SVERIGES RIKSBANK
WORKING PAPER SERIES 218

The Riksbank's Forecasting Performance

Michael K. Andersson, Gustav Karlsson and Josef Svensson

DECEMBER 2007

WORKING PAPERS ARE OBTAINABLE FROM

Sveriges Riksbank • Information Riksbank • SE-103 37 Stockholm
Fax international: +46 8 787 05 26
Telephone international: +46 8 787 01 00
E-mail: info@riksbank.se

The Working Paper series presents reports on matters in
the sphere of activities of the Riksbank that are considered
to be of interest to a wider public.
The papers are to be regarded as reports on ongoing studies
and the authors will be pleased to receive comments.

THE RIKSBANK'S FORECASTING PERFORMANCE

MICHAEL K. ANDERSSON, GUSTAV KARLSSON, AND JOSEF SVENSSON*

SVERIGES RIKSBANK WORKING PAPER SERIES

NO. 218

DECEMBER 2007

ABSTRACT

This paper describes the official Riksbank forecasts for the period 2000-06. The forecast variables are those that are important for monetary policy analysis, i.e. inflation, GDP, productivity, employment, labour force, unemployment and financial variables such as interest rate and foreign exchange rate. The Riksbank's forecasts are presented and analyzed and compared with alternative forecasts, that is, those from other institutions and simple statistical models.

One important message from the study is that macroeconomic forecasts are associated with an appreciable uncertainty; the forecast errors are often sizeable. The forecast memory, defined as how far the forecasts are more informative than the variables unconditional mean, is usually limited to the first year. Furthermore, we find that the inflation forecasts exhibit several appealing features, such as a predictability memory that (possibly) includes the second year, relatively low RMSE and weak efficiency. The forecasts for the investigated real variables are shown to be less precise and they have a shorter forecast memory. The exchange rate predictions demonstrate the least accurate (of the investigated variables) forecasts.

Compared to other forecasters, the Riksbank's predictions are often more accurate. This holds for a comparison with the National Institute of Economic Research, even though the differences are statistically insignificant, as well as for a comparison with the participants in the Consensus Forecasts panel, where the Riksbank's predictions often are among the best.

We also find indications that misjudgements for productivity growth have had effects on forecasts for both inflation and GDP, but the results suggest that the Riksbank has considered available information in an acceptable fashion. This is also true for the undertaken revisions (from one forecast occasion to another) of the published forecasts.

KEYWORDS: Judgements, Forecast Evaluation, Central Bank, Inflation, GDP, RMSE

JEL-CLASSIFICATION: E52, E37, E27

* Correspond to: Monetary Policy Department, Sveriges Riksbank, SE-103 37 Stockholm, Sweden, michael.andersson@riksbank.se and gustav.karlsson@riksbank.se. We are grateful to the National Institute of Economic Research for providing their historical forecasts. We have benefited from discussions with Todd Clark and Paolo Giordani. The views expressed in this paper are solely the responsibility of the authors and should not be interpreted as reflecting the views of the Executive Board of Sveriges Riksbank.

1. Introduction

Developments in the Swedish economy are regularly forecast by commercial banks and other institutions, and media frequently draw attention to their predictions. The Riksbank's forecasts are currently presented three times a year in *Monetary Policy Report*.[1] These forecasts often attract a good deal of interest when they are published but less so after that. However, pausing a moment to analyse earlier forecasts retrospectively is important. This is relevant not least for the Riksbank because the forecasts are a major basis for monetary policy decisions. An evaluation of the Riksbank's forecasting performance is therefore of interest for the Bank's internal development work as well as for those who follow Swedish monetary policy.

The results show that the accuracy of the Riksbank's forecasts is fully on a par with that of other forecasters. They also indicate that the inflation forecasts have appropriate properties. Another finding is that in a statistical sense, the prediction memory is about one year for the investigated variables, which is in line with results from Galbraith and Tkacz (2006). This is what to expect given if the arguments for the Great Moderation are taken seriously. Some implications of the Great Moderation is presented by Bernanke (2004), Campbell (2004) and Stock and Watson (2007). However some variables, such as inflation and possibly also employment, may exhibit a memory that embraces also the second year.

An evaluation of the Riksbank's recent monetary policy, including a relatively brief comparison of forecasts, is presented annually in the year's first *Monetary Policy Report*. Evaluations of various forecasters' assessments of the Swedish economy have been presented earlier by, among others, Blix et al. (2001 & 2002) and Bergvall (2005). More extensive forecast evaluations are relatively rare.

Models used by the Riksbank have been presented and evaluated earlier. CPI and nominal short-term interest rate forecasts with the Riksbank's Bayesian VAR model and structural general equilibrium model (the DSGE model) are evaluated in Adolfson et al. (2008) and compared with the Riksbank's published assessments.[2] The Riksbank's time-series procedures are described and evaluated in Andersson & Löf (2007). In the context of the ongoing development work, the present article looks closely at forecasts of the variables that are most central for the *Monetary Policy Report*.[3]

The article unfolds as follows. The next two sections describe the data studied and the analytical methods that are used to evaluate the forecasts. The forecasting performance of the Riksbank is then compared with that of other observers in the section after that, followed by a more detailed account of various aspects of the accuracy of the Riksbank's forecasts. A summary concludes.

2. The data

The data analysed in this article include forecasts that were used as a fundament for the *Inflation Report* in the period 2000–06. These forecasts are analysed on a quarterly frequency and described in section 2.1. The quarterly forecasts are then used mainly in section 5 on the Riksbank's forecasting performance but also in section 4 for comparisons with the Swedish National Institute of Economic Research (NIER). Consensus Economics' compilation of the forecasts that are used in section 4 to compare the Riksbank's forecasts with those of other observers are described in section 2.2.

2.1 Quarterly forecasts

During the period with an inflation-targeting regime the Riksbank has become increasingly transparent and published more details about its forecasts. Subsequently, the number of variables that are forecast quarterly has increased. This study focuses on the period 2000–06, for which quarterly forecasts are documented for GDP, UND1X (core inflation), CPI, SEK/TCW[4] exchange rate, the Riksbank's policy rate and the oil price. Forecasts for the number employed, the number in the labour force, the number unemployed and the GDP ratio for the number employed, which are documented from 2002 onwards, are also considered.

[1] Prior to 2007, *Monetary Policy Report* was called *Inflation Report* and in that guise was published four times a year prior to 2006.
[2] For an account of the Riksbank's DSGE model, see Adolfson et al. (2007a and 2007b).
[3] A condensed version of this article is given by Andersson et al. (2007).
[4] TCW is short for Total Competitiveness Weights and the SEK/TCW-index is a measure of the Swedish effective exchange rate.

For some of the studied variables, the documented quarterly forecasts go back further than 2000, so certain aspects of the evaluation could have referred to a longer period. In the present article, however, we have chosen to focus on a period that is covered by most of the main variables. Evaluations further into the past would be feasible using annual forecasts of GDP and the CPI, see Berg et al. (2004), but quarterly forecasts permit a more in-depth analysis. Briefly, then, these circumstances have somewhat restricted the evaluation period.

The study includes a separate comparison of quarterly forecasts by the Riksbank and NIER, respectively. Quarterly forecasts have been provided by NIER for GDP, UND1X, CPI, employment, the labour force, the SEK/TCW exchange rate and the Riksbank's policy rate. This data includes NIER's forecasts from 2001 onwards except for GDP, for which the forecasts begin at 2003.

2.2 Annual forecasts for comparisons with other forecasters

The data for the studies of forecasts on annual (average) growth rates come from surveys by Consensus Economics as well as from *Inflation Reports* in the period from October 2000 through October 2006. Thus, the evaluation period is much the same as for the analysis of quarterly forecasts.

The intention is to compare various agents' forecasting accuracies at times when they all have access to the same amount of information. Achieving this in practice is, however, difficult on account of the considerable differences in the timing and regularity of the forecasts. For example, as the June *Inflation Report* usually has had a cut-off date in May, Consensus' May survey is used in those cases. Many of the survey's respondents submit forecasts that may be less than recent; most forecasters produce new assessments on a limited number of occasions each year and by no means every month as a rule. Moreover, both the cut-off date for statistics in the Riksbank's reports and the dates of Consensus surveys tend to vary over time. So in practice it is very hard to achieve a comparison in which all forecasters have access to exactly the same information.

In the present analysis, however, we have striven to minimise this problem. All the comparisons are made with the last Consensus survey to be published before the cut-off date and the consequences of choosing an alternative basis for the comparison are considered in a sensitivity analysis.[5] Still, the fact that forecasts which are reported to Consensus are liable to be somewhat out-dated can at times put other forecasters at a disadvantage.[6]

As the purpose of this article is to evaluate the Riksbank's forecasts, the procedure starts from the dates on which the Riksbank's forecasts have been published. The forecasts available at that time from NIER were then registered, together with an average of those from other observers in the Consensus panel. NIER is treated separately from the rest of the panel because their forecasts are compared with the Riksbank's in greater detail. Moreover, NIER is the observer with the largest resources for forecasting work.

Forecasts from individual observers are generally not considered separately in this study. However, the forecast errors of the individual observers are used to rank the Riksbank's forecasts on each occasion. Moreover, some years that were particularly hard to predict are studied in greater detail and that account shows which observer was the first to foresee the future outcome.

3. Method

The forecasts are studied visually, using charts, as well as with statistical methods. The paper employs a variety of analysis methods and they are briefly summarized in Table 1. Forecast error is defined throughout as the outcome less the forecast. The statistical methods involve calculating the bias or mean

[5] The sensitivity analysis involves repeating the comparison with the forecasts from all the other observers moved forward one month. The other observers then do somewhat better but the qualitative conclusions are unchanged.

[6] In particular this implies that the NIER forecasts, which are compared with those from the first inflation report from each year, are based on National Accounts data for the third quarter of the preceding year. In five of the six forecast occasions, the Riksbank's corresponding forecasts where based on National accounts for the final quarter of the year before. NIER's performance for GDP improves somewhat when the forecast occasion one month forward is used, but the general picture holds.

error (ME), the root mean squared error (RMSE) [7], estimates of forecast memory[8] and regressions for testing weak efficiency. RMSE is a measure of the forecasts' average deviation from outcomes. Forecast memory indicates how far ahead it is meaningful, in a statistical sense, to forecast the path of that particular variable. The frequency with which forecasts manage to predict the correct direction in the coming two quarters is also studied, see Appendix A.2.

The test for weak efficiency examines whether the forecasts deviate systematically from outcomes.[9] The information provided by the regressions for weak efficiency partly corresponds to the analysis of the RMSE and the ME.

Moreover, correlations are used to study whether the available information, in the form of outcome data for the most central variables, has been used efficiently by the Riksbank, see more on this in Appendix A.4. To investigate if the Bank's behaviour in revising forecasts has been reasonable we compare the Riksbank's revisions to an autoregression. The underlying idea is described in Appendix A.5. Correlations are also used to study relationships between forecast errors for different variables. The starting point for this is the hypothesis that the Riksbank's misjudgements of productivity had consequences for forecasts of other variables. For further details, see Appendix A.6.

To put the size of the deviations in perspective, the Riksbank's forecast error is compared with the error associated with a simple autoregressive time-series model (referred to in the following as autoregression or AR model).[10] Two variants of autoregressive specifications have been examined: AR(1), with just the value for the preceding period (first lag) as the explanatory variable, and a variant with the number of lags based on a model choice procedure.[11] The accepted variant is the one that has been most accurate as a benchmark in comparisons with the Riksbank's performance. The models are estimated on data from 1993 up to the latest observation that was available at each forecast occasion. As the results suggest that in most cases the AR(1) specification performs somewhat better, it is mostly this which has been used as benchmark.[12]

Forecast evaluation calls for decisions about some basic matters. One is the form in which the forecasts are to be studied. The choice has generally been to evaluate the quarterly forecasts in the form of quarterly growth rates, the main reason being that such observations do no overlap.[13] Annual growth rates are made up of changes in the latest four quarters, which means that two consecutive quarterly-observed annual growth rates have three quarters (75 per cent) in common. This makes drawing statistical inferences much more complicated than in the case of quarterly growth rates.[14] In general, it is hard to do statistical inferences in such small samples as 28 observations (i.e. seven years of forecasts). Thus, most of our results will be of a descriptive nature.

---

[7] The root mean squared error summarises average misjudgements and the variation in forecast error in the evaluated forecasts. $RMSE(h) = \sqrt{\frac{1}{\tau_2 - \tau_1 + 1}\sum_{t=\tau_1}^{\tau_2}\left(x_{t+h} - x_{t+h}^{pred}\right)^2}$ The mean squared error can be decomposed into forecast error variance and bias (mean error), making it possible to study whether forecast error mostly consists of bias or of a variation around actual outcomes (see also Appendix A.1).

[8] The engaged definition of the forecast memory implies that the memory is used up when the RMSE of the forecasts equals the standard deviation of the predicted series, or stated slightly differently, the forecasts are as accurate as the process unconditional mean. This does not imply that forecasting beyond the memory is not meaningful. Rather, the memory indicates to which horizon information in indicators and the correlation structure (persistence) in data is informative. After that horizon, the forecasts should be guided by the steady state of the processes, and if possible, in a multivariable consistent fashion. This is, essentially, how the Riksbanks' forecasts are produced. The forecast memory is discussed more by Andersson (2000) and Andersson and Löf (2007).

[9] The test of weak efficiency is described in Appendix A.3.

[10] An autoregression involves specifying a model so that a variable's value today is dependent on the variable's earlier outcomes. The autoregression is described in Andersson & Löf (2007).

[11] The procedure for the choice of lags (the information criterion) that is used here is BIC (see Schwartz, 1978). BIC selects a specification on the basis of available data and a statistical criterion. In retrospect the selected model may, of course, turn out not to be the best in terms of average forecast accuracy.

[12] The exceptions are UND1X, CPI and unemployment. For UND1X and CPI an AR(1) with seasonal dummies is used since this performs better than both the AR(1) and the forecasts generated by the automatic procedure. For unemployment the best choice is AR(2), as selected by the automatic procedure.

[13] The evaluation is done for variables that are stationary (stable). Thus, the level is evaluated for variables with a stationary level and the growth rate for variables with a non-stationary level.

[14] The forecast errors will not be independent even for quarterly observations. For a correctly specified model, the h-step ahead forecast errors will follow a moving average of order h-1. However, the forecast error correlation will be much smaller for quarterly growth rates than for yearly growth rates.

Another matter is whether to relate the forecasts to the final (revised) or the preliminary outcome.[15] As the available outcomes at the time of each forecast are revised, historical forecasts tend to be further from the final outcomes than from the preliminary. This makes it natural to compare forecasts with the first available (and unrevised) outcomes. A comparison with the first outcome is usually called a real-time evaluation. Still, the effect of data revisions may also be of interest and is therefore worth studying separately. In general, forecast errors are somewhat larger in relation to the latest outcome than in relation to the real-time outcome. Revisions to GDP data are further discussed in Appendix A.7.

4. Comparisons with other forecasters

To be meaningful, an evaluation calls for something with which the forecasts in question can be compared. Comparisons of accuracy are commonly made with alternative forecasts from models or other observers. Properties of the data can also be used for evaluation. For example, information about forecast memory can be derived from the relationship between a series' standard deviation and the root mean square error of forecasts for the same series.

Evaluating the Riksbank's forecasts by comparing them with model forecasts can be of assistance in internal development work and provide indications about the accuracy and other properties of the official forecasts. Comparing the Bank's forecasts with those of other observers can yield additional lessons. The Riksbank's work stands to benefit from knowledge that a particular observer is notably successful in forecasting certain variables, or that predictability varies between observers. In this section, the Riksbank's forecasts are compared with NIER's and the mean from the Consensus surveys, as well as with forecasts generated with an autoregressive model.[16] The next section (5) elaborates on various aspects of the Riksbank's forecasts.

Unlike the other forecasts, in the greater part of the evaluation period the Riksbank's were based on the assumption of an unchanged policy rate in the forecast period. Comparisons with other forecasts might therefore be misleading. However, the assumption of a constant policy rate did produce relatively accurate policy rate forecasts in this period (though that was not the purpose of the assumption). The accuracy of the policy rate forecasts is presented in Chart 1, which shows that the constant-rate forecast was more accurate than the autoregressive for all horizons and that predictability memory extends six quarters into the future. The constant-rate forecast also compares relatively favourably with NIER's forecasts (see Tables 2–4).[17]

4.1 Comparisons with the National Institute of Economic Research (NIER)

A fair comparison between the Riksbank and NIER calls for calculations for each variable that cover one and the same period. GDP is therefore evaluated from 2003 onwards, employment and the labour force from 2002 and the other variables from 2001. Results from the comparisons are presented for forecast horizons up to eight quarters ahead. A comparison of longer horizons would be less relevant because for this, NIER generally has fewer observations than the Riksbank.

On the whole, in the period for this comparison, forecast accuracy differs only slightly between the Riksbank and NIER (see Table 2). With forecast horizons up to one year, RMSEs for GDP and inflation are somewhat lower for the Riksbank's forecasts, while NIER does somewhat better for other variables. For longer horizons, the Riksbank consistently has marginally lower RMSEs except for the exchange rate, which NIER forecasts somewhat more successfully. Note, however, that the observed differences in forecasting performance are not statistically conclusive (statistically significant to use the common term).[18]

The tendencies to systematic over- or undershooting in the forecasts are broadly similar for the Riksbank and NIER (see Tables 3 and 4). Growth rates for GDP, employment and the labour force are

[15] This naturally does not apply to variables that are not revised, such as financial data and price variables. The most obvious example of data that are revised is the national accounts; for the variables that are analysed in this study, this aspect is therefore particularly important for the GDP forecasts.
[16] For more details about data and methods, see section 3.
[17] During the evaluation period, NIER's policy rate forecasts should be interpreted as policy recommendations and not forecasts of the most likely outcome.
[18] Moreover, we have run encompassing tests and conclude that the Riksbank's forecasts do encompass the NIER's and vice versa. This suggests that the separate forecasts do not benefit from being combined, that is to construct a new forecast as a function of the Riksbank's and the NIER's forecasts. The tests are in our case based on a very small sample and may thus be unreliable. See Chong and Hendry (1986) for details on the test.

underestimated by both for most horizons (though NIER does this to a somewhat greater extent than the Riksbank). Inflation is undershot for short horizons and overshot for longer horizons, the former to a somewhat greater extent by NIER and the latter by the Riksbank. For them both, the overshooting for longer horizons is more pronounced for CPI than for UND1X inflation, to a large extent because interest rates turned out to be lower than was assumed/forecast in this period (all else equal, this holds, ceteris paribus, back the CPI but not UND1X).

For financial variables, the general bias tendencies of the forecast error are similar for the Riksbank and NIER. The exchange rate was generally weaker than had been forecast (particularly by the Riksbank) and the policy rate was lower than forecast/assumed.

4.2 Comparisons with the Consensus data

In this section, the Riksbank's forecast errors for percentage annual changes in GDP and CPI are compared with those for the forecasters in the Consensus surveys. Each forecaster's errors in the period studied are presented graphically for each variable. The forecasters are also ranked in order of forecast accuracy. The section concludes with a discussion of some occasions for which outcomes can be said to have been particularly unexpected. The purpose of the latter exercise is to examine how the various forecasters revised their forecasts (compared with the Riksbank) on these specific occasions.

4.2.1 GDP

Chart 2 shows the forecaster's Q1 forecast errors for same-year GDP growth in the period 2001–06.[19] The forecast errors follow much the same pattern and there does not seem to have been any bias (under- or overshooting on average) of GDP growth. Neither does any forecaster's forecast accuracy seem to have changed decisively in this period. The Riksbank has the lowest absolute forecast errors in three of the six years studied, while NIER, Consensus and the AR model are best on one occasion each.

The development of forecast error over time for GDP forecasts is presented in Charts 3–5 for the Riksbank, NIER and Consensus, respectively. The charts show four forecasting occasions a year (two for 2000) and include next-year forecasts. In general, and not surprisingly, the shorter the interval between the forecast and the publication of the outcome, the smaller the forecast error becomes.

Charts 3–5 suggest that for all the forecasters, forecasting eight quarters ahead was appreciably easier for some years than for others. Examples of years that turned out to have been easy to forecast are 2002 and 2005. Forecast error does vary between the forecasters but for 2002 and 2005 it is never extremely large. For 2001, on the other hand, actual GDP growth was overshot initially by all forecasters. In the final quarters of 2001, however, forecast error is small for all forecasters, no doubt because outcomes were then known for a large part of the year.

This section has focused so far on a graphic analysis of forecast error. A relevant question in this context is how the Riksbank and NIER compare with the other Consensus forecasters as a group. Chart 6 presents the rank order for same-year GDP. The Riksbank performs better than 70 per cent of the other forecasters on almost 60 per cent of all the forecasting occasions, whereas NIER, for example, does so on less than 30 per cent of those occasions. The Consensus average is better than 70 per cent of the panel for just over 20 per cent of the forecasts. The Riksbank is better than 90 per cent of the other forecasters on almost 30 per cent of the occasions, while the corresponding figures for NIER and Consensus are 15 and 8 per cent, respectively. For the autoregressive forecast, 70 per cent of the panel is outperformed for over 40 per cent of the forecasts, 80 per cent of the panel for almost 20 per cent and 90 per cent of the panel for almost 10 per cent.

The corresponding ranking for next-year GDP is shown in Chart 7. Here it is appreciably harder to outperform the average. The Riksbank has an only slightly higher ranking than NIER and Consensus for next-year forecasts. This illustrates that the potential gains of a more elaborate forecasting system (such as

[19] Forecasters refers here to the Riksbank, NIER, the average for other forecasters in the Consensus panel and the autoregression. The average of the Consensus forecasts is referred to in the article as the Consensus forecast. The Q1 forecasts are taken from Consensus Forecasts from March each year, except for 2006 when the February edition was utilized due to the early publication of Inflation Report 2006:1. All forecasts regard annual (average) growth rates.

the Riksbank's and NIERs) may be largest at the shorter horizons. See also the estimates of forecast memory in Section 5.

The following analysis concerns forecast error for GDP growth in 2001 and 2004, years for which the outcome turned out to be surprisingly low and high, respectively (1.2 and 3.5 per cent). These years stand out in the evaluation period as being particularly hard to assess with respect to GDP and may therefore repay a somewhat more detailed study. NIER, Swedbank and SEB were best at catching the GDP slowdown in 2001; their GDP growth forecasts were revised substantially downwards between March and May 2001.

The Riksbank's forecast was revised downwards in October, as were the predictions of most other forecasters. However, the Riksbank's same-year GDP growth assessment was already only 2.4 per cent initially in 2001, while the assessments by NIER, Swedbank and SEB were all in the range 3.0–3.5 per cent and thus initially further from what the outcome proved to be. During 2000, all forecasters were more or less equally unsuccessful at predicting the next year's slowdown.

In the case of GDP growth in 2004, in March Skandiabanken was closest to predicting the year's upswing. At the end of 2003, Skandiabanken's next-year GDP forecast was already around half of a percentage point higher than other forecasts. In March the Riksbank was closer to the outcome than other forecasters but still relatively far away.

4.2.2 CPI

The forecasters' Q1 same-year CPI forecast errors in the period 2001–06 are presented in Chart 8. A year that, looking back, can be said to have been hard to predict is 2001, for which CPI inflation was underestimated by all forecasters.[20] For 2005, the direction of the Riksbank's Q1 forecast error for same-year CPI inflation was the opposite of other forecasts. This was because initially in 2005 the Riksbank's CPI forecast was revised substantially downwards (excessively so, as it turned out). CPI predictions were also revised downwards to varying degrees by other forecasters. The development of each forecaster's CPI forecast error over time is shown in Charts 9–11. Compared with the years discussed above (2001 and 2005), the forecasts for other years were more accurate. Forecast error decreases as new outcomes become available and for same-year predictions it is relatively small in the last two quarters for all forecasters.

Ranking the same-year CPI forecasts shows that both the Riksbank and NIER compare well with the entire Consensus panel (see Chart 12). NIER is in the best decile of forecasters for 50 per cent of the forecasts and the Riksbank for over 40 per cent. The Consensus average generally outperforms half of the individual forecasters. The ranking for next-year CPI is shown in Chart 13. As for GDP, being frequently among the best is harder here. The Riksbank tends to be among the best performers somewhat more often than NIER and the Consensus average.

The following analysis concerns the CPI forecast errors for 2001 and 2004, years when the outcome turned out to be surprisingly high and low (2.6 and 0.5 per cent, respectively). CPI inflation in 2001 was underestimated by all forecasters to varying degrees (from 0.1 to 1.1 percentage points). The mean forecast error for these predictions is just over 0.7 percentage points. On this occasion, the Riksbank was one of the first forecasters to revise upwards. The main adjustment to the CPI forecast was made one quarter earlier than other forecasters'.

Nordea and Öhman Fondkommission also revised upwards early on, but not as markedly as the Riksbank. During 2000, however, Handelsbanken and SEB had been better than the Riksbank in predicting higher next-year inflation.

During 2004, the only forecasters who predicted very low same-year inflation as early as in March were the Riksbank and NIER. The forecast errors for the individual forecasters ranged from 0.1 to 1.2 percentage points (in March, the overestimations of inflation's outcome averaged almost 0.6 percentage points). In the next round of forecasts, in May 2004, all the other forecasters followed in the wake of the Riksbank and NIER and revised their predictions downwards. In 2003, moreover, the Riksbanken and NIER had already revised their next-year CPI forecasts downwards in Q2, while other forecasters only began to do this at least one quarter later but to a greater extent.

[20] In this year there was an exceptional increase in meat prices on account of the mad-cow and foot-and-mouth diseases. Prices of services also rose strongly.

All in all, the results of the comparison with the Consensus panel show that the properties of the Riksbank's forecasts are at least as good and often better than NIER's and the Consensus average. However, the differences are not statistically conclusive. Bearing these results in mind, in the next section the Riksbank's published forecasts are studied in detail. The account largely considers each variable in turn in accordance with the aspects described in section 3.

5. The Riksbank's forecasting performance

5.1 GDP

The paths of Riksbank GDP forecasts since 2000 are presented in Chart 14. Two phases in this period were particularly difficult to foresee: the weak tendency in 2001 and the strong development in recent years. When economic activity slackened early in 2001, forecasts for the coming year were revised downwards relatively markedly initially but the slowdown was judged to be short-lived. Further large downward adjustments of the short run were then made in each of the two autumn 2001 issues of the *Inflation Report* but the dip was still predicted to be relatively brief, with a fairly favourable outlook for 2003. This assessment was maintained during 2002, so the subsequent double-dip was not foreseen. Recently, moreover, the strength of the economic upswing has generally been underestimated. Forecasts have been revised upwards successively in the light of the latest outcomes but have still been on the low side. An exception was the minor dip around the turn of 2004, for which forecasts were generally on the high side. To sum up, these cyclical fluctuations have been hard to foresee and resulted in forecast errors for most observers.

The GDP statistics are revised so much that in a forecast evaluation it is desirable to distinguish effects of revisions from genuine forecast errors. For this reason, the calculations in this section evaluate forecasts in relation to the real-time outcome, which is the first version of each quarter's outcome. The revisions of GDP statistics are studied in more detail in Appendix A.7.

Estimated mean errors (ME) for the Riksbank's GDP forecasts are presented in Chart 15. The MEs for the different horizons range from –0.1 to 0.1, which shows that the bias in the Riksbank's forecasts has been very small.[21] RMSEs for the GDP forecasts are presented in Chart 16. RMSE for forecasts one quarter ahead is 0.3. [22] Measured as RMSE, the uncertainty in the Riksbank's forecasts comes in general mostly (to 85 per cent) from the variation in forecast error, so the contribution from bias is small.[23]

While the estimated biases for quarterly growth rate forecasts are relatively small, for longer horizons they are all negative. This could, of course, imply that next-year growth rate forecasts (for 5–8 quarters ahead) may contain significant bias. RMSEs and MEs for forecasts expressed as annual percentage changes are presented in Table 6. Here, too, there is no significant GDP bias. It is noteworthy that the proportion of bias in MSE is very low.

Compared with the AR model, the Riksbank's GDP forecast errors are somewhat higher for one quarter ahead, somewhat lower for two to four quarters ahead and then relatively similar. The forecast error gradually approaches the series' standard deviation for the evaluation period (0.35) and then exceeds this at the five-quarter horizon. This can be interpreted as the Riksbank's GDP forecasts having a four-step memory, that is, the first year. The circumstance that for longer horizons the forecast errors become somewhat smaller may have to do with the calculations being based on increasingly fewer observations (rather than indicating an increase in accuracy). The direction of the GDP forecasts was correct in just below 70 per cent of the predictions one quarter ahead and in approximately half of those for two quarters ahead (see Table 5).

[21] Given that the forecasts are unbiased, the width of a 50 per cent confidence interval is approximately +/- 0.2. All the estimated MEs lie inside this interval, so there is no reason to suppose that the Riksbank's GDP assessments contain a statistically significant bias.

[22] Given a normal distribution, this means that the width of a 95 per cent confidence interval is 1.2 (that is ± 0.6), which illustrates the uncertainty in the forecasts. How this compares with the uncertainty in, for instance, NIER's forecasts is studied in section 4.1.

[23] The proportion of bias is given by $bias^2/MSE$, where $MSE = bias^2 +$ forecast error variance. This proportion indicates the extent to which average over- or underestimations have contributed to the average forecast errors. For a further discussion see section 3.

The Riksbank's GDP forecasts one quarter ahead is not weakly efficient and the forecast-outcome degree of explanation is approximately 50 per cent (see Table 7).[24] For the one- and two-year horizons the hypothesis that the forecasts are efficient is not rejected but the explanatory power of the regressions decreases.[25] For the longest horizon the explanatory power is about ten per cent. The residuals are autocorrelated, which is an indication that an overestimation at the 4- and 8-quarter horizons on one occasion is often followed by further overestimations for these horizons in the next forecasts. A probable source of the autocorrelation is the autocorrelation in quarterly GDP growth rates.

Another interesting question is whether the revisions that have been made were reasonable. It is suggested in Regeringen (2002) and Riksrevisionen (2006) that an experiment with revisions to the official government forecasts that were larger than those which were actually made entailed an average loss of forecast accuracy.

The issue of the revision behaviour is a non-trivial one. In this article we merely compare the Riksbank's forecasts to an AR for some specific periods where the Riksbank appears to have revised the forecasts too little or too much. Remember that the general forecasting performance is summarized in the RMSE estimates. The reason for using the AR benchmark is that it produces competitive forecasts, on average, and relies on historical data only. Thus, any deviations from the AR may be interpreted as the value added from other information. See also Appendix A.5 for a further discussion.

After the fact, it seems like the Riksbank was too slow to revise the GDP forecasts for 2001, when growth was relatively weak. However, at the time of the forecasts only the available information was used. Chart 17 demonstrate that the autoregression (from 2000 and early 2001) predicted fairly strong growth for 2001, which was well in line with the available observations. The Riksbank's judgements in early 2000 also predicted strong growth. The forecast based on data up to the fourth quarter 2000 predicted a marked slowdown for mid 2001. Even though the outcome was notably lower the Riksbank's forecast was much better than the AR. In light of that we conclude that the Riksbank's revisions for GDP (during 2000 regarding the second quarter 2001 outcome) seem justified given the available information when the forecasts were made.

Tables 8 and 9 contain some interesting correlations between the Riksbank's forecast errors for GDP and the available information (more specifically, outcomes for the most recent quarters for the most central variables, described more fully in Appendix A.4). The only correlation at the one-quarter horizon is a positive one between the forecast error for GDP at t+1 and the exchange rate at t. So a weakening of the exchange rate has been associated on average with GDP being higher than forecast. For GDP forecasts one year ahead the forecast error correlates positively with the oil price, which could indicate an overestimation of the oil price's real impact on the Swedish economy. The GDP forecast errors eight quarters ahead correlate negatively with the policy rate in t–2 and t–3, which could be said to indicate that the Riksbank overestimated the policy rate's effects for GDP. Note, however, that this is just a partial and comparatively simple analysis. It is often quite likely that forecast errors arise from the combined effects of a variety of factors and the approach used here is not designed to detect this.

5.2 CPI and UND1X

In the case of UND1X inflation it is above all the upward movement in 2001–02 and the low level recently that proved hard to foresee in the evaluation period (see Chart 18). Inflation's sharp increase in the spring of 2001 did not feature in the previous year's assessments. When inflation had reached the higher level, upward revisions were made but inflation was expected to fall back again relatively soon. A decline did occur but more slowly than had been predicted. Inflation's drop at the end of 2003 was underestimated in forecasts during 2002 but was predicted from the beginning of 2003. Since 2004 UND1X inflation has been distinctly lower than in the preceding years; this has been foreseen to some extent in the short-run forecasts but the adjustments in predictions of the longer run have been too small.

An evaluation of the inflation forecasts should bear in mind that CPI definitions were changed at the turn of 2004. In general, the procedure adopted here has been to compare forecasts and outcomes in terms of the

[24] The forecasts do not, of course, explain the outcomes in a causal sense but it is still desirable to have a high degree of explanation in the efficiency regression.

[25] The degree of explanation represents the percentage of the outcome that the forecast predicts. A decreasing degree of explanation as the forecast horizon becomes longer is only natural because the uncertainty in the forecasts then tends to increase.

definitions that were used in the relevant issue of *Inflation Report*. There was some variation in this respect around the year-end when the new definitions were introduced.[26] The account in this section refers mainly to UND1X inflation; the tendencies for CPI inflation were broadly similar (see Charts 22 and 23, and Tables 5-7).

The mean error in the Riksbank's UND1X forecasts shows that bias was relatively small (Chart 19). In the two-year forecasts, inflation was overestimated by an average of around 0.2 percentage points (in annual terms), which in a statistical sense is not particularly remarkable. Table 6 also shows that there is no significant bias in the annual percentage changes in inflation. On average, however, CPI inflation two years ahead was overestimated by six tenths of a percentage point, which is not negligible.

Compared with the AR model (see Chart 20), the Riksbank's UND1X forecasts have a consistently lower RMSE. For every studied horizon, the Riksbank's RMSE is smaller than the series' standard deviation for the evaluation period, which may possibly indicate a relatively long forecasting capacity.[27] It is worth noting that the AR model is at something of a disadvantage in the comparison because it is ignorant of known monthly outcomes in the first forecast quarter. Generally, the first forecast quarter includes one or two known monthly outcomes. The number of known same-quarter price outcomes tends to vary; at the time of each forecast in the evaluation period, an average of approximately 45 per cent of the outcome for the first forecast quarter was available (see Table 10). The comparatively low forecast error for one quarter ahead (0.1 percentage point) in Chart 20 should be seen in this light.

The direction of the forecasts was consistently correct one quarter ahead and correct to over 90 per cent two quarters ahead (see Table 5). Direction is partly affected by a seasonal pattern that is allowed for in the forecasts and this no doubt helps to explain the high success rate.

Table 7 shows that the Riksbank's CPI and UND1X forecasts were weakly efficient for all forecast horizons, that is, the correlations between forecasts and outcomes are relatively high. As pointed out earlier, explanatory power is very high for the one-quarter forecasts. For longer horizons, explanatory power is around or just below 50 per cent, which can be considered high in relation to the length of the horizon.

Chart 18 gives the impression that the UND1X forecasts for 2004 and 2005 where constantly above the outcomes. Was the Riksbank too slow to adjust the forecast? Chart 21 presents this period in detail. We can notice that the drop in inflation in the early 2004 was well forecasted (with data up to the last quarter 2003). At the slightly longer horizons (for that particular forecast) the forecasts was actually lower than the outcomes, but that did not render any revisions of the forecast. This was (after the fact) a reasonable action. Moreover, the subsequent forecasts exhibit decent short-range performance and, thus, it was a reasonable behaviour to stick with the longer-term judgements. When UND1X dropped again in the first quarter, the Riksbank adjusted the whole forecast path, resulting in an underestimation of the inflation rate. The corresponding autoregression also predicts the drop in early 2004 (even though the magnitude was underestimated). Thereafter, the AR model overestimates the UND1X, and to a greater extent than the judgements. However, after the drop in 2005 the autoregression produces fairly accurate forecasts for 2005 and 2006.

In addition to the observations above, the RMSE is lower for the judgements, which implies that the average performance for the Riksbank's forecasts is superior to the autoregression. All in all, we find no indications that the Riksbank would have gained from engaging a different revision scheme, especially given data at hand, during the period of interest.

Tables 8 and 9 indicate that, in general, the Riksbank's inflation forecasts seem to have made efficient use of available information (in the somewhat limited sense intended here). The two-year forecast errors correlate negatively with the latest outcomes for actual inflation, the exchange rate and the interest rate. This means that inflation two years ahead was underestimated at times when actual inflation was low and overestimated when it was high. A weak exchange rate outcome coincides with an average overshoot in

[26] The earlier definition was used for forecasts up to an including *Inflation Report* 2004:1. Some of these forecasts extend beyond 2004 and outcome series for them have been extended synthetically in accordance with the earlier definition. This has been done by using detailed price statistics to arrive at simple estimates of what the rate of inflation after 2004 would have been with the earlier definition. The forecasts in *Inflation Report* 2004:2–2004:4 were based on the earlier definition up to the end of 2004 and the new definition after that. As of 2005, all forecasts used the new definition.

[27] However, the seasonality of UND1X adds to the variability (standard deviation) of the series, whereas a relatively stable seasonal pattern aids the forecasts. Compared to the standard deviation of seasonally adjusted UND1X figures the estimated forecast memory amounts to one quarter ahead, see also Chart 19.

eight-quarter inflation forecasts, which could be a sign that the exchange rate's impact was overestimated. Monetary policy's impact, on the other hand, may have been underestimated in the Riksbank's forecasts: an above-average policy rate coincides with actual inflation being lower on average than had been forecast.

5.3 Labour market

As mentioned in the section on data, the evaluation includes forecasts for the numbers in employment and the labour force, as well as unemployment, from 2002 onwards. The main focus is on employment; the labour force and unemployment are considered more briefly. In the evaluation period it is mainly the weak employment trend in 2003–04 and the strength of the recent upswing that proved hard to predict (see Chart 24). During 2002 the future dip was broadly expected to be deeper but shorter than proved to be the case. The increase in employment in the first half of 2005 was smaller than had been predicted the year before. After that, on the other hand, the expansion of employment was underestimated.

The labour market statistics also undergo revisions but these are by no means as extensive as in the case of the GDP statistics. However, the modifications that were introduced in the labour force statistics as of 2005 did tend to alter the picture. The data studied here is seasonally adjusted, so it should also be mentioned that certain revisions occur when a new filtration is done in connection with each new outcome. In order to handle these aspects, the forecasts for labour market variables are compared with real-time outcomes.

In the Riksbank's employment forecasts, the horizons from one to eight quarters ahead have a small bias (see Chart 25). The growth of employment was underestimated in the short run and overestimated to some extent for horizons about one year, though in a statistical sense these tendencies are very marginal. The bias measured as percentage annual change is also small (see Table 6). RMSE for the employment forecasts is less than the series' standard deviation in the evaluation period for all horizons studied (see Chart 26). This means that some forecasting capacity may remain two years ahead. However, the RMSE's are close to the standard deviation, especially for the second year, which implies that the memory in practice may well be shorter than two years.

The forecast errors in the Riksbank's employment forecasts are generally somewhat lower than those of the AR model. As regards the first-quarter forecasts, however, it should be noted that, as in the case of UND1X and CPI inflation, the Riksbank's forecasts have the advantage of often being able to incorporate known outcomes for one or two months. The direction of the one-quarter forecasts was correct in 60 per cent of the cases and in half of the two-quarter forecasts occasions (see Table 5).

The weak efficiency test shows that, like the GDP forecasts, the employment forecasts are weakly efficient for the 4- and 8-quarter horizons but not for the shortest horizon (see Table 7). Explanatory power is somewhat lower than for GDP except for the longest horizon.

Chart 27 presents the Riksbank's employment forecasts for 2005 and 2006. During this period, the outcomes have consistently been stronger than predicted. Again, one may claim that the Riksbank should have been more aggressive in the forecast revisions. Ex post this may seem like a reasonable claim, but the Riksbank's forecasts are relatively accurate compared to the autoregression. In fact, the forecasts resemble the AR, and subsequently the extra information (not contained in employment data) seems to have had no apparent impact on the Riksbank's forecasts. We can not conclude that it is obvious that the Riksbank should (ex ante) have revised more aggressively.

Correlations between the Riksbank's forecast errors and the information that was available at the time of each forecast are presented in Table 8 and 9. The errors one step ahead in the employment forecasts correlate positively with the exchange rate and the Riksbank's policy rate (a strong exchange rate and a low policy rate each accompany an overestimation of employment), negatively with GDP (strong GDP growth coincides with an underestimation of employment) and positively with inflation (high inflation accompanies an underestimation of employment). The opposite broadly applies to the one-year forecasts. For the two-year forecasts, forecast errors correlate negatively with available observations of the exchange rate and the policy rate in t to t-2: a low policy rate and a strong exchange rate has often coincided with an underestimation of employment's growth in the longer run.

In connection with the analysis of employment it may also be of interest to consider the properties of the labour force and unemployment forecasts. For the labour force, the forecast errors are close to the series' standard deviation at all horizons; the results are similar for forecasts with the AR model (see Chart 29). The forecast errors are partly connected with some underestimation at most horizons of the increase in

labour supply (see Chart 28). The underestimation in quarterly growth rates is not significant but shows up as a significant bias in annual growth rates (see Table 6). This is mirrored in the unemployment forecasts, which tend to undershoot outcomes in the period in question (see Chart 30). For the eight-quarter horizon, unemployment is underestimated by an average of approximately one percentage point. The unemployment forecasts, like those for employment, are somewhat more accurate than the AR model (see Chart 31). The direction of the forecasts for the labour force as well as for unemployment was correct in rather more than 60 per cent of the cases one quarter ahead and in rather more than half two quarters ahead (see Table 5).

5.4 Labour productivity

Productivity is a central variable in the Riksbank's monetary policy analysis. In practical work, productivity is mostly measured as labour productivity, calculated as GDP divided by hours worked. The requisite annual forecasts were made throughout the period studied here but it is only recently that hours worked has been forecast on a quarterly basis, which precludes a meaningful quarterly evaluation of labour productivity as defined above. The ratio of GDP to the number of persons in employment (employment) is a related but less conceptually suitable measure; it does not incorporate changes in the employed persons' average working hours, a factor that should be accounted for in the measurement of productivity.

However, a comparison of annual forecasts of GDP/hours and GDP/employment does indicate that the more limited variable catches the general tendencies relatively well in the evaluation period (see Charts 32 and 33). Both measures show that productivity was mainly underestimated in the period and most so in 2003 and 2004. The pictures for 2002 are less congruent, which has to do with the fall in average hours worked in the first years of the new century (see Chart 34).

The comparison of annual forecasts of GDP/hours and GDP/employment also shows that the forecast errors bias point in the same direction – these variables' outcomes were stronger on average than had been predicted. The forecast error correlations between annually observed labour productivity and GDP/employment are about 0.5 for both the 1-year and the 2-year horizon.

For a forecast evaluation, then, GDP/employment should serve as an approximation of the properties of the productivity predictions. This can be done only from 2002 onwards because quarterly employment forecasts were not produced earlier. As in the case of the separate analyses of GDP and employment, the evaluation of the ratio between them is done in real time.

The forecasts contain some positive bias at most horizons (see Chart 35). Measured as the percentage annual change, this bias is relatively large, 0.68, for forecasts two years ahead, see Table 6.[28] However, the bias is not significant. The AR model exhibits a similar bias in this period. RMSE is somewhat lower than or matches the series' standard deviation at all horizons (see Chart 36). The AR model is generally somewhat more accurate than the Riksbank's forecasts in this period. The direction of the forecasts was correct in over two-thirds of the cases one quarter ahead and in roughly half two quarters ahead (see Table 5).

The forecasts of GDP/employment are not weakly efficient one quarter ahead (see Table 7). The same applies to GDP and employment taken separately. For forecasts four and eight steps ahead, there is a virtually zero correlation between forecasts and outcomes. In other words, the Riksbank's forecasts of GDP/employment and thereby, most likely also, of labour productivity were not particularly accurate.

Starting from the hypothesis that the Riksbank's forecast errors for labour productivity had consequences for other forecast errors, one would expect to find correlations between the forecast errors for the different variables.[29] One such expected (negative) correlation concerns the forecasts errors for GDP/employment and UND1X. This is in fact one of the detected correlations (see Table 11). Thus there are indications that the underestimation of GDP/employment entailed an overestimation of UND1X. There is, moreover, a negative correlation between the forecast errors for GDP and UND1X, which are under- and overestimated respectively. This is also congruent with an underestimation of productivity.

[28] GDP/employment evaluation covers the period 2002-06. For that particular sample GDP and employment also exhibit positive bias.

[29] Note that it is correlations between the forecast errors for different variables that are analysed here, not those between forecast errors and outcomes (as in the analysis of efficient use of information).

5.5 Financial variables and oil price

The forecast errors are generally relatively large for the SEK/TCW exchange rate and the oil price. To a large extent this has to do with the forecasts' bias. An expected appreciation of the krona failed to materialise and the oil price was throughout considerably higher than predicted.

The Riksbank's predicted path of the exchange rate in the evaluation period was thus unduly strong. This bias underlies the greater part of the RMSE for the exchange rate (see Chart 37 and Table 6). On the whole, the forecasts with the AR model are more accurate than the Riksbank's assessments but RMSE's are large in this case as well (see Chart 38). For the one- and two-quarter horizons, RMSE's for the TCW-weighted nominal exchange rate are lower than the series' standard deviation. However, in the average case the same-quarter forecasts have good guidance in that data are available on a daily basis. This is also a major reason why the direction in the first quarter is correctly foreseen in almost 90 per cent of the cases, see Table 5. For the forecasts two quarters ahead, however, the direction is correct in only about half of the cases. The tendency for exchange rates to follow a random walk, making forecasting difficult, is a well-known research phenomenon (see e.g. Meese & Rogoff (1983) and Engel & West (2005)).

The level of the oil price was substantially underestimated by the Riksbank during the evaluation period, leading to relatively large forecast errors (see Charts 39 and 40). The results for the direction of the forecasts resemble those for the exchange rate, with correct predictions for a large proportion of the first-quarter forecasts and a considerably lower share for subsequent quarters (see Table 5).

6. Summary

All in all, the study shows that the Riksbank's forecasts perform satisfactorily compared with those of other forecasters and with simple AR models. In particular, the Riksbank's same-year GDP forecasts are often among the most accurate. The same applies to CPI inflation, though in this case the difference compared with NIER and the Consensus panel is smaller.

Compared with NIER's, the Riksbank's quarterly forecasts perform well. Their average forecast errors are somewhat smaller than NIER's for a majority of the variables but the differences are generally small and not statistically significant. As regards indications of bias (average over- or underestimation) in the forecasts, the general tendencies are similar for the Riksbank and NIER.

In comparison with the Consensus panel, the Riksbank's GDP forecasts likewise perform well, besides being often among the better forecasts in this study.

The Riksbank's inflation forecasts (CPI and UND1X) have a forecast memory up to and (possibly) including the second year. The forecasts are weakly efficient throughout and more accurate than the autoregression's. There are no signs of significant bias in the forecast errors. The point estimate of the CPI bias for the second year inflation rate is, however, non negligible.

The prediction memory of the Riksbank's GDP forecasts is likely shorter than that of the inflation forecasts. Forecast memory for GDP extends about four steps ahead. The GDP forecasts show no significant signs of bias. They, too, are weakly efficient at the four- and eight-quarter horizons. The forecast-outcome correlation is on the whole weaker than for the inflation forecasts.

The employment forecast memory is probably limited to the first year, but may even embrace somewhat longer horizons. In other respects their properties broadly resemble those of the GDP forecasts: no clear signs of bias and compliance with the weak efficiency criterion four and eight quarters ahead. At the same time, however, there are certain tendencies to underestimate labour supply and unemployment.

The development of labour productivity in the evaluation period was more favourable than had been predicted. The analysis of the approximate indicator of this, GDP/employment, shows a relatively weak forecasting performance that is partly connected with an average underestimation of the rate of increase. There are also signs that the Riksbank's misjudgements of productivity have played a part in the co-variation of the forecast errors for GDP and inflation.

The effective exchange rate (SEK/TCW) and the oil price have forecast errors that are relatively large. This has to do with the Swedish krona being systematically weaker than had been forecast and the oil price being higher than expected.

Although the revisions to the Riksbank's forecasts for GDP, UND1X and employment, at a first glance, seem to have been too small and too tardy (ex post) in some periods, the study finds no evidence that a different, more aggressive strategy in this respect would have resulted in forecasts that, on average, were more accurate.

7. References

Adolfson, M., M. K. Andersson, J. Lindé, M. Villani and A. Vredin, (2008), "Modern forecasting models in action: Improving analyses at central banks", *International Journal of Central Banking*, forthcoming.

Adolfson, M., S. Laséen, J. Lindé and M. Villani, (2007a), "Evaluating an estimated new Keynesian small open economy model", Sveriges Riksbank Working Paper Series No 203.

Adolfson, M., S. Laséen, J. Lindé and M. Villani, (2007b), "RAMSES – a new general equilibrium model for monetary policy analysis", *Economic Review* 2007:2, Sveriges Riksbank.

Andersson, M.K., (2000), "Do long-memory models have long memory", *International Journal of Forecasting* 16, 121-124.

Andersson, M.K. and M. Löf, (2007), "The Riksbank's new indicator procedures", *Economic Review* 2007:1, Sveriges Riksbank.

Andersson, M.K., G. Karlsson and J. Svensson, (2007), "An evaluation of the Riksbank's forecasting performance", Economic Review 2007:3, Sveriges Riksbank.

Assarsson, B. (2007) , "Riksbank forecasts of import prices and inflation", Economic Review 2007:3, Sveriges Riksbank.

Berg, C., P. Jansson and A. Vredin, (2004), "How useful are simple rules for monetary policy? The Swedish experience", Sveriges Riksbank Working Paper Series No. 169.

Bergvall, A., (2005), "Utvärdering av Konjunkturinstitutets prognoser" (Evaluation of NIER's forecasts), *Specialstudie*, no 5, March 2005, Konjunkturinstitutet.

Bernanke, B.S. , (2004), "The Great Moderation", Remarks at the Meetings of the Eastern Economic Association, Washington, DC.

Blix, M., J. Wadefjord,, U. Wienecke, and M. Ådahl, (2001), "How good is the forecasting performance of major institutions?", *Economic Review*, 2001:3, Sveriges Riksbank.

Blix, M., K. Friberg, and F. Åkerlind, (2002), "An evaluation of forecasts for the Swedish economy", *Economic Review*, 2002:3, Sveriges Riksbank.

Campbell, S.D., (2004), "Macroeconomic Volatility, Predictability and Uncertainty in the Great Moderation: Evidence from the Survey of Professional Forecasters", FEDS Working Paper No. 2004-52.

Chong, Y.Y. and D.F. Hendry, (1986), "Econometric evaluation of linear macroeconomic models", *Review of Economic Studies* 53, no. 4, 671-690.

Engel, C. and K.D. West, (2005), "Exchange Rates and Fundamentals", *Journal of Political Economy* 113, 485-517.

Galbraith, J. and G. Tkacz, (2006), How far can we forecast?: Forecast content horizons for some important macroeconomic time series, Department of Economics Working Paper 13, McGill University.

Engel, C. and K.D. West, (2005), "Exchange Rates and Fundamentals", *Journal of Political Economy* 113, 485-517.

Gu, Z., (2007), "Controlling for actual earnings: Does it mitigate or create spurious relations in forecast efficiency?", manuscript, Tepper School of Business, Carnegie Mellon University.

Meese, R.A. and K. Rogoff, (1983) "The out-of-sample failure of exchange rate models: Sampling error or misspecification", in J.A. Frankel (ed), *Exchange Rate and International Economics*, University of Chicago Press, Chicago, Il.

Mincer, J. and V. Zarnowitz, (1969), “The evaluation of economic forecasts”, in *Economic Forecasts and Expectations*, Mincer, J. (ed.), NBER, New York.

Nordhaus, W. D., (1987), “Forecast Efficiency: Concepts and Applications”, *Review of Economics and Statistics* 69 (4), 667-674.

Regeringen (2002), Vårpropositionen 2002/03:100 Bilaga 1: Svensk Ekonomi (Spring Budget Bill, Annex: The Swedish economy).

Riksrevisionen, (2006), “Det makroekonomiska underlaget i budgetpropositionerna” (The macroeconomic foundation in the draft budgets).

Schwartz, G., (1978), “Estimating the order of a model”, *The Annals of Statistics* 2, 461-464.

Stock, J.H and M. W. Watson, (2007), "Why Has U.S. Inflation Become Harder to Forecast?", *Journal of Money, Credit and Banking* 39, 3-33, Blackwell Publishing.

Appendix: Methods for the analysis

A.1 Average forecasting performance

For the collection of Riksbank forecasts, the average accuracy can be estimated. The engaged measure of accuracy is the root mean squared error (RMSE), which summarises the dispersion of forecast errors and any systematic deviation.[30] The lower estimated RMSE is, the better the forecasting performance. For a forecasting method that is invariably accurate, RMSE is zero. The mean error (ME = Outcome-Forecast) is also studied. RMSE is used to measure the average size of forecast errors, while ME is used to study when the forecasts contain bias, that is, a systematic over- or underestimation.[31] To assess forecast memory, the forecast errors are related to the respective series' standard deviation in the evaluation period.[32]

The starting point as regards the AR models used in the evaluation has been two different variants; a so-called AR(1), with just the value for the preceding period (first lag) as the explanatory variable, and a second variant with the number of lags based on a model choice procedure.[33] The most accurate of the AR variants are chosen as a benchmark. The models are estimated on data from 1993 up to the latest observation that was available for the forecast in question. As the results suggest that in most cases the AR(1) specification performs somewhat better, it is mostly this which has been used as a benchmark[34].

It may be worth noting that the number of observations for an evaluation generally decreases with the length of the forecast horizon. The present database with the Riksbank's quarterly forecasts, for example, contain 27 one-quarter forecasts for GDP but only 18 ten-quarter forecasts. If the average forecast error seems to be remarkable for longer horizons, it is therefore important to be aware that this may be at least partly due to the sample's diminishing size. There are, for instance, some cases where forecast accuracy seems to improve appreciably at the longest horizons which should thus be interpreted with caution. To make this less of a problem, the present study focuses on forecasts that stop at eight quarters ahead. One problem (in a limited sample) that remains is that (possible) outcomes which are hard to predict at the start of an evaluation sample do not affect RMSE estimates in the longer horizons. For example, the unexpectedly high UND1X outcomes in 2001 do not affect the estimated precision in the longest horizons, and the estimated RMSE can, therefore, be lower for longer horizons.

A.2 Direction forecasts

Besides studying the forecasts' average accuracy, it can be worth analysing the extent to which their direction is correct. This can be of interest above all in the short run, for instance to see how often GDP growth is correctly predicted to rise or fall from the latest to the next quarterly outcome, or the exchange rate to strengthen or weaken in the near future from its present level.

The analysis here concentrates on the direction of one- and two-quarter forecasts. More specifically, quarterly changes are studied for GDP, employment, labour force, GDP/employment, UND1X and CPI, and levels for unemployment, the nominal SEK/TCW exchange rate and the oil price. As a good deal of first-quarter information is already known when certain variables are forecast, it is more interesting to study the direction of forecasts two quarters ahead. This is particularly the case for the exchange rate and the oil price, for which forecasters can use daily outcomes up to the cut-off date for new statistics; it also applies to some extent to prices and labour market variables, for which one or two monthly outcomes are normally available.

---

[30] $RMSE(h) = \sqrt{\sum_{t=\tau_1}^{\tau_2} \left(y_{t+h} - y_{t+h}^{pred} \right)^2 / (\tau_2 - \tau_1 + 1)}$, where y_{t+h} is the outcome at time *t+h* and y_{t+h}^{pred} is the forecast at time *t+h* and *h* is the forecast horizon.

[31] $MSE(h) = Bias^2(h) + pf\ \mathrm{var}(h)$, where $pf\ \mathrm{var}$ is the forecast error variance.

[32] See Andersson & Löf (2007) for a detailed review of the forecast memory concept.

[33] The procedure for the choice of lags (the information criterion) that is used here is BIC, (see Schwartz, 1978). BIC selects a specification on the basis of available data and a statistical criterion. In retrospect the selected model may, of course, turn out not to be the best in terms of average forecast accuracy.

[34] The exception is UND1X where an AR(1) with seasonal dummies is used, since this functions better than both a common AR (1) and that which is generated by lag choice procedure.

It may be worth noting in this context that the direction in the second-quarter forecast is analysed on each occasion in relation to the forecast for the preceding quarter (not to the preceding quarter's outcome, which was not known at the time of the forecast).

A.3 Weak efficiency

Weak efficiency refers here to forecast errors that are unsystematic. Forecast errors will hopefully be small and show a random distribution. Put somewhat differently, it is desirable for forecasts' deviations from outcomes to be white noise. This can be written:

(A1) $$Y_t = \alpha + \beta F_t + \varepsilon_t \, .$$

Equation (A1) can be used to define some tests. A forecast is said to be weakly efficient if α =0 and β =1 jointly. The random term ε can be used for additional tests. It is, for example, desirable that the random terms are uncorrelated over time and independent of available information at the time of the forecast. It is worth pointing out that the test for weak efficiency is a rather strong test and that small deviations from the assumed white noise errors have implications on the estimated coefficients α and β. For more information about forecast efficiency, see Nordhaus (1987).

From the estimated equation (A1) it may be interesting also to examine the degree of explanation (coefficient of determination, or R^2) in the equation. This quantity tells how much of the variation in the outcomes that are explained by the variation in the forecasts. This should not be interpreted as a causal explanation, but rather as a (squared) correlation between outcomes and forecasts. Much of the qualitative information in the RMSE can also be found in the R^2 statistic, but the scale (zero to one) of this statistic may be intuitively more appealing.

A.4 Efficient use of information

Besides weak efficiency, efficient use of information is assessed with correlations between forecast errors and data that were available at the time of the forecast. The focus here is mainly on short forecast horizons and outcomes shortly before the forecast period (although lags up to twelve quarters are considered). The idea behind efficient use of information comes from Mincer & Zarnowitz (1969). However, their test is not used directly here because some doubts about it have been voiced; see, for example, Gu (2007).

As the Riksbank has not saved exhaustive sets of the data that was available at each forecasting round, the test of efficient use of information does not include all the relevant information. Important omissions are, for example, variables for the rest of the world. What is referred to here as available data is thus by no means a complete representation of all the information that was available at the time of each forecast. Another important limitation is that we only study bivariate correlations. A more advanced approach would be to study how a more inclusive amount of information (known at the time of the forecast) as a whole affected the forecast errors that were detected later. That, however, lies outside the scope of this article.

The available information included in the test comprises known outcomes for GDP, inflation, number of employed, number in the labour force, interest rate, exchange rate and oil prices (i.e. variables that are included in the forecast evaluation).

A.5 Revision patterns

Economic forecasts are constructed from interpretations of data that are currently available, together with the forecaster's appraisal of the economic situation. The results in this article are a clear illustration of forecasts' considerable uncertainty. The forecasts are revised because forecast errors are observed when new outcomes become available. Revisions can, of course, be made for other reasons, too, such as taking new indicator statistics into account or a macroeconomic reassessment.

To analyse the revision behaviour of, e.g., the Riksbank's forecasts are extremely difficult per se. After the fact (ex post), it is easy to conclude that the forecasts have been revised too little and too late or vice versa. But, that conclusion is trivial and not especially informative. The important thing is whether the forecaster has used available information in an acceptable way. One measure of this is again RMSE, or RMSE relative to a competing forecast. However, RMSE demonstrates the average performance of a forecaster

and says (typically) nothing about the performance during particular periods. Thus, it may be of interest to look deeper into those particular periods.

We have chosen to scrutinize the way the Riksbank has revised their forecasts in occasions when (ex post) the revisions appear not fully satisfying. One such occasion is forecasts for GDP 2001. The forecasts for GDP 2001 where revised downwards in the reports from 2000 and 2001, but never fully reaching the low outcome in mid 2001. To assess the forecasts (and their revisions) we compare those with an AR. The autoregression summarizes information in the modelled series but accounts for no other information.

Moreover, the autoregression produces reasonably competitive forecasts in general. The deviations between the Riksbank's and the AR forecasts may be interpreted as the Riksbank's take on extra information. This information may sometimes aid the forecasts, but in other occasions the extra information may worsen the accuracy, see for instance Adolfson et al (2008). Recently, the employment outcomes have surprised the Riksbank, i.e., employment increases have on average been stronger than the corresponding forecasts since 2005.

One other typical period of seemingly slow revision reactions is the Riksbank's inflation (represented here by UND1X) forecasts during 2004. These forecasts (all) indicated rising inflation that never materialized, or at least not until at a later stage.

It deserves to be emphasized that the occasions above are selected to demonstrate periods when the Riksbank's forecasts appears to have been revised too much or too little. This does not have to be true, and there are other periods when the forecasts seem to be appropriate. However, these occasions are not scrutinized in detail here. We recommend the interested reader to further explore Charts 14, 18, 24 and 33 for a complete picture.

An alternative way of evaluating the Riksbank's revision behaviour is to estimate a function that describes this behaviour, given known outcomes and known forecast errors. The function can then be modified so that only outcomes are taken into account or so that old forecast errors are given more weight (than the Riksbank did). A comparison of GDP forecasts reveals that when only known outcomes are taken into account, the forecasts are almost as accurate as those based on the Riksbank's estimated historical strategy. When longer-term forecasts are revised markedly in response to each registered forecast error, the forecaster will, on average, perform considerably worse than one that follows the Riksbank's estimated historical strategy. An experiment shows that doubling the weight assigned to observed forecast errors means that RMSEs are also doubled. This applies for all horizons. It indicates that the relevant information for forecasts is contained in the data and that earlier forecast errors do not have much to say about future revisions, especially at longer horizons.

A.6 Forecast-error correlations

Relationships between forecast errors (for different variables) are studied in this article in terms of correlations. In this way it is possible to identify signs of inconsistency in forecasting methods. It is reasonable to assume that over a long evaluation period these forecast errors should be uncorrelated. With a short evaluation period, however, economic shocks are likely to have a dominant effect on forecast errors, making it natural to find certain correlations between forecast errors for different variables.

The number of forecast errors available for study is, however, very large and an unconditional examination would not be manageable. One way of tackling this problem is to construct and test hypotheses. Given the sample period, a hypothesis that comes to mind is that the Riksbank's forecast errors were affected by import prices being lower than expected. This issue is discussed by Assarsson (2007). Another natural hypothesis is that misjudgements of productivity growth affected other forecast errors. If this hypothesis holds, there should be a negative correlation between the errors for GDP and inflation.

As in the study of efficient use of information, only bivariate correlations are used. The limitations pointed out in Appendix A.4 are therefore relevant here as well.

A.7 Revisions to GDP data

An evaluation of GDP forecasts may have cause to identify effects of revisions to the outcome data. Preliminary outcome data can be said to mainly mirror the economic situation as it was perceived at the time of each forecast, while later outcomes are presumably closer to what the situation actually amounted

to (the two are not necessarily congruent). This makes it of interest to study both the preliminary and the latest versions of outcome data, which is accordingly done in this forecast evaluation. This is done by focusing first on how the forecasts deviate from the first published outcome and then supplementing this with an analysis of how outcome data are revised from the first version onwards. In this way, both dimensions are covered simultaneously and their effects can be separated.

Revisions of GDP data have many causes. In the first place, they are done as additional information becomes available. The first calculation of a quarterly outcome for GDP has to make do with information that is relatively sparse and needs to be supplemented with various types of estimation; later, as the material becomes more complete, the outcome may have to be revised. Another reason for revisions is, for instance, that definitions and methodological conventions change; this may result in a different outcome even though the underlying information is the same as before. In the present study there is a further source of revisions in that for seasonally-adjusted data, the entire data series is filtered anew before each publication. Revisions to GDP data are often occasioned by combinations of the various factors.

Statistics Sweden's outcomes for seasonally-adjusted quarterly GDP growth rates in the period 2000–06 have been revised by an average of about 0.3 percentage points from the first published figure to the current outcome series. On average, moreover, there was a marginal underestimation (0.05 percentage points) of quarterly growth in this period. The largest single revision of seasonally-adjusted quarterly growth rates is in the region of 0.6 percentage points (see Chart 42). The revisions of annual growth rates are liable to be larger because they include any revisions to the first three quarters. The average revision of the percentage annual change in seasonally-adjusted data amounts to over 0.6 percentage points and the average underestimation measured in this way is about 0.3 percentage points. The largest single revision of the annual growth rates is about 1.5 percentage points.

It may also be of interest to analyse the revisions' temporal distribution – how the adjustment from the first estimate to the latest outcome is spread over the intervening period. Chart 43 presents average differences from the latest outcome for the first twelve versions of GDP outcome in the period 2000–06. For quarterly growth rates, the initial difference of 0.3 percentage points has been more or less halved by the time of the fifth version of the outcome, after which the difference decreases gradually and rather slowly. For annual growth rates, it takes roughly twice as long for the corresponding difference to be halved. The marginal underestimation of quarterly growth persists on average up to the tenth version and then disappears (see Chart 44). For the annual growth rates, the average underestimation decreases more gradually but it, too, is relatively small by the time of the tenth outcome version.

In the evaluation period the average size of the revisions tends to vary somewhat between the quarters (see Table 12). In the case of the annual growth rates, the observed average underestimation is more marked in the year's first two quarters. The average size of the revisions is also somewhat larger in quarters one and two but the differences here are considerably less marked. The corresponding analysis of seasonally-adjusted quarterly growth rates shows a slightly different picture. Here it is quarters one and four that were underestimated, while the tendencies to systematic errors in quarters two and three are negligible. The average size of the revisions is somewhat larger for quarters one and four but the differences from quarters two and three are slight. All in all, there are certain indications that preliminary GDP figures for quarter one can be somewhat more uncertain than those for the other quarters but the picture is not clear-cut.

The Riksbank's average first-quarter forecast error in relation to the latest outcome is 0.42 and thus larger than the evaluation period's average revision of outcome data, 0.30. In other words, the Riksbank's forecasts have, on average, been over one-tenth of a percentage point further away from the final outcome than Statistics Sweden's first estimate of the same outcome. This means that for predicting the final outcome, there seems to be no case for placing greater reliance on an own forecast one quarter ahead compared with Statistics Sweden's first estimate of the same quarter's GDP.

Tests of weak efficiency provide further indications of this. Table 13 shows that the Riksbank's forecasts are not weakly efficient estimates of final GDP outcomes for horizons up to four quarters. Results of tests as to whether preliminary versions of the National Accounts are weakly efficient forecasts of final National Accounts are presented in Table 14. Preliminary National Accounts display better forecasting properties than the Riksbank's forecasts.

Tables and charts

Table 1. Analysis methods

Method	Purpose
• Graphical analysis	To present and roughly describe the Riksbank's forecasts.
• Comparisons with other forecasters	To evaluate the Riksbank's performance compared to competing forecasts.
• Direction of forecasts	To investigate how often the direction is correctly predicted.
• RMSE	To measure the average forecasting performance.
• Encompassing test	To investigate if two forecasts benefit from being combined.
• Mean error (ME, Bias)	To detect if forecasts are biased.
• Forecast memory	To establish the maximum informative forecast horizon.
• Weak Efficiency	To formally test the efficiency of the forecasts.
• Efficient use of information	To investigate if the available information (data) have been utilized in a proper fashion.
• Forecast error correlation	To study if the forecast error of a particular variable co-variate with the error of other variables.
• Revisions vs. AR	To investigate if the Riksbank's revision pattern has been reasonable.

Note. The methods above are applied to each of the investigated forecast variables. Some methods serve as a tool for description, whereas, other methods are employed for statistical inference.

Table 2. Differences between the Riksbank's and NIER's RMSEs for selected variables and horizons

	GDP	Employ-ment	Labour force	UND1X	CPI	SEK/ TCW	Policy rate
1	0.00	0.02	-0.04	-0.01	0.03	1.23	-0.07
2	-0.02	-0.01	0.05	0.08	0.04	0.51	0.06
3	0.01	0.05	0.07	-0.03	-0.05	0.65	0.15
4	-0.03	0.01	-0.05	-0.13	-0.09	0.78	0.14
5	0.00	0.00	-0.04	-0.20	-0.07	0.89	-0.05
6	-0.02	-0.02	-0.08	-0.06	-0.02	0.89	-0.11
7	-0.09	-0.04	0.00	0.00	-0.01	0.69	-0.18
8	-0.06	-0.03	-0.01	0.08	-0.05	0.71	-0.37

Note. Differences are positive (negative) when the Riksbank's RMSE is higher (lower) than NIER's. The calculations are based on the following variables: RMSEs for quarterly changes in GDP, Employment and Labour force; RMSEs for annual percentage changes in UND1X and CPI; RMSEs for percentage deviation in the level of SEK/TCW; RMSEs for changes in the level of the policy rate. Calculations are done for the period 2001-06 for all variables except for GDP where the period is 2003-06 (see also section 2.1).

Table 3. The Riksbank's MEs for selected variables and horizons

	GDP	Employ-ment	Labour force	UND1X	CPI	SEK/ TCW	Policy rate
1	0.11	0.06	0.10	0.02	0.03	0.29	-0.03
2	0.05	0.07	0.18	0.13	0.06	1.29	-0.13
3	0.05	0.01	0.11	0.15	0.00	2.01	-0.23
4	0.10	-0.05	0.12	0.16	-0.10	2.55	-0.36
5	0.11	-0.03	0.08	0.09	-0.25	2.45	-0.42
6	0.08	0.04	0.06	-0.12	-0.49	3.29	-0.51
7	0.08	0.00	0.06	-0.31	-0.70	3.31	-0.69
8	0.06	0.03	0.10	-0.45	-0.87	3.32	-0.86

Note. See the note to Table 2.

Table 4. NIER's MEs for selected variables and horizons

	GDP	Employ-ment	Labour force	UND1X	CPI	SEK/ TCW	Policy rate
1	-0.01	0.12	0.10	0.00	0.02	0.08	0.05
2	0.04	0.09	0.13	0.09	0.11	0.46	0.14
3	0.06	0.04	0.12	0.16	0.14	0.72	0.03
4	0.08	0.03	0.16	0.25	0.15	0.89	-0.14
5	0.12	0.04	0.17	0.26	0.03	1.00	-0.40
6	0.13	0.05	0.15	0.03	-0.37	1.22	-0.71
7	0.14	0.04	0.11	-0.15	-0.67	1.41	-1.08
8	0.14	0.06	0.12	-0.25	-0.93	1.55	-1.45

Note. See the note to Table 2.

Table 5. Results from the evaluation of direction

	One quarter ahead		**Two quarters ahead**	
	Proportion correct (%)	*p*-value	Proportion correct (%)	*p*-value
GDP	68	0.03	50	0.50
UND1X	100	0.00	96	0.00
CPI	100	0.00	93	0.00
Employment	60	0.19	50	0.50
Labour force	63	0.13	53	0.41
Unemployment	61	0.17	56	0.31
GDP/Employment	68	0.05	53	0.41
SEK/TCW	89	0.00	56	0.28
Oil price	85	0.00	41	0.83

Note. The *p* denotes the probability of the null hypothesis that the direction is correct in not more than 50 per cent of the forecasts (in a given population).Thus, a low *p*-value indicates that more than 50 per cent of the forecasts point in the correct direction. The evaluation period is 2002:1–2006:3 for employment, labour force, unemployment and GDP/employment and 2000:1-2006:3 for the other variables.

Table 6. RMSEs, MEs and distribution of mean squared errors for quarterly forecasts 2000–06
All variables in percentage annual changes except exchange rate (log level).

				Mean error		Prop. of MSEs (%)	
Variable	Horizon	RMSE	S.D.	ME	*p*-value	Bias	Var. (pf)
GDP	4	0.96	1.04	-0.05	0.86	0	100
	8	1.20	1.04	-0.37	0.34	10	90
CPI	4	0.65	0.95	0.06	0.75	1	99
	8	1.13	0.95	-0.60	0.08	28	72
UND1X	4	0.62	0.80	0.26	0.10	18	82
	8	0.82	0.80	-0.19	0.47	5	95
Employment	4	0.66	0.95	0.27	0.32	17	83
	8	0.92	0.95	0.20	0.65	5	95
Labour force	4	0.84	0.78	0.69	0.01	67	33
	8	0.98	0.78	0.49	0.12	25	75
GDP/Emp.	4	1.13	0.96	0.12	0.77	1	99
	8	1.21	0.96	0.68	0.11	38	62
Exch. rate	4	5.37	4.59	3.61	0.01	45	55
	8	6.60	4.59	5.10	0.01	60	40

Note. RMSE is the estimated root mean squared error and ME is the estimated mean error (bias). The *p*-value indicates the probability of not rejecting the null hypothesis that the forecasts are unbiased (null hypothesis: bias=0). The test statistics are adjusted for oversampling/overlapping information. The proportion of mean squared errors (Prop. of MSEs) denotes the percentages of the mean squared error that come from bias and variation in forecast error, respectively. Employment and GDP/employment computations are based on data for the period 2002-06.

Table 7. Test of weak efficiency of the Riksbank's forecasts

		Estimated equation			Efficiency test			Residuals	
Variable	Hor.	α	β	R^2	F stat	fg	*p*-value	AK (p-v)	NF (p-v)
Exch. r.	1	-8.13	1.07	0.784	0.62	2; 25	0.548	0.809	0.000
	4	128.96	0.01	0.000	16.15	2; 23	0.000	0.001	0.345
	8	258.11	-1.05	0.298	51.25	2; 19	0.000	0.092	0.870
Policy r.	1	-0.03	1.01	0.996	1.64	2; 25	0.214	0.260	0.000
	4	-0.15	0.97	0.692	2.64	2; 23	0.092	0.001	0.567
	8	0.90	0.54	0.212	8.06	2; 19	0.003	0.001	0.426
GDP	1	0.33	0.50	0.493	12.26	2; 25	0.000	0.528	0.989
	4	-0.03	0.96	0.373	0.52	2; 23	0.601	0.015	0.377
	8	0.08	0.80	0.094	0.42	2; 18	0.661	0.038	0.284
Empl.	1	0.08	0.50	0.393	6.14	2; 17	0.010	0.421	0.736
	4	-0.30	2.23	0.258	0.95	2; 15	0.408	0.135	0.573
	8	-0.03	1.33	0.182	0.10	2; 11	0.903	0.290	0.738
GDP/Em.	1	0.37	0.30	0.140	7.26	2; 17	0.005	0.432	0.437
	4	0.73	-0.28	0.012	1.78	2; 14	0.207	0.027	0.648
	8	0.57	0.40	0.007	2.68	2; 10	0.109	0.274	0.591
UND1X	1	0.03	0.96	0.974	1.19	2; 25	0.320	0.400	0.734
	4	0.06	0.97	0.490	0.17	2; 23	0.838	0.481	0.000
	8	-0.05	0.90	0.552	1.02	2; 19	0.248	0.424	0.674
CPI	1	0.04	0.98	0.971	2.00	2; 25	0.157	0.321	0.328
	4	-0.01	0.94	0.400	0.13	2; 23	0.879	0.548	0.000
	8	-0.12	0.86	0.433	3.07	2; 19	0.070	0.404	0.367

Note. The forecasts are compared with the first available outcomes. The table shows estimated coefficients and explanatory power from equation (A4), tests of the weak efficiency hypothesis and residual diagnostics.

Table 8. Correlations between forecast errors and data
First and fourth forecast horizons, short lags in data

Forecast errors for	Horizon	Data for	Lag	Sign	Significance level (%)
GDP	1	TCW	t	+	10
	4	Policy rate	t	-	10
	4	WDPOQ	t : t-3	+	1,5,10
	4	GDP	t-4,t-5	+	1
	4	Employment	t-6. t-7	-	10
UND1X	1	TCW	t : t-5	+	1,5,10
	4	TCW	t : t-4	+	1,5,10
	1	Policy rate	t : t-5	+	1,5,10
	4	Policy rate	t, t-1	+	5
	1	WDPOQ	t: t-6	-	1,5,10
	1	GDP	t : t-3	-	1,10
	1	GDP	t-4	+	5
	4	GDP	t-1 : t-3	-	5
	1	UND1X, CPI	t : t-8	+	1, 5, 10
	4	UND1X, CPI	t-3,t-5,t-7	+	5,1,1
	1	Employment	t-4	+	10
Employment	1	TCW	t : t-3	+	5,10
	4	TCW	t-1 : t-6	-	1,5,10
	1	Policy rate	t, t-1	+	5
	4	WDPOQ	t : t-6	+	1
	1	GDP	t, t-1	-	5,10
	4	GDP	t-1, t-3, t-5	+	10, 10, 1
	1	UND1X	t	+	5
	4	UND1X	t-3	-	10
	1	CPI	t	+	5
	4	CPI	t-3	-	10
	4	Employment	t, t-1	+	10,5

Note. The table shows variables that correlate at the specified horizons with the same and other variables, with lags. The sign indicates whether the correlation is positive or negative and the level of significance is that at which the the hypothesis that the correlation is zero is rejected.

Table 9. Correlations between forecast errors 8 quarters ahead and data

Forecast errors for	Data for	Lag		Signifi-cance (%)	Data for	Lag		Signifi-cance (%)
GDP	TCW	t-11	+	10	GDP	t,t-1	+	1
	Policy r.	t-2 : t-3	-	5	GDP	t-4	-	5
	WDPOQ	t	+	10	Employ.	t-11	-	10
UND1X	UND1X	t	-	5	GDP	t-9	-	10
	TCW	t : t-2	-	5,10	Employ.	t-2,t-3	-	10,10
	TCW	t-10,t-12	+	10,1	Employ.	t-8	-	10
	Policy r.	t : t-3	-	1,5				
	Policy r.	t-8, t-9	+	5				
Employment	TCW	t : t-2	-	5	GDP	t, t-1	+	5,10
	TCW	t-8 : t-11	+	5,10	GDP	t-5,t-9,t-11	-	5,10
	Policy r.	t : t-2	-	10	UND1X	t	-	5
	Policy r.	t-8 : t-10	+	1,5,10	CPI	t	-	10
	WDPOQ	t	+	10	Employ.	t-4	-	1
					Employ.	t-8	-	10

Note. See the note to Table 8.

Table 10. Number of known monthly CPI outcomes for next forecast quarter at the time of the *Inflation Report* forecast

	2000	2001	2002	2003	2004	2005	2006
IR1	2	2	1	1	2	1	1
IR2	1	1	1	1	1	2	1
IR3	2	2	2	2	2	2	0
IR4	1	1	1	1	1	1	-

Table 11. Forecast error correlations

Forecast errors for		Horizon	Sign	Significance (%)
GDP	UND1X	2	-	5
GDP	Employment.	1	-	5
		4	-	10
Employment.	UND1X	1	+	10
		5	-	10
UND1X	GDP/Employm..	8	-	10
GDP	GDP/Employm..	2-5	+	1-5
Employment.	GDP/Employm..	3	-	5

Note. The table shows the direction of correlations between forecast errors for different variables. Significance denotes the level of significance at which the zero correlation hypothesis is rejected.

Table 12. GDP revisions
Percentage points

Quarter	Annual growth rate		Quarterly growth rate	
	RMSE	ME	RMSE	ME
1	0.70	0.52	0.35	0.13
2	0.66	0.43	0.20	-0.01
3	0.50	0.12	0.32	-0.01
4	0.50	0.09	0.38	0.05
1–4	0.60	0.29	0.32	0.04

Table 13. Test of weak efficiency of Riksbank forecasts (compared with the final outcome)

		Estimated equation			Efficiency test			Residuals	
Variable	h.	α	β	R^2	F stat	d.f.	*p*-value	AK (p-v)	NF (p-v)
GDP	1	0.44	0.40	0.218	8.51	2;25	0.002	0.073	0.596
	4	0.71	-0.06	0.001	3.77	2; 23	0.039	0.002	0.606
	8	0.25	0.75	0.109	0.79	2; 18	0.467	0.012	0.445
NA prel.	1	0.05	1.01	0.706	0.96	2; 25	0.398	0.004	0.596
Employ.	1	0.09	0.49	0.302	4.33	2; 17	0.030	0.645	0.866
	4	-0.26	2.15	0.197	0.59	2; 15	0.569	0.108	0.569
	8	0.04	1.06	0.103	0.09	2; 11	0.914	0.399	0.697
NA prel.	1	0.00	1.07	0.886	0.32	2; 17	0.732	0.115	0.894

Note. See the note to Table 6. The table presents results of efficiency tests of Riksbank forecasts and of preliminary compared with revised National Accounts outcomes.

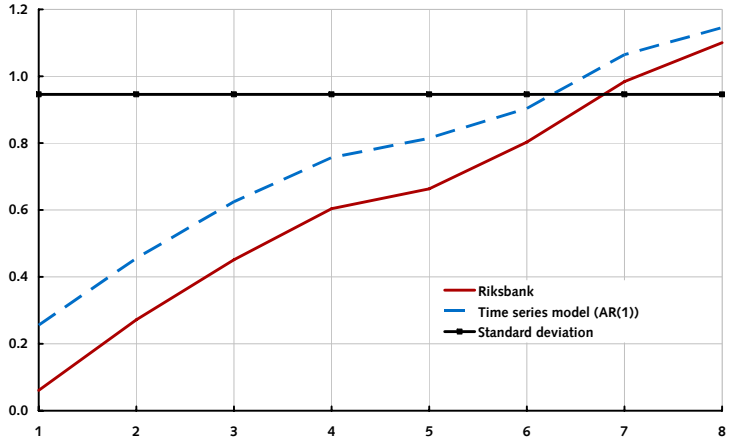
Table 14. Tests of weak efficiency of the first outcome of National Accounts and Riksbank forecasts

	Estimated equation			Efficiency test			Residuals	
NA version	α	β	R^2	F. stat	d.f.	*p*-value	AK (p-v)	NF (p-v)
1	0.06	1.07	0.79	3.59	2; 16	0.051	0.671	0.547
2	0.05	1.05	0.73	1.38	2; 15	0.281	0.288	0.043
4	0.02	1.07	0.87	2.59	2; 12	0.117	0.486	0.637
8	-0.01	1.06	0.96	0.60	2; 6	0.579	0.129	0.788
Riksbank forecast								
first	0.369	0.638	0.14	1.99	2; 16	0.169	0.047	0.637
second	0.319	0.783	0.41	4.38	2; 16	0.030	0.131	0.539

Note. See the note to Table 6. The table shows efficiency tests for the first National Accounts release as a forecast for the final outcome, as well as for the Riksbank's forecasts. As the same NA version has frequently been used for two consecutive Riksbank reports, two forecast series have been produced for the Riksbank: the first is used to test the forecasts for which an NA version was used for the *first* time, the second to test those for which the same NA version was used for the *second* time.

Chart 1. RMSE for policy rate forecasts

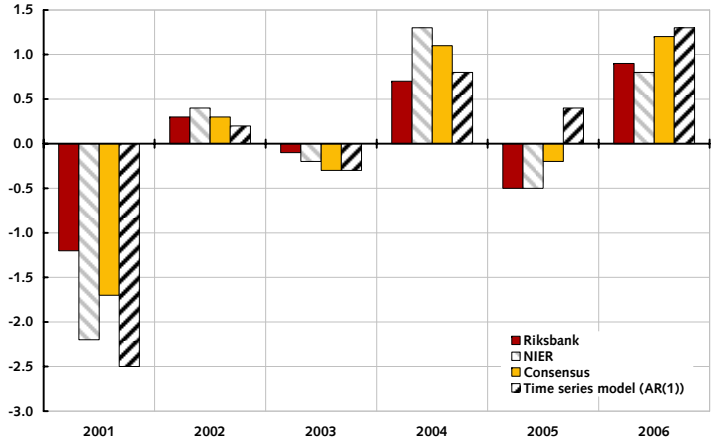
Percentage points



Note. The chart plots the root mean squared error in the Riksbank's forecasts and for forecasts with an autoregression, at forecast horizons from one to eight quarters ahead together with the standard deviation of the policy rate data.

Chart 2. GDP forecast errors for same-year average made in Q1, 2001–06

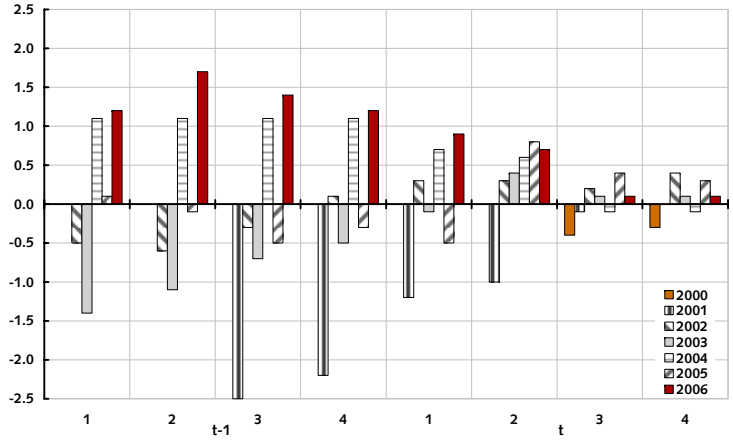
Per cent



Note. The bars represent the errors for each forecaster. Forecast error is defined as outcome less forecast. The forecasts are those constructed at the first quarter (Q1) of the year in question.

Chart 3. The Riksbank's GDP forecast errors, 2000-06

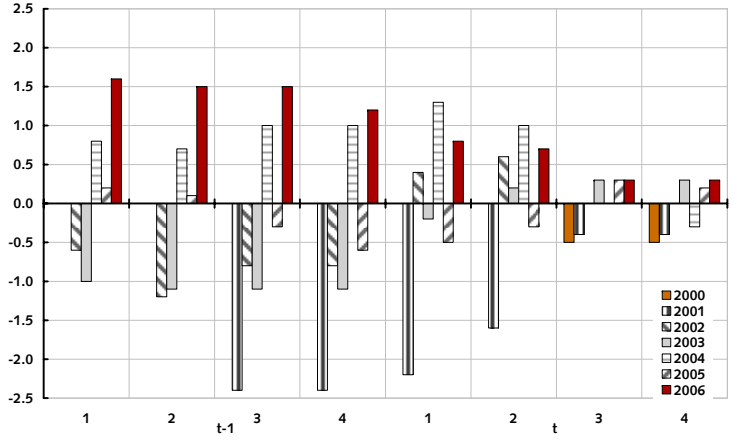
Per cent



Note. The bars represent the forecast errors for each year that were made in the four quarters (1,2,3,4) of the previous year (t-1) and in the four quarters of the year being forecast (t).

Chart 4. NIER's GDP forecast errors, 2000-06

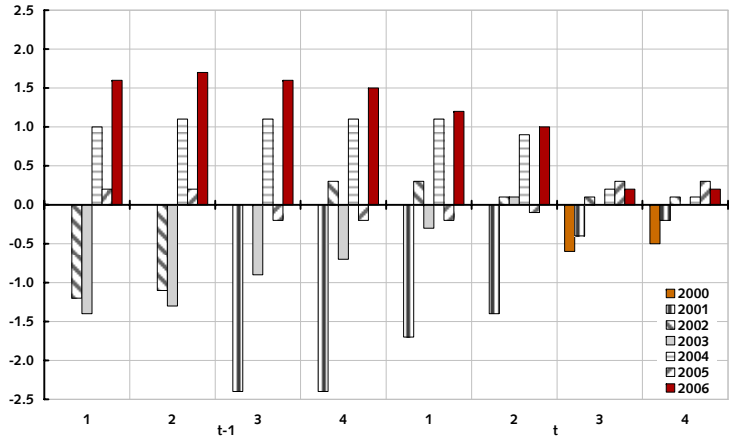
Per cent



Note. See the note to Chart 3.

Chart 5. Consensus panel forecasts, 2000-06

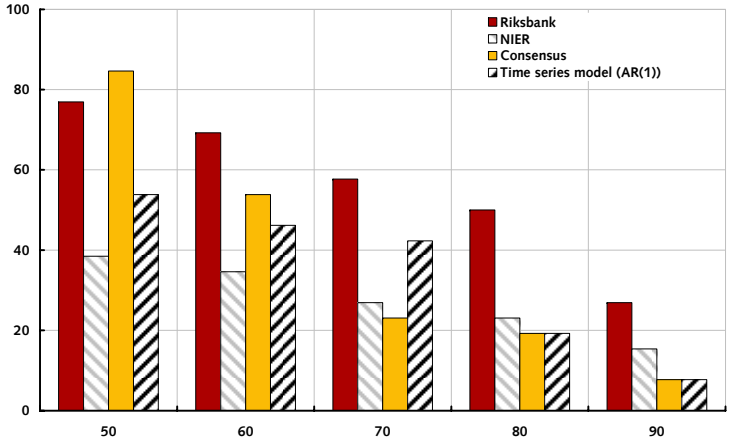
Per cent



Note. See the note to Chart 3. The errors reported are the average of the participants of the Consensus Economics panel, excluding the NIER.

Chart 6. Ranking for same-year GDP

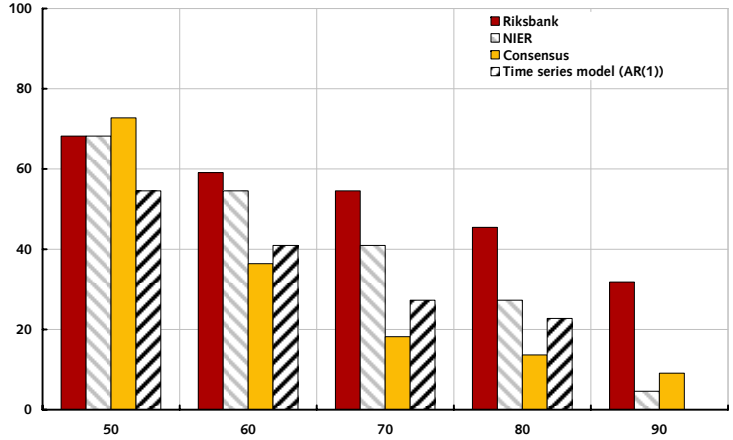
Per cent



Note. The vertical scale denotes the percentage of all forecasting rounds in which a forecaster performed better than the horizontal scale's different percentages of all forecasters. For example, to find out how frequently the Riksbank performed better than 50 per cent of the other forecasters, go first to 50 on the horizontal axis and read off the red bar against the vertical scale; in this particular case, the Riksbank accordingly did better than 50 per cent of the other forecasters in almost 80 per cent of the forecasting rounds.

Chart 7. Ranking for next-year GDP

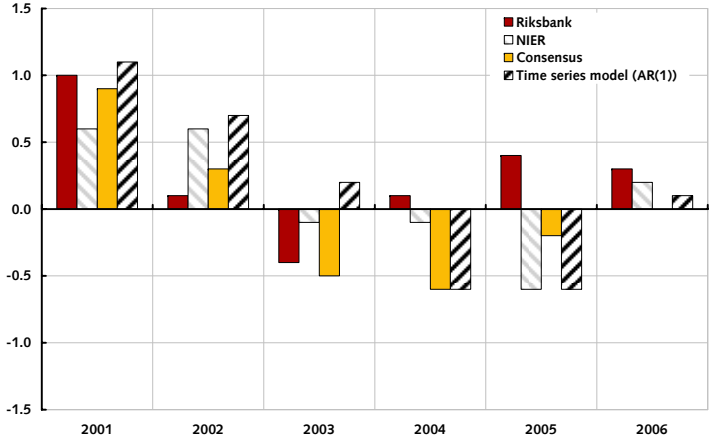
Per cent



Note. See the note to Chart 6.

Chart 8. CPI forecast error for same-year average made in Q1,

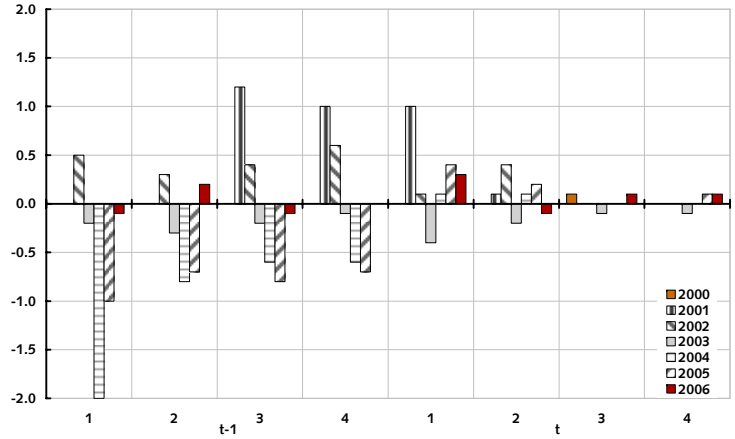
Per cent



Note. See the note to Chart 2.

Chart 9. Riksbank CPI forecast error, 2000-06

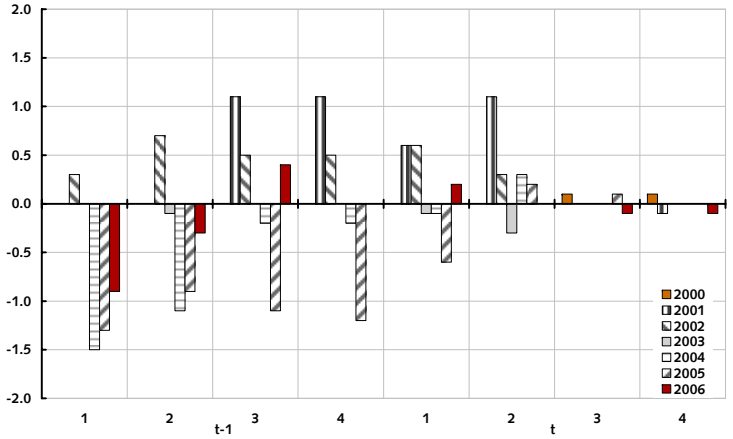
Per cent



Note. See the note to Chart 3.

Chart 10. NIER CPI forecast error, 2000-06

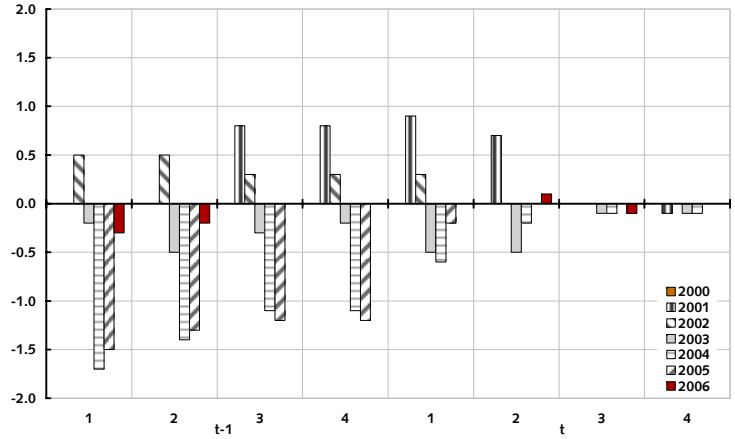
Per cent



Note. See the note to Chart 3.

Chart 11. Consensus panel forecast error, 2000-06

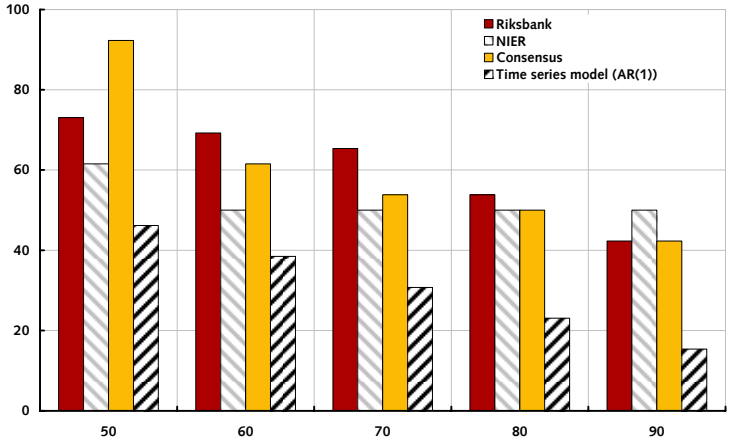
Per cent



Note. See the note to Chart 3 and 5.

Chart 12. Ranking for same-year CPI

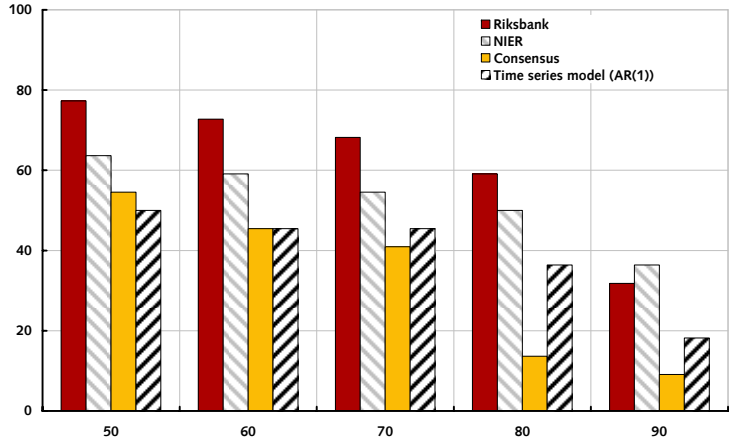
Per cent



Note. See the note to Chart 6.

Chart 13. Ranking for next-year CPI

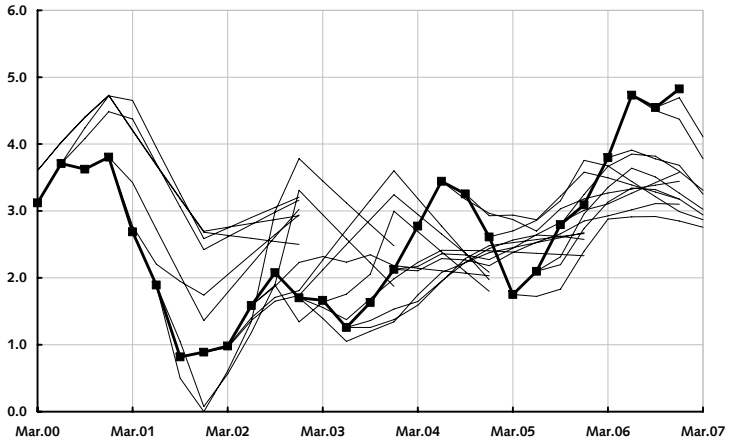
Per cent



Note. See the note to Chart 6.

Chart 14. GDP, outcomes and forecasts

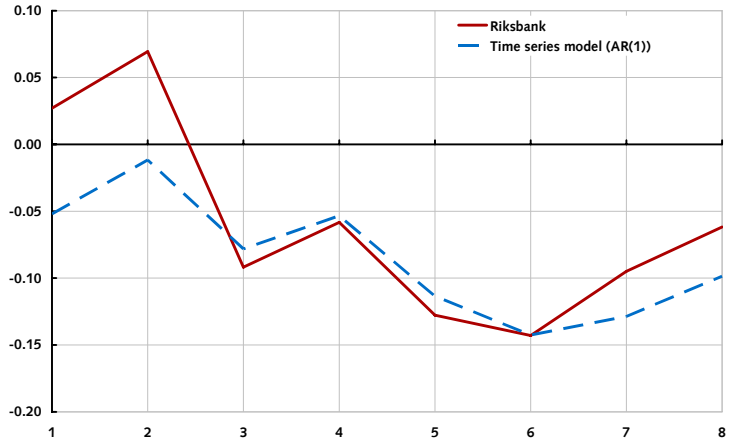
Annual percentage change



Note. The curve marked with squares is made up of each quarter's first available outcome and the other curves represent the Riksbank's forecasts at each forecasting round.

Chart 15. ME for GDP forecasts

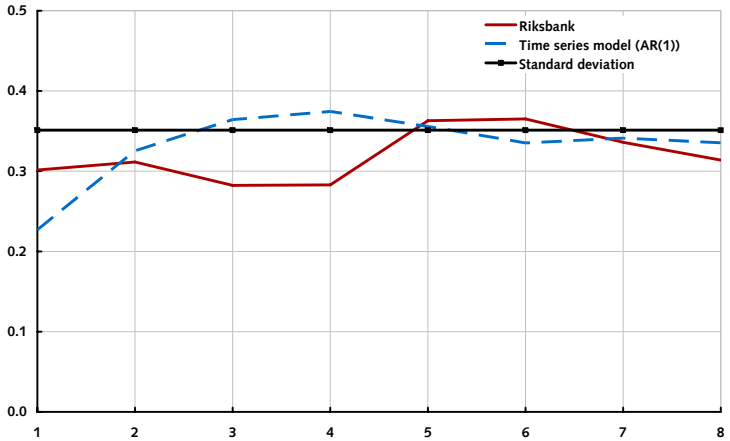
Percentage points (calculated on quarterly changes)



Note. The chart shows the mean error in the Riksbank's forecasts and in forecasts from an auto-regression at forecast horizons from one to eight quarters ahead. Error is outcome less forecast.

Chart 16. RMSE for GDP forecasts

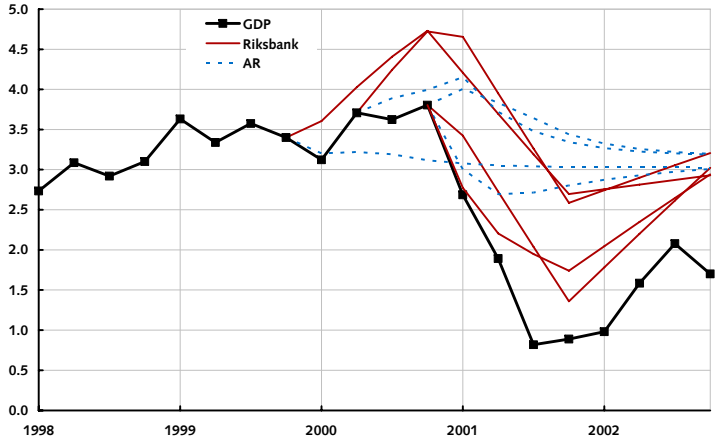
Percentage points (calculated on quarterly changes)



Note. See the note to Chart 1.

Chart 17. GDP, outcomes and forecasts

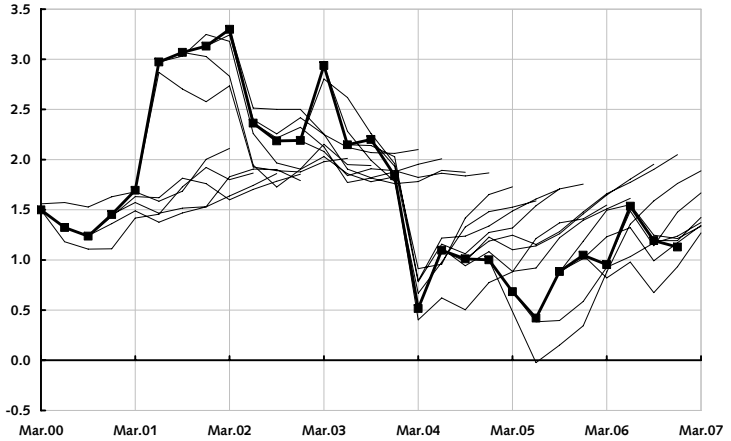
Annual percentage change



Note. The curve marked with squares is made up of each quarter's first available outcome and the other solid curves represent the Riksbank's forecasts at each forecasting round. Dotted lines represent the corresponding AR forecasts.

Chart 18. UND1X, outcomes and forecasts

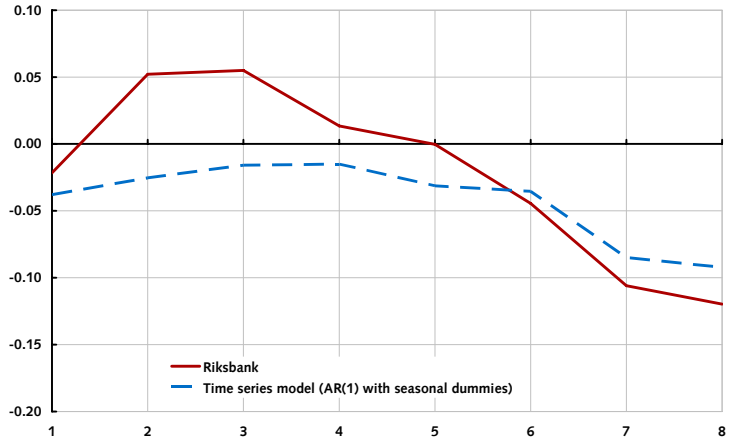
Annual percentage change



Note. The curve marked with squares is based on the earlier definition of UND1X up to the turn of 2004 and on the new definition after that. The other curves represent the Riksbank's forecasts at each forecasting round.

Chart 19. ME for UND1X forecasts

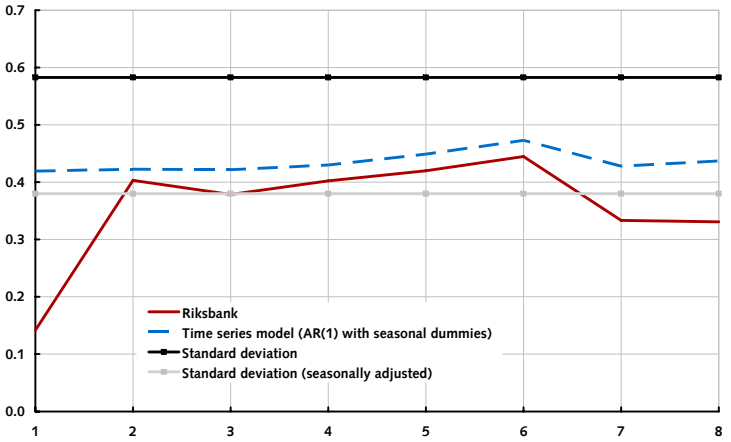
Percentage points (calculated on quarterly changes)



Note. See the note to Chart 15.

Chart 20. RMSE for UND1X forecasts

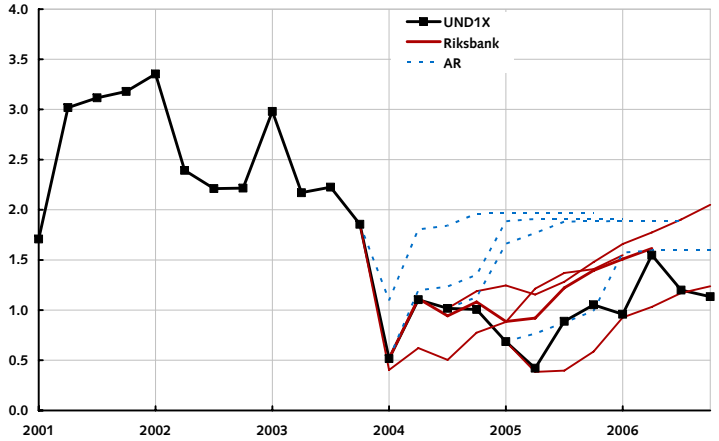
Percentage points (calculated on quarterly changes)



Note. See the note to Chart 1 The standard deviation of UND1X is also provided (as a further reference) for seasonally adjusted data.

Chart 21. UND1X, outcomes and forecasts

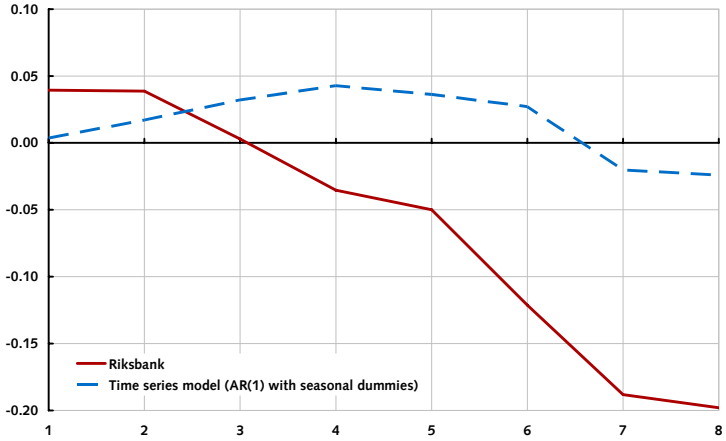
Annual percentage change



Note. See note to Chart 17.

Chart 22. ME for CPI forecasts

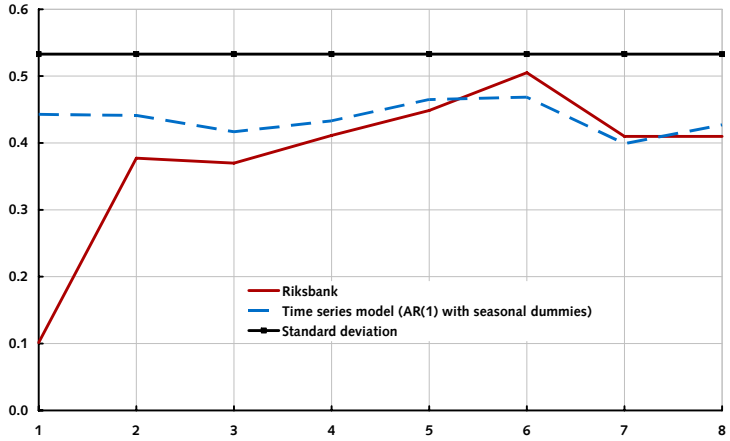
Percentage points (calculated on quarterly changes)



Note. See the note to Chart 15.

Chart 23. RMSE for CPI forecasts

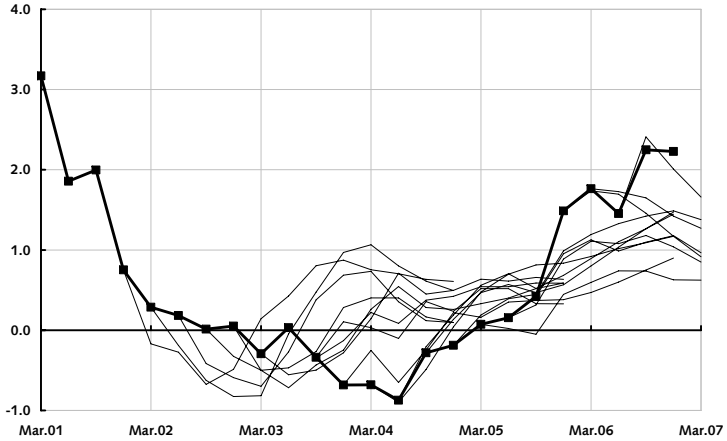
Percentage points (calculated on quarterly changes)



Note. See the note to Chart 1.

Chart 24. Employment, outcomes and forecasts

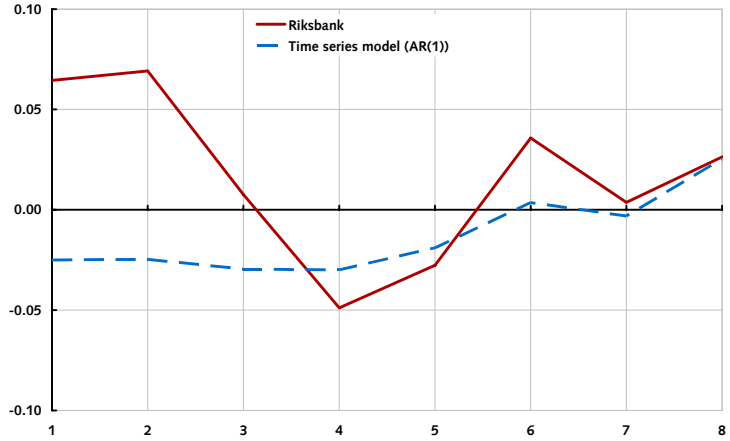
Annual percentage change



Note. The curve marked with squares is the outcome and the black curves are the Riksbank's forecasts at each forecasting round.

Chart 25. ME for forecasts of employment

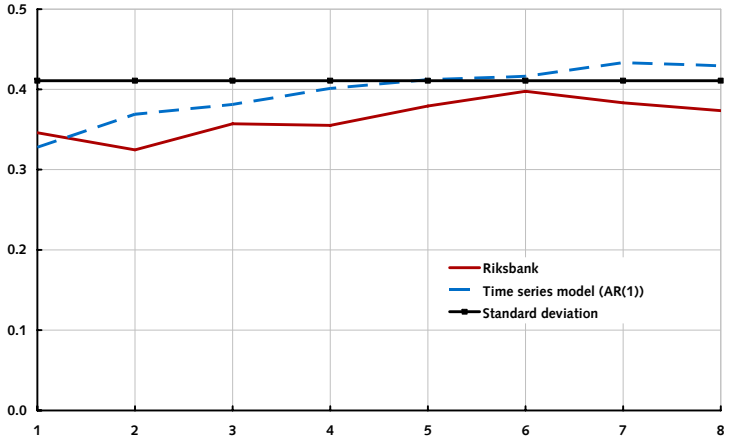
Percentage points (calculated on quarterly changes)



Note. See the note to Chart 15.

Chart 26. RMSE for forecasts of employment

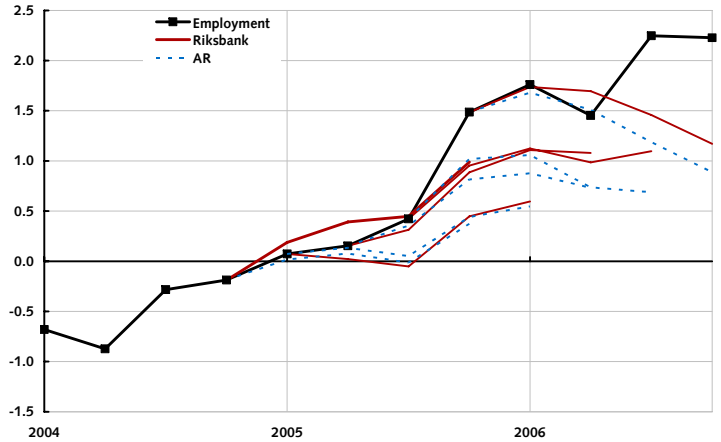
Percentage points (calculated on quarterly changes)



Note. See the note to Chart 1.

Chart 27. Employment, outcomes and forecasts

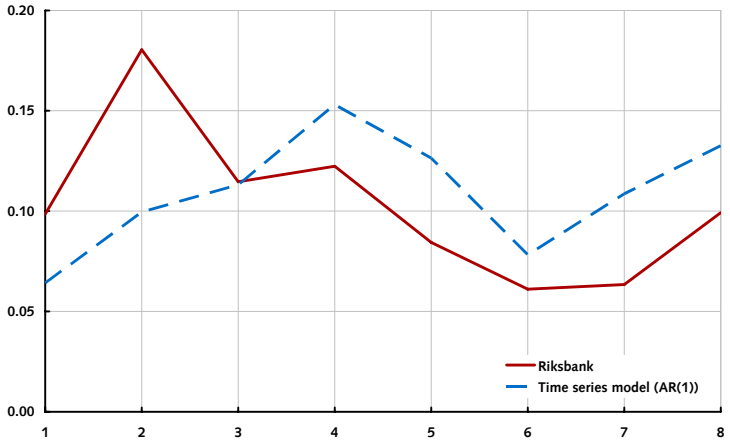
Annual percentage change



Note. See note to Chart 17.

Chart 28. ME for labour force forecasts

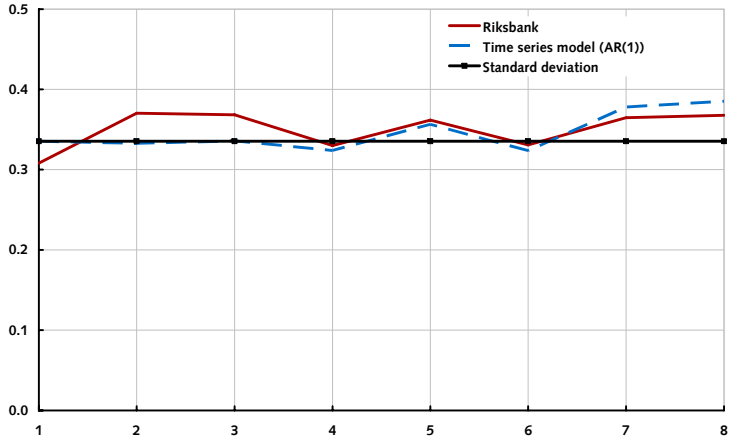
Percentage points (calculated on quarterly changes)



Note. See the note to Chart 15.

Chart 29. RMSE for labour force forecasts

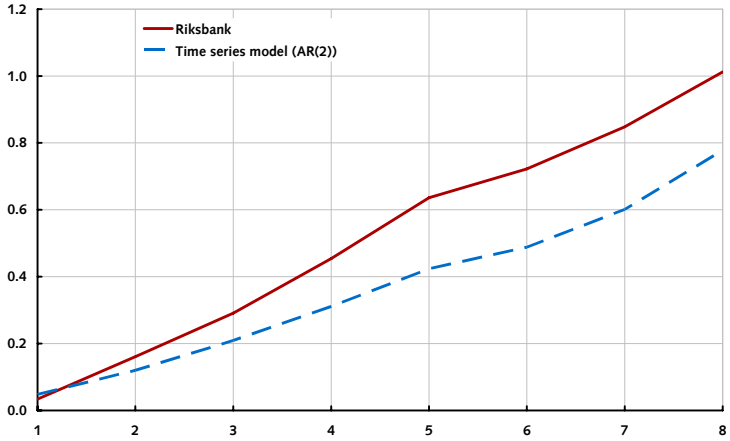
Percentage points (calculated on quarterly changes)



Note. See the note to Chart 1.

Chart 30. ME for unemployment forecasts

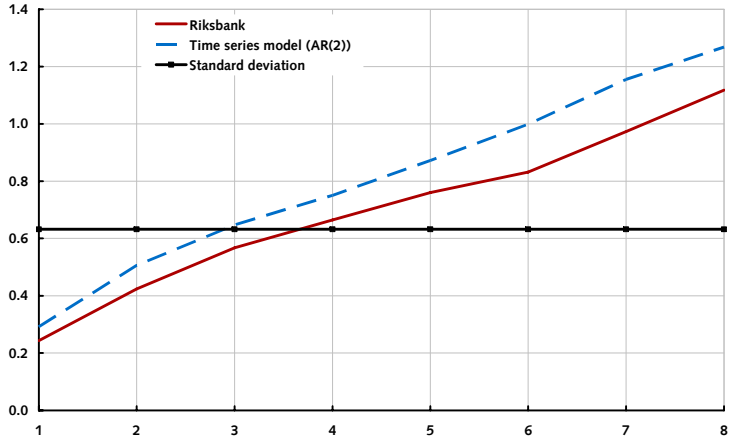
Percentage points (calculated on quarterly changes)



Note. See the note to Chart 15.

Chart 31. RMSE for unemployment forecasts

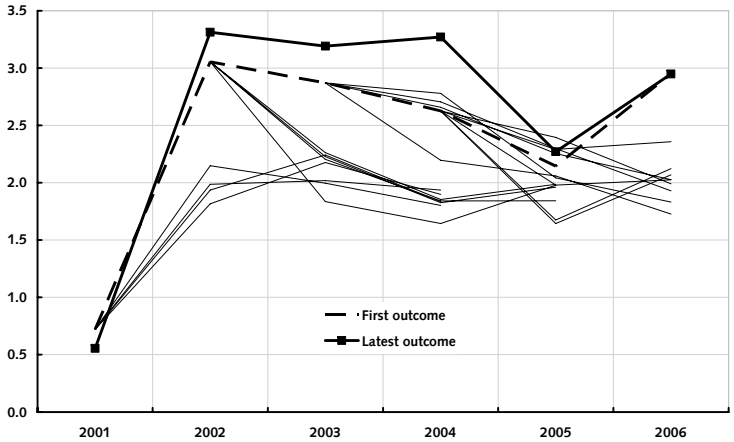
Percentage points (calculated on quarterly changes)



Note. See the note to Chart 1.

Chart 32. Annual forecasts for GDP/hours worked 2002–06

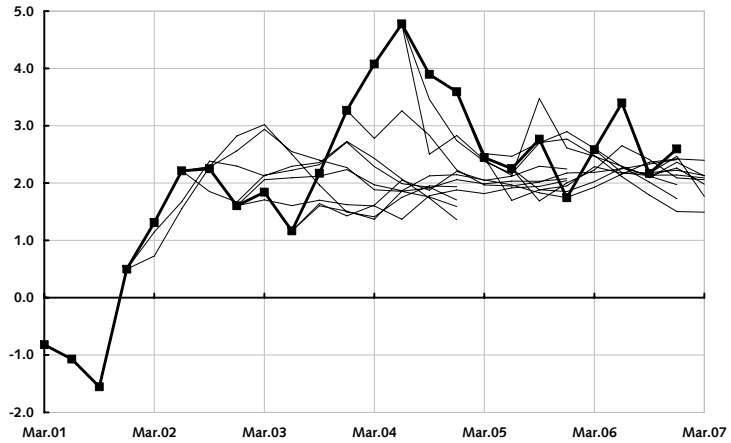
Annual percentage change



Note. Outcomes are shown as the preliminary (first) figures as dashed line and the latest outcome series as solid line with squares. The remaining curves represent the Riksbank's annual forecasts at each forecasting round.

Chart 33. Quarterly forecasts for GDP/number employed, 2002–06

Annual percentage change



Note. See the note to Chart 24.

Chart 34. Outcomes for hours worked, number employed and average working hours

Annual percentage change

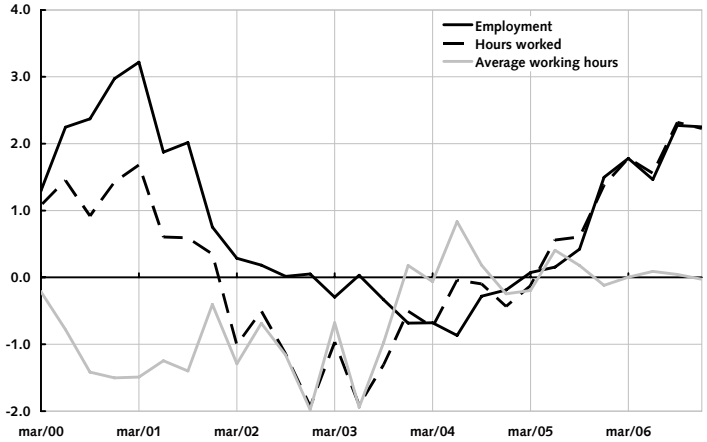
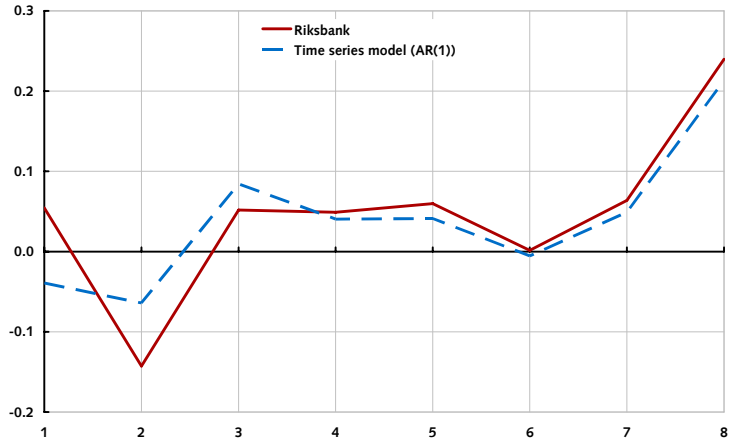


Chart 35. ME for forecasts of GDP/number employed

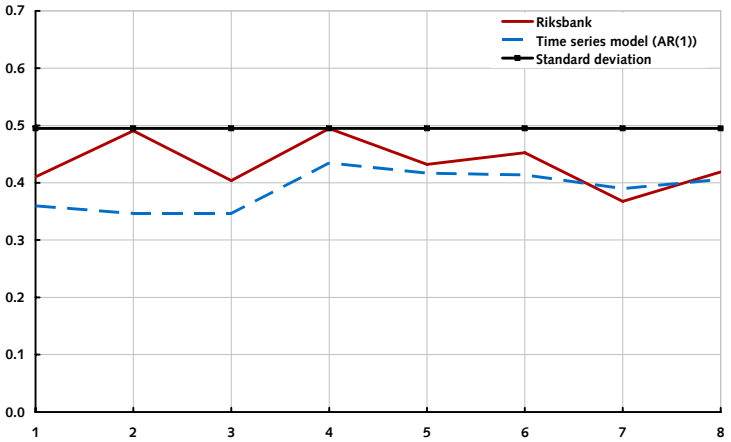
Percentage points (calculated on quarterly changes)



Note. See the note to Chart 15.

Chart 36. RMSE for forecasts of GDP/number employed

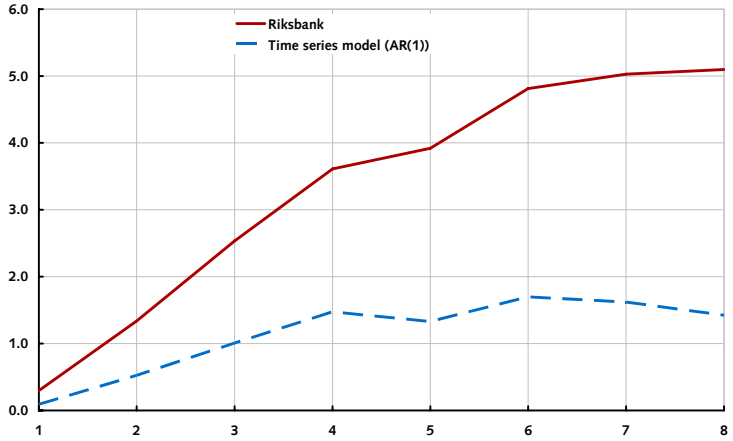
Percentage points (calculated on quarterly changes)



Note. See the note to Chart 1.

Chart 37. ME for forecasts of nominal SEK/TCW exchange rate

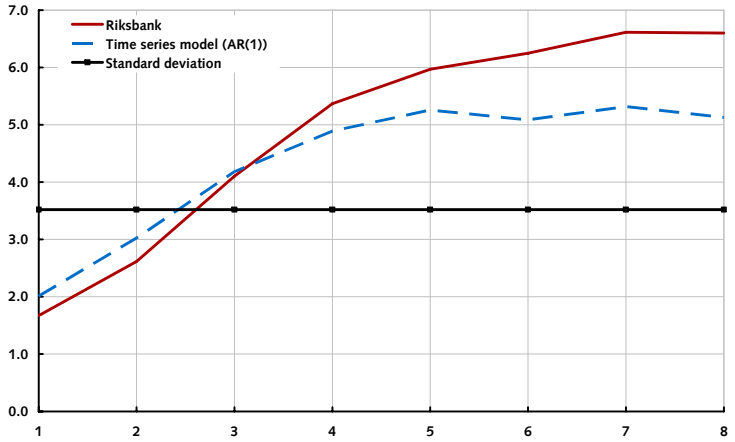
Per cent (calculated on the percentage deviation in level)



Note. See the note to Chart 15.

Chart 38. RMSE for forecasts of nominal SEK/TCW exchange rate

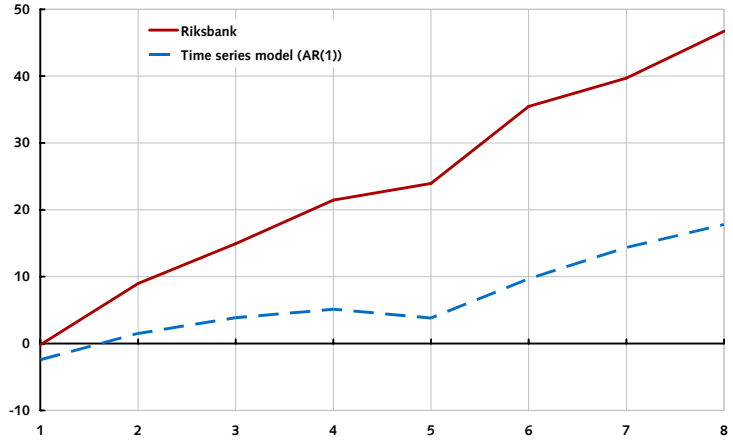
Per cent (calculated on the percentage deviation in level)



Note. See the note to Chart 1.

Chart 39. ME for oil price forecasts (Brent)

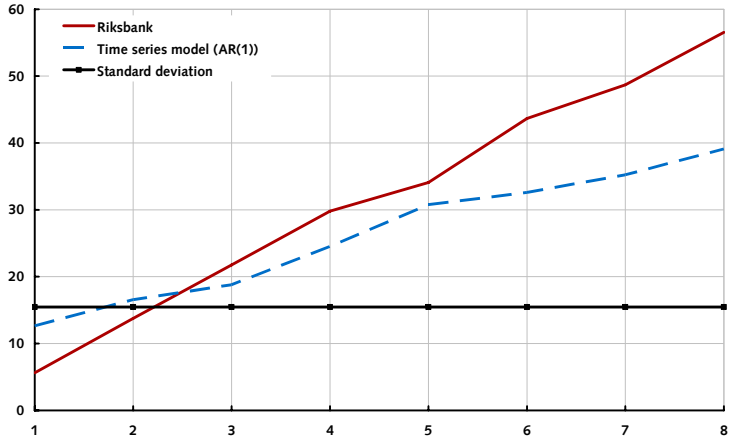
Per cent (calculated on the percentage deviation in level)



Note. See the note to Chart 15.

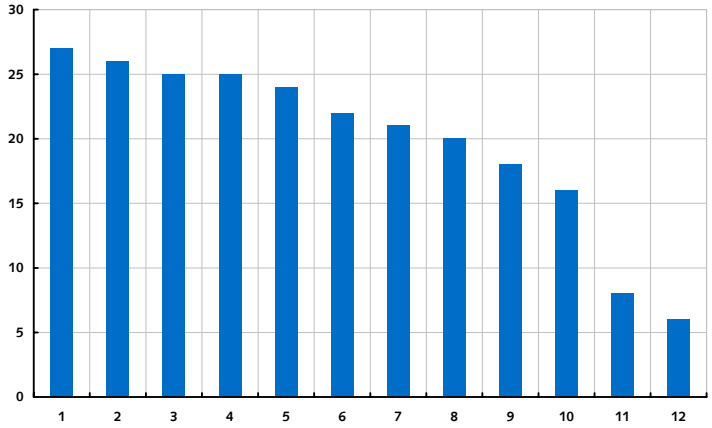
Chart 40. RMSE for oil price forecasts (Brent)

Per cent (calculated on the percentage deviation in level)



Note. See the note to Chart 1.

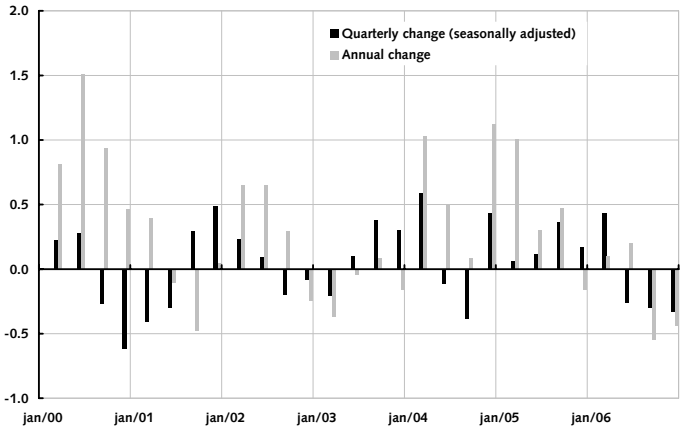
Chart 41. Number of quarterly GDP forecasts broken down by forecast horizons



Note. The forecasts are those in the Riksbank's *Inflation Report* 2000–06, i.e. the period for the forecast evaluation.

Chart 42. Differences between latest and first versions of GDP outcomes

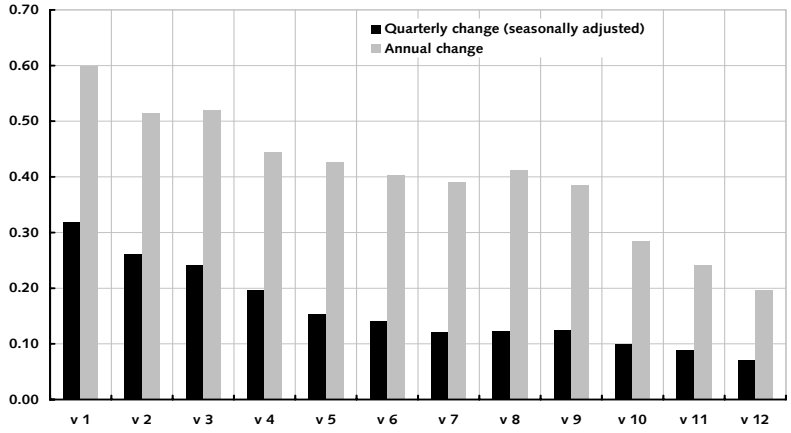
Percentage quarterly annual changes



Note. The chart illustrates the size of revisions from the time of the first preliminary outcome to the latest available figure for each quarter.

Chart 43. RMSE for different versions of GDP outcomes, 2000–06

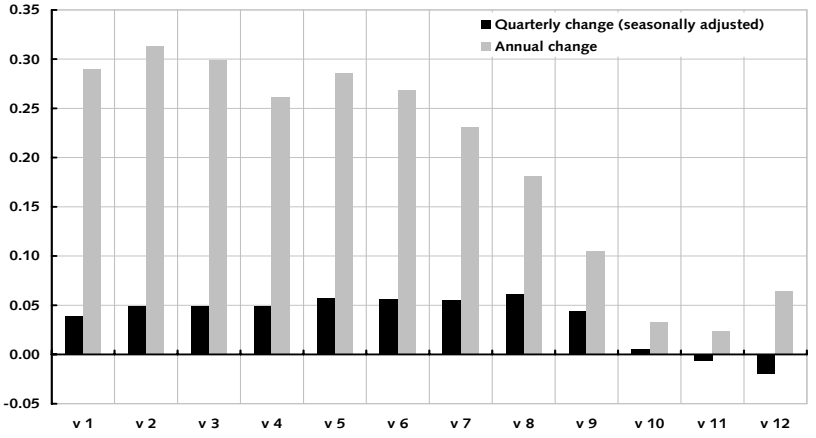
Percentage points



Note. The chart illustrates average differences between preliminary outcomes and the latest available figures and shows how this changes as the data are revised. The first preliminary version of data is denoted v1, the first revisions of this v2 and so on

Chart 44. ME for different versions of GDP outcomes, 2000–06

Per cent



Note. The chart illustrates the bias for preliminary outcomes in relation to the latest available outcomes and shows how this changes as the data are revised. The first preliminary version of data is denoted v1, the first revisions of this v2 and so on.

Earlier Working Papers:

For a complete list of Working Papers published by Sveriges Riksbank, see www.riksbank.se

Sveriges Riksbank
Visiting address: Brunkebergs torg 11
Mail address: se-103 37 Stockholm

Website: www.riksbank.se
Telephone: +46 8 787 00 00, Fax: +46 8 21 05 31
E-mail: registratorn@riksbank.se

ISSN 1402-9103 E-Print, Stockholm 2007