4. How Good Are the Growth and Inflation Forecasts for Slovenia?

Timotej JAGRIČ[1]
Jani BEKŐ[2]

Abstract

This paper examines the quality of the macroeconomic forecasts of six institutions that regularly publish forecasts for Slovenia. The analysis focuses on an evaluation of the quality of forecasts for the real and nominal growth of GDP and for the average annual inflation rate for the period from 1997 to 2009. The quality of forecasts for selected macroeconomic variables was evaluated based on five groups of criteria: statistical measures of accuracy, comparison with the results of naive models, trace test, sign test and statistical tests of the unbiasedness and efficiency of forecasts. The results of the analysis do not provide an "absolute winner", but they do indicate the features of particular forecasts. It is also clear that the developers of models have until now most likely given priority to reducing forecasting errors, while neglecting the congruence in the direction of trend between the forecast and the actual result. The latter criterion in particular is very important for effective economic policy making.

Keywords: forecasts of macroeconomic variables, loss functions, measures of quality and unbiasedness, evaluation, economic policy

JEL Classification: C53, E37

1. Introduction

Any future-oriented measure taken by economic policy makers or decision made by economic agents under conditions of uncertainty is based on expectations, and hence based directly or indirectly on economic forecasts. The responses of economic agents to changes in expectations can be highly varied. Thus, for instance, the forecast of a cyclical downturn in households prompts greater caution in personal and investment consumption. The response of economic policy makers depends on the goals defined. Anti-cyclical measures may be triggered and the operation of automatic stabilizers accelerated, which increases public consumption while concurrently decreasing tax

[1] Faculty of Economics and Business, University of Maribor, Razlagova 14, 2000 Maribor, Slovenia. Phone: +38622290343; fax: +38622516141; e-mail: timotej.jagric@i-mb.si.
[2] *Faculty of Economics and Business, University of Maribor, Razlagova 14, 2000 Maribor, Slovenia. Phone: +38622290266; fax: +38622516141; e-mail: jani.beko@uni-mb.si.*

revenues. If economic policy makers wish to preserve the balance in public finances, they will try to compensate for the loss of tax revenue by cutting back on budget items. In this way they will strengthen the cycle in the short term but this can have the effect of contributing to greater growth over the long term. To the extent that reactions among economic agents due to these changes in expectations regarding future economic activities are synchronized, a credible forecast can have an important effect: depending on the reaction of economic agents, the forecast can be either self-fulfilling or self-destroying (Baumgartner, 2002a).

It is thus clear that it is very difficult to evaluate the quality of a forecast due to the way it is incorporated into the behavior of economic agents (Diebold and Mariano, 1995). The simplest way is of course to make a direct comparison of the forecast with the actual result: a good forecast is one which displays a high degree of congruence with what actually happened. In cases where the forecast does not have an effect on the outcome, such an approach is suitable. However, in economic forecasting this is not always so: indeed, in extreme cases the forecast can have such a strong influence on economic agents that the realization is based on completely different conditions. For this reason an ex-post comparison does not show a high degree of congruence. Nevertheless, this does not necessarily mean that such a forecast was bad or unreasonable, since the alternative change in the direction of events could not have been foreseen (Baumgartner, 2002a)..

Furthermore, as suggested by Baumgartner (2002a), in judging the quality of a forecast we need to consider that we are usually dealing with a conditional forecast which is valid only under certain assumptions. For a small, open economy such as Slovenia's, these assumptions are usually dependent on an external environment (economic trends in the EU and USA, changes in exchange rates, costs of raw materials and energy products, etc.). Determining the external environment is thus one of the most significant sources of uncertainty in forecasts for a small open economy.

Likewise, the forecast result is also dependent on the structure of the econometric model, which includes assumptions about the performance of the economy. If these assumptions are incorrect, the forecast will be poor even if in individual cases the error is equal to zero. At the same time, differences between the actual result and a forecast based on suitable information about the external environment show weaknesses in the structure of the econometric model (Baumgartner, 2002a).

These difficulties often arise due to divergent goals of forecasting. Economic policy makers and institutions which prepare forecasts may pursue different goals. This is reflected in the selection of the loss function – the selection of the criterion for evaluating the divergence of the forecast from the realization. For a rational (independent) forecast it is important that errors are minimal, and that positive and negative errors are equally weighted, which implies a symmetric loss function. In practice, it is frequently the case that errors are differently weighted. Thus a late forecast of a downturn can be a greater cost to economic policy makers than a late forecast of an upturn. In such a case more cautious forecasts are favored. This would point to an asymmetric loss function carrying a higher sanction for the overestimation of the actual outcome (i.e., a negative forecasting effort). Thus monetary policy, which

pursues the goals of price stabilization, usually has an asymmetric loss function. However, in the case of real GDP (or potential output) this function is asymmetric in the opposite direction: an underestimated forecast of real GDP growth (and also inflation) has greater weight in the loss function and is for this reason more heavily sanctioned. In the cases mentioned it is necessary to consider different loss functions in the evaluation of the quality of forecasts. Since economic policy makers have very different loss functions for each of the key macroeconomic variables, this means that in a discussion of the evaluation of the quality of forecasting, arguments in favor of selecting a suitable loss function dominate (Diebold, 2001 and Baumgartner, 2002a).

The situation is even more complicated since the statistical properties of errors estimated based on asymmetrical loss functions are sometimes indeterminate (Christoffersen and Diebold 1996; 1997). This does not, however, apply to the evaluation of forecasts considering symmetric loss functions, which has been statistically sufficiently well analyzed, and which allow a comparison over time and between countries. Due to the challenges cited above, we will base our evaluation of macroeconomic forecasts solely on a symmetrical loss function. At the same time, discussion to date indicates that there is no absolute measure of quality of a forecast (Baumgartner, 2002a). Thus as a reasonable approach we offer an analysis in which we attempt to evaluate the forecasts of different institutions over different time periods, taking into account a large number of statistical criteria for the quality of a forecast.

This paper follows a series of documented studies (Andersson et al., 2007; Ash et al., 1998; Batchelor, 2001; Baumgartner, 2002a and 2002b; Blix et al., 2001; Lenain, 2002; Öller and Barot, 2000) but also brings two new findings as result of a project prepared for Institute of Macroeconomic Analysis and Development (Jagric, 2010). First, it verifies the quality of forecasting economic growth and inflation trends using a wide range of statistical criteria for the quality of forecasting in the case of Slovenia, which in studies to date has not been the subject of systematic examination. And second, in this article we look at the evaluation of economic forecasts from six institutions also taking into account the influence of the global financial and economic crisis on the quality of forecasts. The article consists of four sections in addition to the introduction. In the second section of the article the basic characteristics of the data base are described. The third section is devoted to a concise survey of the testing methods used. Empirical results are presented in the fourth section. The fifth and final section provides the main findings of the analysis.

2. Available Data Base

This analysis was performed on forecasts covering the period from 1997 through 2009. We decided to start with the year 1997 since from that year onwards domestic analytical institutions in Slovenia have officially and systematically published forecasts of economic trends for Slovenia. At the same time, the time period is one in which the Slovenian economy has passed through an entire business cycle (the actual number of cyclical turns and their timing is a matter of discussion, since we do not yet have official dating of the cyclical turns). Because the data also include the years 2008 and

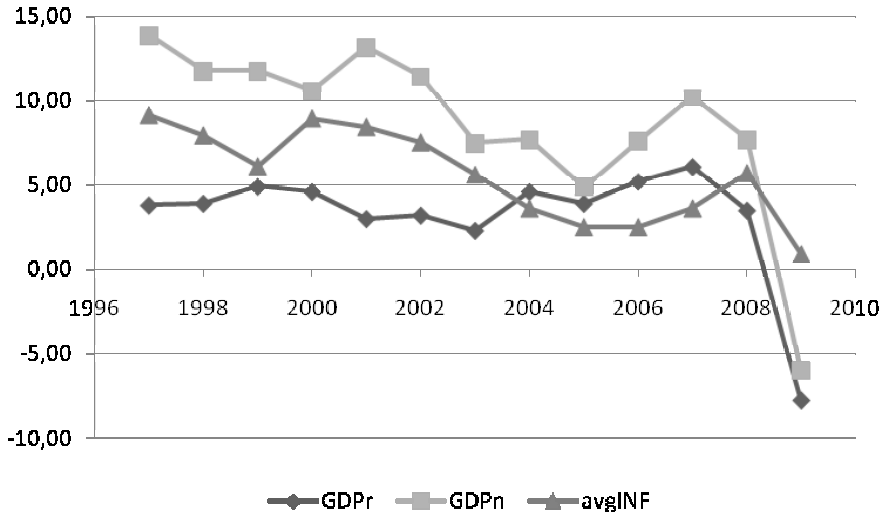
2009, it is possible to study the effects of the global financial crisis and resultant recession.

Data were collected for several institutions which do economic forecasting and forecasting of particular key macroeconomic variables for Slovenia. In this way it was possible to take into account the influence of the goals pursued by various analytical institutions, the following of which are included in the data base: the Institute of Macroeconomic Analysis and Development (IMAD), the Bank of Slovenia (BS), Economic Outlook, Analysis and Forecasts at the Chamber of Commerce and Industry of Slovenia (SKEP), the European Commission (EC), the International Monetary Fund (IMF) and the Vienna Institute for International Economic Studies (WIIW). It should be pointed out that we do not have data for all these institutions for the whole of the period from 1997 to 2009. The actual time frame will be specified in the presentation of the results of the analysis.

Since the goals of forecasting can vary among different economic variables, in this analysis we include two key variables for the Slovenian economy: gross domestic product (GDP) and inflation (Figure 1). In the framework of the analysis of forecasts of GDP we look at forecasts for real and nominal growth of GDP. Some problems appear in the evaluation of the accuracy in forecasting inflation, since different institutions forecast this variable based on different definitions of price increases. In particular cases in the period observed there have also been changes in the definition of variables, which makes performing all the statistical tests impossible.

Figure 1

Changes in Selected Macroeconomic Variables for the Slovenian Economy (in %)



Note: GDPr – real rate of growth in gross domestic product, GDPn – nominal rate of growth in gross domestic product, avgINF – average annual rate of growth of inflation.

The quality of the forecast is of course also dependent on the timing of its release. Institutions which release forecasts later have a certain informational advantage. All institutions prepare forecasts twice a year, but at different times: IMAD and BS usually prepare forecasts in March and September, SKEP in May and November, the EC and IMF in April and October, and WIIW in February and July. The source of the forecasts are the official publications of the institutions cited:

- Spring Report (1997-2006), Autumn Report (1997-2006), Spring Forecast of Economic Trends (2007-2009) and Fall Forecast of Economic Trends (2007-2009) for IMAD;
- Realization of Short-Term Direction of Monetary Policy (2002-2004), Direction of Monetary Policy of the Bank of Slovenia (2001), Report on Monetary Policy (2004-2006) and Report on Price Stabilization (2007-2009) for the BS;
- Economic Trends (1997-2009) for Economic Outlook, Analysis and Forecasts at the Chamber of Commerce and Industry of Slovenia (SKEP);
- Country Report (2003-2004) and World Economic Outlook (1999-2009) for the IMF;
- Economic Forecasts for the Candidate Countries (2000-2002) and Economic Forecasts (2002-2009) for the EC;
- WIIW Monthly Report (2003-2009) for the Vienna Institute for International Economic Studies.

Based on the publications of these institutions we performed an analysis of the accuracy or quality of four types of forecasts: for the spring forecast for the current year (Spring (t)), for the fall forecast for the current year (Fall (t)); for the spring forecast for the following year (Spring (t+1)) and for the fall forecast for the following year (Fall (t+1)).

3. Methods

As already mentioned in the introduction, the users of forecasts as well as those who prepare them pursue different goals. For this reason evaluations of the quality of a forecast for a particular variable may differ significantly. In order to avoid bias, we evaluate the quality (accuracy) of the forecast based on a number of criteria which are divided into five groups:

- statistical measures of accuracy which take into account error of forecast;
- a comparison with naive models;
- trace test;
- sign test, which tests the congruence between the forecast and the realization in terms of the direction of the trend;
- statistical tests of unbiasedness and efficiency of the forecast.

In the continuation, a significance level $\alpha = 0.05$ is usually used for statistical tests. This means that there is a 5% probability of error that the null hypothesis is rejected even if correct. Precise significance levels are given under results. In order to show the methods used we will first provide a definition of basic variables:

P_t - forecast for year t,

R_t - realization for year t,

T - number of observations.

The mean error represents the mean deviation of the forecasted value from the realization, which indicates the bias in the forecast:

$$ME = \frac{1}{T}\sum_{t=1}^{T}(R_T - P_T) \quad (1)$$

If the errors – defined as the difference between the realization and the forecasted value – are distributed evenly, then this value is close to zero. A positive value signals a tendency for the forecast on average to underestimate the realization. The method of testing the statistical properties of this type of bias is presented below.

An optimal forecast should exhibit a lower variance than the realizations, since the forecast, unlike the actual value, includes no irregular component (Granger and Newbold, 1977). In this regard, Aiginger (1979) also points to a revealed smoothing tendency – if information is scarce, it seems plausible to expect an average development. The more unclear the trend of developments, the closer the expectations will be to the average realization in the past. The smoothing tendency is supported by the fact that it is stronger in forecasts with an extended forecast horizon than in short term forecasts. Generally speaking we can assert that a good forecast is one in which the ratio between the standard deviation of the forecast and the realization is smaller than one and decreases with a longer forecast horizon.

In calculating the mean error, positive and negative errors may cancel each other out. Therefore based on this criterion we cannot provide a more detailed analysis of the quality. The mean absolute error (MAE) and the mean square error (MSE) are measurements which can also be used for calculating the accuracy of the forecast:

$$MAE = \frac{1}{T}\sum_{t=1}^{T}(|R_T - P_T|)$$

$$MSE = \frac{1}{T}\sum_{t=1}^{T}(R_T - P_T)^2$$

$$MSE = (\bar{R} - \bar{P})^2 + (s_R - s_P)^2 + 2(1 - r)s_R s_P \text{ - decomposition 1}$$

$$MSE = (\bar{R} - \bar{P})^2 + (s_R - rs_P) + 2(1 - r^2)s_R{}^2 \text{ - decomposition 2}$$

$$s_P^2 = \frac{1}{T}\sum_{t=1}^{T}(P_t - \bar{P})^2 \text{ - variance of a forecast}$$

$$s_R^2 = \frac{1}{T}\sum_{t=1}^{T}(R_t - \bar{R})^2 \text{ - variance of realization} \quad (2)$$

$$r = \frac{\frac{1}{T}\sum_{t=1}^{T}(R_t - \bar{R})(P_t - \bar{P})}{s_R s_P} \text{ - correlation coefficient between forecast and realization}$$

$$\bar{P} = \frac{1}{T}\sum_{t=1}^{T}P_t \text{ - mean of forecasts}$$

$$\bar{R} = \frac{1}{T}\sum_{t=1}^{T}R_t \text{ - mean of realizations}$$

The two measurements differ according to the weighting of the errors entered in the calculation: with MAE they are weighted linearly, with MSE to the power of 2. Thus, in the case of MSE we give a greater emphasis to greater errors. Often instead of MSE the root mean square error (RMSE) is used.

$$RMSE = \sqrt{MSE} \quad (3)$$

The advantage of the measurement presented lies in the fact that the calculated value is in the same units as the forecast and the realized values of the variable observed. Independent of the selected measurement it holds that the smaller the calculated value of the measurement, the better the forecast.

The mean square error can be broken down into several parts (Equation 2). In this way it is possible to estimate the accuracy of the forecast in more detail. Here it holds that (Baumgartner, 2002a):

$$UM + US + UC = 1$$
$$UM + UR + UD = 1$$
$$UM = \frac{(\bar{R}-\bar{P})^2}{MSE}$$
$$US = \frac{(s_R - s_P)^2}{MSE} \quad (4)$$
$$UC = \frac{2(1-r)s_R s_P}{MSE}$$
$$UR = \frac{(s_P - r s_R)^2}{MSE}$$
$$UD = \frac{(1-r^2)s_R^2}{MSE}$$

A good forecast is characterized by a small value for the indicator of bias (UM), variance (US) and regression (UR), while the indicators of covariance (UC) and distribution (UD) are close to 1 (Theil, 1966; 1971).

In the examples cited there is no indicator of the difficulty of the forecast in the calculation. Here it is known that it is easier to forecast variables which have a low variability than variables which have a high variability. Thus, we may calculate the standardized value of the mean square error and the mean absolute error. We carry out standardization if both measurements are divided by the standard deviation of the actual result. This kind of correction allows a comparison of forecasts for different variables (Baumgartner, 2002a).

Theil (1966; 1971) developed several measurements which have a function similar to those already presented. Their advantage is that they are standardized to a value of 1, enabling easier interpretation. Here we use the general principle according to which we compare the RMSE of different forecasts. In our particular case we will compare the forecasts of different institutions with the forecasts of two naive models proposed by Theil (1966; 1971).

The test statistic W assumes a naive forecast based on the hypothesis that there will be no change in the rate of growth of the observed phenomenon: the rate of growth of the observed variable from the previous period is used for the current forecast (Equation 5). This approach also applies to forecasts with a longer time horizon (Baumgartner, 2002a).

$$W = \sqrt{\frac{\sum_{t=h+1}^{T}(R_t - P_t)^2}{\sum_{t=h+1}^{T}(R_t - R_{t-h})^2}} \quad (5)$$

A naive forecast in the case of the test statistic U is somewhat different (Equation 6): it assumes that there will be no change in the value of the variable (level) which is being

forecasted. This means that the level of growth of the phenomenon in question for the period of the forecast is equal to zero.

$$U = \sqrt{\frac{\sum_{t=1}^{T}(R_t - P_t)^2}{\sum_{t=1}^{T}R_t^2}} \quad (6)$$

For users of forecasts it is sometimes more important that the forecast indicate the correct direction of development of the phenomenon in question than to minimize error. In evaluating the congruence in direction we can make use of a contingency table (Table 1).

Table 1

Contingency Table

		Forecasts		
Realizations		≥ 0	< 0	Frequencies
	≥ 0	a	b	a+b
	< 0	c	d	c+d
Frequencies		a+c	b+d	a+b+c+d

Using a non-parametric test we find that the direction of the forecast is congruent with the direction of the realization (Ash, 1998). The null hypothesis test is that the signs of the direction of the forecast and the signs of the direction of the realization are mutually independent. It makes sense for the signs of the direction of the forecast to be as similar as possible to those which are obtained for the direction of the realization, hence the goal of the test is the rejection of the null hypothesis. The ratio of congruence (ER) is an indicator of the degree of congruence within the observed period:

$$ER = (a+d)/(a+b+c+d) \quad (7)$$

The measurement covers values between 0 and 1. A good forecast should have a value of more than 0.5 and a *p* value below 0.05. Probability is estimated based on the test statistic (Bleymüller *et al*., 1994):

$$\chi^2 = \frac{(a+b+c+d)(ad-bc)^2}{(a+b)(c+d)(a+c)(b+d)} \quad (8)$$

An unbiased forecast should have the same mean value as the realization and the mean value of error, equal to zero. The standard procedure for checking unbiasedness and efficiency is an estimate of the regression model in which the dependent variable is the value of the realization, and the explanatory variable is the value of the corresponding forecast (Mincer and Zarnowitz, 1969):

$$R_t = \beta_0 + \beta_1 P_t + u_t \quad (9)$$

Using the F test we can test the null assumption that the forecast is unbiased. Holden and Peel (1990) showed that this criterion is a sufficient, but not a necessary condition for an unbiased forecast. For this reason they proposed an upgrading of the test. The key problem lies in the fact that we need a sufficiently large sample for its execution. Since in our case this condition is not fulfilled, we will carry out only the basic test and only for data from IMAD.

4. Explanation of Results

We begin the presentation of results with an analysis of the success of the spring and fall forecasts of real GDP growth for the current year (Table 2). All analyses in the continuation were prepared according to the same principle: test statistics and criteria of quality were calculated for three different periods:

- from 2002 to 2008 (eliminating the effect of the global financial and economic crisis);
- from 2002 to 2009;
- from 1997/98 to 2009 (this period was used only for data from IMAD, which has the longest tradition of systematic forecasting of the variables covered within the sample selected).

The first criterion is mean error. Due to the method of calculation of this criterion we had to be cautious in its use. In principle we find that the indicator for the IMAD data shows an underestimation in forecasting the current real growth in the period from 2002 to 2008 for spring forecasts. The effect of the crisis is very illustrative: there was clearly an overestimation of current economic growth in the spring forecast, which was not the case for the fall one. Taking into account this criterion, IMAD forecasts rank among the better ones, especially if we consider the timing of the forecasts and the available information about the state of the economy.

The criteria of MAE and MSE show that the spring forecast of IMAD was the best in the two comparable periods. The fall forecast was also among the better ones according to these two criteria. The forecast of the EC stands out, but it has an informational advantage as a result of the timing of its release. The same results are also shown for the criterion RMSE. The effect of the economic crisis is shown in all three criteria, since their value rose significantly for the period from 2002 to 2009. This effect is visible in all the institutions analyzed.

The measures MAE/SD and RMSE/SD have a value less than one in the spring and fall forecasts. IMAD's spring forecast is the best in this instance as well, and the fall forecast is among the better ones. They are directly comparable to the forecasts of the BS, which are released synchronously with those of IMAD. Here it is interesting that IMAD achieves better results for the period including the economic crisis than the BS.

Table 2

Analysis of the Accuracy of Spring and Fall Forecasts of Real GDP Growth for the Current Year

Period	Institution	Spring forecast						Fall forecast					
	Criterion	IMAD	BS	SKEP	EC	IMF	WIIW	IMAD	BS	SKEP	EC	IMF	WIIW
02-08		0.24	0.39	0.33	0.37	0.42	0.23	-0.03	0.10	-0.04	0.16	0.34	0.43
02-09		-0.26	-0.39	-0.38	-0.23	-0.27	-0.78	-0.09	-0.05	-0.14	0.09	-0.09	-0.10
97-09	ME	-0.20						0.04					
02-08		0.76	0.79	0.79	0.89	0.88	0.94	0.43	0.50	0.59	0.41	0.57	0.71
02-09		1.14	1.41	1.35	1.33	1.41	1.80	0.44	0.58	0.61	0.41	0.89	1.10
97-09	MAE	0.95						0.49					
02-08	MSE	0.78	0.93	0.89	1.11	0.97	1.17	0.35	0.34	0.49	0.25	0.45	0.63
02-09		2.48	5.02	4.29	3.39	4.10	8.63	0.34	0.45	0.51	0.24	1.59	2.35

Period	Institution	Spring forecast						Fall forecast					
	Criterion	IMAD	BS	SKEP	EC	IMF	WIIW	IMAD	BS	SKEP	EC	IMF	WIIW
97-09		1.84						0.38					
02-08		0.88	0.96	0.94	1.05	0.99	1.08	0.60	0.59	0.70	0.50	0.67	0.79
02-09		1.58	2.24	2.07	1.84	2.03	2.94	0.58	0.67	0.71	0.49	1.26	1.53
97-09	RMSE	1.35						0.62					
02-08		0.64	0.66	0.66	0.74	0.74	0.79	0.36	0.42	0.49	0.35	0.48	0.60
02-09		0.28	0.34	0.33	0.32	0.34	0.44	0.11	0.14	0.15	0.10	0.22	0.27
97-09	MAE/SD	0.29						0.15					
02-08		0.74	0.81	0.79	0.88	0.83	0.91	0.50	0.49	0.59	0.42	0.56	0.66
02-09		0.38	0.55	0.51	0.45	0.49	0.72	0.14	0.16	0.17	0.12	0.31	0.37
97-09	RMSE/SD	0.41						0.19					
02-08		0.58	0.64	0.63	0.70	0.65	0.72	0.40	0.39	0.46	0.33	0.44	0.53
02-09		0.37	0.53	0.49	0.43	0.48	0.69	0.14	0.16	0.17	0.12	0.30	0.36
97-09	W	0.39						0.18					
02-08		0.21	0.22	0.22	0.25	0.23	0.25	0.14	0.14	0.16	0.12	0.16	0.18
02-09		0.32	0.46	0.43	0.38	0.42	0.60	0.12	0.14	0.15	0.10	0.26	0.32
97-09	U	0.30						0.14					
02-08		0.08	0.16	0.12	0.12	0.18	0.04	0.00	0.03	0.00	0.10	0.26	0.29
02-09	UM	0.03	0.03	0.03	0.01	0.02	0.07	0.02	0.01	0.04	0.03	0.00	0.00
02-08		0.54	0.37	0.59	0.45	0.35	0.40	0.11	0.13	0.06	0.00	0.07	0.24
02-09	US	0.84	0.89	0.91	0.84	0.88	0.86	0.12	0.42	0.16	0.12	0.81	0.89
02-08		0.39	0.47	0.28	0.43	0.47	0.55	0.88	0.84	0.93	0.90	0.67	0.46
02-09	UC	0.14	0.08	0.06	0.14	0.10	0.07	0.86	0.58	0.80	0.85	0.18	0.11
02-08		0.18	0.05	0.23	0.04	0.04	0.00	0.01	0.01	0.00	0.06	0.00	0.07
02-09	UR	0.77	0.81	0.85	0.74	0.80	0.73	0.08	0.35	0.11	0.09	0.74	0.83
02-08		0.75	0.78	0.65	0.83	0.77	0.95	0.99	0.96	0.99	0.84	0.73	0.64
02-09	UD	0.21	0.16	0.12	0.24	0.18	0.20	0.90	0.64	0.86	0.88	0.25	0.17
02-09	ER	0.75	0.86*	0.63	0.63	0.63	0.71*	0.88	0.86*	0.75	0.88	0.75	1.00*
	χ^2	2.00	3.94*	0.53	0.53	0.53	2.10*	4.80	3.94*	2.00	4.80	2.00	7.00*
	p	0.16	0.05*	0.47	0.47	0.47	0.15*	0.03	0.05*	0.16	0.03	0.16	0.01*
2002		-0.58	0.07	0.81	0.08	0.45	-0.17	2.56	0.35	-0.87	-0.23	0.76	0.36
2003		-1.42	-0.42	-0.09	-0.85	-0.28	-0.72	1.94	0.00	-1.59	0.22	0.86	0.18
2004		-0.37	0.64	0.90	0.34	0.61	-0.06	3.17	1.39	-0.58	1.58	1.62	0.91
2005		-0.26	0.64	0.81	0.51	0.53	-0.06	3.17	1.04	-1.01	1.80	1.62	1.36
2006		0.79	1.35	1.81	1.27	1.51	0.67	4.20	2.26	0.00	2.70	2.70	2.45
2007		2.26	2.41	3.25	2.80	2.81	1.56	4.81	2.96	0.29	2.93	3.46	3.45
2008		1.32	1.91	2.80	2.21	2.32	0.89	2.15	1.22	-1.73	0.90	2.59	2.73
2009	TS	-2.68	-2.19	-1.99	-1.53	-1.83	-3.44	1.13	-0.70	-2.89	0.00	-0.76	-0.73

*Note: * The sample is smaller than indicated. Shaded fields highlight the best/worst result. Where a field is empty of data, the sample did not allow for calculation.*

Theil's W and U statistics are less than 1 in the IMAD spring and fall forecasts. IMAD has the best results for both statistics in the spring forecasts. In the fall forecasts the results from the EC are slightly better, but IMAD does not significantly lag behind. The effect of the economic crisis on the results with respect to the W statistic is interesting: values fell with the inclusion of the year 2009. This is a consequence of the method of designing the forecast in the naive forecasting strategy, which assumes the maintenance of growth from the previous period.

Decomposition of the mean square error shows that the UC and UD components do not reach a level which would indicate a very good forecast in the case of spring forecasts. This is true not only of IMAD but also of all the other institutions. Results for

the fall forecast are better, since both components have a value close to one. IMAD also has the best results in the period that includes the crisis year, though the effect of the crisis year had a significant influence on the quality of forecasts in all the institutions.

The results of the sign test for testing directional forecast accuracy are quite interesting. Only the BS achieves satisfactory results in the spring forecasts for the observed case. The fall forecasts are somewhat better, though SKEP and the IMF do not achieve a satisfactory level of congruence of the direction of the forecast with the direction of the actual result.

The trace test was done for each institution for all forecasts, but results in the tables are shown only from the year 2002. For this reason results are not directly comparable. A critical value of the test was exceeded by IMAD for the years 2006 and 2007, but otherwise the results are not of particular significance.

Table 3 shows the results for the analysis of the success of the spring and fall forecasts for real GDP growth for the following year. Based on the ME criterion the best results for the 2002 to 2008 time frame were achieved by IMAD. If we also include the data for 2009, the EC and IMF stand out for the spring forecast and the BS and EC for the fall one. Inclusion of the year 2009 negatively affected the quality of the forecasts for all cases.

The MAE and MSE criteria in the case of the spring forecast identified the IMF as the best. The result is different for the fall forecasts. For the 2002-2008 period IMAD stands out based on the MAE criterion, while SKEP is the best according to the MSE criterion. With the inclusion of the year 2009 SKEP stands out according to the MAE criterion and the EC according to the MSE criterion. Taking into account the timing of the release of the forecasts, IMAD forecasts are among the best.

The MAE/SD and RMSE/SD criteria were best for the IMF for the spring forecasts, and for SKEP and the EC in the fall ones. Here the same finding as for the MAE and MSE holds. If we take into account the timing, IMAD achieves good results. A somewhat negatively surprising fact is that both parameters were close to or even more than one for all institutions.

Theil's statistics, compared to the values in Table 2, are significantly higher. For the 2002-2009 period, W exceeds a value of 1 for IMAD, the BS, SKEP, the EC and WIIW (spring forecasts) and for the IMF (fall forecast). This is reflective of the mistaken estimate of the effect of the crisis. It is clear that the models of the institutions listed do not include the mechanisms that caused the crisis.

Table 3

Analysis of the Accuracy of Spring and Fall Forecasts of Real GDP Growth for the Following Year

Period	Institution	Spring forecast						Fall forecast					
	Criterion	IMAD	BS	SKEP	EC	IMF	WIIW	IMAD	BS	SKEP	EC	IMF	WIIW
02-08		0.03	0.11	0.14	0.23	0.19	0.32*	0.17	0.43	0.24	0.34	0.39	0.29
02-09		-1.46	-1.36	-1.38	-1.25	-1.25	-1.49*	-1.21	-1.04	-1.19	-1.04	-1.10	-1.26
97-09	ME	-1.00						-0.73					
02-08	MAE	1.14	1.06	1.03	1.03	0.99	1.18*	1.00	1.03	0.93	1.06	1.01	1.14
02-09		2.49	2.39	2.40	2.35	2.28	2.77*	2.24	2.31	2.21	2.26	2.33	2.51

Period	Institution	Spring forecast						Fall forecast					
	Criterion	IMAD	BS	SKEP	EC	IMF	WIIW	IMAD	BS	SKEP	EC	IMF	WIIW
97-09		1.98						1.73					
02-08		1.71	1.42	1.36	1.41	1.30	1.79*	1.29	1.41	1.17	1.47	1.43	1.73
02-09		19.20	18.35	19.19	18.05	17.10	23.15*	15.98	17.20	16.70	15.60	17.78	19.81
97-09	MSE	13.14						10.88					
02-08		1.31	1.19	1.17	1.19	1.14	1.34*	1.14	1.19	1.08	1.21	1.19	1.31
02-09		4.38	4.28	4.38	4.25	4.14	4.81*	4.00	4.15	4.09	3.95	4.22	4.45
97-09	RMSE	3.63						3.30					
02-08		0.96	0.89	0.86	0.86	0.83	0.99*	0.84	0.86	0.78	0.89	0.85	0.96
02-09		0.61	0.58	0.59	0.57	0.56	0.68*	0.55	0.56	0.54	0.55	0.57	0.61
97-09	MAE/SD	0.60						0.52					
02-08		1.10	1.00	0.98	1.00	0.96	1.13*	0.95	1.00	0.91	1.02	1.00	1.10
02-09		1.07	1.05	1.07	1.04	1.01	1.18*	0.98	1.01	1.00	0.96	1.03	1.09
97-09	RMSE/SD	1.10						1.00					
02-08		0.87	0.79	0.77	0.79	0.76	0.89*	0.75	0.79	0.72	0.81	0.79	0.87
02-09	W	1.03	1.01	1.03	1.00	0.98	1.14*	0.94	0.98	0.96	0.93	1.00	1.05
02-08		0.31	0.28	0.27	0.28	0.27	0.31*	0.27	0.28	0.25	0.28	0.28	0.31
02-09	U	0.90	0.88	0.90	0.87	0.85	0.99*	0.82	0.85	0.84	0.81	0.87	0.92
02-08		0.00	0.01	0.01	0.04	0.03		0.02	0.13	0.05	0.08	0.10	0.05
02-09	UM	0.11	0.10	0.10	0.09	0.09		0.09	0.06	0.08	0.07	0.07	0.08
02-08		0.55	0.65	0.62	0.74	0.82		0.54	0.35	0.48	0.34	0.43	0.26
02-09	US	0.79	0.82	0.76	0.86	0.88		0.84	0.77	0.80	0.81	0.77	0.65
02-08		0.45	0.34	0.37	0.23	0.16		0.44	0.52	0.46	0.58	0.47	0.70
02-09	UC	0.10	0.08	0.14	0.06	0.02		0.07	0.17	0.11	0.12	0.16	0.27
02-08		0.34	0.01	0.00	0.01	0.12		0.00	0.00	0.01	0.02	0.00	0.13
02-09	UR	0.03	0.01	0.07	0.03	0.49		0.36	0.01	0.13	0.21	0.00	0.14
02-08		0.66	0.98	0.99	0.96	0.86		0.98	0.87	0.94	0.90	0.89	0.82
02-09	UD	0.85	0.89	0.83	0.88	0.42		0.55	0.92	0.79	0.72	0.93	0.78
02-09	ER	0.38	0.43*	0.88	0.71*	0.50	0.67*	0.38	0.43*	0.63	0.43*	0.50	0.57*
	χ^2	0.53	0.06*	4.80	1.22*	0.00	1.20*	0.53	0.06*	0.53	0.06*	0.00	0.19*
	p	0.47	0.81*	0.03	0.27*	1.00	0.27*	0.47	0.81*	0.47	0.81*	1.00	0.66*
2002		-0.68	-0.29	-0.05	-0.26	-0.99		0.29	0.13	-0.24	-0.04	-0.76	-0.12
2003		-1.69	-0.96	-0.70	-0.98	-1.63	-0.61	-0.52	-0.39	-0.79	-0.62	-1.19	-0.60
2004		-1.34	-0.54	-0.38	-0.60	-1.24	-0.40	0.06	0.22	-0.19	0.04	-0.43	-0.32
2005		-1.24	-0.59	-0.27	-0.47	-1.34	-0.25	0.12	0.26	-0.19	0.18	-0.52	-0.16
2006		-0.68	-0.08	0.38	0.04	-0.74	0.22	0.81	0.86	0.46	0.71	0.05	0.56
2007		0.38	0.71	1.50	0.89	0.30	1.01	1.85	1.69	1.44	1.55	1.05	1.39
2008		-0.08	0.34	1.07	0.68	0.05	0.69	1.21	1.30	0.84	1.06	0.90	0.80
2009	TS	-6.09	-4.57	-5.36	-4.26	-5.54	-3.75	-5.08	-3.59	-5.23	-3.67	-4.57	-4.02

*Note: * The sample is smaller than indicated. Shaded fields highlight the best/worst result. Where a field is empty of data, the sample did not allow for calculation.*

Decomposition of the mean square error signals the structural deficiencies of the models. The effect of the economic crisis is very clear, which is also shown in the results of the sign test. Only SKEP achieves acceptable results for the spring forecasts.

As for the forecasts for the current year, we also did a trace test for forecasts for the following year. The results show that the models worked satisfactorily up until the year 2009. With this crisis year the values for almost all institutions exceeded a critical threshold. The exceptions are the spring forecast of WIIW, and the fall forecasts of the BS and EC. In all three cases the values of the test statistics were close to a value of four.

The unbiasedness of IMAD's forecast for real growth of GDP was analyzed using a regression model. Results are presented in Table 4. In an ideal forecast the value of the regression constant (β_0) would be zero, and the value of the coefficient for the variable representing the forecast value (β_1) would be one. Results show that this holds only for fall forecasts for the current year. This can be reinforced (but not confirmed) by the Wald test, which does not reject the null hypothesis in the case of the fall forecast for the current year. A similar result is obtained for the spring forecast for the following year, but taking into account the other results the model does not indicate unbiasedness.

The results of the regression analysis therefore indicate the possible presence of bias in the forecasts. Here it should be noted that the analysis was performed on a small sample and hence it was not possible to conduct all the tests that would be required to confirm the findings cited.

Table 4

Regression Analysis of the Spring and Fall Forecasts of Real Growth of GDP for the Current and Following Year

	Spring (t)		Fall (t)		Spring (t+1)		Fall (t+1)	
	Coefficient	Prob.	Coefficient	Prob.	Coefficient	Prob.	Coefficient	Prob.
β_0	-1.75	0.0044	-0.11	0.6841	10.24	0.5665	-20.92	0.0388
β_1	1.46	0.0000	1.05	0.0000	-1.73	0.6881	6.24	0.0207
R-squared	0.93	-	0.97	-	0.02	-	0.43	-
Adjusted R-squared	0.92	-	0.96	-	-0.08	-	0.37	-
S.E. of regression	0.96	-	0.65	-	3.74	-	2.85	-
Sum squared resid	10.15	-	4.67	-	139.67	-	81.06	-
Log likelihood	-16.84	-	-11.79	-	-31.75	-	-28.49	-
F-statistic	143.35	0.0000	324.79	0.0000	0.17	0.6881	7.52	0.0207
Wald Test:	β_0=0, β_1=1		β_0=0, β_1=1		β_0=0, β_1=1		β_0=0, β_1=1	
F-statistic	7.42	0.0091	0.37	0.7023	0.65	0.5448	3.05	0.0924
Chi-square	14.85	0.0006	0.73	0.6942	1.29	0.5242	6.10	0.0473
Durbin-Watson stat	2.16	-	2.15	-	1.09	-	0.66	
Sample	1997-2009		1997-2009		1998-2009		1998-2009	
Included observations	13		13		12		12	

Table 5 presents the results of the analysis of accuracy of the spring and fall forecasts for the nominal growth of GDP for the current and following year. The results are presented only for IMAD, since we did not have data for the other institutions in the sample.

Table 5

Analysis of the Accuracy of Spring and Fall Forecasts for Nominal GDP Growth for the Current and Following Year

Period	Institution	IMAD			
	Criterion	Spring (t)	Fall (t)	Spring (t+1)	Fall (t+1)
02-08		0.50	-0.05	0.02	0.08
02-09		-0.14	-0.37	-1.74	-1.54
97-09	ME	-0.21	-0.11	-0.87	-0.87
02-08	MAE	1.15	1.00	1.32	1.39
02-09		1.58	1.20	2.92	2.83

Period	Institution	IMAD			
	Criterion	Spring (t)	Fall (t)	Spring (t+1)	Fall (t+1)
97-09		1.36	1.17	2.43	2.09
02-08		1.95	1.21	2.99	2.48
02-09		4.35	1.90	27.47	22.97
97-09	MSE	3.24	1.72	19.51	15.50
02-08		1.40	1.10	1.73	1.58
02-09		2.09	1.38	5.24	4.79
97-09	RMSE	1.80	1.31	4.42	3.94
02-08		0.59	0.51	0.67	0.71
02-09		0.32	0.24	0.58	0.56
97-09	MAE/SD	0.28	0.24	0.49	0.43
02-08		0.71	0.56	0.88	0.80
02-09		0.41	0.27	1.04	0.95
97-09	RMSE/SD	0.37	0.27	0.90	0.80
02-08		0.54	0.42	0.67	0.61
02-09	W	0.39	0.25	0.97	0.89
02-08		0.17	0.13	0.21	0.19
02-09	U	0.26	0.17	0.65	0.59
02-08		0.13	0.00	0.00	0.00
02-09	UM	0.00	0.07	0.11	0.10
02-08		0.17	0.12	0.10	0.18
02-09	US	0.72	0.44	0.50	0.61
02-08		0.70	0.88	0.90	0.82
02-09	UC	0.28	0.49	0.39	0.28
02-08		0.00	0.00	0.06	0.01
02-09	UR	0.57	0.34	0.00	0.06
02-08		0.87	1.00	0.94	0.99
02-09	UD	0.42	0.59	0.88	0.83
02-09	ER	0.75	0.75	0.38	0.63
	χ^2	1.74	1.74	0.53	0.18
	p	0.19	0.19	0.47	0.67
2002		-0.43	1.66	1.72	1.50
2003		-0.86	1.27	0.96	0.58
2004		-0.50	1.87	0.71	0.46
2005		-1.38	0.59	-0.08	-0.35
2006		-0.35	1.44	0.13	0.22
2007		1.77	2.47	1.61	1.51
2008		1.41	1.01	1.53	1.18
2009	TS	-1.96	-1.21	-4.28	-4.99

The following conclusions can be drawn from the results in Table 5. First, the year 2009 has the same effect as in the case of forecasting the real GDP growth rate. Second, RMSE/SD in the forecasts for the following year in samples which include the year 2009 reach or exceed the value of one. Third, Theil's W statistic approaches a value of one in forecasts for the following year when the year 2009 is included. Fourth, the sign test shows insignificant congruence between the forecast and the realization. Fifth, the trace test in 2009 in forecasts for the following year exceeded the critical threshold.

An analysis of the unbiasedness of the forecasting of the nominal rate of GDP growth is presented in Table 6. Due to the small size of the sample, the results do not give clear answers to the question as to whether the forecasts are unbiased. In all cases the Wald test is unable to reject the null hypothesis, that the regression constant is equal to zero and the regression coefficient is equal to one, but the other results do not necessarily support these conclusions.

Table 6

Regression Analysis of Spring and Fall Forecasts of Nominal GDP Growth for the Current and Following Year

	Spring (t)		Fall (t)		Spring (t+1)		Fall (t+1)	
	Coefficient	Prob.	Coefficient	Prob.	Coefficient	Prob.	Coefficient	Prob.
β_0	-1.75	0.1812	-1.49	0.0890	-2.85	0.6842	-6.80	0.2352
β_1	1.17	0.0000	1.16	0.0000	1.22	0.1286	1.65	0.0172
R-squared	0.89	-	0.95	-	0.22	-	0.45	-
Adjusted R-squared	0.88	-	0.94	-	0.14	-	0.39	-
S.E. of regression	1.80	-	1.23	-	4.72	-	3.96	-
Sum squared resid	35.49	-	16.64	-	223.18	-	156.93	-
Log likelihood	-24.97	-	-20.05	-	-34.57	-	-32.45	-
F-statistic	86.36	0.0000	196.64	0.0000	2.74	0.1286	8.13	0.0172
Wald Test:	β_0=0, β_1=1		β_0=0, β_1=1		β_0=0, β_1=1		β_0=0, β_1=1	
F-statistic	1.04	0.3871	1.88	0.1986	0.25	0.7868	0.92	0.4280
Chi-square	2.07	0.3550	3.76	0.1527	0.49	0.7822	1.85	0.3966
Durbin-Watson stat	1.74	-	2.87	-	1.16	-	1.10	-
Sample	1997-2009		1997-2009		1998-2009		1998-2009	
Included observations	13		13		12		12	

Tables 7, 8, and 9 show the results of analyses of the accuracy in forecasting inflation. Since the institutions do not all use the same definition for the indicator of the growth in the general price index, a direct comparison is possible only among the forecasts of IMAD, SKEP, the IMF and WIIW. In the cases mentioned we observe the forecast of the average growth of inflation.

The BS up until 2007 used the price growth in the form of growth rate of the prices of the current quarter compared to the same quarter of the previous year. In 2007 it shifted to using the definition of harmonized indices of consumer prices (HICP). Due to these changes it is not possible to make a comparison with the other institutions, nor does it make sense to use certain tests/criteria. The EC forecasts the HICP throughout the whole of the observed period. In this case as well a direct comparison is not possible.

In Table 7 we first analyze the accuracy of the spring and fall forecasts of inflation for the current year. According to the first criterion (ME), forecasting by SKEP was the best in the period 2002 to 2008 for the spring forecasts, while for the period from 2002 to 2009 forecasting by WIIW was the best. For IMAD both values are positive, which indicates an underestimation in forecasting. For the fall forecasts WIIW and the IMF had superior results.

Table 7

Analysis of the Accuracy of Spring and Fall Forecasts of Inflation for the Current Year

Period	Institution	Spring forecast						Fall forecast					
	Criterion	IMAD	BS+	SKEP	EC++	IMF	WIIW	IMAD	BS+	SKEP	EC++	IMF	WIIW
02-08		0.47	0.21	0.24	0.16	0.58	0.44	-0.11	-0.21	-0.06	-0.14	-0.07	0.01
02-09		0.48	0.25	0.20	0.16	0.56	0.19	-0.11	-0.21	-0.04	-0.13	-0.01	-0.06
97-09	ME	0.29						-0.14					
02-08		0.47	0.30	0.41	0.27	0.64	0.81	0.17	0.36	0.09	0.23	0.19	0.30
02-09		0.48	0.33	0.38	0.26	0.61	0.91	0.16	0.34	0.09	0.20	0.21	0.34
97-09	MAE	0.51						0.18					
02-08		0.40	0.20	0.28	0.22	0.80	1.01	0.05	0.15	0.01	0.10	0.05	0.19
02-09		0.39	0.21	0.24	0.20	0.72	1.20	0.05	0.14	0.01	0.09	0.06	0.21
97-09	MSE	0.43						0.07					
02-08		0.64	0.45	0.53	0.47	0.89	1.00	0.23	0.39	0.11	0.31	0.22	0.43
02-09		0.62	0.46	0.49	0.45	0.85	1.10	0.22	0.37	0.11	0.29	0.25	0.46
97-09	RMSE	0.66						0.27					
02-08		0.27	0.19	0.24	0.16	0.36	0.47	0.10	0.22	0.05	0.13	0.11	0.17
02-09		0.24	0.17	0.19	0.13	0.30	0.46	0.08	0.18	0.04	0.10	0.11	0.17
97-09	MAE/SD	0.19						0.07					
02-08		0.37	0.28	0.30	0.28	0.51	0.58	0.13	0.24	0.06	0.18	0.13	0.25
02-09		0.31	0.25	0.25	0.23	0.42	0.55	0.11	0.20	0.05	0.15	0.13	0.23
97-09	RMSE/SD	0.25						0.10					
02-08		0.43	0.15	0.36	0.33	0.61	0.68	0.16	0.13	0.07	0.22	0.15	0.29
02-09		0.28	0.14	0.23	0.21	0.39	0.50	0.10	0.11	0.05	0.14	0.12	0.21
97-09	W	0.32						0.13					
02-08		0.13	0.10	0.11	0.10	0.19	0.21	0.05	0.09	0.02	0.07	0.05	0.09
02-09		0.14	0.11	0.11	0.10	0.19	0.25	0.05	0.09	0.02	0.07	0.06	0.10
97-09	U	0.11						0.04					
02-08		0.55	0.23	0.21	0.11	0.42	0.19	0.24	0.30	0.29	0.21	0.10	0.00
02-09	UM	0.59	0.30	0.16	0.13	0.43	0.03	0.26	0.33	0.12	0.18	0.00	0.02
02-08		0.00	0.12	0.02	0.10	0.11	0.12	0.14	0.08	0.03	0.13	0.26	0.02
02-09	US	0.00	0.15	0.06	0.07	0.09	0.31	0.10	0.05	0.17	0.15	0.50	0.02
02-08		0.45	0.65	0.77	0.79	0.47	0.68	0.62	0.62	0.69	0.66	0.64	0.97
02-09	UC	0.41	0.55	0.78	0.80	0.48	0.66	0.64	0.62	0.71	0.66	0.50	0.97
02-08		0.01	0.19	0.00	0.17	0.04	0.01	0.17	0.12	0.03	0.18	0.30	0.07
02-09	UR	0.01	0.20	0.02	0.12	0.04	0.09	0.12	0.08	0.18	0.19	0.54	0.00
02-08		0.44	0.59	0.79	0.72	0.55	0.80	0.59	0.58	0.68	0.61	0.60	0.92
02-09	UD	0.41	0.50	0.82	0.74	0.53	0.88	0.62	0.59	0.69	0.63	0.46	0.98
02-09	ER	0.88	1.00*	0.88	0.75	1.00	0.71*	1.00	1.00*	0.88	0.88	0.75	0.86*
	χ^2	4.44	7.00*	4.44	1.74	8.00	1.22*	8.00	7.00*	4.44	4.44	1.74	3.73*
	p	0.04	0.01*	0.04	0.19	0.00	0.27*	0.00	0.01*	0.04	0.04	0.19	0.05*
2002		1.18	0.31	4.51	3.09	6.08	1.64	-5.42	-1.19	-3.79	0.00	7.58	
2003		1.38	-0.31	4.33	2.25	5.97	1.21	-6.50	-1.78	-4.33	-1.06	6.95	0.30
2004		1.97	0.31	4.16	2.53	5.87	0.77	-6.50	-2.07	-4.88	-2.12	6.74	-0.89
2005		1.97	0.62	3.47	2.25	6.08	0.22	-6.50	-2.96	-5.42	-2.65	6.53	-1.48
2006		2.76	1.54	3.81	2.53	6.19	0.33	-7.58	-4.15	-5.42	-2.65	6.53	-1.78
2007		5.52	4.92	5.72	5.91	7.16	1.42	-6.50	-2.67	-4.88	-1.06	7.37	1.19
2008		6.50	4.62	6.59	6.19	8.99	3.40	-9.21	-4.44	-5.96	-4.76	6.95	0.30
2009	TS	7.48	6.15	6.41	6.75	9.42	1.64	-9.75	-5.04	-5.42	-4.76	7.79	-1.48

*Note: * The sample is smaller than indicated. Shaded fields highlight the best/worst result. Where a field is empty of data, the sample did not allow for calculation. + and ++ - a different definition of the variable was used.*

According to the criteria MAE, MSE and RMSE, SKEP has the best forecasts among the four compared (IMAD, SKEP, IMF, WIIW). This is true of both spring and fall forecasts. IMAD is in second place in most cases. If we take into account the informational advantage of SKEP, forecasting by IMAD is better. The year 2009 brought about lower values of individual criteria in all cases.

We can come to similar conclusions if we observe the criteria MAE/SD in RMSE/SD. These also allow a comparison between the BS and the EC. Here we find that the BS achieves better results in the spring forecasts than IMAD, while for the fall forecasts the opposite is true.

Theil's W and U are less than one in all cases. The year 2009 even reduced the value of W, while it did not have a significant influence on U. SKEP achieves the lowest values for both statistics among the four institutions compared, followed by IMAD, as the previous results already showed. The BS values for the test statistic W in the spring forecasts stand out, since it achieves by far the lowest values, but a direct comparison is not possible.

Decomposition of the mean square error shows an unfavorable ratio between the components for IMAD. This holds for the spring as well as fall forecasts. Relatively good results are achieved by the forecasts of SKEP (spring forecast) and WIIW (fall forecast).

The sign test shows that the congruence between the direction forecasted and the direction of the actual result is not significant for the spring forecasts of the EC and WIIW or the fall forecasts of the IMF. IMAD achieves a very good result in the fall forecast: the ratio of congruence (ER) is equal to one. This is also true of the BS, but it is necessary to take into account a different definition of inflation.

The trace test reveals structural problems in forecasting models. It exceeds the critical threshold in the spring forecasts of IMAD (2007 to 2009), the BS (2007 to 2009), SKEP (2002 to 2004 and 2007 to 2009), the EC (2007 to 2009) and the IMF (2001 to 2009). The situation is similar for the fall forecasts. The critical threshold is exceeded by IMAD (1999 to 2009), the BS (2006, 2008, 2009), SKEP (2003 to 2009), the EC (2008, 2009) and the IMF (2000 to 2009). It is interesting that the critical threshold is never exceeded in the case of WIIW.

An analysis of the accuracy of the spring and fall forecasts of inflation for the following year is shown in Table 8. According to the ME criterion the best results for spring forecasting is achieved by SKEP and for fall forecasting by IMAD (sample from 2002 to 2008) and SKEP (sample from 2002 to 2009). According to the MAE criterion the best results are achieved by IMAD. This assertion holds true to the spring and fall forecasting and for both time periods sampled.

Based on the MSE criterion the best results for the sample from 2002 to 2008 are achieved by IMAD. With the inclusion of the year 2009 the IMF (spring forecasting) and SKEP (fall forecasting) are somewhat better. This is true also of the RMSE criterion.

The MAE/SD criterion favors IMAD in all forecasts. This holds also for RMSE/SD with the exception of fall forecasting in the sample from 2002 to 2009, where the best result is achieved by SKEP. Considering the informational advantage, IMAD has the best results of the four compared. It is interesting that a similar performance is

achieved by the BS, but it must be excluded from the comparison due to a different definition of inflation.

Table 8

Analysis of the Accuracy of Spring and Fall Forecasts of Inflation for the Following Year

Period	Institution	Spring forecast						Fall forecast					
	Criterion	IMAD	BS+	SKEP	EC++	IMF	WIIW	IMAD	BS+	SKEP	EC++	IMF	WIIW
02-08		0.76	0.67*	0.63	0.39	0.66	0.53*	0.36	0.57	0.43	0.13	0.63	0.90
02-09		0.38	0.20*	0.23	0.04	0.39	0.09*	-0.06	0.18	0.04	-0.24	0.25	0.28
98-09	ME	0.70						0.17					
02-08		1.10	1.00*	1.20	1.41	1.31	1.23*	0.87	0.77	0.91	1.04	1.23	1.19
02-09		1.25	1.23*	1.38	1.54	1.34	1.43*	1.14	1.00	1.14	1.26	1.38	1.55
98-09	MAE	1.42						1.10					
02-08		2.10	1.82*	2.33	2.73	2.50	2.58*	1.26	1.18	1.31	1.38	2.45	2.50
02-09		2.50	2.53*	2.89	3.11	2.47	3.17*	2.23	1.88	2.05	2.18	2.87	4.29
98-09	MSE	3.34						2.16					
02-08		1.45	1.35*	1.53	1.65	1.58	1.61*	1.12	1.08	1.14	1.17	1.57	1.58
02-09		1.58	1.59*	1.70	1.76	1.57	1.78*	1.49	1.37	1.43	1.48	1.69	2.07
98-09	RMSE	1.83						1.47					
02-08		0.63	0.62*	0.69	0.83	0.75	0.71*	0.50	0.48	0.52	0.61	0.71	0.68
02-09		0.62	0.66*	0.69	0.77	0.67	0.71*	0.57	0.54	0.57	0.63	0.69	0.77
98-09	MAE/SD	0.54						0.42					
02-08		0.83	0.84*	0.88	0.96	0.91	0.92*	0.64	0.67	0.66	0.68	0.90	0.91
02-09		0.79	0.86*	0.85	0.89	0.78	0.89*	0.74	0.74	0.72	0.74	0.84	1.03
98-09	RMSE/SD	0.70						0.56					
02-08		0.98	0.45*	1.03	1.15	1.07	1.09*	0.76	0.36	0.77	0.82	1.06	1.07
02-09		0.72	0.49*	0.78	0.84	0.72	0.81*	0.68	0.42	0.66	0.70	0.77	0.95
98-09	W	0.89						0.71					
02-08		0.30	0.30*	0.32	0.35	0.33	0.34*	0.24	0.24	0.24	0.25	0.33	0.33
02-09		0.35	0.38*	0.38	0.39	0.35	0.40*	0.33	0.33	0.32	0.33	0.38	0.46
98-09	U	0.30						0.24					
02-08		0.27		0.17	0.05	0.17		0.10	0.28	0.14	0.01	0.16	0.32
02-09	UM	0.06		0.02	0.00	0.06		0.00	0.02	0.00	0.03	0.02	0.02
02-08		0.02		0.05	0.01	0.08		0.09	0.08	0.02	0.01	0.11	0.10
02-09	US	0.13		0.16	0.07	0.21		0.21	0.22	0.13	0.10	0.25	0.13
02-08		0.70		0.78	0.94	0.74		0.81	0.64	0.84	0.98	0.73	0.58
02-09	UC	0.81		0.82	0.93	0.73		0.79	0.76	0.86	0.87	0.73	0.85
02-08		0.03		0.04	0.17	0.03		0.00	0.00	0.02	0.07	0.02	0.00
02-09	UR	0.01		0.02	0.07	0.00		0.00	0.00	0.00	0.01	0.00	0.12
02-08		0.70		0.79	0.77	0.80		0.90	0.72	0.83	0.92	0.82	0.67
02-09	UD	0.94		0.97	0.93	0.94		1.00	0.98	1.00	0.96	0.98	0.87
02-09	ER	0.50	0.67*	0.50	0.43*	0.63	0.50*	0.75	0.86*	0.63	0.57*	0.38	0.29*
	χ^2	0.04	0.38*	0.04	0.19*	0.53	0.00*	1.74	3.73*	0.53	0.19*	0.53	2.10*
	p	0.85	0.54*	0.85	0.66*	0.47	1.00*	0.19	0.05*	0.47	0.66*	0.47	0.15*
2003		4.94	0.57	4.79	0.78	5.24	0.42	3.36	0.80	4.35	0.16	5.41	1.68
2004		4.45	0.33	4.26	-0.39	4.41	-0.21	2.18	0.70	3.62	-1.03	4.64	1.74
2005		4.09	0.08	3.84	-0.85	3.88	-0.91	1.73	0.70	3.30	-1.74	4.26	1.10
2006		4.24	-0.24	3.96	-0.91	4.18	-1.05	1.73	0.60	3.46	-1.74	4.26	1.10
2007		5.29	0.98	4.61	-0.07	4.88	-0.14	2.55	1.80	4.35	-0.71	4.97	1.87
2008		7.55	3.26	6.44	1.76	6.41	2.24	4.55	4.00	6.28	0.71	6.39	4.06
2009	TS	5.93	1.14	4.91	0.20	5.53	0.42	1.82	1.40	4.11	-1.50	5.08	1.42

*Note: * The sample is smaller than indicated. Shaded fields highlight the best/worst result. Where a field is empty of data, the sample did not allow for calculation. + and ++ - a different definition of the variable was used.*

Theil's test statistics once again favor IMAD. In the case of SKEP, the EC, the IMF and WIIW (spring forecast, sample from 2002 to 2008) and the IMF and WIIW (fall forecast, sample from 2002 to 2008), the test statistic W even exceeds the value of one. Very good values are also achieved by the BS, but these are not comparable to the values for IMAD.

None of the four institutions ensured a significant congruence between the direction of the forecast and the actual realization. Only the BS achieved a satisfactory degree of congruence of direction in the fall forecasting. Results were also poor for the trace test, which signals structural difficulties in the models.

Table 9 presents the results for the analysis of unbiasedness. As with Tables 4 and 6, we performed the analysis just for IMAD. The fall forecast for the current year and the fall forecast for the following year stand out, since we can say with a high degree of certainty that they are unbiased. It is clear that IMAD has a better quality information base for the preparation of the fall forecasts, which results in better quality forecasting.

Table 9

Regression Analysis for Spring and Fall Forecasts of Inflation for the Current and Following Year

	Spring (t)		Fall (t)		Spring (t+1)		Fall (t+1)	
	Coefficient	Prob.	Coefficient	Prob.	Coefficient	Prob.	Coefficient	Prob.
β_0	0.69	0.0795	-0.07	0.6662	0.77	0.5904	-0.29	0.8315
β_1	0.92	0.0000	0.99	0.0000	0.98	0.0058	1.09	0.0012
R-squared	0.96	-	0.99	-	0.55	-	0.67	-
Adjusted R-squared	0.95	-	0.99	-	0.50	-	0.63	-
S.E. of regression	0.60	-	0.25	-	1.85	-	1.59	-
Sum squared resid	3.96	-	0.68	-	34.17	-	25.29	-
Log likelihood	-10.72	-	0.75	-	-23.31	-	-21.50	-
F-statistic	237.50	0.0000	1438.59	0.0000	12.20	0.0058	20.00	0.0012
Wald Test:	β_0=0, β_1=1		β_0=0, β_1=1		β_0=0, β_1=1		β_0=0, β_1=1	
F-statistic	2.34	0.1427	2.12	0.1666	0.86	0.4514	0.13	0.8768
Chi-square	4.67	0.0967	4.24	0.1202	1.72	0.4223	0.27	0.8752
Durbin-Watson stat	2.32	-	2.52	-	1.62	-	2.02	-
Sample	1997-2009		1997-2009		1998-2009		1998-2009	
Included observations	13		13		12		12	

5. CONCLUSION

In this paper we analyzed the quality of macroeconomic forecasting by six institutions which regularly release forecasts for Slovenia. The analysis focuses on the quality of forecasting for the real and nominal growth of GDP and the average annual rate of inflation for the period 1997-2009. We analyzed four types of forecasts separately: spring and fall forecasts for the current year and spring and fall forecasts for the following year.

Results of the analysis do not provide us with an "absolute winner" but they do indicate the main features of particular forecasts. If we attempt to summarize the key findings of the analysis, we see that the forecasts of IMAD are among the better ones. Also taking into account the timing of the releases of forecasts by the institutions analyzed – IMAD in addition to the BS releases its forecasts of trends in GDP and inflation roughly a month earlier than the other institutions in the sample – attests to the superiority of IMAD's estimates.

At the same time the tests performed showed that the models analyzed have structural deficiencies. The global financial and economic crisis has had a major influence on the quality of the forecasts. It has also confirmed indications of those parameters which pointed to the structural problems of the forecast models in the sample before the crisis.

Regardless of the order which we can form based on the measures of quality presented, we can still assert that there is room for improvement in the quality of the models. It can also be observed that the developers of models up until now have most likely given priority to reducing error in the forecasts while giving insufficient attention to congruence between the direction of the trend in the forecast and the actual result. This criterion is very important for effective economic policy making.

References

Aiginger, K.,1979. Das Element 'Vorsicht' in Zukunftsdaten.*WIFO-Monatsberichte*, 52(4), pp. 167-184.

Andersson, M.K., Karlsson, G. and Svenssson, J., 2007. The Riksbank's Forecasting Performance.*Sveriges Riksbank Working Paper Series*, December, 218.

Ash, J.C.K., Smyth, D.J. and Heravi, S.M.,1998. Are OECD Forecasts Rational and Useful? A Directional Analysis.*International Journal of Forecasting*, 14(3), pp.381-391.

Batchelor, R.,2001. How Useful are the Forecasts of Intergovernmental Agencies? The IMF and OECD Versus the Consensus.*Applied Economics*, 33(2), pp. 225-235.

Baumgartner, J.,2002a.Evaluation of Macro-Economic Forecasts for Austria in the 1980s and 1990s.*Austrian Economic Quarterly*, 4, pp. 191-206.

Baumgartner, J., 2002b. Die Wirtschaftsprognosen von WIFO und IHS. Eine Analyse für die achtziger und neunziger Jahre.*WIFO-Monatsberichte*, 75(11),pp. 701-716.

Bleymüller, J., Gehlert, G. and Gülicher, H.,1994. *Statistik für Wirtschaftswissenschaftler*, *9th Edition.*Munich: VahlenVerlag.

Blix, M., Wadefjord, J., Wienecken, U. and Adahl, M., 2001. How Good is the Forecasting Performance of Major Institutions? *SverigesRiksbank Economic Review*, 3, pp. 38-68.

Christoffersen, P.F., Diebold, F.X., 1996. Further Results on Forecasting and Model Selection under Asymmetric Loss.*Journal of Applied Econometrics*, 11(5), pp.561-571.

Christoffersen, P.F. and Diebold, F.X., 1997. Optimal Prediction under Asymmetric Loss.*Econometric Theory*, 13(6),pp. 808-817.

Diebold, F.X., 2001. *Elements of Forecasting, 2.Edition*.Cincinnati: Southwestern College Publishing.

Diebold, F.X., Mariano, R.S., 1995. Comparing Predictive Accuracy.*Journal of Business and Economic Statistics*, 13(3), pp.253-265.

Granger, C.W.J. and Newbold, P., 1977. *Forecasting Economic Time Series.*New York:Academic Press.

Holden, K. and Peel, D.A., 1990. On Testing for Unbiasedness and Efficiency of Forecasts.*The Manchester School*, 58(2), pp.120-127.

Jagric, T., 2010. OvrednotenjemakroekonomskihnapovedizaSlovenijo v obdobju 1997-2009.*IB revija*, 2,pp. 41-55.

Lenain, P., 2002. What is the Track Record of OECD Economic Projections?. Paris: OECD, Available at: http://www.oecd.org/pdf/M00024000/M00024465.pdf [8.9.2010].

Mincer, J. and Zarnowitz, V., 1969. The Evaluation of Economic Forecasts. In:J. Mincer,ed., *Economic Forecasts and Expectations: Analysis of Forecasting Behavior.* NBER Studies in Business Cycles, 19.

Öller, L.-E.andBarot, B., 2000. The Accuracy of European Growth and Inflation Forecasts.*International Journal of Forecasting*, 16(3), pp.293-315.

Theil, H., 1966. *Applied Economic Forecasting.*Amsterdam:North Holland.

Theil, H., 1971. *Principles of Econometrics*. New York: Wiley.