Uncertainty of USA GDP Forecasts Determined by The Variables Aggregation

Mihaela Bratu[1]

The aggregation of the variables that compose an indicator, as GDP, which should be forecasted, is not mentioned explicitly in literature as a source of forecasts uncertainty. In this study based on data on U.S. GDP and its components in 1995-2010, we found that GDP one-step-ahead forecasts made by aggregating the components with variable weights, modeled using ARMA procedure, have a higher accuracy than those with constant weights or the direct forecasts. Excepting the GDP forecasts obtained directly from the model, the one-step-ahead forecasts resulted form the GDP components' forecasts aggregation are better than those made on an horizon of 3 years . The evaluation of this source of uncertainty should be considered for macroeconomic aggregates in order to choose the most accurate forecast.

Keywords: source of uncertainty, forecasts, accuracy, disaggregation over variables, strategy of prediction, DM test

JEL Classification: E21, E27,C51, C53

1. Introduction

One of the sources of forecast uncertainty less depth in the literature is the aggregation of variables that compose the indicator that will be forecasted. Interestingly, no author identify this source together with other sources of uncertainty of forecasts that are based on models. In literature there are studies where the forecasts accuracy is evaluated when the interest variable is modeled using its components. In these studies the variables is also forecasted by aggregating the forecasts of

---

[1] Academy of Economic Studies, Faculty of Cybernetics, Statistics and Economic Informatics, Bucharest,. *Email: mihaela_mb1@yahoo.com*

its components.

The forecasts of macroeconomic aggregates are of interest not only for government, but also for private sector. The accuracy can be improved for forecasts obtained by forecasting aggregate's components, followed by the aggregation of these predictions. The conclusion was stated in literature, but it remains valid only in the context of knowledge of data series used to draw up estimates of the models. Hubrich, K. (2005) showed that the aggregation of forecasts components does not necessarily help in annually forecasting.

2. Literature

There are various uncertainty sources, Vega, M. (2003) recalling the measurement of errors, structural changes in the economy, the uncertainty that is intrinsically generated by the model, subjective adjustments of the models, the exogenous variables. Ericsson, N. (2001) considers that the uncertainty sources are: the forecasted variable, the economic process, based on available data, the model type used to develop forecasts, forecast horizon length.

Clements, M. P., Hendry, D. F. (1998) identify five sources of uncertainty for predictions based on model: the inaccuracy of parameter estimates, the Incorrect specification of the model, the errors in data measurement, the future structural changes in the economy and the future shocks. Clements and Henry (1999) show that structural breaks (the slope or the level breaks) of the data series are a factor of based on model uncertainty forecasts growth. Lanser, D. , Kranendonk, H. (2008) identify four sources of uncertainty of forecasts that are based on models: the uncertainty in the data provided by the institution that collected them, the uncertainty in the series of exogenous variables, the uncertainty in the parameters of behavioral equations and the uncertainty in error terms.

Lanser, D. , Kranendonk, H. (2008) modeled the four sources of uncertainty first theoretically, for each model specifying the corresponding disturbance by probability density. After the theoretical

presentation, the authors assess the sources of uncertainty for Saffier model, the quarterly macroeconomic model of the Dutch Bureau for Economic Policy Analysis. This institution assessed since 1991 the quality of its macroeconomic forecasts based on simulations, producing many works about the exogenous variables, parameters and error models uncertainty. Hendry and Hubrich (2009) consider that one of the causes of forecast failure is the inconsistence of parameters generated by the use of disaggregated data in the absence of structural shocks. Therefore, the aggregation / disaggregation of variables can be considered as a source of forecast uncertainty.

In last years, due to the aggregation of geographical areas, the problem of calculating and forecasting the aggregate indicators was put for each region or member state in case of the Euro zone.

Clements, M. P., Hendry, D. F. (2006) propose instead of the forecasting of an aggregate's components, followed by the forecasts aggregation, to include in a model the variables that compose the aggregate, because the forecasts would be more accurate.

Hendry, DF and Hubrich, K. (2006) lists the authors as Espasa, Senra and Albacete (2002), Hubrich (2005) and Benalal, Diaz del Hoyo, Land, Rome and Skudelny (2004) with important contributions to preview inflation in the euro area. Fair, R., Shiller, C. (1990) performed an analysis similar but for the U.S. GDP. About aggregation and disaggregation in related activity forecasting few authors have important contributions, being recalled by Hendry, DF and Hubrich, K. (2006): Grunfeld and Griliches (1960), Kohn (1982), Lutkepohl (1984, 1987), Pesaran, Pierse and Kumar (1989), Van Garderen, Lee and Pesaran (2000). Granger (1990) puts the issue of aggregation from the time series variables and Lutkepohl (2005) takes into account aggregate forecasts based on VARMA models.The concept of predictability used by Hendry, DF and Hubrich, K. (2006) refers to the connection between variables analyzed and the appropriate data set and was previously used by Diebold and Kilian. Clements, M. P., Hendry, D. F. (2010) and Hendry D.F. and Hubrich, K. (2006, 2009)

are concerned with the assessment obtained by aggregating indicators forecast accuracy of other variables. The data used by them refers to the rate of inflation in the euro area and U.S.

3. Forecasting strategies

Clements, M. P., Hendry, D. F. (2010) specify two forecasting strategies: aggregating forecasts for disaggregate components and direct forecasting of the aggregates.

First, we assess the modification effects of the information set by adding the aggregates of the analyzed macroeconomic indicator. Lack of predictability depends on available information. We consider the variable over which predictions are made having an evolution as: $x_t = f_t(I_{t-1}) + u_t$. In this case, u_t is a non-degenerate and unpredictable vector of random variables in relation to the information set from the past (I_{t-1}). By reducing the information set from I_{t-1} to J_{t-1} forecasts with a lower degree of accuracy will result, even if they remain unbiased, as Clements and Hendry (2005) showed. So, a larger set of information is preferred to be used in order to improve the accuracy. If we start from the conditioned distribution $\left(D_{x^a_{T+1}}(x^a_{T+1} / \cdot) \right)$, the amount of information from the original set increases by disaggregating the variable x^a_T in variables $x^a_{i,T}$ and by adding the aggregates in the set of information. Practically, the new set of information, having a longer length, is J_{t-1}, including the old aggregates and disaggregates. It is considered a scalar x^a_T on which the forecasts are made and this will be split into the following form:

$$x^a_{T+1} \equiv x_{1,T+1} \cdot g_{1,T+1} + x_{2,T+1} \cdot g_{2,T+1} \quad (a)$$

$g_{1,T+1}, g_{2,T+1}$ - Specific weights or ponders which can be fix or may change in time. We know that: $g_{1,T+1} \equiv 1 - g_{2,T+1}$ Assuming that GDP is the aggregated variable, then z_T is the set of variables that contains: aggregate variable with lags, disaggregated components and other variables. Aggregate variable and

its components are represented by: $\begin{array}{l} x_t^a = z_{t-1}^{'}\delta_t + v_t \\ x_{i,t} = z_{t-1}^{'}\gamma_{i,t} + e_{i,t} \end{array}$. Conditional expectation of each component can vary over time and it is equal to the minimum value of square error of predictors: $E_{T+1}[x_{i,T+1}/z_T] = z_T^{'} \cdot \gamma_{i,T+1}$. (b) Introducing the relation (a) in (b) it will result:

$$E_{T+1}[x_{T+1}^a/z_T] = \sum_{i=1}^{2} g_{i,T+1}E_{T+1}[x_{i,T+1}/z_T] = \sum_{i=1}^{2} g_{i,T+1}z_T^{'}\gamma_{i,T+1} = \sum_{i=1}^{2} z_T^{'}\gamma_{i,T+1}^{*} = \lambda_{T+1}^{'}x_T$$

(c)

x_{T+1}^a is predicted starting from $: z_T: E_{T+1}[x_{T+1}^a/z_T] = \delta_{T+1}^{'}z_T$ (d)

The above two relations, (c) and (d) are equivalent, fact that implies the same prediction error: $\delta_{T+1}^{'}z_T = \lambda_{T+1}^{'}z_T, \forall z_T \Rightarrow x_{T+1} - E_{T+1}[x_{T+1}^a/z_T] = v_{T+1}$. In conclusion, the direct prediction of x_{T+1} components is equivalent to forecasts aggregation.

In practice, even if the coefficients of models components or the specific weights change, forecasting the aggregate directly on its components have a higher degree of accuracy than if we aggregate the forecasts components. The explanations of this situation can be related to the fact that certain components of the aggregate can be volatile or that the covariance between them provide stability to the aggregate indicator. Disaggregates can be easily predicted under of an increased stability of the models coefficients or weights. Clements, M. P., Hendry, D. F. (2010) concluded that the aggregation of forecasts through disaggregates is a better solution in terms of accuracy than forecasting the aggregate directly. For forecasting the aggregate it is not indicated the forecasting of its changes, but the inclusion of the lags of disaggregates, which shows that the specific weights of predictions are not necessary in order to aggregate the components forecasts.

4. The evaluation of forecasts performance

Forecast accuracy is a large chapter in the literature aimed at assessing forecast uncertainty. There are two methods used to compare the quality of forecasts: vertical methods (for example, the mean square error of prediction) and horizontal methods (such as distance in time). A comprehensive coverage of the issue taking into account all the achievements of the literature is impossible, but we will outline some important conclusions.

To assess the forecast performance, as well as their ordering, statisticians have developed several measures of accuracy. For comparisons between the MSE indicators of forecasts, Granger and Newbold proposed a statistic. Another statistic is presented by Diebold and Mariano for comparison of other quantitative measures of errors. Diebold and Mariano test proposed in 1995 a test to compare the accuracy of two forecasts under the null hypothesis that assumes no differences in accuracy. The test proposed by them was later improved by Ashley and Harvey, who developed a new statistic based on a bootstrap inference. Subsequently, Diebold and Christoffersen have developed a new way of measuring the accuracy while preserving the cointegrating relation between variables.

Armstrong, J. S. , Fildes, R. (1995) showed that the purpose of measuring an error of prediction is to provide information about the distribution of errors form and they proposed to assess the prediction error using a loss function. They showed that it is not sufficient to use a single measure of accuracy.

Since the normal distribution is a poor approximation of the distribution of a low-volume data series, Harvey, Leybourne, and Newbold improved the properties of small length data series, applying some corrections: the change of DM statistics to eliminate the bias and the comparison of this statistics not with normal distribution, but with the T-Student one. Clark evaluated the power of equality forecast accuracy tests , such as modified versions of the DM test or those used by or Newey and West, based on Bartlett core and a determined

length of data series.

Clements, M. P., Hendry, D. F. (2010) presented the most used accuracy measures in literature, which are described below.

       1.   The specific loss function

Diebold, M., Gunther, R., Tay, S. (1998) started from a loss function $L(a_t, x_{t+1})$, where:

a_t -specific action

$x_{t+1} \rightarrow f(x_{t+1})$ - the future value of a random variable whose distribution is known and f (.)-density forecast

The optimal condition involves minimizing the loss function when the density forecast is $p_{t,1}(x_{t+1})$:

$$a_{t,1}^* = \underset{a_{t,1} \in A}{\arg\min} \int L(a_{t,1}, x_{t+1}) p_{t,1}(x_{t+1}) dx_{t+1}$$

The expected value of loss function is:

$$E[L(a_{t,1}^*, x_{t+1})] = \int L(a_{t,1}^*, x_{t+1}) f(x_{t+1}) dx_{t+1}$$

The density forecast will be preferred above any other density for a given loss function if the following condition is accomplished:

$$E[L(a_{t,1}^*(p_{t,1}(x_{t+1})), x_{t+1})] < E[L(a_{t,2}^*(p_{t,2}(x_{t+1})), x_{t+1})] \,,$$

where $a_{t,i}^*$ − the optimal action for the following forecast: $p_{t,i}(x)$.

Making decisions based on forecast accuracy evaluation is important in macroeconomics, but few studies have focused on this. Notable achievements on forecasts performance evaluation were made in practical applications in finance and in metrology. Recent improvements refer to the inclusion of disutility that is presented in actions in the future states and take into account the entire distribution of forecast. Since an objective assessment of prediction errors cost can

not be made, only general absolute loss functions loss or loss of error squares can be used.

    2. *Mean square forecast error (MSFE) and the second error of the generalized forecast (GFESM)*

The most used measure to assess the forecasts accuracy is the mean square forecast error (MSFE). In case of a vector of variables, a MSFE matrix will be built: $V_h \equiv E[e_{T+h} e'_{T+h}] = V[e_{T+h}] + E[e_{T+h}]E[e'_{T+h}]$, where

e_{T+h} - vector of errors with h steps- ahead-forecast

The trace and the determinant of the mean square errors matrix are classical measures of forecast accuracy.

Generalized forecast error second moment (**GFESM**) is calculated according to Clements, M. P., Hendry, D. F. (1993) as a determinant of the expected value of the forecast errors vector for future moments up to the horizon of interest. If forecasts up to a horizon of h quarters present interest, this indicator is calculated as:

$$GFESM = \left| E \left| \begin{pmatrix} e_{t+1} \\ e_{t+2} \\ ... \\ e_{t+h} \end{pmatrix} \right| \cdot \left| \begin{pmatrix} e_{t+1} \\ e_{t+2} \\ ... \\ e_{t+h} \end{pmatrix}^{T} \right| \right|.$$

e_{t+h} -n-dimensional forecast error of n variables model on horizon h

It is considered that GFESM is a better measure of accuracy, because it is invariant to elementary operations with variables, unlike the MSFE trace and it is also a measure that is invariant to basic operations of the same variables on different horizons of prediction, in contrast with MSFE matrix trace and determinant. Clements, M. P., Hendry, D. F. (1993) showed that the MSFE disadvantages related to invariance models are determined by the lack of invariance indicator non singular linear transformations, that preserves the scale. MSFE comparisons determined inconsistent ranks of forecast performance

of different models with several steps along the variables transformations.

3. *Measures of relative accuracy*

Relative measure for assessing forecast accuracy suppose the comparison of forecast with one of reference, called in literature as „ benchmark forecast" or "naïve forecast ". However, it remains a subjective approach the choice of forecast used for comparison. Problems that may arise in this case are related to: the existence of outliers or inappropriate choice of models on which forecasts are developed, and the emergence of shocks. A first measure of relative accuracy is Theil's U statistic, for which the reference forecast is the last observed value recorded in the data series. Collopy and Armstrong proposed a new indicator instead of U statistics similar (RAE). Thompson improved MSE indicator, proposing a statistically determined MSE (mean squared error log ratio).

Relative accuracy can also be measured by comparing predicted values with those based on a model built using data from the past. The tests of forecast accuracy compare an estimate of forecast error variance derived from the past residue and the current MSFE. To check whether the differences between mean square errors corresponding to the two alternative forecasts are statistically significant the tests proposed by Diebold and Mariano, West, Clark and McCracken, Corradi and Swanson, Giacomini and White are used.

Starting from a general loss function based on predictive ability tests, the accuracy of two alternative forecasts for the same variable is compared. The first results obtained by Diebold and Mariano were formalized, as showed R. Giacomini and H. White (2006), by West, McCracken, Clark and McCracken, Corradi, Swanson and Olivetti, Chao, Corradi and Swanson. Other researchers started from the particular loss function (Granger and Newbold, Leitch and Tanner, West, Edison and Cho, Harvey, Leybourne and Newbold).

5. The assessment of U.S. GDP forecast accuracy using two forecasting strategies

From FRED database (Federal Reserve Economic Database) I downloaded data on the U.S. economy for variables such as GDP, final private consumption, government consumption and investment, net exports. The indicators are expressed in constant prices (billion dollars, 100 = 2005) and the period of registration is 1995-2007. The linear regression models were developed and they are used to make forecasts. There are two types of forecasts: one-year-ahead forecasts and forecasts for 3 years.

Each of forecasts was developed in two specific versions, regarding the specific weights used to aggregate the forecasts of GDP components:

- o With constant weights;
- o With variable weights.

In the version with constant weights, structures of the year chosen as forecast origin, the last year in data series, are used as weights. These weights show the share of consumption, investment and government spending, net exports respectively in GDP of that year.

The evolution of components weights in GDP is described using the autoregressive moving average processes. Forecasts of weights based on these models are presented in **Appendix A**. The models used to make one-step-ahead forecasts were built using EViews and these are presented in **Table 1.** Using data from the period 1995-2007, models for GDP and its components were obtained and used to predict the value of indicator in 2008. Using data from 1995-2008 series models used to forecast GDP in 2009 were developed.

Table 1

Models used for one-year-ahead forecasts

Year for which the forecast is made	The model used for direct forecasting
2008	$PIB_t = -1{,}104 \cdot PIB_{t-3} + 2{,}232 \cdot consum_{t-1} + 3{,}063 \cdot gi_{t-1} + 1{,}25 \cdot \exp\_net_{t-1} + e_t$
2009	$PIB_t = -1{,}354 \cdot PIB_{t-3} + 2{,}239 \cdot consum_{t-1} + 4{,}185 \cdot gi_{t-1} + 0{,}935 \cdot \exp\_net_{t-1} + e_t$
2010	$PIB_t = -1{,}158 \cdot PIB_{t-3} + 2{,}247 \cdot consum_{t-1} + 3{,}319 \cdot gi_{t-1} + 1{,}345 \cdot \exp\_net_{t-1} + e_t$

Year for which the forecast is made	The models used to develop forecasts that will be aggregated
2008	$PIB_t = 1{,}503 \cdot consum_{t-1} + e_{1,t}$ $PIB_t = 5{,}326 \cdot gi_{t-1} + e_{2,t}$ $PIB_t = -21{,}85 \cdot \exp\_net_{t-1} + e_{3,t}$
2009	$PIB_t = 1{,}494 \cdot consum_{t-1} + e_{1,t}$ $PIB_t = 5{,}334 \cdot gi_{t-1} + e_{2,t}$ $PIB_t = -21{,}649 \cdot \exp\_net_{t-1} + e_{3,t}$
2010	$PIB_t = 1{,}483 \cdot consum_{t-1} + e_{1,t}$ $PIB_t = 5{,}311 \cdot gi_{t-1} + e_{2,t}$ $PIB_t = -21{,}929 \cdot \exp\_net_{t-1} + e_{3,t}$

Source: own calculations using EViews.

In **Figure 1** it can be observed large deviations of directly forecasted GDP values to those actually recorded or forecasted by aggregation.

<div align="right">**Figure 1**</div>

<div align="center">

The effective **GDP and the forecasted GDP using the two forecasting strategies (2008-2010)**
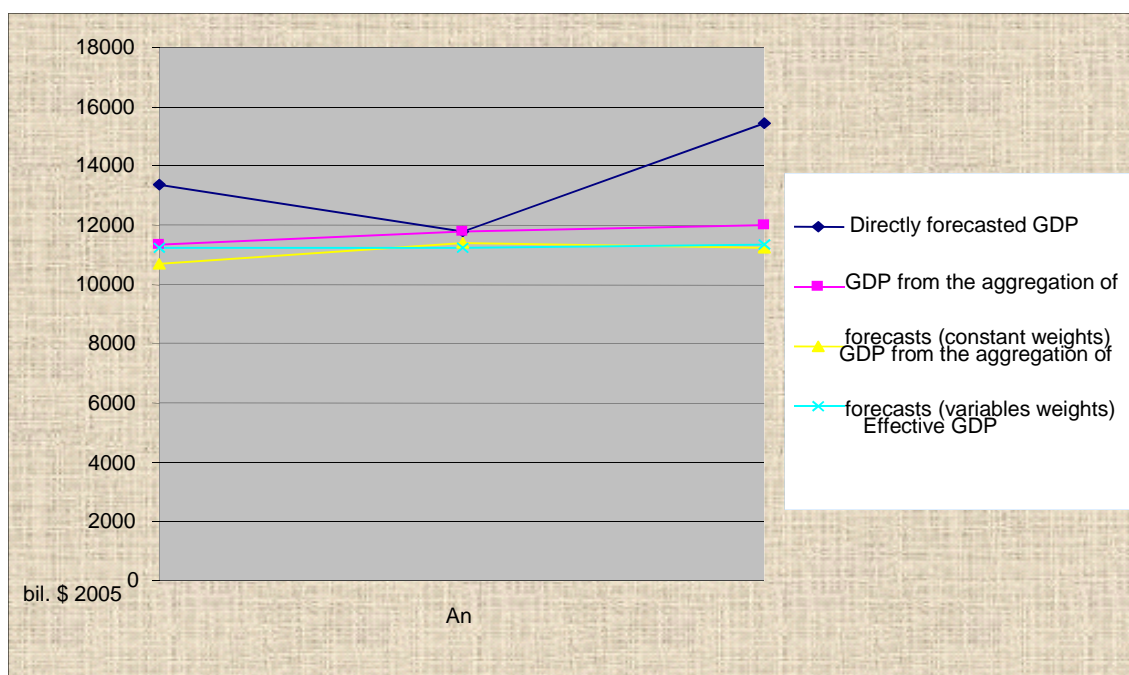
</div>

Chart 1. Gross Domestic Product (billions $ 2005)

Accuracy is assessed by a relative error used in making comparisons between predictions, the percentage error:

$$e_r = \frac{GDP_{effectiv} - GDP_{forecssted}}{GDP_{effectiv}} \cdot 100 \,.$$

RJ Hyndman and AB Koehler (2005) showed that the

percentage error can be used to calculate several indicators, including mean absolute percentage error-MAPS. For a one-step-ahead forecasts made on the horizon 2008-2010, the smallest mean absolute square error registers the GDP forecasts obtained by aggregation with variable weights. For forecasts on 3 years, the ones with constant weights have the highest degree of accuracy, achieving a value of 10, 69% for MAPE, unlike a value around 17% for the other forecasts.

As for the one-step-ahead forecasts and those on 3 years, the value of directly forecasted GDP is higher than the one of forecasts obtained from aggregating the GDP components. However, the higher mean square error for one-step-ahead forecasts is registered for directly predicted GDP and the lower for forecasted GDP using variable weights. The GDP forecasted values resulted applying the two strategies for one step ahead forecasts and those on three years, and the values of RMSM and MAPS are presented in Table 2 and Table 3.

Table 2

One-step-ahead forecasts of USA GDP in 2008-2010

Year	Directly forecasted GDP (bil. dolars 2005)	Forecasted GDP by aggregating the components' forecasts (constant weights)	Forecasted GDP by aggregating the components' forecasts (variable weights)
2008	13383,94	11333,09	10684,82
2009	16089,43	11756,92	11386,06
2010	15459,61	12009,33	11210,66

RMSM	3884,819	493,6839	331,5005
MAPE	32,97 %	3,87 %	2,5 %
rel_RMSE	1,95	0,35	0,16

Source: own calculations using EViews.

Table 3

The forecasts on 3 years of U.S. GDP (2008-2010)

Year	Directly forecasted GDP (bil. dolars 2005)	Forecasted GDP by aggregating the components' forecasts (constant weights)	Forecasted GDP by aggregating the components' forecasts (variable weights)
2008	13383,94	11333,09	10684,82
2009	13285,47	9985,936	8354,037
2010	12922,05	9079,8	8962,4
RMSM	1992,714	1423,552	2123,93
MAPE	17,19 %	10,69 %	17,11 %

Source: own calculations using EViews.

For forecasted GDP by aggregating its components with variable weights there is a tendency of underestimation, while the directly forecasted GDP is overestimated. For forecasts developed on a three years horizon, the GDP forecasts resulted by aggregation of components forecasts with constant weights have the lowest accuracy, because the RMSM is the lowest. Percentage error values are

presented in **Table 4** and **Table 5.** The calculated relative errors are large for direct forecasts of GDP and smaller in other cases. The lowest relative error was registered in 2008 for predicted GDP by aggregating the forecasts of GDP components (constant weights) and the largest one for directly forecasted GDP in 2009.

Table 4

Relative errors (errors percentages) of one-step-ahead forecasts (%)

	2008	2009	2010
Directly forecasted GDP (billions dollars 2005)	-19,35 %	-43,42 %	-36,14 %
Forecasted GDP by aggregating the components' forecasts (constant weights)	-1,06 %	-4,80 %	-5,75 %
Forecasted GDP by aggregating the components' forecasts (variable weights)	4,72 %	-1,50 %	1,28 %

Source: own calculations using the data from **Table** *2.*

Table 5

Relative errors (errors percentages) of 3-years horizon forecasts (%)

	2008	2009	2010
Directly forecasted GDP (billions dollars 2005)	-19,35 %	-18,43 %	-13,79 %
Forecasted GDP by aggregating the components' forecasts (constant weights)	-1,06 %	10,99 %	20,04 %
Forecasted GDP by aggregating the components' forecasts (variable weights)	4,72 %	25,53 %	21,08 %

Source: own calculations using the data from **Table 3.**

The relative errors of forecasted GDP by aggregating the forecasts components (variable weights) became increasingly smaller. Most relative errors values are negative, showing a tendency of overestimation of forecasted values from those actually registered.

A generalization of Diebold-Mariano test (DM) is used to determine whether the MSFE matrix trace of the model with aggregation variables is significantly lower than that of the model in which the aggregation of forecasts is done. If the MSFE determinant

is used, the DM test can not be used in this version, because the difference between the two models MSFE determinants can not be written as an average. In this case, a test that uses a bootstrap method is recommended. The DM statistic is calculated as:

$$DM_t = \frac{\sqrt{T} \cdot [tr(MSFE_{aggregated\_model})_h - tr(MSFE_{aggregated\_forecasts\_model})_h]}{s} =$$

$$\frac{1}{s} \cdot \sqrt{T} \cdot [\frac{1}{T} \sum_{t=1}^{T} (em_{1,h,t}^2 - er_{1,h,t}^2 - er_{2,h,t}^2 - er_{3,h,t}^2)]$$

T-number of years for which forecasts are developed

$em_{i,h,t}$ – the h-steps-ahead forecast error of variable i at time t for the aggregated model

$er_{i,h,t}$ – the h-steps-ahead forecast error of variable i at time t for the model with aggregated forecasts

s- the square root of a consistent estimator of the limiting variance of the numerator

The null hypothesis of the test refers to the same accuracy of forecasts. Under this assumption and taking into account the usual conditions of central limit theorem for weakly correlated processes, DM statistic follows a standard normal asymptotic distribution. Pentru varianţă se utilizează estimatorul Newey-West cu lagul de trunchiere a parametrului de h-1. For the variance the Newey-West estimator with the corresponding lag-truncation parameter set to $h - 1$ is used.

The DM test was applied both for the version with constant specific weights of GDP components and for the one with variable weights for one-step-ahead forecasts. In the first case, the value of DM statistic (0.77) is lower than the critical one, so it if we use constant weights in the forecasts aggregation model we get the same accuracy as if we directly forecast the GDP. If we use variable weights, the DM statistic value (29.0704) is greater than the critical value, so the accuracy of direct forecasts differs significantly from the one obtained by aggregating the forecasts with variable weights. The forecasts based

on aggregated model have a lower degree of accuracy than those obtained by aggregating the forecast with variable specific weights.

Another possibility is to apply the CPA test in MatLab, which leads to the same result. DM test statistic is modified so that another measure of forecasts accuracy is used instead of MSFE, namely GFESM. The results are the same.

Hyndman, R. J., Koehler, A.B. (2005) proposed in comparisons the use of relative measures of accuracy which are independent of measurement scale of the indicator, namely, the use of relative RMSE, which is calculated as: $rel\_RMSE = \dfrac{RMSE}{RMSE_b}$, where $RMSE_b$ is the RMSM of the benchmark model.

A subunit value of indicator shows that the forecast to compare is better than the compared one, in terms of accuracy. This indicator is used to compare the h steps-ahead-forecasts and the ones on 3 years, which are chosen as reference forecasts. **Table 2** shows that only for directly forecasted GDP on 3 years horizon the forecasts are better than the one-step-ahead ones. In case of GDP forecasts obtained by aggregating the components, the one-step-ahead forecasts are the most accurate.

6. Conclusions

After the empirical study of GDP forecasts the following conclusions resulted:

GDP forecasts obtained by aggregating the components' forecasts with variable weights using ARMA models have the highest degree of accuracy.

Moreover, one-step-ahead forecasts obtained by components aggregation with variable weights are better than the 3-years horizon forecasts.

For forecasts of indicators resulted from aggregation the evaluation of aggregation as a source of uncertainty and the choice of most accurate forecasting strategy are recommended.

References

Armstrong, J. S. , Fildes, R. (1995), "On the selection of Error Measures for Comparisons Among Forecasting Methods", Journal of Forecasting, 14 (1): 67-71

Clements, M. P., Hendry, D. F. (1995), "Forecasting in cointegrated systems", Journal of Applied Econometrics,10 (1): 127–146

Clements, M.P., Hendry, D.F. (2003), "Evaluating a Model by Forecast Performance", Cambridge University Press, Cambridge

Clements M. P., Hendry D.F. (2010), "Forecasting from Mis-specified Models in the Presence of Unanticipated Location Shifts", Discussion Paper Series of Department of Economics, 484 (1): 134-140

Ericsson N. (2001), "Forecast Uncertainty in Economic Modeling", MIT Press, Cambridge

Giacomini, R., White, H. (2006), "Tests of conditional predictive ability", Econometrica, 74 (6): 1545-1578

Hendry, D. F., Hubrich, K. (2006), "Forecasting economic aggregates by disaggregates", Working Paper Series of European Central Bank, 589 (1):1-32

Hendry, D. F., Hubrich, K. (2009), "Combining disaggregate forecasts versus disaggregate information to forecast an aggregate", Journal of Business and Economic Statistics, DOI:10.1198/jbes.2009.07112

Hubrich, K. (2005), "Forecasting euro area inflation: Does aggregating forecasts by HICP component improve forecast accuracy?", International Journal of Forecasting, 21(1): 119–136

 Hyndman, R. J., Koehler, A.B. (2005), "Another Look at Measures of Forecast Accuracy", International Journal of Forecasting, 22(4):679-688

 Lanser, D., Kranendonk, H. (2008), "Investigating uncertainty in macroeconomic forecasts by stochastic simulation", CPB Netherlands Bureau for Economic Policy Analysis, Amsterdam

Vega M. (2003), "Policy Makers Priors and Inflation Density Forecasts", Working Paper of Bank of Peru, 12(2):1-32

FRED (2011), Federal Bank of Saint Louise Data base, http://research.stlouisfed.org/fred2/categories/18 , Accessed in July 2011.

APPENDIX A

Models used to predict variable weights

Period	Variable weights
1995-2007	$g_{C_t} = 0{,}9268 \cdot e_{t-1} + e_{t1}$ $g_{GI_t} = 0{,}9895 \cdot g_{GI_{t-1}} + e_{t2}$ $g_{\exp\_net_t} = 1{,}048 \cdot g_{\exp\_net_{t-1}} + e_{t3}$

Source: own calculations using EViews.