system has too many parameters compared to the sample size. Therefore, we have to restrict the model representation to a very parsimonious VAR. In most cases, the VAR order is 2 or 3 (for comparison, in models designed for developed countries the order ranges from 4 to 16).

The forecasting procedure consists of the following steps:

1) Estimate VAR in Levels:

$$Y_t = \Sigma_{s=1}^{q} \Pi_s Y_{t-s} + \beta Z_t + A_0 + \epsilon_t \qquad (LS) \ ,$$

VAR in Differences:

$$\Delta\, Y_t = \Sigma_{s=1}^{q} \Pi_s \Delta Y_{t-s} + \beta Z_t + A_0 + \epsilon_t \qquad (LS),$$

where Y_t is a vector of endogenous variables, i.e.

$Y_t = \{$real GDP, CPI, Unemployment rate, Exchange rate, M2 money aggregate$\}$;

Z_t is a vector of exogenous variables, i.e.

$Z_t = \{$Euro area real GDP, EURO/USD exchange rate, dummies for "special" events$\}$;

Π_s and β are coefficient matrices; q is the VAR order (2 or 3); and ϵ_t is a vector of innovations that may be contemporaneously correlated but is uncorrelated with its own lagged values and uncorrelated with all of the right-hand side variables.

2) Compute forecasts - point estimates and standard errors. In EViews, the stochastic dynamic forecast of the endogenous variables can be obtained by solving the model that describes the joint relationship between the variables in the system and is created from the estimated VAR. The corresponding commands are *makemodel* and *model.solve* .

In addition to the standard VAR-based forecasts, we employ exponential smoothing methods to compute the predictions of the endogenous variables. This technique calculates forecasts that, unlike predictions from regression models with fixed coefficients, are adjusted for past forecast errors. We use the Holt-Winters smoother (filter). For a detailed description see the EViews5 Users Guide. The value of the smoothing parameters for mean, trend, and seasonal factors are estimated by minimizing the sum of squared errors. The exponential smoothing procedure is performed in Eviews by choosing the appropriate time series and selecting *Proc/Exponential Smoothing.*

The advantages of the exponential smoothing methodology are that it does not require long time series and allows a nonlinear structure of the problem and drift in parameter values (though the specification is very parsimonious). The disadvantage is that forecasting for more distant periods is often not satisfactory. Moreover, standard errors tend to be high if the series is persistent.

1.2 State Space Model

State Space models allow the estimation of a dynamic system with unobservable factors. The technique is particularly useful in modeling unstable processes with the parameters of the system changing over time. We employ the state space methodology in order to capture the process of economic transformation typically facing all transition economies. One example may be the shift from oil-based growth in Russia to a healthier growth based on light industry, manufacturing, and the development of the non-tradable sector.

We build on the model as a simple application of the state space methodology to the VAR setting. In practice, we can allow only a few coefficients to vary because of a small sample size. We choose the intercept to vary over time due to some unobserved component (variable representing the pace of reforms, economic transformation, etc.).

The dynamic system is expressed in the so-called state space representation, which, in general terms, can be represented by the following system of equations:

$$\xi_{t+1} = F\xi_t + v_{t+1} \qquad (1)$$

$$Y_t = \Sigma_{s=1}^{q} \Pi_s Y_{t-s} + H\xi_t + \beta Z_t + w_t \ , \qquad (2)$$

where Y_t is an $(n \times 1)$ vector of variables observed at date t; ξ_t is a possibly unobserved $(r \times 1)$ vector - time varying intercepts, called the state vector; F is a transition matrix for intercepts, Π is a matrix of coefficients corresponding to the $q-$th lag, and Z_t is a $(k \times 1)$ vector of exogenous or predetermined variables. Equation (1) is known as a state equation, and (2) is known as an observation/signal equation. The $(r \times 1)$ vector v_t and the $(n \times 1)$ vector w_t are vectors of white noise:

$E(v_t v_t') = Q$, $E(v_t v_\tau') = 0$; $E(w_t w_t') = R$, $E(w_t w_\tau') = 0$; $E(v_t w_\tau') = 0$ - uncorrelated at all lags. Moreover, v_t is uncorrelated with all lagged values of ξ.

We define the vector of endogenous and exogenous variables similarly as in the previous subsection. We use the first LOG differences (natural log), which gives the percentage change and enables us to resolve the non-stationarity problem. Variables that are already in percentage terms (e.g. unemployment rate) are used in simple differences.

For simplicity we assume the random walk process for the intercept: $\xi_t = \xi_{t-1} + v_t$.

The estimation procedure includes Kalman filtering. An important problem is the sensitivity of the estimation results to starting values. We choose OLS estimates on a small subsample as starting values for parameters. Alternatively, Eviews may do it automatically.

The task is to estimate the values of the unknown (unobservable) parameters of the system on the basis of the observed information contained in $(Y_{1,}Y_2, ...Y_T, and\ Z_1, Z_2, ...Z_T)$. The Kalman filter computes recursively the linear least squares forecasts of the state vector on the basis of the data observed through date t, $\widehat{\xi}_{t+1/t} \equiv \widehat{E}(\xi_{t+1}/Y_t, Z_t)$, where $\widehat{E}(\xi_{t+1}/Y_t, Z_t)$ denotes the linear projection of ξ_{t+1} on Y_t, Z_t . The derivation of the Kalman Filter is presented in detail in Hamilton (1994), Chapter 13.

On the basis of the estimated state vector we can forecast the next period values of the endogenous variables:

$$Y_{t/t-1} = \Sigma_{s=1}^{q} \Pi_s Y_{t-s} + \beta Z_t + H' \widehat{E}(\xi_t / Z_{t-1,} Y_{t-1}) = \Sigma_{s=1}^{q} \Pi_s Y_{t-s} + \beta Z_t + H' \widehat{\xi}_{t/t-1}.$$

On the basis of the information about Y_t and the forecasting error $(Y_t - Y_{t/t-1})$ (evaluate its variance), we update our inference about the current value of ξ_t and thus can obtain the estimate $\widehat{\xi}_{t/t}$. Eventually, on the basis of this estimate we can forecast the next period value of the unobservable state vector,

$$\widehat{\xi}_{t+1/t} = \widehat{E}(\xi_{t+1}/Z_t, Y_t) = F\widehat{E}(\xi_t/Z_t, Y_t) + \widehat{E}(v_{t+1}/Z_t, Y_t) = F\widehat{\xi}_{t/t} + 0.$$

In order to perform Kalman filter computations we need to have the numerical values of parameters Q, Π, H, R, which might be unknown. The estimation of the parameters of the state space model is performed by MLE under the assumption that disturbances are Gaussian.

To perform the estimation in EViews, we proceed as follows:

a) construct five VAR State Space objects for each of the endogenous variables : select *Object/New Object/Sspace* from the main toolbar;

b) estimate each State Space model separately by MLE. In order to estimate the model click on the Estimate button on the toolbar or select *Proc/Estimate*;

c) make the model from the estimated State Space equations in order to obtain the forecasts of the endogenous variables, use the command *var_ss.make model(model_ss).* A similar procedure is performed for all VAR state space equations;

d) merge all models into the final model - *model_final ;*

e) the command *model_final.solve* solves the final model and enables obtaining joint forecasts of the endogenous variables as well as standard errors.

1.3 Bayesian Vector Autoregressions (BVAR)

An important attempt to improve the forecasting performance of unrestricted VAR is the application of Bayesian techniques which take account of any prior information available.

The Bayesian approach provides a general method for combining our beliefs with the evidence contained in the data. In contrast to the classical approach to estimating a set of parameters, the Bayesian approach presupposes a set of prior probabilities about the underlying parameters to be estimated.

We concentrate on Litterman's (1986) approach to Bayesian estimation of VAR that considers a single equation in isolation. In general, the model specification takes the following form:

$$Y_t = A_0 + \Pi_1 Y_{t-1} + \Pi_2 Y_{t-2} + ... + \Pi_q Y_{t-q} + \varepsilon_t.$$

In matrix notation:

$$Y = (X \bigotimes I_q)\Pi + \epsilon,$$

where X is stacked vector of $Y_{t-1}, Y_{t-2}, ..., Y_{t-q}$.

The objective of Bayesian estimation is to produce coefficient estimates which combine the evidence from the data with the information contained in our beliefs. The information contained in the data is summarized in the sample probability density function, $f(Y_t \mid \Pi)$. Formally, the non-sample prior information about the set of parameters $\Pi = (\widetilde{\Pi}_1, ..., \widetilde{\Pi}_q)$ that is to be estimated is assumed to be represented by a prior multivariate normal distribution with the density function

$$f(\Pi) = \left(\frac{1}{2\pi}\right)^{q^2p/2} |V_\Pi|^{-1/2} \exp\left\{-\frac{1}{2}(\Pi - \Pi^*)' V_\Pi^{-1} (\Pi - \Pi^*)\right\},$$

where Π^* is the prior of the mean of Π, and V_Π is the variance of the prior.

The likelihood function for the Gaussian process is given by

$$L(\Pi|Y) = \left(\frac{1}{2}\right)^{qT/2} \left|I_T \bigotimes \Sigma\right|^{-1/2} \exp\left\{-\frac{1}{2}\left(Y - (X \bigotimes I_k)\Pi'\right)' (I_T \bigotimes \Sigma^{-1}) \left(Y - (X \bigotimes I_k)\Pi'\right)\right\}.$$

Therefore, the posterior conditional density $f(\Pi|Y)$ is proportional to
$\exp\left\{-\frac{1}{2}(\Pi - \bar{\Pi})'\Sigma_\Pi^{-1}(\Pi - \bar{\Pi})\right\}$,
where the posterior mean is

$$\bar{\Pi} = \left(V_\Pi^{-1} + (XX' \bigotimes \Sigma^{-1})\right)^{-1} (V_\Pi^{-1}\Pi^* + (X' \bigotimes \Sigma^{-1})Y), \tag{3}$$

and the posterior covariance matrix is

$$\Sigma_\Pi = \left(V_\Pi^{-1} + (XX' \bigotimes \Sigma^{-1})\right)^{-1}. \tag{4}$$

Priors. In practice, the prior mean Π^* and and the prior variance V_Π need to be specified. According to Litterman (1986), the prior for the variance of the coefficient on lag l of variable j in equation i can be given by:

$$V_{ij}(l) = (\lambda/l)^2 \quad \text{if} \quad i = j,$$

$$V_{ij}(l) = (\theta\lambda/l)^2(\sigma_{ii}^2/\sigma_{jj}^2) \quad \text{if} \quad i \neq j,$$

where θ, λ are hyperparameters. θ governs the strength of cross-variable responses (small θ imposes weak responses); λ determines how distant lags affect current values (small λ imposes a weak response to distant lags).

Prior for Π^*: typically $\Pi_{ii} = 1$ (random walk) or $\Pi_{ii} = 0$ (stationary/iid) and $\Pi_{ij} = 0$ for $i \neq j$.

In our case, the model specification takes the following form: Endogenous variables:

$Y_t = \{$real GDP, CPI, Unemployment rate, Exchange rate, M2 money aggregate$\}$;

Exogenous variables:

$Z_t = \{$Euro area real GDP, EURO/USD exchange rate, dummies for "special" events$\}$;

We estimate the following equation:

$$Y_t = A_0 + \Sigma_{s=1}^{q} \Pi_s Y_{t-s} + \beta Z_t + \varepsilon_t. \qquad (5)$$

The forecasting procedure includes the following subroutines:

a) endogenous selection of optimal hyper-parameters θ and λ (estimate BVAR for each θ and λ in the loop, set the interval for θ and λ on the basis of intuition and previous estimations). In fact, we estimate BVAR on the subsample for different values of hyperparameters, i.e at this stage we perform the model selection. The in-sample forecast is produced for the period $[T-t, T]$. Thereafter, we evaluate the performance of different models and on the basis of the MSE criterion select the best model, which is used to produce the out-of-sample final forecast. More specifically we:

i) estimate the VAR system (5) by OLS to get the 1st stage estimates of covariances σ;

ii) compute priors: mean and variance. The priors for coefficients Π_s are given by: $\Pi_{ii} = 1$ and $\Pi_{ij} = 0$ for $i \neq j$, i.e. we assume the unit root in every series. According to Litterman (1986), we compute the prior for variance in the form: $V_{ij}(l) = (\lambda/l)^2$ if $i = j$, $V_{ij}(l) = (\theta\lambda/l)^2(\sigma_{ii}^2/\sigma_{jj}^2)$ if $i \neq j$;

iii) compute Bayesian estimates (posterior) for coefficients Π_s and variance of the estimated parameters according to formulas (3) and (4);

iv) compute ex ante forecasts and prediction errors for the estimated BVAR (for certain values of θ, λ) and evaluate the performance of the model;

b) estimation of the Bayesian VAR for optimal θ and λ (optimal values are selected on the basis of the MSE criterion) and computation of the forecast;

c) computation of the confidence bounds (95% confidence interval) using the parametric bootstrap method.

Unfortunately, Eviews doesn't provide estimation tools for Bayesian VAR and the subroutines have to be programmed manually.

1.4 Diffusion index models

Diffusion index models are based on the idea that a few common forces are governing the dynamics of all macroeconomic series. The method suggests the forecasting of a single time series (inflation, real GDP, etc.) on the basis of so-called dynamic factors (diffusion indexes), which are averages of contemporaneous values of a large number of macroeconomic time series. These factors usually are not directly observable and therefore have to be extracted (estimated) from the economic data. Diffusion indexes summarize the information contained in a large number of economic time series and, as a result, allow considering a much wider set of predictors compared to other forecasting techniques (for example, VAR). In our analysis we closely follow the methodology developed by Stock and Watson (1998).

The forecasting procedure involves two major steps.

The first step is the preparation of the data set and factor extraction. At this stage, we collect series with a sufficient variation and number of observations, convert the data into a set of stationary series by differencing sufficiently many times (use ADF to determine the order of integration) and log transformation. In most cases, we consider more than 100-200 series. The series are selected judgmentally to represent the main macro categories: real output and income; employment and hours; real retail, manufacturing, and trade sales; consumption; stock prices; exchange rates; interest rates; price indexes; etc. These series are taken from the IMF's Webstract, which is the most complete source of available macroeconomic series. Thereafter, the underlying factors are extracted from the large set of economic data with the use of principal components analysis. Other tools are also available but they either impose unrealistic assumptions or are more computationally demanding (factor analysis).We focus on the 5-6 first principal components of the data set. Further principal components do not contribute much to explaining the variance of dependent variables as 5-6 components explain up to 95% of the total variance in the data. Principal components are computed on the basis of the correlation matrix because variables are in different units. In turn, the correlation matrix is computed for the first differences of the endogenous variables. We do not include lags of principal components in forecasting regression but we do include lags of the left hand-side variable to take into account the possibility of partial adjustment.

Technically, the model can be specified in the following way.

Let Y_t be an N-dimensional multiple time series variable. In our case, this is the vector of endogenous variables defined in the previous subsections. It is assumed that Y_t can be represented by a factor structure:

$$Y_t = A_0 + \Lambda_t * PC_t + \varepsilon_t \ ,$$

where PC_t is the $r \times 1$ common factor, Λ_t is the factor loading, and ε_t is $N \times 1$ idiosyncratic disturbance. Diffusion indexes are interpreted as estimates of the unobserved factors. The estimation procedure is quasi-MLE, in the sense that the estimator is motivated by making strong parametric assumptions: $\Lambda_t = \Lambda_0$ and $\varepsilon_{i,t}$ are i.i.d and independent across series. The total number of 5-6 factors are estimated. In Eviews, the command *df_pc.pcomp pc1 pc2 pc3 pc4 pc5* performs the extraction of the components.

At the second stage, the estimated values of factors PC_t are used in order to construct the forecast. Given the estimated factors, one just needs to regress the variable of interest on these factors (current values and if necessary their lags) and the q-th own lags of the variable of interest. Unfortunately, one cannot use diffusion index models for a q-step ahead forecast directly. To get a q-step ahead forecast, we project the variable of interest on the q-th lags of the estimated principal components. The forecast is then compiled as a stack of q-step ahead forecasts. The standard error is the vector of stacked standard errors of each of the q-step forecasts. To forecast q periods ahead ($q = 1, .., 8$ - two years) we estimate:

$$
\begin{aligned}
Y_t &= A_0 + \alpha_1 PC^{(1)}_{t-q} + \alpha_2 PC^{(2)}_{t-q} + ... + \alpha_5 PC^{(5)}_{t-q} + \beta Y_{t-q} + \\
&\quad + \beta_1 d\log(rgdp\_eu_t) + \beta_2 d\log(ex\_eurousd_t) + \varepsilon_t \ ,
\end{aligned}
$$

where PC_t are the estimated factors and $rgdp\_eu_t$ and $ex\_eurousd_t$ are exogenous variables - the Euro area real GDP and EURO/USD exchange rate, respectively. The number of factors is recursively selected by BIC. Given q, the coefficients of the above-specified equation are estimated by OLS. In fact, we estimate q regressions in order to obtain q-period ahead forecasts.

In addition, we estimate the Diffusion index model with the correction for the intercept (add factoring). We include an add factor in the equation in order to reduce the persistence of the error. In other words, we model the path of the residuals by hand. The procedure includes the following steps:

- estimate the equation of interest;
- take the average of the last four residuals;
- add the average to the intercept;
- forecast with the model that uses the updated intercept;
- repeat the same steps for all eight equations to obtain the forecast for eight periods ahead.

1.5 Discounted Least Squares

The structure of transition economies is evolving at a much faster rate than that of developed countries. As a result, the distant past may reveal no valuable information about the present state of the economy. This argument motivates the estimation of our model by Discounted Least Squares (DLS). Unlike Ordinary Least Squares, this approach assigns higher weights to the most recent data.

The model specification takes the following form:

$$
\{\beta, \alpha_1, ..., \alpha_k\} = \arg\min \left\{ \Sigma_{t=0}^{T} \lambda^{T-t} (y_t - \alpha_1 y_{t-1} - ... - \alpha_k y_{t-k} - \beta z_t)^2 \right\},
$$

where λ is the discount factor.

The forecasting procedure includes the following steps:

a) select an optimal lag structure for the specific discount factor on the basis of BIC;

b) estimate the equation by Weighted Least Squares and compute the forecast;

c) iterate over the range of the discount factor values in order to choose weights minimizing MSE;

d) re-estimate the equation with the optimal discount factor and lag structure; produce the final forecast.

1.6 Forecast on forecast

It is extremely complicated to track all news, changes, and developments happening within a wide array of transition countries. At the same time, other organizations may be able to commit larger resources for forecasting purposes and thus possess additional valuable information. Moreover, these agencies may have insider information, which apparently is not reflected in the historical time series, or some subjective judgments from specialists in the field. Ignoring this additional information may result in significant forecasting error. We include the Economist Intelligence Unit (EIU) and OECD forecasts (if available) in our model. Because forecasts are likely to be highly correlated with other explanatory variables, we run a separate regression. To add dynamics, we include the lags of the dependent variable in the set of regressors.

For each of the endogenous variables we estimate the following "rational forecast" models :

$$y_t = A_0 + \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \beta_1 \hat{y}_{t-1}^{EIU} + \beta_2 \hat{y}_t^{EIU} + \beta_3 \hat{y}_{t+1}^{EIU} + \epsilon_t \qquad (LS),$$

$$y_t = A_0 + \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \beta_1 \hat{y}_{t-1}^{OECD} + \beta_2 \hat{y}_t^{OECD} + \beta_3 \hat{y}_{t+1}^{OECD} + \epsilon_t \qquad (LS).$$

1.7 Univariate

ARIMA models. For some macroeconomic series, e.g. CPI and unemployment rate, the data sources provide both monthly and quarterly observations. The possibility to expand the data set (more observations and higher frequency) enables us to build on a more flexible model representation. As a result, we may enhance the dynamic characteristics of the model and the forecasting outcomes can be improved.

The models specification is presented as follows:

$$y_t = \sum_{s=1}^{OPTIMAL} \alpha_s y_{t-s} + \beta D_t + A_0 + \epsilon_t,$$

where y_t is the stationary series of endogenous variables; D_t is the dummy variable for country-specific events; and the number of lags is selected optimally so as to maximize Schwartz IC. The equation is estimated with LS and the forecast is produced. Finally, we aggregate the monthly predictions into the quarterly series.

Rolling regressions. We also use rolling regressions to forecast the series available at a monthly frequency. The idea and motivation behind this method is similar to DLS. We focus only on very recent subsamples of macroeconomic time series. The disadvantage of this model is that it discards a portion of the information that may be still valuable.

For each country we consider CPI and unemployment rate.

Rolling regression is presented in the form :

$$\{\beta, \alpha_1, ..., \alpha_k\} = \arg\min \left\{ \Sigma_{t=T-Q}^{T} (y_t - \alpha_1 y_{t-1} - ... - \alpha_k y_{t-k} - \beta D_t)^2 \right\},$$

where the optimal lag structure is selected on the basis of BIC; window $Q =$ the last 4 years.

1.8 Merge all forecasts - point estimates and standard errors

We have just described the variety of methods employed in our forecasting procedure. Each model tends to emphasis a few, but not all, aspects of the economy. By pooling results from a variety of techniques, the problem of model misspecification can be reduced. To aggregate the predictions based on different models into the final forecast, we compute a weighted average of forecasts obtained from each model. The weights are the inverses of standard errors. More sophisticated pooling algorithms are possible, but the one we use is the most robust according to Monte Carlo simulations.

2 Data Sources

At present we use a variety of sources. The most important are international organizations such as the IMF and, to lesser extent, the OECD, national statistical offices, and central banks. We can sometimes observe data discrepancy across alternative sources and a choice has to be made. Whenever a series is available from the IMF statistics, we prefer using this series in order to keep the consistency of data across countries. Usually the IMF makes an additional check of the data and employs a unified methodology, so that the overall quality of data tends to be better than those from statistical offices. In addition, some series (for example the Russian unemployment series) have to be reconstructed from external information. Clearly, such reconstructed series are subject to measurement error but they are preferable to highly imprecise official statistics.

Regarding the international exogenous variables, the oil price is taken from the website www.economagic.com and the values for Euro area GDP, the harmonized euro area CPI, and the EURO/USD exchange rate are taken from the ECB.

3 References

1. Stock, James H. , Watson, Mark W., "Diffusion Indexes" NBER Working Paper No. 6702, Aug., 1998.

2. Hamilton, J.D. , "Time Series Analysis" Princeton University Press, 1994.

3. Litterman, Robert B., "Forecasting with Bayesian Vector Autoregression: Five years of Experience", Journal of Business and Economic Statistics, Vol.4, No. 1, Jan., 1986.

4. "EViews 5: User's Guide" Quantitative Micro Software, LLC.