WORKING PAPER No. 45
March, 1995

THE ACCURACY OF THE SWEDISH NATIONAL BUDGET FORECASTS 1955-92

REINHOLD BERGSTRÖM

Svensk resumé på sid 41.

The Accuracy of the Swedish National Budget Forecasts 1955-92

by

Reinhold Bergström
Department of Statistics
Uppsala University
P.O. Box 513
S-751 20 Uppsala
Sweden

Abstract

In connection with the presentation of the budget in early January each year, forecasts are presented for a number of important macroeconomic variables, the PNB (Preliminary National Budget) forecasts. A revised set of forecasts is given about three months later, the Revised National Budget (RNB) forecasts. The accuracy and rationality of these forecasts is analyzed for the period 1955-92 using descriptive measures and formal tests based on regression models. With two exceptions the PNB forecasts of the variables considered are found to be unbiased and efficient. They outperform naive forecasts. The variability of the forecasts is smaller than that of the outcomes indicating cautiousness in the forecasting process. The two variables where the PNB forecasts are less accurate are wages and prices, where a systematic underestimation of changes is noticeable. More pronounced changes over time in the forecast accuracy are not found, although there is a tendency towards improved forecasts. The RNB forecasts are based on more information than the PNB forecasts, which is reflected in more accurate forecasts.

1. INTRODUCTION

During the greater part of the post-war period, the Ministry of Finance in cooperation with the National Institute of Economic Research (NIER) has produced short-term forecasts for the Swedish economy. These forecasts have been published as a Preliminary National Budget (PNB) and a Revised National Budget (RNB) and primarily covered the development one year ahead. These forecasts are an extremely important basis for many decisions made in Sweden. Consequently the accuracy of the forecasts is of great importance.

Although forecasts are constantly produced in great numbers, too little effort is in general spent on an evaluation of the forecasts. Much valuable insight can be gained from such a process, insights that may possibly improve future forecasts. With these considerations in mind the PNB and RNB forecasts for the period 1955-92 have been evaluated. The present work is an extension of an earlier study published in Swedish, Bergström (1988), B88[1]. Hultcrantz (1971) contains a detailed study of the forecast accuracy for the period before 1970. Kim (1988) analyzes the accuracy of certain NIER forecasts during the period 1977-86 and Carling and Kim (1989) contains further results on the NIER forecasts.

Forecasts of the following important macroeconomic variables will be analyzed: Private Consumption (CP), Total Fixed Investment (I), Exports (X), Imports (Z), Central Government Consumption (CCG), Local Government Consumption (CLG), Gross Domestic Product (GDP), Wages (W) and Consumer Prices (P). All variables, except wages and prices, are in constant prices.

2. THE SWEDISH NATIONAL BUDGETS

The National Budget concept cannot be said to be uniquely defined. In general, the term is used for annual statements about the future in cases, where those responsible for the budgets only to a limited extent can influence most of the variables that are budgeted. The purpose of a National Budget is to serve as a basis for the decision-making of various agents in the economy, including both central government and individual firms. In addition, the forecasts included in the National Budget in Sweden have been widely used by other producers of short-term forecasts.

National Budgets have been published in Sweden since the late 1940s. During the period we

[1] **I am grateful to Alek Markowski and Cheong -Han Kim at the NIER for valuable comments on an early version of B88, from which the present paper also has benefited. I am also grateful to Lars-Erik Öller at the NIER for valuable comments on a previous version of the present paper.**

are looking at, both the NIER and the Ministry of Finance (or the Ministry of Economic Affairs) have been involved in the production of the forecasts. The NIER was established already in 1937 as an independent research institute, which was to analyze the economic development in Sweden and abroad. As part of the work, different types of economic data were to be collected. After the Second World War, work in connection with the National Budgets became one of the most important tasks for the NIER.

As mentioned above, both the Ministry of Finance and the NIER have been involved in the National Budget process. Up to 1985, the Ministry was responsible for a general summary and certain chapters on individual sectors (variables), while the NIER was responsible for the chapters dealing with the remaining sectors. Cederwall (1987) contains a description of how the budget process developed in cooperation between the NIER and the Ministry. During the greater part of the period, the published accounts include a statement of exactly how the responsibility was devided. In some cases, two different (although in general not very different) forecasts were explicitly presented for some major variables.

There was some variation in exactly how the results were published, but during most of the period, publication regularly occurred in volumes entitled "The Swedish Economy" (Swedish: "Konjunkturläget") with subtitles of the type "Preliminary National Budget" and "Revised National Budget". A further report made entirely by the NIER was issued in the autumn. These autumn reports will not be analyzed here (except as sources of certain outcome data).

The Preliminary National Budget (PNB) was published early in January in connection with the presentation of the budget for the forthcoming fiscal year (starting July 1), while the Revised National Budget appeared three to four months later. The forecasts included in the PNB for a given year thus were completed at the end of the preceding year, while the forecasts in the RNB included information on the development during the first few months of the forecast year.

For the period from 1986 and onwards, the NIER and the Ministry have published separate documents, in which the respective forecasts are reported. In December each year the NIER now publishes a third economic report which primarily contains forecasts for the following year. It is stated that the document is intended as a starting-point for the work on the national budget. The Ministry of Finance is then entirely responsible for the PNB published in January as an appendix to the "Fiscal Plan", which is the main policy document of the budget proposition. The Ministry is also responsible for the RNB, which is an appendix to the "Revised Fiscal Plan" published in April. The spring report of the NIER, published in May reports (in Swedish) on studies done at the NIER and gives a very brief analysis of the state of the economy. The situation during the spring as seen by the NIER is also presented in a rather short press release issued in March and is available as a basis for the subsequent RNB. The NIER still publishes an autumn report.

In many countries forecasts similar to those in the PNB and the RNB are produced using (at least partly) formal macroeconomic models in the form of simultaneous equation systems, although the model results in general are more or less modified before yielding the final forecast. For much of the period covered by the present study this has not been the case in Sweden. The earlier overall methodology is described in Pettersson (1988) and characterized as "iterative". For some variables, forecasts were obtained using special relationships between variables. These were more or less explicitly formulated, from exactly estimated econometric relationships to simple "rules of thumb". Information gathered by surveys formed the basis for the forecasts of several variables. Fixed investment in manufacturing is such a variable, where plan data gathered from a stratified sample of Swedish enterprises were used. Business surveys based on qualitative variables ("Konjunkturbarometern") was a further source of information. Obviously the forecasts for individual variables had to be internally consistent. This consistency was reached by the iterative method described by Pettersson. Further information on the earlier forecasting process used can be found in Kragh (1964), Hultcrantz (1971) and Markowski (1979). Since 1987 an econometric macromodel (KOSMOS) has been one of the tools used. Although the model support has much increased since the start, the forecasts are still mainly judgemental. For more information on KOSMOS, see Ernsäter and Rosenberg (1989) and Markowski and Persson (1993).

3. MEASURING FORECAST ACCURACY

In general forecasts are made "given the conditions" at the time of the forecast. In economics this usually means that the forecasts assume an "unchanged economic policy"[2]. Since 1986 the main difference between the NIER and the Ministry forecasts is that the former assume an unchanged policy, while the latter add the estimated impact of future government measures. When comparing forecasts and outcomes it should therefore be borne in mind that forecast errors are due to the following fundamental causes among other things

1) The wrong forecasting model

2) Changed conditions, such as changes in economic policy.

The producer of the forecast should obviously not be blamed for differences between predicted

[2] In our case this has somewhat different meaning for different variables before 1986, as those variables that came from "the general summary" included information on the policy changes proposed at the same time, while the remaining variables were not based on such information.

and actual values that are caused by point 2) above. It might even be the case that the forecast itself caused the change in economic policy that made the forecast wrong. An analysis should correct for these types of changes. However, this is hardly possible in practice and only in the most favourable case one may be able to give some assessments. As in most other prediction/realization studies, we are therefore going to analyze total errors. The analysis will be restricted to variables in the form of growth rates, which is natural in view of how the forecasts are presented.

There are several different methods available for the analysis of forecast errors. The simplest method is to compute the mean error (ME) and/or the mean absolute error (MAE). If P denotes the predicted value and A the actual value this means that (unless otherwise stated, P and A will refer to percentage changes compared with the preceding year)

$$ME = \frac{1}{n}\Sigma(P_t - A_t) \tag{3.1}$$

and

$$MAE = \frac{1}{n}\Sigma|P_t - A_t| \tag{3.2}$$

Another often used measure is the Root-Mean-Square Error (RMSE)

$$RMSE = \sqrt{\frac{1}{n}\Sigma(P_t - A_t)^2} \tag{3.3}$$

ME shows whether systematic over- or underprediction is present, while MAE and RMSE more directly measure forecast accuracy. Large errors influence RMSE more than MAE.

It is well-known that certain variables are more erratic than others. Consequently they should be more difficult to predict. One way of checking this is to compute the standard deviation of the actual series. When using this measure account should be taken of a possible trend in the series. If a strong trend is present, a large standard deviation need not imply that the series is irregular and difficult to predict.

To obtain a certain standardization, Theil (1966) has suggested the following measure of forecast accuracy

$$U = \frac{\sqrt{\frac{1}{n}\Sigma(P_t - A_t)^2}}{\sqrt{\frac{1}{n}\Sigma A_t^2}} \tag{3.4}$$

U is 0 when the forecasts are perfect, but there is no upper limit for U. The naive forecast "no change" implies $P_t = 0$ for all t, so that U = 1. The value of U consequently can be interpreted as the accuracy relative to the error obtained when the most naive prediction method is used.

The square of the numerator in equation (3.4) can be divided into the following three components

$$MSE = (\overline{P} - \overline{A})^2 + (S_P - S_A)^2 + 2(1-r)S_P S_A \tag{3.5}$$

where $\overline{P}$ and $\overline{A}$ are the means of P and A, S_P and S_A the standard deviations of P and A (defined with n in the denominator instead of the usual n-1) and r the correlation coefficient between P and A. Theil has proposed the following additional coefficients

$$U^M = \frac{n(\overline{P} - \overline{A})^2}{\Sigma(P_t - A_t)^2} \tag{3.6}$$

$$U^S = \frac{n(S_P - S_A)^2}{\Sigma(P_t - A_t)^2} \tag{3.7}$$

$$U^C = \frac{2n(1-r)S_P \cdot S_A}{\Sigma(P_t - A_t)^2} \tag{3.8}$$

Obviously $U^M + U^S + U^C = 1$. U^M shows whether the predictions on average are equal to the actual values. U^S shows whether the variability of the predictions is the same as that of the actual values. U^C, finally, is equal to 0 if r = 1. This term measures the degree of linear relationship between predicted and actual values. Use of the three coefficients U^M, U^S and U^C means that we decompose the total error.

The Theil-coefficients described above are not always easy to interpret. An alternative is the model

$$P_t = \alpha + \beta A_t + \varepsilon_t \tag{3.9}$$

In this model both the parameters α and β and the size of the residuals are important. Perfect forecasts imply $P_t = A_t$ for all observations. i.e. $\alpha = 0$, $\beta = 1$ and all residuals equal to 0. In a prediction-outcome diagram with the actual value on the horizontal and the predicted value on the vertical axis, all observations fall on the 45°-line through the origin. In this case there are neither systematic nor random errors.

If we still lack systematic errors but random errors are present, the observations will show some dispersion around the 45°-line. If in addition systematic errors are present, the true regression line does not coincide with the 45°-line through the origin. This implies $\alpha \neq 0$ and/or $\beta \neq 1$. On the basis of historical data, it is possible to test whether $\alpha = 0$ and/or $\beta = 1$.

If systematic errors are present, a linear correction might improve the forecasts. To achieve this, it is necessary to use the reverse regression

$$A_t = \gamma + \delta P_t + \varepsilon'_t \tag{3.10}$$

Using estimated values of γ and δ, it is possible to correct the originally systematically wrong forecasts P_t and use modified forecasts $\hat{P}_t = \gamma + \delta P_t$

Equation (3.10) can also be used to test, whether the forecasts can be considered as rational expectations or more precisely that part of the rationality that consists of unbiasedness (see Holden and Peel 1985). Unbiased forecasts require that $\gamma = 0$ and $\delta = 1$. This can be tested using (3.10) either by separate t-tests of H_0: $\gamma = 0$ and H_0: $\delta = 1$ or by a simultaneous F-test of the joint hypothesis. In both cases strong assumptions have to be made about the errors.

A further model of interest is

$$P_t = \beta_0 + \beta_1 A_t + \beta_2 A_{t-1} + \varepsilon_t \tag{3.11}$$

This model can be used to test if there is any tendency for the forecasts to "lag behind" the outcomes. This would show up as a significant effect of the lagged variable in the model. Such an effect has been noted in the case of survey data, Bergström (1992).

A second property of rational expectations is that they use all available information.

Mullineaux (1978) and Holden and Peel (1985) propose that this should be tested by measuring the correlation between the forecast errors and the information available at the time of the forecast. When this information consists of earlier values of the variable itself, the property tested is termed the degree of efficiency. The number of earlier observations to be used in this test is not obvious. We have used three observation as we are dealing with yearly data. The estimated model is

$$P_t - A_t = \beta_0 + \beta_1 A_{t-1} + \beta_2 A_{t-2} + \beta_3 A_{t-3} + u_t \qquad (3.12)$$

The hypothesis to be tested is H_0: $\beta_0 = \beta_1 = \beta_2 = \beta_3 = 0$. This can be accomplished by an F(4; n-4)-test using standard assumptions for the linear model.

At the time of the forecast, A_{t-1} is not fully known which means that (3.12) need not be a completely true representation of the available information. Particularly in the analysis of quarterly data, it is common to substitute the right-hand variables of (3.12) by the latest available (preliminary) data for the variables. With yearly data, this effect has been considered small and not worth the cost of the extra collection of data. (The development during the base year is to a large extent known at the time of the forecast).

A third property of the rationality, the consistency of the forecasts, is not possible to test with the kind of data we are using.

A further property of forecasts that can be looked upon as rational expectations is that we can write $A_t = P_t + \varepsilon_t$, where P_t and ε_t are uncorrelated and the expectation of ε_t is 0, Maddala (1988, p.363). This implies that the variance of P_t should be smaller than or equal to that of A_t.

In general, it is desirable to compare the forecasts with some alternative in order to further elucidate the quality of the forecasts. As we have seen, the definition of U implies one such naive forecast principle, namely $P_t = 0$. In many cases this is "too naive" a method and we are going to use the following naive forecasts (compare Sims 1967) as comparisons

$$PNI_t = \frac{1}{I}\sum_{k=1}^{I} A_{t-k} \qquad I = 1,3 \qquad (3.13)$$

The naive forecast PN1, means that the forecast is always equal to the outcome of the previous period (note that we are still discussing variables measured as percentage changes), while PN3 uses the mean of the outcomes of the last three periods as the forecast.

To be of any value we must reasonably require that a forecast has an accuracy that is greater than that of naive methods of the type discussed above.

To test whether there is a difference between the MSEs of two sets of forecasts, it is possible to use the methodology suggested by Ashley et al (1980) and recently described by Kolb and Stekler (1993). The difference between the MSEs of two forecasting methods can be written as

$$MSE(e_1) - MSE(e_2) = V(e_1) - V(e_2) + \bar{e}_1^{\,2} - \bar{e}_2^{\,2} \qquad (3.14)$$

where e_1 and e_2 are the observed errors of the two methods (e_1 is assumed to be a naive method and e_2 a more ambitious method). Introducing the further notation $d = e_1 - e_2$ and $s = e_1 + e_2$, we can rewrite (3.14) in the following way

$$MSE(e_1) - MSE(e_2) = cov(d, s) + \bar{e}_1^{\,2} - \bar{e}_2^{\,2}$$

where cov (d,s) is the sample covariance of d and s.

If we further consider the regression model

$$d_t = \beta_1 + \beta_2(s_t - \bar{s}_t) + \varepsilon_t \qquad (3.15)$$

it can be shown that tests of the hypotheses $\beta_1 = 0$ and $\beta_2 = 0$ can be used to test the hypotheses that the MSEs of the two methods are equal.

4. THE SWEDISH ECONOMY 1955-92: A BROAD OUTLINE

As a basis for the evaluation, the period 1955-92 has been chosen. The motivation for the starting point is that a reasonable stability in the forecasting procedures should have been achieved by then. Before turning to the analysis of the forecast accuracy, the economic development should be described broadly. To show the general cyclical development (primarily in industry, however) the NIER series on composite resource utilization ("Resursutnyttjandet inom industrin") is often used. (Figure 1).

The resource utilization series is available from 1964 and in combination with other series showing the pre-1964 development it reveals a stable cyclical pattern in Sweden after the

Second World War. During the period 1955-92 cycles of a length of approximately five years can be discerned with peaks and throughs in the following quarters according to Sundberg (1992):

Peaks: 55:1, 61:2, 65:1, 70:1, 74:2, 80:1, 84:4, 89:3
Throughs: 58:3, 63:1, 68:1, 72:2, 77:4, 82:1, 87:1

A direct use of the resource utilization series in Fig.1 yields very similar results. This series also shows that the intensity at the peaks was lower during the earlier part of the 1980s than in the 1960s and 1970s, while the corresponding throughs were deeper.

GDP, as shown in Figure 2, shows a cyclical pattern that agrees well with the resource utilization series for most of the period, although the exact dating of peaks and throughs in some cases differs slightly. The GDP series further shows the trendwise reduction in growth rate of GDP, which has been present from the early 1960s. This can also be seen in Table 2. During the period 1955-64, the average growth rate of GDP was 4.2% compared with 1.4%, during the period 1975-84 and the even lower 0.9% in the last period. The decrease in the trendwise growth rate of GDP is about 1.1 percentage points per decade.

GDP is not the only variable showing a decreasing growth rate. The same is true for all the other variables except wages and prices in the sense that a model linear in time yields significantly negative estimates for the trend parameter. P shows a trendwise increasing growth rate at least until the early 1980s, while W does not show any marked changes except for low values during the 1950s and very high values 1974-76 .

It is possible to discern a cyclical pattern for several of the other variables, too, although it is less pronounced than for GDP. The growth rate of X peaks in the years 56, 60, 64, 69, 73, 78/79 and 83, which in general is about a year earlier than GDP. There is no pronounced peak during the late 1980s. Thus the export cycle has tended to lead the general cycle, a fact which can be utilized by a forecaster.

The development during the later part of the 1980s does not completely follow the earlier very regular pattern. Following the peak in 1984/85, there was hardly any downturn at all before yet another high peak in 1988 according to the resource utilization series. The level reached then was about the same as that at the peaks of 1965 and 1969/70. The following recession has been unusually severe and the value reached in 1992 is the lowest yet recorded.

There are a few further features of the development after 1985, which should be mentioned.

Although the resource utilization series showed exceptionally high values around 1988 and the economy was considered to be severely "overheated", the growth rates of several of the major macroeconomic variables were not exceptionally high. Thus GDP peaked at growth rates of 2.3-2.4 %, which can be compared with 1960s peaks of more than 5 %, but at similar resource utilization levels. The growth rate in private consumption reached high levels during 1985-1988 following the deregulation of the credit market. The growth rate of total investment reached both a new high of +11 % in 1989 and a new low of -11 % in 1992. The inflation rate was greatly influenced by the tax reform in 1990-1991, otherwise it was comparatively low during the period after 1985.

We mentioned earlier that when judging the accuracy of the forecasts for different variables (and periods), it is important to take into account the difficulty of forecasting in each case. One way of getting some information on this is a computation of standard deviations. As shown by Tables 2 and 3, the variability in the growth rates of the foreign trade variables is large. The same is true for investment, where in addition, the variability is around a lower average level. Private consumption and GDP show the smallest variability. More pronounced changes in variability over time have occured for private investment (an increase towards the end) and public consumption (the largest variability in the period 1965-74).

5. FORECASTS AND OUTCOMES

5.1 The choice of outcome values

It is important that outcome figures refer to variables that are defined in the same way as those which were forecast. This means that the outcome values should not be gathered from sources that are too far away in time from the forecast time-point. Otherwise there is a considerable risk that data revisions make the comparability between forecasts and outcomes doubtful. At the same time, it is necessary to wait long enough for the figures to be reasonably close to final.

Outcome figures are available already in the PNB of the next year, but these are obviously preliminary as part of the year still remains when the data in the PNB are finalized. More definite figures are available in the RNB and the autumn report of the NIER later in the year. Using the RNB as the source for outcome data has the advantage that data, that are directly comparable with the forecasts can be found, while this is not fully possible starting from the autumn report.

In addition to later editions of National Budgets, the official national accounts statistics that are regularly published offer further alternative data sources. Here the difficulties with data revisions and finding data for exactly those variables which were forecast cannot be avoided.

In order to study the effect of the choice of outcome figures some alternatives were investigated in B88 for the period 1970-85. Outcome figures from the RNB of the following year, the NIER autumn report of the following year and the NIER autumn report yet another year later were compared. A general conclusion was that these three sources produced very similar outcome figures.

In view of these results the NIER autumn report from the following year was chosen as the source for outcome data (compare also Markowski 1979) for the period 1970-85, while for the period 1986-92 the NIER December report of the following year was employed. For the period 1955-69, Hultcrantz (1971) was used as source for both forecast and outcome data. His general principle for selection of outcome data is similar to the one described above, although the actual implementation is not quite so consistent, as it is for the period 1970-92. In Appendix A, certain comments on the choice of data are made.

5.2 The NIER forecasts compared with those of the Ministry of Finance.

As mentioned earlier the NIER and the Ministry of Finance have produced separate forecasts for the period 1986-1992. An interesting question is whether there is any difference between the two forecasts. Some results are shown in Table 1. On average, the forecasts are almost identical. A possible exception is wages and prices, where the NIER forecasts on average are half a percentage point higher. As regards forecast accuracy, the Ministry achieves a slightly smaller absolute error for several of the variables, prices and wages again being exceptions. As the differences in general are small and the available evidence is limited, it is not possible to say that there is a systematic difference in forecast accuracy. A possible exception is private consumption, where the PNB forecasts are slightly better for every single year of the period. To find such a difference is not unexpected as the PNB forecasts are produced about a month later and also include changes in economic policy introduced in connection with the budget. Thus in the following analysis of the complete period 1955-1992, it would not greatly matter which of the two sets of forecasts that we used. We have chosen the PNB forecasts as this seems to be most consistent with the forecasts for the earlier time period.

5.3 The PNB forecasts

Table 3 and Figure 2 summarize results for the whole period. The average PNB forecast does not differ much from the average outcome for most of the variables. Exceptions from this pervasive pattern are W and P, where the underestimation is considerable.[3] For X and Z there is a certain underestimation. The table also clearly shows that the forecaster tends to be cautious, as the standard deviations of the forecasts are smaller than those of the actual values. Thus there is a distinct reluctance to produce forecasts that deviate from the "normal". This is a property that we expect from rational forecasts, but a large difference between the variances of outcomes and forecasts do of course imply large forecast errors. The average absolute forecast error is largest for the foreign trade variables but considerable for investment, too. The accuracy is greater for CLG than for CCG. For the important variables CP and GDP, the MAE is 1.3 and 1.2 percentage points, respectively.

It is of special interest to look at the forecast accuracy during the last few years (Table 4). For GDP, the MAE achieved is lower for the period 1985-1992 than for any of the earlier periods. The forecasts were almost perfect for the years 1986-1990. The recession in 1991 and 1992 was anticipated, but not the extent of the recession, forecasts of 0.2 % contrasting with outcomes of about -1.7 %. The overall accuracy of CP is the same as that observed earlier. The consumption boom of the middle 1980s was only anticipated with some delay and the great decrease in 1992 was unexpected. The error in the investment forecast is greater than that in any earlier time period, which is not surprising as the variability in investment was very large during the period. Both the great increase in 1987-89 and the following decrease in 1991-1992 were unexpected. In the latter case there was a reaction in the correct direction with some delay. The foreign trade forecasts are very good, which partly is due to the unusually smooth behaviour of these variables during the period. Wages were underestimated as usual except for the final year, when they finally came down to the low level expected. The price forecasts showed a marked improvement compared with those of the previous ten years. Both the increase in the inflation rate in 1990-1991 associated with the tax-reform and the subsequent dramatic fall were very well anticipated.

Table 4 makes possible a study of changes in the forecast accuracy over time. Using the MAE as the criterion, the table does not give an impression of more marked changes in accuracy. A formal test could be based on the model

[3] As mentioned in the Appendix, alternative forecasts have in some cases been given for prices and wages. The main results are based on what is called Alternative I (or some corresponding expression). If Alternative II is used, when available, the mean error of W is slightly reduced.

$$AE = \alpha + \beta t + \varepsilon \quad (5.1)$$

A tendency towards an improved forecast accuracy should imply a negative value of β. No clear pattern emerges from the estimated values of β (detailed numerical results not shown). Both positive and negative values are obtained. Four variables have significant β estimates, that of I being positive and those of X, CCG and CLG negative. The analysis based on (5.1) does not correct for the "difficulty of forecasting" in different periods. In connection with the discussion of naive forecasts later, such a test will be presented. This should be more reliable than the present one.

In contrast to the MAE, the average forecast error, ME, shows systematic variations over time. Exceptions from this pattern are W and P, where the underestimation is stable and persistent. In several subperiods ME and MAE coincide for these variables implying underestimation every year during the subperiod. During the latter part of the 1970s, i.e. between the first and second oil crises, the forecasts for variables such as GDP, CP, I, X and Z were often rather high. The slowing-down in the growth rate which took place during this period was not reflected in the forecasts (at least not completely) and this lack of adaptation persisted for a number of years. During the 1980s, on the other hand, when growth rates have been low, too, the average forecast accuracy, ME, has improved implying that such an adaptation has finally taken place.

Table 5 shows that the RMSE provides the same picture as the MAE. As usual the RMSE values are slightly higher. The Theil U-coefficient is smaller than one in all cases, which shows that the PNB-forecasts are considerably better than the naive forecast "no change". The U-values are smallest for W and P, which , however, rather shows one of the deficiencies of the U measure. This measure has the property that, given a certain RMSE or MAE value, the larger the absolute values of the variable considered the lower is the U value. On the other hand, large absolute values often imply greater forecast difficulties, but need not always do so, e.g. when the variability of the variable is small or regular.

The U measure should be interpreted with care. The decomposition of U into U^M, U^S and U^C on the other hand, produces easily interpretable information. U^M is close to 0 for all variables except W and P, confirming that on average forecasts equal outcomes. In several cases U^S is of the order 0.2-0.3, which reflects underestimation of the variability. For some variables, in particular CP, the forecast error is almost completely due to a lack of correlation between forecast and outcome (U^C is close to 1).

Estimation of model (3.9) by OLS produces the results shown in Table 6. The pattern is very consistent. The constant term is always positive and the slope is smaller than one. For all variables except P the constant term differs significantly from 0, while the slope is significantly different from 0 throughout but also significantly different from 1. This means that in years with low growth rates the forecast is too high and vice versa.

A further inspection of the results in Table 6 shows that for model (3.9) there is positive autocorrelation in the residuals of all estimated equations except those for GDP and P (the lower critical value for the Durbin- Watson test is 1.43 at the 5 %-level). To obtain more reliable estimates, model (3.9) therefore has been estimated by the maximum likelihood method (ML) under the assumption that the residuals follow an AR(1) process, Beach and MacKinnon (1978). This eliminates the serial correlation in all cases. In general the differences compared with the OLS estimates are small and the conclusions are unchanged . Where notable differences occur (CP, CCG, CLG and W), these all go in the same direction. The constant term is even larger than with OLS and the slope closer to 0.

Model (3.10), where the forecast is the explanatory variable, can be used to test if the forecasting process leads to forecasts that are unbiased. In Table 6 results of the simultaneous test H_0: $\gamma = 0$, $\delta = 1$ are given in addition to results of the simple tests H_0: $\gamma = 0$, H_0: $\delta = 1$. With the exception of W and P all variables have values of $\hat{\gamma}$ and $\hat{\delta}$ that are close to 0 and 1, respectively. In none of these cases is there a significant difference compared with the value postulated under H_0, but $\hat{\delta}$ always differs significantly from 0. For W and P the simultaneous test leads to a rejection of the null hypothesis implying biased forecasts. On average there is an underestimation of the inflation rate by 1.4 percentage points in the sample, while the forecast error for W varies with the forecast value.

The result that the forecasts for most of the variables are unbiased differs from that obtained in several other studies, notably that of McNees (1978) who for three US econometric models found many examples of biased forecasts.

For model (3.10), there is only one variable, I, with positive serial correlation in the residuals. Re-estimation of this model by ML assuming AR(1)-residuals hardly changes the estimates at all. The estimated constant term is 0.03 (standard error 0.91) and the estimated slope 0.94 (0.19).

Results of an estimation of the distributed lag model (3.11) are shown in Table 7. There is considerable autocorrelation in the residuals when OLS is used. Therefore we present ML-AR1 results. In most cases the difference between the parameter estimates of the two methods are not large. The parameter estimate of the lagged outcome variable is positive for all the nine variables considered and it is significant for all but two of the variables (CP and I). For two of the variables, GDP and W, the parameter estimate of the lagged variable is numerically larger than that of the current form of the variable. In conclusion we can say that the forecasters tend to show a delayed reaction with regard to changes and seem to be influenced by reality they are just observing when making their forecasts.

Is it possible to improve the forecasts using knowledge of the systematic errors? A test of this for the period 1986-92 using the results for the period 1955-85 reported in B88 is based on the following correction equations

$$P^K = 1.58 + 1.01\ P_{PNB}$$
$$W^K = 4.97 + 0.62\ W_{PNB}$$

For 1986-92 this gives values of the MAE that are 1.48 for P and 1.35 for W, which can be compared with the uncorrected values of 1.39 and 1.96, a considerable improvement in the latter case.

Table 8 shows results of an efficiency test. The low quality of the forecasts for W and P are again in evidence, while in all other cases the hypothesis of efficiency is not rejected.
The wage and price forecasts merit some further discussion. Obviously they are systematically too low for much of the period. It is natural to ask whether there is a specific reason for this. The fact that alternative wage forecasts are made in some cases is one indication of special problems. Another is that the word "forecast" seems to be avoided in the case of wages, which could mean that the forecaster uses a value that differs from what is actually believed to be the most realistic one. The reason for this could be an effort to avoid "pushing wages upwards" by showing a high wage figure, as for much of the period too high wage increases have been a constant feature of the Swedish economy. A quote from Cederwall (1987) gives a hint to the difficulties.

"The treatment of wages at times when wage contracts had not been reached on the labour market constituted a constant problem. The solutions varied.... The question was always considered sensitive. The labour market organizations did not want guidelines from the minister of finance.... The problem must still be considered as unsolved".

Pettersson (1988) gives further information regarding the wage forecasts. Obviously the wage forecasts influence the price forecasts to some extent. However, errors in wages explain only a minor part of the errors in prices. This means that there are remaining errors in the price forecasts, and a more systematic analysis of the causes of these should be worthwhile.

We know that in level terms there is an identity connecting GDP and a number of the other variables that we are considering. This means that the forecast errors of GDP should be correlated with the forecast errors of certain other variables. There are relationships between a number of the other variables, too, that should produce tendencies towards similar forecast errors.

Table 9 shows the correlation matrix of the forecast errors of PNB for the complete period 1955-92. For several of the variables the forecast errors are highly correlated with those of GDP. The highest figure is obtained for X, but CP and I also show correlations that are higher than 0.6. The forecast errors of public consumption (CCG and CLG) on the other hand are almost uncorrelated with those of GDP. As expected, the correlation between the errors of Z and CP are high and the same is true of Z and X, which means that the errors in the foreign trade variables at least partly cancel when the trade balance is computed.

The errors in the wage forecasts are weakly correlated to those of most other variables, including prices. This conclusion holds even if subperiods are analyzed (not shown). Thus the greater part of the forecast errors in prices remains even after correction for the forecast errors in wages. The forecast errors in prices are negatively correlated with those of many other variables, although the correlation in most cases is weak. However, the correlation between the errors in prices and GDP is as strong as -0.46.

5.4 A comparison with naive forecasts

Table 10 shows that among the naive forecasts, PN3 is in general preferable, although the difference compared with PN1 and PN2 in many cases is rather moderate. With W as the only exception, we find that the PNB forecasts are more accurate than naive forecasts. For GDP the MAE is 1.18 for PNB compared with 1.73 for PN3. A simple paired t-test of the hypothesis that the MAE of the PNB forecast equals that of a naive method can be carried out, but the results should be treated with some caution as the construction of PN3 implies that a certain dependence between observations close in time exists[4]. The negative differences are significant for all variables except CP, X and P (Table 11). Thus on average the PNB forecasts are systematically better than naive forecasts.

---

[4] PN3 is a moving average, which means that succesive terms are dependent. However, this need not imply that the forecast errors of PN3 are dependent or what is important for the test, that the difference between the absolute forecast errors of PNB and PN3 is dependent. The D-W values in Table 11 indicate that in general positive serial correlation should not be a problem.

It is also possible to test if the PNB forecasts improve over time after correction for "the difficulty of forecasting", when this difficulty is measured by the precision of the naive forecasts. The model is

$$y = \alpha + \beta t + \varepsilon \qquad (5.2)$$

with $y = AE_{PNB} - AE_{PN3}$. We use the difference between AE_{PNB} and AE_{PN3} instead of the ratio (as proposed by Smyth 1983) as this reduces the problems caused by small values of AE_{PN3}, even if these problems can be reduced if rank correlation measures such as Kendall's τ are used.

The estimated β parameter is negative for all variables except CLG and P, which indicates a tendency towards improved forecasts. The only variable where significance is reached according to the OLS estimates is I. However, in this case there is positive autocorrelation in the residuals and ML estimation assuming an AR(1) process in the residuals leads to an insignificant β-estimate of -0.0835 (standard error: 0.0546).

The results of a more formal comparison of the MSEs of the naive method PN3 and PNB are shown in Table 12. To facilitate the interpretation, the table also includes further descriptive results. A decomposition of the MSE into bias and variance reveals that the latter term almost completely dominates in all cases except for the PNB estimates of W and P. With one exception the bias of PN3 is positive, which is due to the negative trend observed in most of the variables.

OLS estimation of the model (3.15) in many cases leads to serially correlated errors. In view of this ML-AR1 estimates are given in Table 12. We notice that the estimated parameter β_1 is close to the difference between the biases (based on OLS it is exactly equal to this difference). Except for W and P, it is not significantly different from 0. The estimates of β_2 are always positive and also significant for all variables except W and P (and marginally CP). The overall conclusion to be drawn from this is that with the exception of W and P, the PNB forecasts have significantly smaller MSEs than the naive PN3 forecasts.

In many cases the interpretation of the results from the present type of analysis is far from straightforward as will be discussed elsewhere. Some of the complications can be seen in the results for W and P. Here the estimates of both β_1 and β_2 are significantly positive, which might be interpreted as smaller MSEs in the case of PNB. The result for the parameter β_1 does in fact mean that the bias of the naive method PN3 is significantly larger than the bias of PNB. However, in the present context this cannot be interpreted as a "strengthening" of the evidence in favour of PNB obtained from the positive parameter estimate of β_2. The reason is that the PNB biases are negative. Thus in the case of W and P the significant differences in the biases actually imply a worse performance of PNB with regard to this criterion. Thus in the overall comparison of the MSE, we have to add a smaller variance of PNB to the larger bias of this method. For both variables the differences more or less cancel out.

5.5 The RNB forecasts

Table 3 shows that the average forecast error of RNB is markedly lower than that of PNB for the variables P and W. There has been a reduction in the large underestimation of these variables. For all other variables except CLG the accuracy has also improved. A general pattern is that, the greater the PNB error, the greater is the improvement. Thus whereas CP and GDP show only marginal error reductions, I, X and Z show marked reductions. For many variables the standard deviation has increased compared with PNB, but it is still lower than the standard deviation of the outcomes.

Table 13 gives results of the regression-based analysis of RNB. The simple correlation between forecast and outcome has increased compared with PNB for all variables except CLG. The basic pattern obtained is the same as for PNB. With the forecasts as the dependent variable all constant terms are positive and all slopes are considerably smaller than one. The residuals are positively autocorrelated for three of the variables. On the basis of the more reliable ML-AR1 estimates the constant terms are always significantly positive, while the slopes are significantly larger than zero and also significantly smaller than one. Application of OLS to the reverse model (3.10) produces residuals that are not significantly positively correlated except for W. The RNB forecasts are unbiased for all variables except W and P.

6. THE FORECAST ACCURACY COMPARED WITH THAT IN SOME OTHER STUDIES

There has been considerable discussion in the forecasting literature on the merits of econometric methods compared with simpler methods. As mentioned in Section 2, the National Budget methodology can only partly be characterized as "econometric". Whatever the exact term used to describe the PNB/RNB forecasting procedure, the present study offers one further piece of evidence on the forecasting accuracy obtained by a more ambitious method compared with naive methods. We have found that the National Budget forecasts in general are superior to naive methods. Even if formal statistical significance is not always reached, the tendency in favour of the PNB forecasts is very clear.

Ridker (1963) and Sims (1967) analyzed the Dutch and Norwegian National Budget forecasts finding that the Dutch forecasts were better than naive forecasts, while the Norwegian forecasts were of a quality comparable to that of naive forecasts. The rather short and not very up-to-date observation periods, however, make these results of limited interest. Holden and Peel (1983) analyze data for the UK 1969-80 as regards the two variables inflation rate and growth rate of GDP. Their conclusion is that econometrically based forecasts are better than naive "no change" and "same change"-forecasts. Armstrong (1978) in a review of 16 studies finds that econometrically based forecasts are better than various naive methods in 6 cases, worse in 7 cases, while there is no difference in 3. The conclusion reached by Armstrong is that econometric forecasts have not proved themselves superior.

A comparison of the accuracy reached in Sweden with that in other countries is only possible to a limited extent as comparable studies do not exist. The study that is most closely comparable with the present one it that by Smith (1983) analyzing the yearly OECD forecasts of GDP, inflation rate and current foreign balance. The December forecasts for seven countries are analyzed during the rather short period 1968-79. As regards information available, these forecasts should be comparable to the PNB forecasts. On the other hand it seems reasonable to assume that a forecaster within a certain country should have an advantage as to detailed knowledge of the economic condition compared with the OECD, although OECD starts from national forecasts.

In a comparison, it is necessary to adjust for the forecast difficulty, i.e. how smooth the development thas been. While the average GDP growth rate varies from 7.1% (Japan) to 2.4% (the UK), the variability as measured by the standard deviation of the growth rate is more homogenous. With the exception of Japan (4.3 percentage points) all countries have values between 1.8 and 3.0. For the period 1968-79 the Swedish figures are 2.5% (only fractionally higher than the UK) and 2.2 percentage points. For this period the MAE of the PNB forecasts is 1.6, which is a greater accuracy than for all countries except the US and France.

The average inflation rate varies between 4.8% (Germany) and 12.2% (the UK) compared with Sweden's 7.7%. The same countries are extremes as regards standard deviation; 1.9 and 6.6 percentage points with Sweden on 3.2. The Swedish MAE is 2.3, which is worse than the OECD accuracy for four countries. Two of the countries where the OECD forecasts are bettered are Italy and the UK, countries with high and variable inflation rate where forecasting consequently should be difficult.

8. CONCLUSIONS

In connection with the presentation of the budget in early January each year, forecasts are presented for a number of important macroeconomic variables, the PNB (Preliminary National Budget) forecasts. A revised set of forecasts is given about three months later, the Revised National Budget (RNB) forecasts. The accuracy and rationality of these forecasts is analyzed for the period 1955-92 using descriptive measures and formal tests based on regression models. With two exceptions the PNB-forecasts of the variables considered are found to be unbiased and efficient. They outperform naive forecasts. The variability of the forecasts is smaller than that of the outcomes, indicating cautiousness in the forecasting process. The two variables where the PNB-forecasts are less accurate are wages and prices, where a systematic underestimation of changes is noticeable. More pronounced changes over time in the forecast accuracy are not found, although there is a tendency towards improved forecasts. The RNB forecasts are based on more information than the PNB forecasts, which is reflected in more accurate forecasts.

REFERENCES

Armstrong, J.S. (1978). Forecasting with Econometric Methods: Folklore versus Fact. Journal of Business, 51, pp 549-564.

Ashley, R., Granger, C.W.J. and Schmatensee, R. (1980). Advertising and aggregate consumption: An analysis of causality. Econometrica, 48, pp 1149-1167.

Beach, C.M. and MacKinnon, J.G. (1978). A Maximum Likelihood Procedure for Regression with Autocorrelated Errors. Econometrica, 46, pp 51-58.

Bergström, R. (1988). The Forecast Accuracy of the Preliminary (PNB) and Revised National Budgets (RNB) 1955-85 (in Swedish). Research Report 88-4, Dept of Statistics, Univ of Uppsala , Sweden.

Bergström, R. (1992). The Relationship Between Manufacturing Production and Different Business Survey series in Sweden. NIER Working Paper 12.

Carling, A. and Kim Cheong-Han (1989). Forecast and Outcome II. Special Studies, NIER, May 1989, pp 61-78.

Cederwall, G. (1987). The National Budget and the Long Range Report (in Swedish) in Henriksson, R. (ed) The National Institute of Economic Research during Erik Lundberg's Time. The National Institute of Economic Research, Stockholm.

Ernsäter, L. and Rosenberg, I. (1989). Kosmos - A Macroeconomic Model for Sweden (in Swedish). Special Studies, NIER, May 1989, pp 99-133.

Holden, K. och D.A. Peel (1983). Forecasts and Expectations: Some Evicence for the UK. Journal of Forecasting, 2, pp 51-58.

Holden, K. och D.A. Peel (1985). An Evaluation of quarterly National Institute Forecasts. Journal of Forecasting, 4, pp 227-234.

Hultcrantz, G. (1971). Forecasts and Economic "Reality" (in Swedish) in Lundberg, E. (red) Fiscal Policy in Theory and Practice in Sweden. Aldus, Stockholm.

Kim Cheong-Han (1988). Forecast and outcome (in Swedish). Special studies, NIER. May 1988, pp 27-47.

Kolb, R.A. and Stekler, H.O. (1993). Are Economic Forecasts significantly better than naive predictions? An appropriate test. Int J of Forecasting, 9, pp 117-120.

Kragh, B. (1964). Assessing the state of the Economy (in Swedish). Prisma, Stockholm.

Maddala, G.S. (1988). Introduction to Econometrics. MacMillan, New York.

Markowski, A. (1979). An Unformalised Forecasting Model. Occasional Paper 11 National Institute of Economic Research, Stockholm.

Markowski, A and T. Persson (1993). Capital Rent Cost and the Adjustment for the Effects of the Investment Fund system in the Econometric Model KOSMOS. NIER Working Paper 33.

McNees, S.K. (1978). The "rationality" of economic forecasts. American Economic Review, 68, pp 301-305.

Mullineaux, D.J. (1978). On testing for rationality: another look at the Livingston price expectations data. Journal of Political Economy, 86, pp 329-336.

Pettersson, B. (1988). The Short-Range Forecasts of the NIER (in Swedish). Special Studies, NIER, May 1988, pp 15-26.

Ridker, R.G. (1963). An evaluation of the forecasting ability of the Norwegian national budgeting system. Review of Economics and Statistics, 45, pp 23-33.

Sims, C.A. (1967). Evaluating Short-term Macroeconomic Forecasts: The Dutch Performance. Review of Economics and Statistics, 49, pp 225-236.

Smyth, D.J. (1981). Short-Run macroeconomic forecasting: The OECD performance. Wayne State Economic Papers, 135.

Smyth, D.J. (1983). Short-Run Macroeconomic Forecasting: The OECD performance. Journal of Forecasting, 2, pp 37-49.

Sundberg, O. (1992). Economic Cycles in Sweden 1946-1991 (in Swedish). Special Studies, NIER, May 1992, pp 53-78.

Theil, H. (1966). Applied Economic Forecasting. North-Holland, Amsterdam.

APPENDIX A

Comments regarding data

Data for the period up to 1969 are taken from Hultcrantz (1971). Outcome figures have been obtained from the autumn report of the NIER of the following year for the period 1970-85 and from the NIER December report of the following year for the period 1986-92.[5] The collection of the data is not completely straightforward. In many cases information is available on several different variables, the definitions of which differ very little. Information can be collected both from tables and text and from chapters dealing with certain sectors (variables) and a summarizing chapter. The figures thus obtained are not always quite identical. As could be expected over such a long time, the variables included in the published reports in some cases have changed. However, consistency over time has been aimed at, including close agreement with the variables that Hultcrantz worked with.

Before 1986, information on Private Consumption, Imports, Exports and GDP are taken from the summarizing chapter, which means that the foreign trade variables include services 1981-85. Wages and prices come from the chapter "De enskilda konsumenternas ekonomi" and for wages include changes in hours worked. Price changes are measured as changes December-December. The investment figures come from the chapter "Investeringarna" and include investment in housing. Public consumption figures, finally, are taken from the chapter "Den offentliga verksamheten".

For the period 1986-92 data on all variables except wages and prices are taken from the summarizing chapter. Information on wages and prices comes from the section "Löner och konsumentpriser".

In a few cases alternative forecasts have been given. The primary cause of this in general is the wage-variable where different assumptions in some cases have been made. As a consequence of this, different alternatives for prices are also shown. The effect on the other variables considered here is small. We have used what is called Alternative I (or some corresponding expression) except for PNB77 where the middle of three alternatives is chosen. Alternative I in general is the lower of the two alternatives given, which means that the forecast accuracy with this alternative is less good than with Alternative II. Under the impact of the first oil crises and the War in the Middle East, PNB 1974 contained an alternative forecast. We have used the main forecast which was considerably more accurate.

---

[5] For 1992 the latest available data at the time of writing have been used (those in the NIER press release in March 1993).

TABLE 1 NIER forecasts made in December of the previous year compared with Ministry of Finance (F) forecasts in the PNB for the years 1986-92

	Mean		St dev		MAE	
	NIER	F	NIER	F	NIER	F
CP	1.6	1.6	0.5	0.8	1.7	1.3
I	-0.2	-0.0	3.2	3.6	5.0	4.5
X	2.5	2.8	1.1	0.9	1.3	1.1
Z	3.6	3.6	1.3	1.4	2.7	2.4
CCG	0.4	0.2	1.2	1.3	1.5	1.5
CLG	1.5	1.6	0.4	0.4	0.5	0.6
GDP	1.1	1.2	0.7	1.0	0.9	0.7
W	6.1	5.6	1.2	1.3	1.6	2.0
P	5.6	5.1	2.2	2.2	1.2	1.4

TABLE 2 Actual values (outcomes) in different subperiods.

	Mean value				Standard deviat			
	55-64	65-74	75-84	85-92	55-64	65-74	75-84	85-9
CP	3.6	2.7	0.8	1.7	1.4	1.7	1.9	2.1
I	4.7	2.6	-0.7	1.1	2.1	2.6	3.6	7.8
X	7.9	7.5	3.6	2.0	3.8	3.9	6.0	1.8
Z	7.3	6.9	1.6	3.6	4.9	6.0	6.8	4.6
CCG	3.5	3.3	0.6	1.2	3.3	4.5	2.6	1.8
CLG	5.0	6.6	3.6	1.3	1.7	3.9	1.0	0.[illegible]
GDP	4.2	3.3	1.4	0.9	1.7	1.5	1.9	1.8
W	8.9	9.7	10.2	7.4	2.7	1.8	3.2	2.[illegible]
P	3.5	5.9	10.0	5.9	1.3	2.5	1.9	2.[illegible]

TABLE 3 Forecasts according to PNB and RNB and outcomes (A) for the period 1955-92

	Mean value			Standard deviation			ME		MAE	
	A	PNB	RNB	A	PNB	RNB	PNB	RNB	PNB	RNB
CP	2.2	2.1	2.4	2.0	1.7	1.8	-0.2	0.2	1.3	1.2
I	2.0	1.9	2.8	4.6	3.2	3.4	-0.0	0.8	2.8	2.3
X	5.4	4.8	5.3	4.8	2.4	3.0	-0.6	-0.1	3.1	2.7
Z	4.9	4.1	4.3	6.0	2.8	3.5	-0.8	-0.6	4.2	3.5
CCG	2.2	2.2	2.5	3.4	3.0	2.9	0.1	0.3	2.1	2.0
CLG	4.3	3.9	3.8	2.9	1.9	2.1	-0.3	-0.5	1.2	1.3
GDP	2.5	2.6	2.8	2.1	1.4	1.6	0.1	0.3	1.2	1.0
W	9.1	7.1	7.7	2.6	2.2	2.3	-2.0	-1.4	2.3	1.7
P	6.3	4.9	5.7	3.2	2.5	2.8	-1.5	-0.6	1.8	1.3

TABLE 4 Forecast errors in different subperiods, PNB

	55-69	60-64	65-69	70-74	75-79	80-84	85-92
				ME			
CP	-1.1	-0.5	-0.7	1.3	0.5	-0.5	-0.1
I	-1.2	-1.5	0.7	1.8	2.5	-0.8	-1.1
X	-4.7	-4.3	-0.1	-0.3	3.3	0.5	0.9
Z	-5.1	-3.3	-1.9	0.8	2.8	0.1	0.3
CCG	0.1	1.1	1.4	-0.3	-0.1	-0.1	-1.1
CLG	0.8	0.2	-2.5	-0.3	-0.9	-0.2	0.3
GDP	-1.0	-1.1	0.3	0.6	1.0	0.5	0.4
W	-2.2	-3.4	-0.7	-0.9	-2.7	-2.6	-1.8
P	-1.6	-0.5	-0.1	-1.5	-3.0	-2.9	-1.1
				MAE			
CP	1.1	0.7	1.3	2.2	1.2	1.3	1.3
I	2.6	2.3	0.7	2.2	2.5	3.4	4.7
X	4.9	4.4	2.1	2.6	4.6	3.3	1.1
Z	5.1	4.9	3.6	4.4	7.8	2.4	2.3
CCG	3.0	2.2	2.4	2.5	2.3	1.0	1.5
CLG	1.1	1.6	2.5	1.1	1.0	0.4	0.6
GDP	1.4	1.3	0.9	1.2	2.0	1.1	0.7
W	2.2	3.4	1.6	2.0	2.7	2.6	2.1
P	1.8	0.6	0.7	1.6	3.4	2.9	1.5

TABLE 5 U-coefficients for the PNB-forecasts

	RMSE	U	U^M	U^S	U^C
CP	1.59	0.53	0.01	0.05	0.94
I	3.41	0.68	0.00	0.19	0.81
X	4.23	0.59	0.02	0.31	0.67
Z	5.17	0.67	0.02	0.37	0.60
CCG	2.62	0.66	0.00	0.02	0.98
CLG	1.75	0.34	0.03	0.28	0.69
GDP	1.54	0.47	0.01	0.22	0.77
W	2.96	0.31	0.47	0.02	0.51
P	2.42	0.34	0.37	0.08	0.55

TABLE 6 Estimated regression models for PNB. Models (3.9) and (3.10) with standard errors in parentheses. The F-test is a test of the simultaneous hypothesis H_0: $\gamma = 0$ and $\delta = 1$. For significance $F > F_{0,95}(2;36) = 3.27$ is required. R denotes the simple correlation between forecast and outcome.

	Model (3.9)							Model (3.10)		
	OLS				ML-AR1			OLS		
	α	β	DW	R	α	β	ρ	γ	δ	F
CP	0.91 (0.31)	0.52 (0.11)	1.29	0.63	1.14 (0.40)	0.41 (0.11)	0.40 (0.15)	0.61 (0.41)	0.78 (0.16)	1.18
I	1.03 (0.42)	0.46 (0.08)	1.20	0.67	1.04 (0.59)	0.44 (0.09)	0.38 (0.16)	0.06 (0.67)	0.98 (0.18)	0.01
X	3.51 (0.53)	0.23 (0.07)	0.94	0.46	3.47 (0.69)	0.22 (0.06)	0.52 (0.14)	1.01 (1.56)	0.92 (0.29)	0.45
Z	2.93 (0.52)	0.24 (0.07)	1.43	0.51	2.99 (0.60)	0.22 (0.06)	0.28 (0.16)	0.46 (1.52)	1.09 (0.31)	0.50
CCG	0.96 (0.44)	0.58 (0.11)	1.32	0.66	1.56 (0.85)	0.27 (0.08)	0.64 (0.13)	0.50 (0.53)	0.75 (0.14)	1.52
CLG	1.62 (0.34)	0.55 (0.07)	1.18	0.81	2.55 (0.55)	0.33 (0.07)	0.66 (0.13)	-0.47 (0.64)	1.20 (0.15)	1.59
GDP	1.51 (0.26)	0.44 (0.08)	1.94	0.68	1.53 (0.27)	0.44 (0.08)	0.02 (0.17)	-0.20 (0.56)	1.03 (0.19)	0.11
W	2.33 (1.06)	0.52 (0.11)	1.37	0.62	3.72 (1.22)	0.36 (0.13)	0.40 (0.17)	3.98 (1.15)	0.73 (0.15)	19.26
P	0.94 (0.56)	0.62 (0.08)	2.23	0.79	0.85 (0.49)	0.64 (0.07)	-0.17 (0.17)	1.39 (0.71)	1.02 (0.13)	10.57

Table 7 Estimates by ML-AR1 of model (3.11)

Variable	β_0	β_1	β_2	$\overline{R}^2$	ρ
CP	0.75	0.40	0.17	0.52	0.35
	(0.44)	(0.11)	(0.11)		(0.16)
I	0.70	0.42	0.16	0.54	0.40
	(0.62)	(0.09)	(0.09)		(0.15)
X	2.53	0.25	0.13	0.45	0.59
	(0.84)	(0.06)	(0.06)		(0.14)
Z	1.95	0.26	0.15	0.39	0.43
	(0.77)	(0.06)	(0.06)		(0.15)
CCG	0.80	0.42	0.24	0.61	0.39
	(0.58)	(0.09)	(0.09)		(0.16)
CLG	1.54	0.35	0.21	0.78	0.49
	(0.47)	(0.07)	(0.07)		(0.15)
GDP	0.96	0.26	0.37	0.67	0.30
	(0.29)	(0.07)	(0.07)		(0.16)
W	0.89	0.24	0.43	0.56	0.33
	(1.25)	(0.11)	(0.12)		(0.17)
P	0.47	0.45	0.24	0.68	-0.11
	(0.50)	(0.10)	(0.10)		(0.17)

TABLE 8 Test of efficiency. $F_{obs} > F_{0.95}(4;34) = 2.65$ is significant at the 5% level

	F_{obs}	Variabel	F_{obs}
CP	0.76	CLG	1.59
I	0.31	GDP	0.26
X	0.30	W	11.86
Z	0.40	P	6.39
CCG	3.09		

TABLE 9 Correlation between the forecast error of PNB for different variables 1955-92. $|r| \geq 0.32$ is significant at the 5 % level (two-tailed test).

	CP	I	X	Z	CS	CK	GDP	W	P
CP	1.00								
I	0.39	1.00							
X	0.36	0.27	1.00						
Z	0.64	0.35	0.66	1.00					
CCG	-0.08	0.08	-0.11	-0.12	1.00				
CLG	0.04	0.06	-0.27	-0.20	-0.23	1.00			
GDP	0.61	0.55	0.68	0.49	0.07	-0.04	1.00		
W	0.29	0.01	0.00	0.27	-0.09	-0.21	0.13	1.00	
P	-0.36	-0.13	-0.32	-0.15	-0.08	-0.04	-0.46	0.23	1.00

TABLE 10 PNB-forecasts compared with naive forecasts. Mean Absolute Errors 1955-92.

	PN1	PN2	PN3	PNB
CP	1.77	1.70	1.62	1.29
I	3.64	3.59	3.72	2.77
X	4.73	4.29	3.88	3.10
Z	6.39	5.83	5.23	4.19
CCG	3.17	3.09	2.77	2.08
CLG	1.53	1.65	1.64	1.15
GDP	1.76	1.74	1.73	1.18
W	1.89	2.18	2.20	2.35
P	2.17	2.09	2.03	1.78

TABLE 11 The MAE error of PNB compared with that of PN3

	Paired t-test			Estimate of model (5.2)[1]		
	Difference in MAE	t	DW	α	β	DW
CP	-0.33	-1.76	1.94	-0.08	-0.0128 (0.0172)	1.97
I	-0.94	-2.12	1.08	0.62	-0.0799 (0.0389)	1.21
X	-0.78	-1.48	2.12	0.84	-0.0831 (0.0466)	2.30
Z	-1.04	-2.02	2.29	-0.75	-0.0147 (0.0475)	2.30
CCG	-0.69	-2.27	1.98	-0.51	-0.0091 (0.0281)	1.99
CLG	-0.48	-2.37	1.98	-0.77	0.0149 (0.0186)	2.02
GDP	-0.55	-3.77	2.59	-0.30	-0.0126 (0.0133)	2.65
W	0.15	0.40	1.47	0.87	-0.0370 (0.0345)	1.51
P	-0.25	-0.70	1.87	-0.11	-0.0072 (0.0333)	1.87

[1] t in model (5.2) is defined as 1,2, ... for the years 1955,56 .... , which means that α refers to the year 1954.

Table 12 The MSE of the PNB forecasts compared with the MSE of the naive forecasts PN3. Estimates based on model (3.15) (variance and MSE computed with n in the denominator).

	Bias		Variance		MSE		ML-AR1 estimates of model (3.15)	
Variable	PN3	PNB	PN3	PNB	PN3	PNB	β_1	β_2
CP	0.25	-0.16	4.01	2.52	4.07	2.54	0.41	0.132
							(0.33)	(0.068)
I	0.87	-0.03	22.64	11.60	23.39	11.60	1.02	0.209
							(0.90)	(0.063)
X	0.27	-0.63	27.52	17.50	27.59	17.90	0.85	0.170
							(1.05)	(0.042)
Z	0.38	-0.81	48.57	26.06	48.72	26.72	1.16	0.170
							(1.05)	(0.039)
CCG	0.11	0 .05	12.52	6.87	12.53	6.88	0.06	0.163
							(0.37)	(0.056)
CLG	0.31	-0.32	5.30	2.95	5.39	3.05	0.59	0.168
							(0.42)	(0.051)
GDP	0.31	0.11	4.36	2.38	4.45	2.39	0.20	0.139
							(0.28)	(0.055)
W	0.10	-2.04	7.73	4.59	7.74	8.74	2.19	0.119
							(0.45)	(0.078)
P	-0.17	-1.47	6.56	3.69	6.59	5.85	1.36	0.157
							(0.50)	(0.093)

TABLE 13 Regression models relating RNB forecasts and outcomes

	Model (3.9) OLS			ML-AR1			Model (3.10)	
	α	β	DW	α	β	R	γ	δ
CP	0.92 (0.30)	0.68 (0.10)	1.06	1.34 (0.43)	0.52 (0.10)	0.74	0.24 (0.37)	0.82 (0.12)
I	1.59 (0.36)	0.61 (0.07)	1.48	1.63 (0.44)	0.57 (0.08)	0.82	-1.09 (0.57)	1.10 (0.13)
X	3.23 (0.59)	0.38 (0.08)	1.23	3.13 (0.70)	0.39 (0.08)	0.61	0.20 (1.28)	0.98 (0.21)
Z	2.19 (0.50)	0.43 (0.07)	1.66	2.28 (0.56)	0.41 (0.06)	0.74	-0.53 (1.06)	1.26 (0.19)
CCG	1.20 (0.41)	0.59 (0.10)	1.07	1.77 (0.91)	0.29 (0.07)	0.69	0.17 (0.54)	0.81 (0.14)
CLG	1.54 (0.45)	0.52 (0.09)	1.45	2.26 (0.61)	0.35 (0.11)	0.70	0.74 (0.69)	0.94 (0.16)
GDP	1.29 (0.23)	0.60 (0.07)	1.85	1.35 (0.25)	0.58 (0.07)	0.81	-0.58 (0.42)	1.11 (0.13)
W	2.04 (0.96)	0.62 (0.10)	1.99	2.04 (0.96)	0.62 (0.10)	0.71	2.81 (1.08)	0.82 (0.13)
P	1.38 (0.63)	0.69 (0.09)	1.89	1.40 (0.65)	0.69 (0.09)	0.79	1.12 (0.74)	0.91 (0.12)

FIG 1 The NIER series on composite resource utilization ("Resursutnyttjandet inom industrin")1964-1993.
Source: The NIER Business Tendency Survey, June 1993.

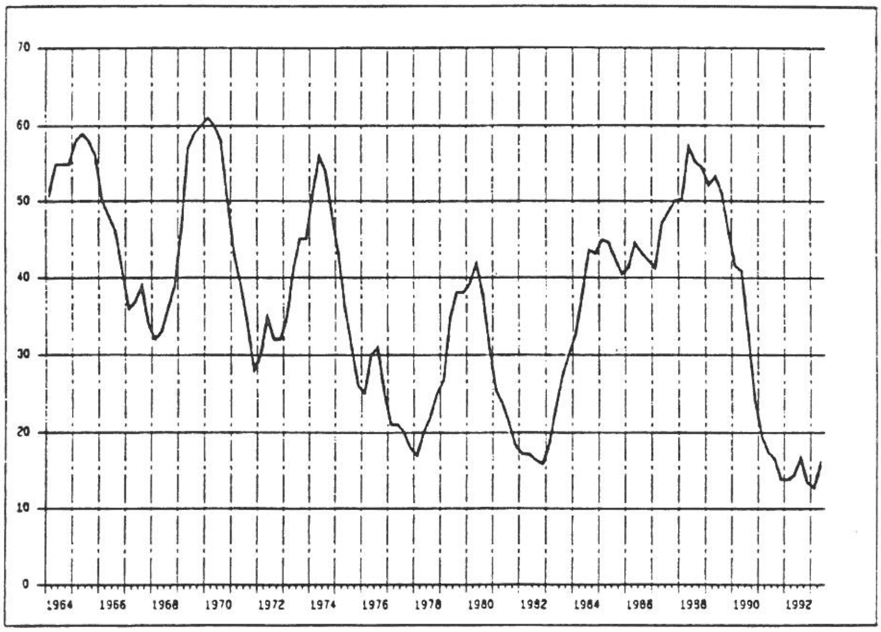
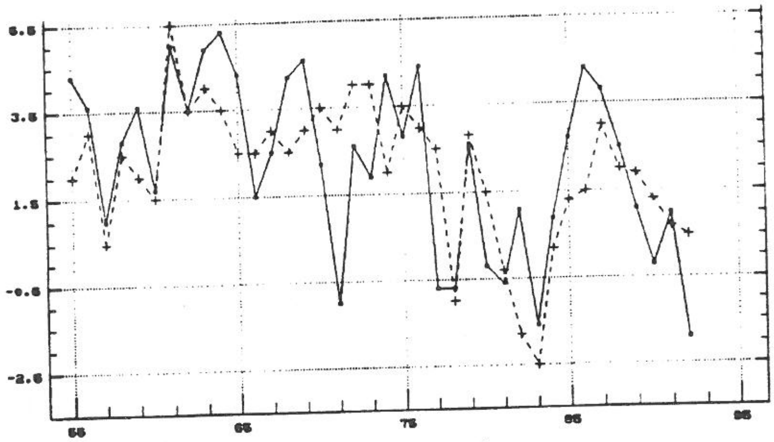
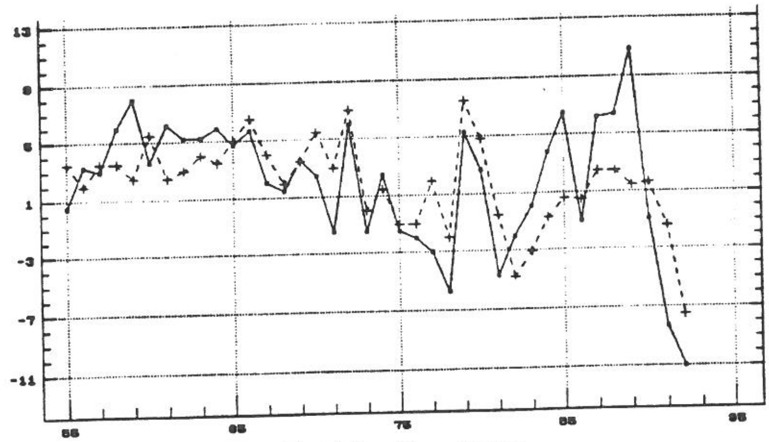


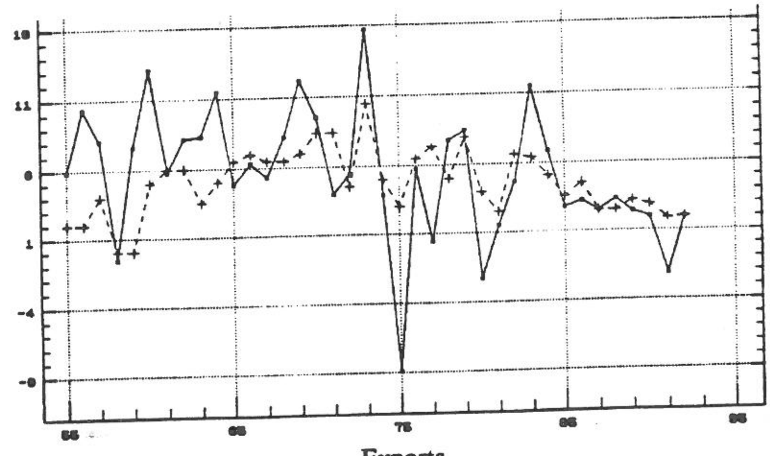
FIG 2 PNB forecasts (- - -) compared outcomes (——) 1955-1992. Percentage changes compared with previous year.



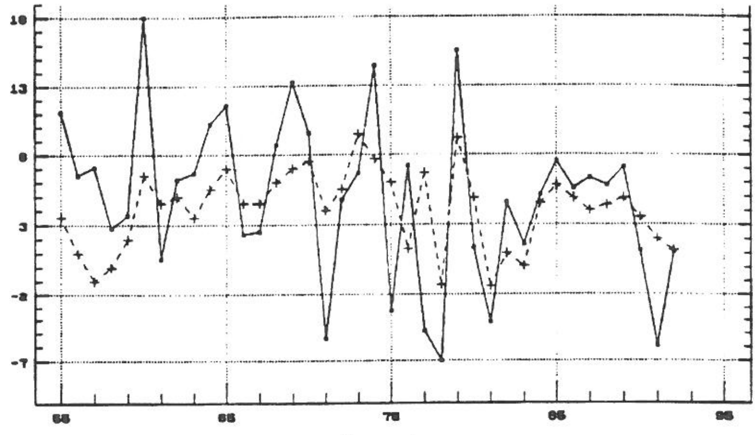
Private consumption



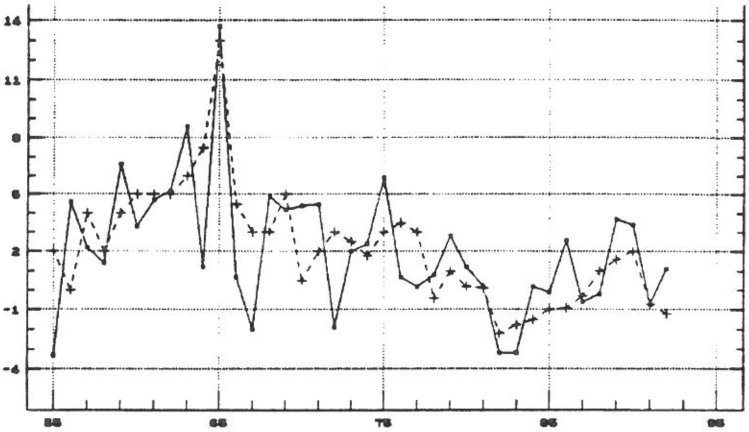
Total fixed investment



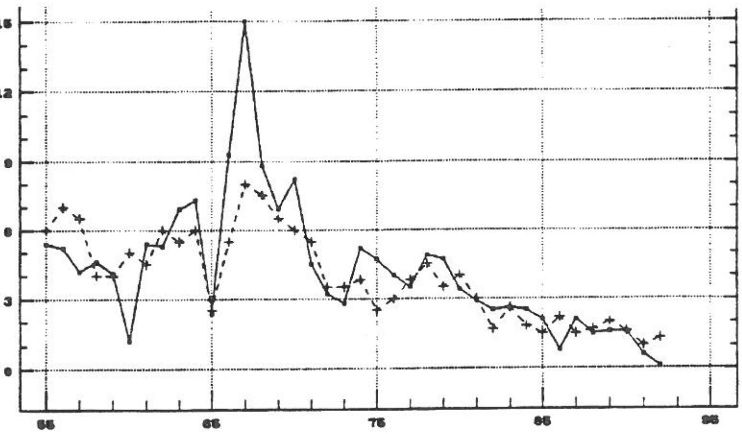
Exports



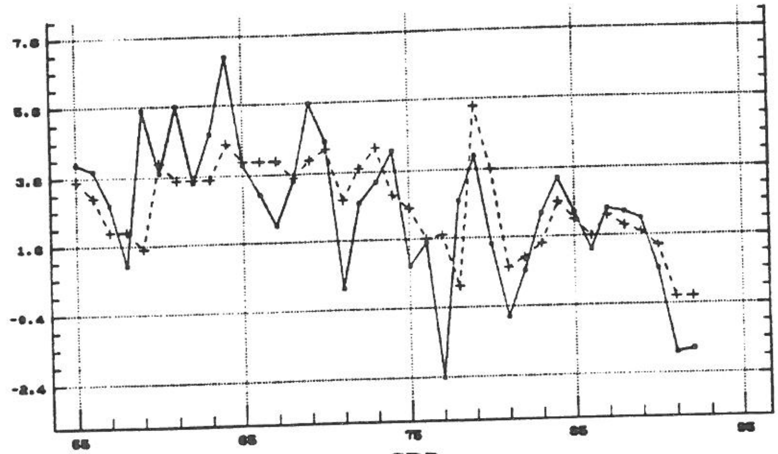
Imports



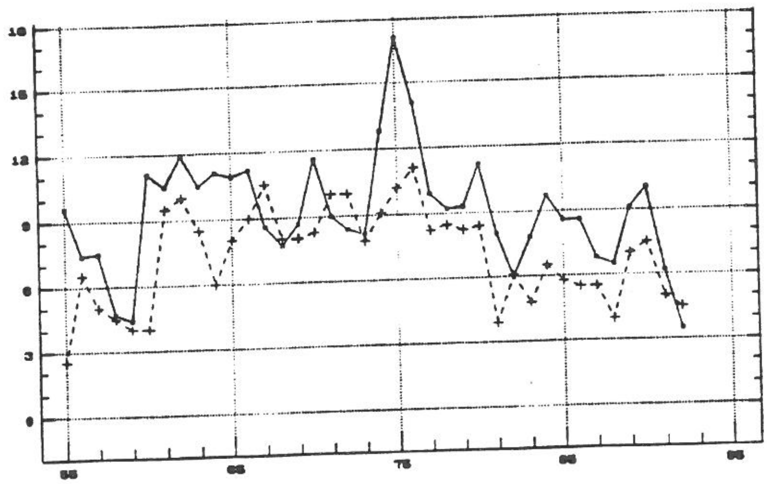
Central Government Consumption



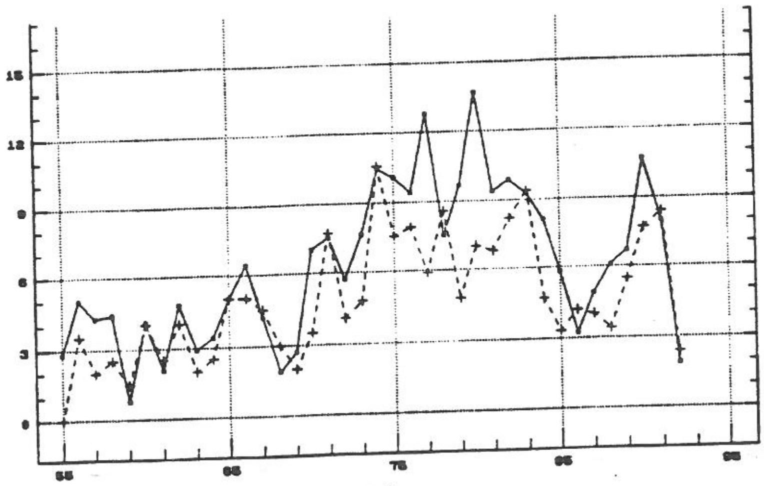
Local Government Consumption



GDP



Wages



Prices

Precisionen i de svenska Preliminära och Reviderade Nationalbudgetarna under perioden 1955-1992.

De prognoser som nu under en mycket lång tid har presenterats i form av den Preliminära nationalbudgeten (PNB) i januari och den Reviderade nationalbudgeten (RNB) senare på våren är den kanske viktigaste och mest använda källan för korttidsprognoser i Sverige. Ansvariga för dessa prognoser har varit och är Konjunkturinstitutet (KI) och Finansdepartementet. Den exakta arbetsfördelningen har varierat under den långa tidsperiod som nu föreliggande studie täcker. En översiktlig beskrivning av hur prognoserna tagits fram ges i den första delen av uppsatsen.

Utvärderingen av prognoserna sker mot utfallsdata som de föreligger i slutet av år t+1 (prognosåret är år t). Olika alternativ till detta val av utfallsdata diskuteras och för- och nackdelar av olika val anges. En mängd olika metoder används för att analysera prognoskvaliteten, alltifrån enkla deskriptiva metoder till regressionsmodeller som visar relationen mellan utfall och prognos på olika sätt. Jämförelser sker dessutom med olika naiva prognosmetoder.

Analyserna av prognosprecisionen för enskilda variabler visar att oberoende av analysmetod har kvaliteten i pris- och löneprognoserna i allmänhet varit sämre än precisionen hos övriga variabler. Priser och löner har nästan genomgående underskattats, även om en viss förbättring tycks ha skett mot periodens slut. Orsakerna till de sämre prognoserna för speciellt lönerna diskuteras.

Sett över hela perioden föreligger ingen klar systematisk över- eller underskattning av variablerna med undantag av priser och löner. Under kortare delperioder kan systematiska prognosfel däremot iakttas, t ex i ansluning till oljekriserna på 1970-talet. Prognosprecisionen mätt som genomsnittligt absolut fel (MAE) eller RMSE har inte genomgått någon markant förändring under den studerade perioden. Denna slutsats ändras inte påtagligt om "svårigheten att göra prognoser" vägs in. Det genomsnittliga absoluta prognosfelet för BNP över hela perioden är 1.2 procentenheter enligt PNB-prognoserna och 1.0 procentenheter enligt RNB-prognoserna. På basis av femårsperioder varierar MAE för PNB-prognoserna mellan 2.0 procentenheter 1975-79 och 0.7 procentenheter under den sista perioden 1985-92.

PNB prognoserna är bättre än de naiva prognoserna, åter med undantag av prognoserna för löner. Över hela tidsperioden kan PNB-prognosens MAE för BNP på 1.2 procentenheter jämföras med de naiva prognosernas motsvarande värde på 1.7 procentenheter.

Regressionsmodeller har skattats både med prognos och utfall som beroende variabel och eventuellt också med olika typer av laggade variabler ingående i modellen. Dessa modeller belyser olika aspekter av prognosprecisionen. Bland annat testas om prognoserna kan anses vara rationella förväntningar och speciellt de två aspekter av rationella förväntningar som brukar betecknas som unbiased prognoser och effektiva prognoser.

Återigen med priser och löner som undantag visar det sig att prognoserna i allmänhet uppfyller dessa båda kriterier. I en modell med prognosvärdet som den beroende variabeln uppnås som genomgående resultat att konstanta termen är positiv och β-koefficienten mindre än 1, vilket visar på en underskattning vid höga utfallsvärden och en överskattning vid låga utfallsvärden. I många fall är OLS-estimaten inte optimala vid denna typ av modeller, vilket föranleder användning av ML-estimat av modeller som tillåter autokorrelation i residualerna.

Som man skulle kunna vänta är RNB-prognoserna bättre än PNB-prognoserna. Förbättringen är speciellt markant för de variabler där PNB-prognoserna var minst lyckade, dvs speciellt för löner och priser.

Titles in the Working Paper Series

Number	Author	Title	Date
1	Warne, Anders and Anders Vredin	Current Account and Business Cycles: Stylized Facts for Sweden	12/89
2	Östblom, Göran	Change in Technical Structure of the Swedish Economy	12/89
3	Söderlind, Paul	Mamtax. A Dynamic CGE Model for Tax Reform Simulations	12/89
4	Kanis, Alfred and Aleksander Markowski	The Supply Side of the Econometric Model of the NIER	11/90
5	Berg, Lennart	The Financial Sector in the SNEPQ Model	2/91
6	Ågren, Anders and Bo Jonsson	Consumer Attitudes, Buying Intentions and Consumption Expenditures. An Analysis of the Swedish Household Survey Data	4/91
7	Berg, Lennart and Reinhold Bergström	A Quarterly Consumption Function for Sweden 1979-1989	10/91
8	Öller, Lars-Erik	Good Business Cycle Forecasts - A Must for Stabilization Policies	2/92
9	Jonsson, Bo and Anders Ågren	Forecasting Car Expenditures Using Household Survey Data	2/92
10	Löfgren, Karl-Gustaf, Bo Ranneby and Sara Sjöstedt	Forecasting the Business Cycle Not Using Minimum Autocorrelation Factors	2/92
11	Gerlach, Stefan	Current Quarter Forecasts of Swedish GNP Using Monthly Variables	2/92
12	Bergström, Reinhold	The Relationship Between Manufacturing Production and Different Business Survey Series in Sweden	2/92
13	Edlund, Per-Olov and Sune Karlsson	Forecasting the Swedish Unemployment Rate: VAR vs. Transfer Function Modelling	3/92
14	Rahiala, Markku and Timo Teräsvirta	Business Survey Data in Forecasting the Output of Swedish and Finnish Metal and Engineering Industries: A Kalman Filter Approach	3/92

15	Christoffersson, Anders, Roland Roberts and Ulla Eriksson	The Relationship Between Manufacturing and Various BTS Series in Sweden Illuminated by Frequency and Complex Demodulate Methods	4/92
16	Jonsson, Bo	Sample Based Proportions as Values on an Independent Variable in a Regression Model	10/92
17	Öller, Lars-Erik	Eliciting Turning Point Warnings From Business Surveys	11/92
18	Forster, Margaret M.	Volatility, Trading Mechanisms and International Cross-Listing	12/92
19	Jonsson, Bo	Prediction With a Linear Regression Model and Errors in a Regressor	12/92
20	Gorton, Gary and Richard Rosen	Corporate Control, Portfolio Choice, and the Decline of Banking	1/93
21	Gustafsson, Claes-Håkan and Åke Holmén	The Index of Industrial Production - A Formal Description of the Process Behind It	2/93
22	Karlsson, Tohmas	A General Equilibrium Analysis of the Swedish Tax Reforms 1989-1991	2/93
23	Jonsson, Bo	Forecasting Car Expenditures Using Household Survey Data - A Comparison of Different Predictors	2/93
24	Gennotte, Gerard and Hayne Leland	Low Margins, Derivative Securities, and Volatility	2/93
25	Boot, Arnoud W.A. and Stuart I. Greenbaum	Discretion in the Regulation of U.S. Banking	4/93
26	Spiegel, Matthew and Deane J. Seppi	Does Round-the-Clock Trading Result in Pareto Improvements?	4/93
27	Seppi, Deane J.	How Important are Block Trades in the Price Discovery Process?	4/93
28	Glosten, Lawrence R	Equilibrium in an Electronic Open Limit Order Book	4/93
29	Boot, Arnoud W.A., Stuart I. Greenbaum and Anjan V. Thakor	Reputation and Discretion in Financial Contracting	4/93
30a	Bergström, Reinhold	The Full Tricotomous Scale Compared with Net Balances in Qualitative Business Survey Data - Experiences from the Swedish Business Tendency Surveys	5/93

30b	Bergström, Reinhold	Quantitative Production Series Compared with Qualitative Business Survey Series for Five Sectors of the Swedish Manufacturing Industry	5/93
31	Lin, Chien-Fu Jeff and Timo Teräsvirta	Testing the Constancy of Regression Parameters Against Continuous Change	5/93
32	Markowski, Aleksander and Parameswar Nandakumar	A Long-Run Equilibrium Model for Sweden. The Theory Behind the Long-Run Solution to the Econometric Model KOSMOS	9/93
33	Markowski, Aleksander and Tony Persson	Capital Rental Cost and the Adjustment for the Effects of the Investment Fund System in the Econometric Model KOSMOS	9/93
34	Kanis, Alfred and Bharat Barot	On Determinants on Private Consumption in Sweden	10/93
35	Kääntä, Pekka and Christer Tallbom	Using Business Survey Data for Forecasting Swedish Quantitative Business Cycle Variables. A Kalman Filter Approach	11/93
36	Ohlsson, Henry and Anders Vredin	Political Cycles and Cyclical Policies. A new test approach using fiscal forecasts	11/93
37	Markowski, Alexander and Lars Ernsäter	The Supply Side in the Econometric Model Kosmos	1/94
38	Gustafsson, Claes-Håkan	On the Consistency of Data on Production, Deliveries, and Inventories in the Swedish Manufacturing Industry	8/94
39	Rahiala, Markku and Tapani Kovalainen	Modelling Wages Subject to Both Contracted Increments and Drift by Means of a Simultaneous-Equations Model With Non-Standard Error Structure	8/94
40	Öller, Lars-Erik and Christer Tallbom	Hybrid Indicators for the Swedish Economy Based on Noisy Statistical Data and the Business Tendency Survey	10/94
41	Östblom, Göran	A Converging Triangularization Algorithm and the Intertemporal Similarity of Production Structures	10/94
42a	Markowski, Aleksander	Labour Suply, Hours Worked and Unemployment in the Econometric Model Kosmos	10/94
42b	Markowski, Aleksander	Wage Rate Determination in the Econometric Model Kosmos	10/94

43	Ahlroth, Sofia, Anders Björklund and Anders Forslund	The Output of the Swedish Education Sector	12/94
44a	Markowski, Aleksander	Private Consumption Expenditure in the Econometric Model Kosmos	12/94
44b	Markowski, Aleksander	The Input-Output Core: Determination of Inventory Investment and Other Business Output in the Econometric Model Kosmos	12/94