Forecasting Swiss inflation using VAR models

Caesar Lack

No. 2 2006

SCHWEIZERISCHE NATIONALBANK
BANQUE NATIONALE SUISSE
BANCA NAZIONALE SVIZZERA
BANCA NAZIUNALA SVIZRA
SWISS NATIONAL BANK

Swiss National Bank Economic Studies

Copies of Swiss National Bank Economic Studies may be obtained from:

Swiss National Bank, Library, Fraumünsterstrasse 8, P.O. Box, CH-8022 Zurich
Fax: +41 44 631 81 14
E-mail: library@snb.ch

This publication is also available on the SNB website (www.snb.ch).

ISSN 1661-142X

Swiss National Bank Economic Studies

No. 2 2006

Forecasting Swiss inflation using VAR models

Caesar Lack*

* Swiss National Bank, P.O. Box, 8022 Zürich, Switzerland. E-mail: caesar.lack@snb.ch. The views expressed in this paper do not necessarily reflect those of the Swiss National Bank. I would like to thank Enzo Rossi, Samuel Reynard, Andreas Fischer, Marco Huwiler and an anonymous referee for their helpful discussions and comments.

Contents

Abstract

A procedure that has been used at the Swiss National Bank for selecting vector-autoregressive (VAR) models in order to forecast Swiss consumer price inflation is presented. In order to examine and improve the quality of the procedure, it is submitted to several modifications and the results are compared with one another. Combining forecasts substantially improves the quality of the forecasts. Models specified with respect to levels of variables are superior to those specified with respect to differences in variables. Bank loans and the monetary aggregate M_3 are the most important variables for inflation forecasting. The optimized procedure reduces the root mean squared error (RMSE) of the inflation forecast to one third of the RMSE of a naive "no change" forecast over the period from 1987 to 2005.

JEL classification: C32, C52, C53, E37
Keywords: Inflation forecasting, VAR models, model selection, model evaluation

Zusammenfassung

Diese Untersuchung präsentiert eine von der Schweizerischen Nationalbank verwendete Methode für die Erstellung von unbedingten Inflationsprognosen für die Schweiz mittels vektorautoregressiven (VAR) Modellen. Um die Güte der Methode zu untersuchen und zu verbessern, wird die Methode auf verschiedene Weise modifiziert, und die Ergebnisse werden miteinander verglichen. Es stellt sich heraus, dass eine Kombination von Prognosen die Prognosequalität deutlich erhöht. Modelle mit Variablen in Niveaus sind Modellen mit Variablen in ersten Differenzen überlegen. Die wichtigsten erklärenden Variablen für die Inflation sind die Bankkredite und die Geldmenge M_3. Die Wurzel aus dem mittleren quadratischen Fehler (RMSE) der optimierten Prognosemethode beträgt in der Periode 1987–2005 nur einen Drittel des RMSE einer Benchmark-Prognose, die eine unveränderte Inflationsrate prognostiziert.

Résumé

Une procédure de sélection de modèles vectoriels autorégressifs (VAR) utilisée à la Banque nationale suisse, ayant comme but la prévision d'inflation, est présentée. Afin d'analyser et d'améliorer cette procédure, elle est soumise à différentes variantes et les résultats sont comparés entre eux. Il ressort qu'une combinaison de prévisions améliore nettement la qualité. De plus, les modèles utilisant des variables en niveau sont supérieurs à ceux calculés à l'aide de variables en première différence. Les crédits bancaires et la masse monétaire M_3 sont les facteurs les plus importantes pour la prévision d'inflation. La procédure optimale, pour la période 1987–2005, réduit la racine de l'erreur quadratique moyenne (RMSE) de la prévision d'inflation à un tiers de celle d'une prévision d'inflation naïve assumant un taux d'inflation constant dans le temps.

1. Introduction

In 2000, the Swiss National Bank (SNB) introduced a new monetary policy framework. Monetary targeting, which was pursued from 1974 until 1999, was abandoned in favour of a new strategy based on an inflation forecast.[1] Since late 1999, inflation forecasts have been regularly calculated and published by the SNB. The inflation forecast which constitutes the focus of the new monetary policy concept is a conditional inflation forecast, i.e. it is conditional upon the assumption that the reference interest rate will remain unchanged over the forecast horizon. The conditional inflation forecast is published quarterly, extends over a horizon of 12 quarters, and is derived mainly from structural models of the economy.[2]

The SNB also calculates unconditional inflation forecasts. They are derived from both structural and non-structural models and are obtained under the assumption that the reference interest rate evolves endogenously. VAR models are one of several different methods used to calculate unconditional inflation forecasts at the SNB.

As the inflation forecast is the cornerstone of the current Swiss monetary policy strategy, knowledge about the degree of reliability of the forecast is essential for policy-making. This study's objective is twofold: First, we want to convey a comprehensive picture of the quality of VAR inflation forecasts in Switzerland. Second, the simulation results will also make it possible to improve the quality of inflation forecasts in the future.[3] Our objectives are achieved by repeatedly simulating the VAR forecasting procedure of the SNB retroactively since 1987.[4] Comparing the out-of-sample forecast performance of different specifications of the forecasting procedure yields a host of insights into the predictability and determinants of Swiss inflation.

The paper is organized as follows: Section 2 describes the procedure used for unconditional inflation forecasting in more detail. Section 3 shows how modifications to the procedure affect the forecasting performance. Section 4 provides a summary and draws some conclusions.

---

1 See Rich (2000) for an analysis of the SNB's experience with monetary targeting, the reasons for the change towards the new strategy, and a description of the new strategy.

2 See Jordan and Peytrignet (2001) for an overview of inflation forecasting at the SNB, Stalder (2001) and Jordan et al. (2002) for a description of structural models used for inflation forecasting.

3 As VAR models can also be viewed as reduced forms of other linear model classes, many of the conclusions obtained by VAR modeling also apply to other models and to conditional forecasts.

4 The starting point of the simulation was chosen such that the simulation period contains a full inflationary cycle.

2. The forecasting procedure

The SNB does not stick to one particular VAR model, preferring to use a model selection procedure which, each quarter, chooses new models for forecasting. This section is entirely devoted to a description of the forecasting procedure, because it is important that this be understood when interpreting the results presented in section 3.

An initial feature of the procedure followed at the SNB is that forecasts from several VAR models are combined to produce one final forecast.[5] It is a well-known fact that forecast accuracy can be improved by combining multiple forecasts.[6] The main source of the improvement is a diversification effect. If individual forecast errors from different models are not perfectly correlated, the forecast variance of a combination of forecasts is smaller than the average forecast variance of the individual forecasts. Furthermore, combining forecasts also reduces the problem of model uncertainty. By combining forecasts from different models, reliance upon one single (possibly misspecified) model is avoided. Eventually, by combining forecasts from models with different variables, information can be aggregated. This argument is particularly important in the case of VAR models, because they may contain a limited number of variables only.

When combining several forecasts into one forecast, several weighting schemes can be used.[7] The SNB's VAR forecasting procedure gives an equal weight to each forecast. Equal weights have some advantages over more complex weighting schemes insofar as they are robust, easy to calculate and understand, and often perform equally well or even better than more complicated weighting schemes.[8]

A second feature of the SNB's VAR forecasting procedure is that it draws its variables from a pool of variables. In its initial setup, the pool contains 10 variables, chosen on the basis of a priori knowledge drawn from empirical studies on the Swiss economy. The pool contains the consumer price index (CPI), the trade-weighted nominal and real exchange rate index of the Swiss franc, real GDP, the monetary aggregates M_1, M_2 and M_3, outstanding domestic bank loans, the three month LIBOR (which has been the monetary policy reference rate of the SNB since 2000) and the ten-year government bond yield. All series are available from 1974Q1 to 2005Q2. With the exception of the interest rates, all variables are transformed with the natural logarithm. The series are not seasonally adjusted, because seasonal adjustment procedures usually apply two-sided filters and thus, for any given point in the past, provide future information not available at the time of measurement.

5 The forecasting procedure has also been described in Jordan et al. (2002).

6 For a review of the literature on combining forecasts see Clemen (1989) or Newbold and Harvey (2002).

7 For an overview of weighting schemes for the combination of forecasts see Newbold and Harvey (2002).

8 Makridakis et al. (1982), Makridakis and Winkler (1983), and several studies discussed in Clemen (1989) find that simple averages often outperform alternative weighting schemes. For Switzerland, Jordan and Savioz (2003) find that simple averages perform either best or not much worse than other weighting schemes for forecasting consumer price inflation.

The problems of data availability and of data revisions are set aside in this investigation. Thus, it is assumed that all series become available simultaneously in their final revised form. In reality, however, some series are available only with a considerable lag, or are likely to be revised significantly during the quarters that follow their initial publication.[9]

The VAR models are estimated either in first differences, which are found to be stationary, or in levels, which allows us to take advantage of potential cointegrating relationships among variables.[10] All models contain a constant term, but no time trend.

The VAR models evaluated in the selection procedure contain up to five variables. This is usually assumed to be sufficient to describe the most important dynamics in an economic system. In the economic literature, VAR models normally consist of three to five variables. The limited number of observations does not allow for the estimation of VAR models consisting of more than five variables when using quarterly data.

Clearly, all models need to include the CPI (either in levels or first differences), in order to forecast inflation. The other variables are determined in a search and evaluation process. First, all possible combinations of one to five pool variables are set up as VAR models.[11] This yields 512 models (256 in levels and 256 in first differences). All models are then estimated in rolling regressions and ranked according to their out-of-sample forecasting performance over a certain forecast horizon during an evaluation period. The best models are then re-estimated up to the latest available observation and the final inflation forecast is calculated as the combination of their forecasts.[12]

Graph 1 illustrates the process of calculating an inflation forecast at time t. First, the 512 VAR models are estimated for an estimation period with data ranging through to j. For each model, the CPI forecast is calculated, and the forecast error of the cumulated change in the CPI

---

9 Ignoring the problem of data revisions and data availability is less daring than might appear at first. As we ascertain in this study, the variable most likely to be misrepresented by this simplified approach, namely GDP, has almost no explanatory power, even in its revised form. Most variables which prove to be good predictors of inflation are available with minor time lags only and are subject to minor revisions only, or no revisions at all.

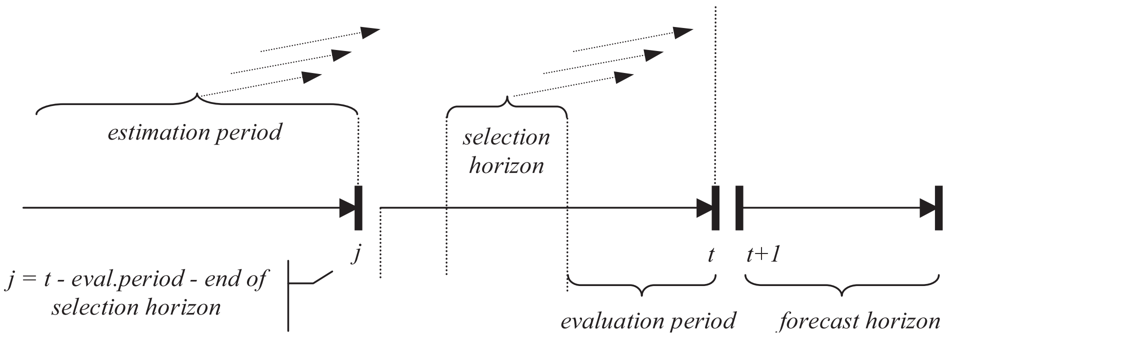
10 Applying the Augmented Dickey Fuller and the Phillips Perron unit root tests to first differences, we find that the unit root hypothesis can be rejected at very high significance levels for almost all time series. Exceptions are the CPI and bank loans, where the unit root hypothesis can only be rejected at the 5% or even 10% levels. However, the KPSS stationarity test does not reject stationarity in these two cases. See Sims, Stock and Watson (1990) on estimation of VAR models when unit roots are present.

11 In the case of only one variable, this is equivalent to a univariate AR-process for the CPI (in levels or first differences).

12 The procedure runs under RATS Version 6.02. Computing time for the simulation of one inflation forecast on a Pentium 4 CPU 2.40 GHz is about 30 seconds, a full simulation for the period 1986Q4–2005Q1 contains 74 forecasts and takes about 35 minutes with a pool of 10 variables.

for the selection horizon is stored.[13] These calculations are performed repeatedly for each quarter by shifting the estimation period and the selection horizon forward by one quarter, until the end of the selection horizon reaches *t*. The 512 models are then ordered according to the root mean squared error (RMSE) of their forecast of the cumulated changes in the CPI for the selection horizon. The best models are re-estimated with data until *t* and the final forecast is calculated as the forecast of the best models, using equal weights for each model. In a full forecasting simulation, *t* runs from 1986Q4 to 2005Q1, which yields 74 inflation forecasts.

Graph 1
The forecasting procedure



The procedure is best explained by means of an example. Let the estimation period be 15 years, the evaluation period 20 quarters, the selection horizon 7 to 10 quarters, and the forecast horizon 16 quarters.[14] These choices approximately reflect the parameters currently in use at the SNB. An inflation forecast at t = 1986Q4 over the subsequent 16 quarters is calculated as follows:

First, all 512 models are estimated with data up to 1979Q2.[15] The forecasts for the annual inflation rate in 1981Q4 are calculated, compared with the actual inflation rate in this period, and the forecast error is stored for each model. This process is repeated 20 times, each time shifting the estimation period and the selection horizon forward by one quarter, until the models are estimated with data up to 1984Q2 and the forecast errors for annual inflation in 1986Q4

13 The selection horizon indicates at which forecast horizon the models are evaluated. For example, a selection horizon of 7–10 quarters corresponds to the annual inflation rate, two and a half years ahead.

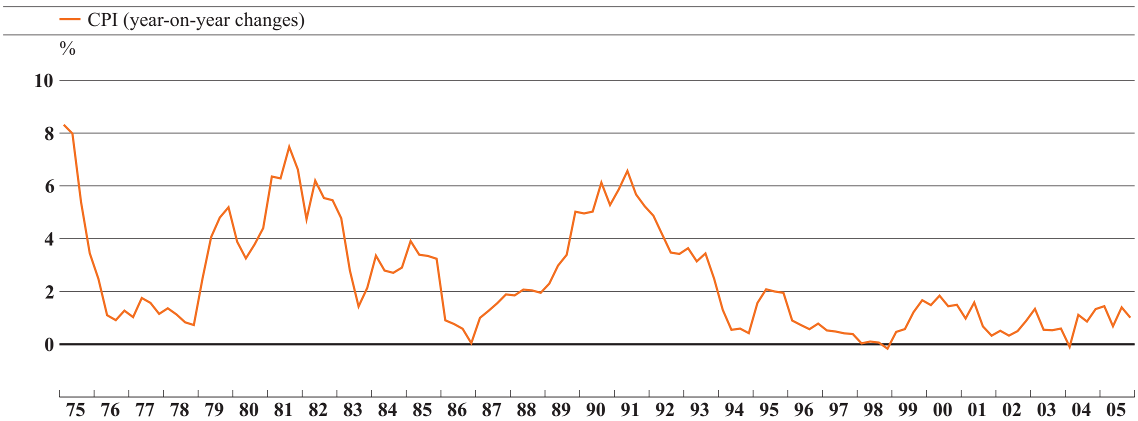
14 Although the target horizon of Swiss monetary policy is three years, a shorter selection horizon of 7–10 quarters was chosen for two reasons: first, in order to save observations during the simulation, and second, to reflect the fact that the SNB is also interested in inflation for horizons shorter than 12 quarters. The forecast horizon (16 quarters) is set above the target horizon of monetary policy (12 quarters) in order to learn more about the forecasting properties of VAR models for longer horizons.

15 The models should be estimated over the full estimation period of 15 years. However, at the beginning of the simulation period and with long estimation periods, not enough observations are available, so that the estimation period is determined by the availability of the data. All time series start in 1974Q1.

have been determined. Subsequently, the RMSE of the inflation forecast for a horizon of ten quarters is calculated for all models. The models with the lowest RMSE are then estimated using data to 1986Q4. The inflation forecast for the period from 1987Q1 to 1990Q4 is finally calculated as the equally weighted inflation forecast of the models with the lowest RMSE. In a full forecasting simulation, this procedure is applied 74 times for t from 1986Q4 to 2005Q1.

The start of the simulation is set to 1987 in order to cover both the rise and the fall of inflation during the last significant inflationary cycle in Switzerland. Graph 2 shows inflation rising from below 1% in 1987 to above 6% in 1991 and going back to below 2% after 1994. Starting the simulation later would be less informative because inflation and inflation volatility have been very low since 1994. Starting the simulation earlier to include periods with even more inflation variation is not possible, due to a lack of observations.

Graph 2
Swiss consumer price inflation 1975–2005



The forecasting procedure was first applied by the SNB in the year 2000. In the following section, several simulations – each consisting of 74 inflation forecasts – will be run, using modified versions of this procedure. A comparison of forecast errors for different modifications of the procedure will then allow us to draw conclusions on the optimal specification of the procedure and the determinants of inflation.

3. Results

The quality of the inflation forecasts derived from a full simulation of the forecasting procedure from 1986Q4 to 2005Q1 will be measured by a variant of Theil's U-statistic, which compares the ratio of the RMSE of the procedure's inflation forecast to the RMSE of a "naive" no-change forecast of the inflation rate. The "naive" model implies $E(\pi_{t+h}|\pi_t)=\pi_t$ at all forecast horizons h, for annual consumer price inflation π_t. A U-statistic above 1 means that the VAR models perform worse than a no-change forecast for the annual inflation rate. Since the "naive" forecast would represent the optimal forecast if inflation followed a random walk process, it will also be referred to as the "random walk forecast" below. The U-statistic is calculated at each forecast horizon h from 1 to 16 quarters as follows:

$$U_h = \frac{\sqrt{\sum_{t=1986Q4}^{2005Q2-h} \left(\hat{\pi}_{t+h} - \pi_{t+h}\right)^2}}{\sqrt{\sum_{t=1986Q4}^{2005Q2-h} \left(\pi_t - \pi_{t+h}\right)^2}}$$

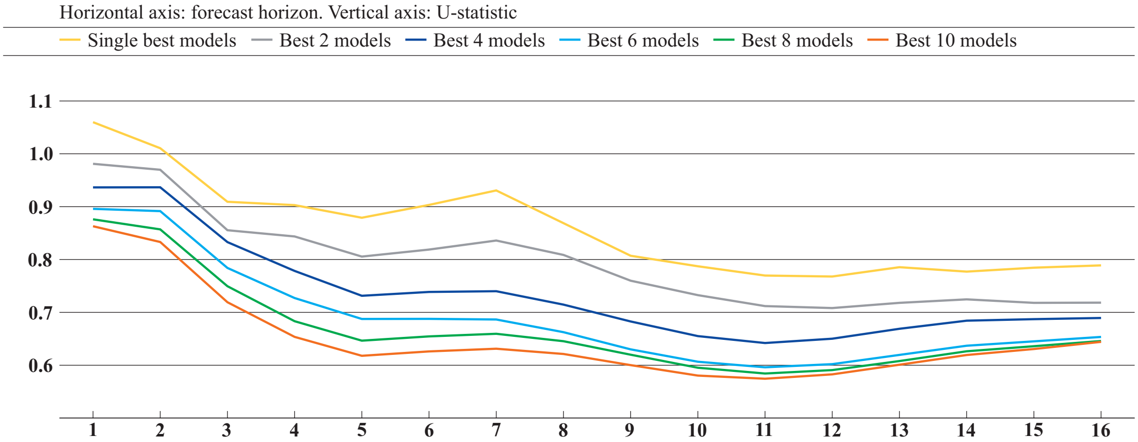
$h=1..16$ quarters, $\pi_t = log(CPI_t) - log(CPI_{t-4})$. $\hat{\pi}_{t+h}$ denotes the inflation forecast resulting from a run of the forecasting procedure at horizon h, calculated at time t.

Of course, apart from the RMSE, many other evaluation criteria are conceivable, such as the mean error or the absolute mean error. However, we will concentrate on the RMSE for several reasons: first, it is a common measure, second, it is also the criterion which is minimized in-sample in the estimation, third, the square of the inflation deviation is probably quite close to the loss function of a central bank, and fourth, to save space in the presentation of the results that follows.

3.1 Benefits of combining forecasts

In a first round of simulations, the benefits of combining forecasts are explored. The simulation is run ten times. The final forecast is calculated as the average forecast of the n best models, for $n = 1..10$. All other parameters of the procedure are held constant at values which are approximately those currently in use at the SNB: the lag length is 4 lags, the estimation period is 15 years, the evaluation interval is 20 quarters, and the selection horizon ranges from 7–10 quarters. At all forecast horizons, an increase in the number of models reduces the U-statistic, i.e. the RMSE of the forecast (Graph 3). Averaging the ten best models, the U-statistic at a horizon of 12 quarters can be reduced to 0.58, while the U-statistic of the single best model amounts to 0.77. Taking into account the strong benefits of combining forecasts, ten models will be averaged to obtain the inflation forecast in all further simulations.

Graph 3
Combining forecasts



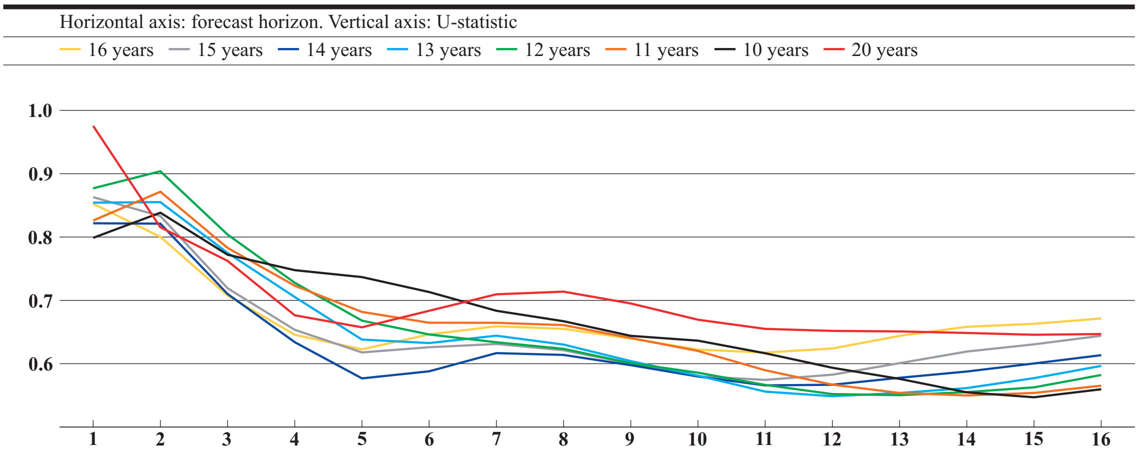
It is interesting to note that the U-statistic decreases with an increase in the forecast horizon. This, however, must not be interpreted as a decrease in the RMSE of the VAR forecasts. Rather, the decline in the U-statistic is due to the fact that the RMSE of the random walk forecasts increases faster than the RMSE of the model forecasts. Thus, it is only the relative performance of the model forecasts that increases compared to the random walk forecast. For short horizons, VAR models are only minimally better than a random walk forecast.

3.2 Optimizing the estimation period

The decision on the length of the estimation period is a trade-off. On the one hand, the longer the estimation period, the more precise the coefficient estimates and therefore probably also the forecasts. On the other hand, if the structure of the economy evolves over time, a short estimation period allows the model to adjust faster to structural changes and will therefore probably yield better results. Graph 4 shows the U-statistic for estimation periods from 10 to 20 years.[16] Too short an estimation period results in a poor forecasting performance. Estimation periods from 12 to 14 years perform about equally well, a further increase again worsens the forecast performance. An estimation period of 14 years is superior to the estimation period of 15 years at all forecast horizons. Therefore, in all further simulations, the estimation period will be set to 14 years.

16 All other parameters are unchanged from the previous simulation (the lag length is 4 quarters, the evaluation interval is 20 quarters, and the selection horizon ranges from 7 to 10 quarters).

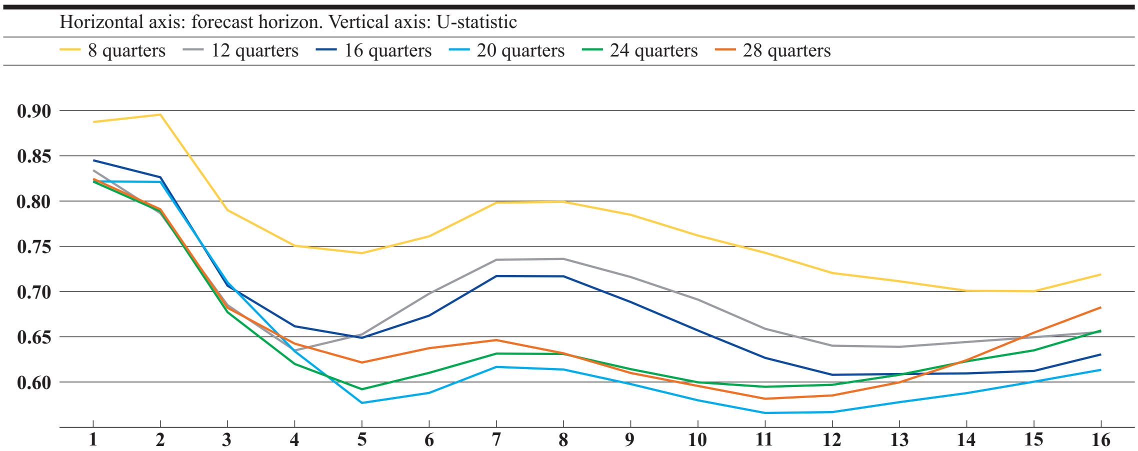
Graph 4
Varying the estimation period



3.3 Optimizing the evaluation period

There are arguments both for and against a long evaluation period. Extending the evaluation period allows for a more thorough evaluation and may thus improve the model choice and the forecast quality. However, the longer the evaluation period, the farther back it extends. Again, this might be a problem if the economic structure changes over time: the longer the evaluation period, the greater the chances that models will be selected which, although they performed well in the past, will probably perform less well for forecasting in the future, due to shifts in the structure of the economy which may have occurred in the meantime. As can be seen from Graph 5, for evaluation periods from 8 to 20 quarters, the marginal benefits from increasing the evaluation period are larger than the marginal costs.[17] For all further simulations, the evaluation period will be set to 20 quarters.

Graph 5
Varying the evaluation period



17 The sample period is very short, especially at the beginning of the simulation in 1987. Extending the evaluation period means reducing the estimation period. Therefore no evaluation periods longer than 28 quarters were considered.

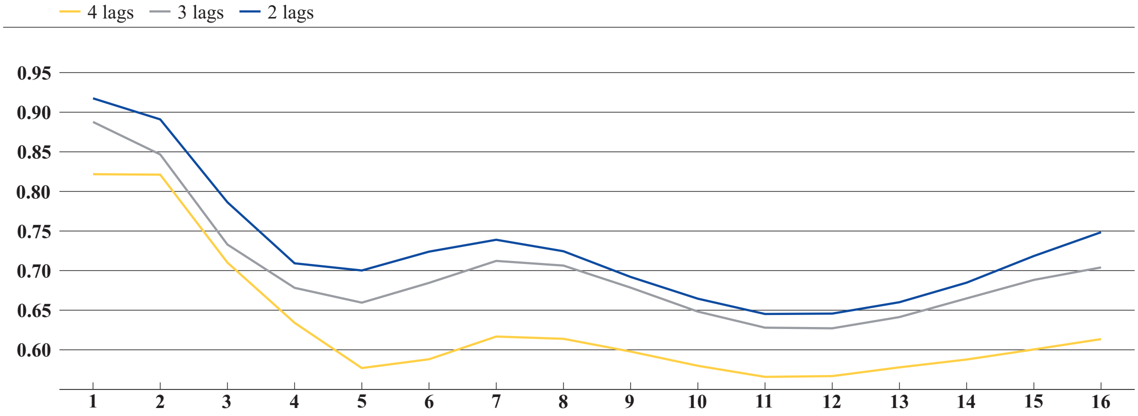
3.4 Optimizing the lag length

The decision on the number of lags is again a trade-off: on the one hand, a short lag length restricts potential intertemporal dynamics and may not remove all autocorrelation in the residuals. On the other hand, a long lag length reduces the precision of the coefficient estimates due to a reduction of degrees of freedom.

Our results show that four lags outperform two or three lags (Graph 6). However, this result is probably mainly due to the strong seasonalities in the data, probably in the CPI itself.[18] In all further simulations, the lag length is left at four lags.

Graph 6
Varying the lag length

Horizontal axis: forecast horizon. Vertical axis: U-statistic



3.5 Which variables predict inflation?

Table 1 shows the occurrence of the pool variables in the top 10 models, the average size of the models, and the shares of models in differences and in levels for a full run of the optimized procedure.

Occurrence of pool variables in top 10 models used for forecasting **Table 1**

Variable	Models in differences	Models in levels	All models
GDP	21%	9%	12%
M_1	43%	15%	22%
M_2	20%	33%	29%
M_3	10%	51%	41%
Outstanding loans	41%	57%	53%
Real exchange rate	28%	24%	25%
Nominal exchange rate	26%	22%	23%
10-year bond yield	0%	37%	28%
3-month interest rate	29%	47%	42%
Share of total	25%	75%	100%
Average size of models	3.2 variables	4.0 variables	3.7 variables

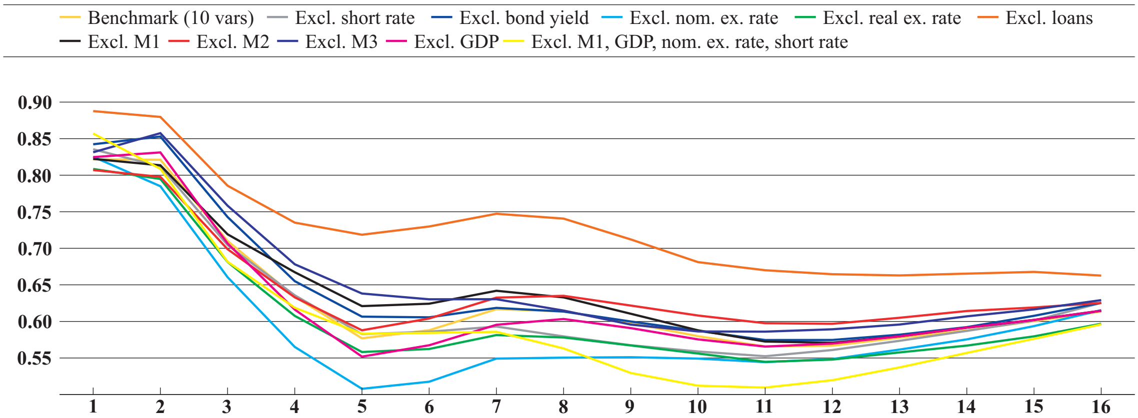
18 Indeed, when seasonally adjusted data is used for forecasting, two lags perform just as well as four lags.

Models in levels are three times more likely to be selected than models in first differences. The average model contains 3.7 variables (4.0 variables for models in levels, 3.2 variables for models in differences). Large models (four or five variables) are therefore not necessarily better suited to forecasting inflation than smaller ones. Indeed, it is not uncommon that models with only two variables and even univariate processes for inflation are among the selected top 10 models. For models in differences, M_1, bank loans and exchange rates seem to be the most informative variables. For models in levels, the main determinants of inflation are M_3 and outstanding loans. Somewhat surprisingly, GDP – which is usually considered an essential component of any macroeconomic model and the only measure of real activity in the pool of variables – is contained in only 12% of the best forecasting models. Thus it is the most rarely selected variable.

To obtain an alternative view on the importance of the variables, the full simulation is run nine times, each time excluding one of the pool variables, except the CPI. Almost without exception, the exclusion of variables improves the forecast compared to the complete benchmark pool (see Graph 7). This means that most variables do not contribute information not already contained in other pool variables. Instead they add additional noise to the forecast and thus increase the forecast error. Only the exclusion of outstanding loans and, to a lesser extent, the monetary aggregates and the bond yield, increase the forecast error. These variables therefore appear to contain information which is not contained in any other pool variable. Reducing the pool by excluding several variables can reduce the forecast error even more. It turns out that by excluding real GDP, M_1, the nominal exchange rate and the short-term interest rate, the forecast error can be reduced considerably for longer horizons. The remaining pool, containing the CPI, M_2, M_3, the real exchange rate, outstanding loans and the bond yield, will therefore serve as the new benchmark pool.

Graph 7
Excluding pool variables

Horizontal axis: forecast horizon. Vertical axis: U-statistic



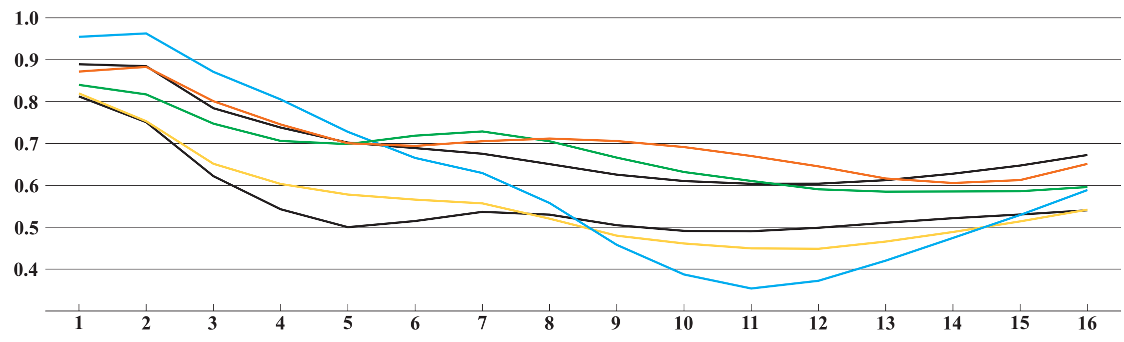
While excluding pool variables can shed light on the information content of the current pool variables, including new variables may shed light on the information content of variables not yet included in the pool. For this purpose, a wide selection of additional variables which might be associated with inflation are added to the new six-variable benchmark pool. The procedure is run with the six benchmark pool variables and an additional variable, and the U-statistic is calculated. Around thirty variables are thus investigated with regard to their additional explanatory power for inflation. These variables include several measures of real activity (exports, industrial production, capacity utilization, cement deliveries, household consumption expenditure, German GDP, the manpower index), asset prices (indices for rental prices for old and new apartments, for office space and for industrial space, price indices for single homes and owner-occupied apartments, a Swiss stock-market index, the oil price, the gold price and a commodity-price index), the Swiss consumer confidence index with all of its sub-components, some additional measures for the money supply (bank note circulation, time deposits) and the two sub-components of bank loans (mortgage loans and non-mortgage loans).

Most of these variables do not improve the forecast. On the contrary, they actually increase the forecast error. Graph 8 shows the two best and the two worst outcomes of this exercise. For longer forecast horizons, the inclusion of rents for new apartments as well as mortgage loans (a sub-component of total loans, which has already been included) significantly improves the forecast, while the inclusion of banknote circulation or the German real GDP significantly increases the forecast error. Most of the outcomes for all of the other variables lie between the confidence bounds and are not depicted.[19]

Graph 8
Including new pool variables

Horizontal axis: forecast horizon. Vertical axis: U-statistic

— 95% confidence bound — Incl. mortgage credit — Incl. new apartment rents — Incl. banknote circulation
— Incl. German GDP



19 A confidence interval is obtained by simulating the procedure 40 times, based on the 6-variable benchmark pool and including a seventh variable which is a random walk constructed from a normally distributed random variable with an autocorrelation of 0.9. Confidence bounds are then calculated by approximating a normal distribution to the distribution, for each horizon, of the U-statistics obtained in this manner.

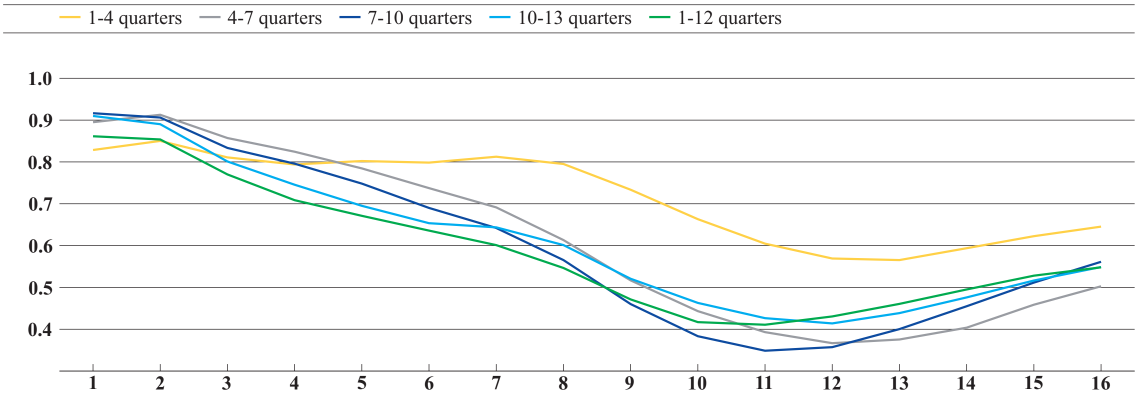
Mortgage loans and the rental index for new apartments seem to contain additional information for longer forecast horizons which is not captured by the 6 benchmark variables. They will therefore be added to the benchmark pool (thereby slightly reducing the forecast quality for shorter horizons). The underperformance of German real GDP and banknote circulation for inflation forecasting is rather unexpected; Swiss business cycles are usually assumed to be heavily influenced by the German business cycle, as Germany is the main buyer of Swiss exports. The banknote circulation is the main component of the monetary base, which served as the intermediate target for monetary policy in the period 1980–1999. If these variables were simply uninformative for inflation forecasting, their U-statistic should not be significantly worse than the U-statistic of the benchmark model. The significant underperformance could indicate structural changes in the relationship between inflation, on the one hand, and German GDP and the banknote circulation, on the other hand. Let us suppose that there were structural changes in the 1980s or in the 1990s. Thus, during the evaluation phase of the procedure, models are selected and estimated which conform to the old structure. However, in the forecasting period starting in 1987, the changing structure results in forecasts that, because of the changing coefficients, are even worse than they would have been if banknote circulation and the German GDP had been omitted. In the case of banknote circulation, at least, this explanation makes sense, as the instability of money demand in the 1990s was one of the main reasons for the introduction of the new monetary policy framework in 2000. However, it remains puzzling that none of the obvious candidates such as German GDP, as well as all measures for domestic real activity, such as Swiss GDP, exports, industrial production and household consumption, improve the inflation forecast.

3.6 Varying the selection horizon

Up to now the selection horizon, which indicates the forecast horizon for which models are evaluated, has been set to 7–10 quarters. However, it can be argued that different models should be used for inflation forecasting at different horizons. For example, it is reasonable to assume that, for short horizons, the exchange rate is an important determinant of inflation, while for longer horizons, monetary variables such as credit or money aggregates will determine inflation. If this argument is true, the selection horizon should influence the quality of the forecast for different forecast horizons. Models selected on the basis of their forecasting performance for a short selection horizon should perform better for a short forecast horizon, while models selected on the basis of their performance for a longer selection horizon should perform better at longer forecast horizons.

Graph 9
Varying the selection horizon

Horizontal axis: forecast horizon. Vertical axis: U-statistic
Selection horizon:



However, simulations with different selection horizons do not support this argument (Graph 9).[20] It is true that a selection horizon of 1–4 quarters appears to be somewhat better than other specifications for very short horizons, while performing much worse at longer horizons. All other specifications, however, perform about equally well at longer horizons. These results suggest that it is not necessary to match the selection horizon to the forecast horizon for forecasting inflation with VAR models.

Although varying the selection horizon does not improve the forecasts, it does have interesting consequences for the variables in the models selected for forecasting (see Table 2).

Occurrence of pool variables in top 10 models used for forecasting at different selection horizons **Table 2**

Models in differences

Variable	Selection horizon				
	1–4	4–7	7–10	10–13	1–12
M_2	36%	35%	38%	35%	26%
M_3	26%	35%	45%	44%	26%
Mortgage loans	29%	72%	76%	76%	73%
Real exchange rate	38%	34%	36%	39%	40%
10-year bond yield	49%	19%	6%	18%	12%
Rental prices of new apartments	42%	35%	46%	53%	44%
Models in first differences	69%	44%	33%	39%	39%
Average size of model	3.2	3.3	3.5	3.6	3.2

20 Based on the results from the previous section, rents of new apartments are added to the former 6 variable pool and outstanding bank loans are replaced with mortgage loans. The new 7 variable benchmark pool thus consists of the CPI, M_2, M_3, the real exchange rate, outstanding mortgage loans, the bond yield and rents of new apartments.

Models in levels

Variable	Selection horizon				
	1–4	4–7	7–10	10–13	1–12
M_2	24%	39%	31%	41%	43%
M_3	29%	54%	65%	55%	59%
Mortgage loans	32%	48%	37%	53%	41%
Real exchange rate	31%	25%	36%	32%	34%
10-year bond yield	31%	39%	45%	35%	45%
Rental prices of new apartments	42%	22%	50%	53%	26%
Models in levels	31%	56%	67%	61%	61%
Average size of model	2.9	3.3	3.6	3.7	3.5

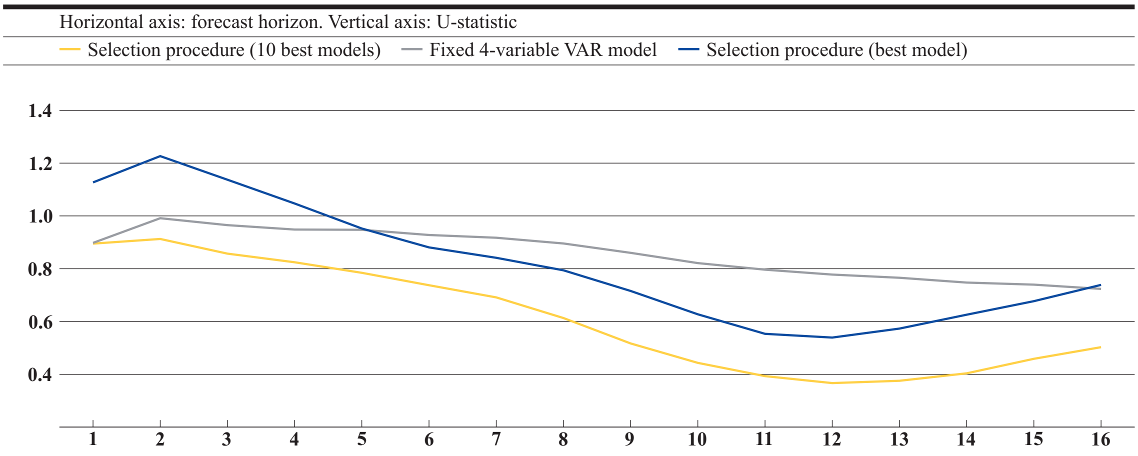
The more distant the selection horizon, the smaller the share of models specified in differences. The superiority of models in levels at long horizons suggests that there may indeed be cointegrating relationships among the CPI and other pool variables, which increase the forecast quality for more distant horizons. The average model size increases with the selection horizon. The longer the selection horizon, the higher the prevalence of monetary aggregates and mortgage loans. Differences in the bond yield do not appear to contain much information for longer horizons, as the bond yield is contained in fewer than 20% of models in differences for horizons longer than four quarters. The variable most often included in the models selected, and thus probably the most informative variable with respect to future inflation, is mortgage loans, followed by M_3 and rental prices for new apartments.

3.7 Comparison with a benchmark VAR model

Due to the quarterly re-evaluation, the inflation forecasting procedure is costly in terms of manpower and computing power. The question arises whether the results are worth the effort. To answer this question, the U-statistic of the optimized procedure is compared to the U-statistic of a four-variable VAR model containing the CPI, GDP, M_1 and the short-term rate (all in first differences).[21]

21 This four-variable VAR model is used at the SNB for conditional inflation forecasting and is described in Jordan et al. (2002).

Graph 10
Comparison with fixed four-variable VAR model



The forecasts of the optimized selection procedure yield a considerably lower U-statistic than the fixed VAR model at all horizons. The reduction in the forecast error is not only due to the diversification effect of combining forecasts. This can be seen from the U-statistic of the procedure using only the single best model for forecasting, which is below the U-statistic of the four-variable VAR model at longer horizons. Some of the reduction in the RMSE must therefore also be due to either the quarterly re-evaluation of the models, the use of levels of variables, or the larger underlying pool of variables. Thus, the significant improvement in forecasting performance justifies the application of the procedure.

3.8 Bayesian estimation

Unrestricted VAR models may suffer from overparameterization, as the number of available observations is often inadequate for a precise estimation of the coefficients. This can cause large out-of-sample forecast errors. One way to handle the problem of overparameterization is the use of a priori restrictions on the coefficients. This section explores whether simple Bayesian VAR estimation techniques are superior to unrestricted classical VAR estimation for forecasting inflation.

The prior distribution used here is a variant of the Minnesota Prior or Litterman Prior.[22] As a specification of a complete normal prior on a VAR would be almost intractable, because of its size, a general form for the prior involving only a few hyperparameters is used.[23] The idea of the Minnesota or Litterman Prior is that most economic time-series are well described as random walks, which implies that the prior mean of the first own lag is one, while the prior means of all

22 See Doan, Litterman and Sims (1983).

23 The RATS procedure SPECIFY is used for estimation. See also the RATS manual for further information on the priors and the estimation technique.

other coefficients are zero. In our case, a mean of one is assumed for the first own lags of the models in levels, while for the models in differences, a mean of zero for all lags is assumed. The prior distributions on the lags of the endogenous variables are assumed to be independent normal, while the priors for the constant are flat.

In order to simplify the specification of the prior, the standard deviation function *S(i, j, l)* of the prior distribution for lag *l* of variable *j* in equation *i* is restricted to the form:

$$S(i,j,l) = \frac{\{\gamma\, g(l)\, f(i,j)\}s_i}{s_j}\;;\; f(i,i) = g(l) = 1.0$$

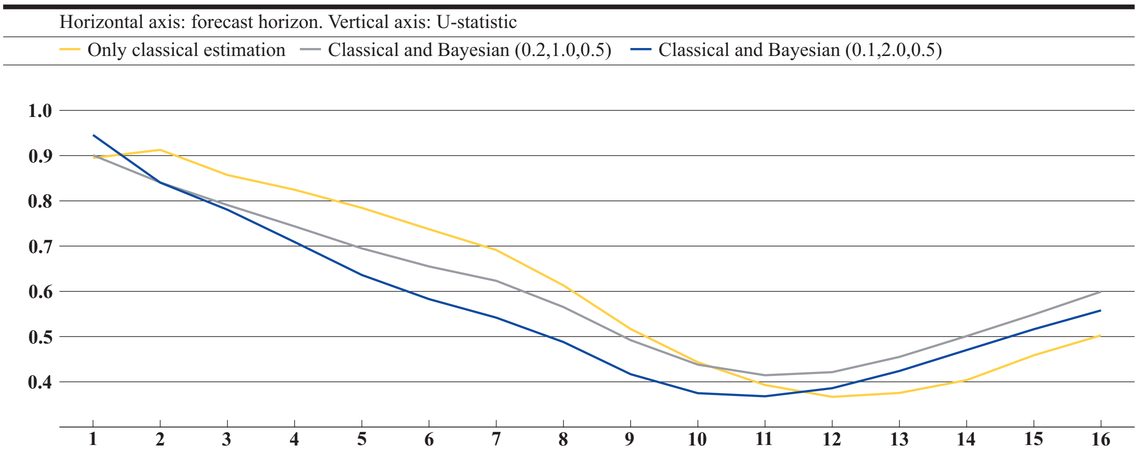
where s_j is the standard error of a univariate autoregression on equation *i*. The part in brackets is the tightness of the prior with respect to coefficient *i, j, l*. It is the product of three elements: the overall tightness γ, the tightness on lag *l* relative to lag 1 *[g(l)]*, and the tightness on variable *j* in equation *i* relative to variable *i [f(i, j)]*. Since γ directly controls the important own lags, they are forced too close to the prior mean if γ is set too small; in practice, it is usually set to something of the order of 0.1 or 0.2. The function *g(l)* defines how the standard deviation changes with increasing lags; in order to tighten up the prior with increasing lags *l*, a harmonic decay of the form: $g(l) = l^{-d}$, with *d* in the proximity of 1.0, is a common choice. The function *f(i, j)* yields 1.0 if *i = j* and ω otherwise, where ω = 0.5 is a common choice.

In order to find out whether Bayesian VAR models improve the inflation forecast, Bayesian models are added to the forecasting procedure, i.e. every possible combination of one to five variables out of the seven pool variables is estimated and evaluated, first as a classical VAR and then as a Bayesian VAR, both in levels and in differences. This is done for several variations of the three hyperparameters (γ, *d*, ω).

For a common combination of values for the hyperparameters (γ, *d*, ω) = (0.2, 1, 0.5), the RMSE of the forecast can be slightly improved for a forecast horizon up to 10 quarters, after which, however, Bayesian models increase the forecast error (Graph 11). Bayesian models represent 60% of all top 10 models selected for forecasting. By slightly modifying the hyperparameters (γ, *d*, ω), the forecasts can be improved even more. It turns out that by tightening the prior distribution to (γ, *d*, ω) = (0.1, 2, 0.5), a further improvement can be achieved for a forecast horizon of up to 11 quarters. Increasing the prior means for the first lags of the models in differences from 0 to values slightly below 1 (to account for possible autocorrelation) does not further improve the forecast.

The fact that Bayesian estimation improves the inflation forecast indicates that there may indeed be an overparameterization problem. Bayesian estimation with the hyperparameters (γ, *d*, ω) = (0.1, 2, 0.5) will therefore be added to the benchmark procedure for further simulations.

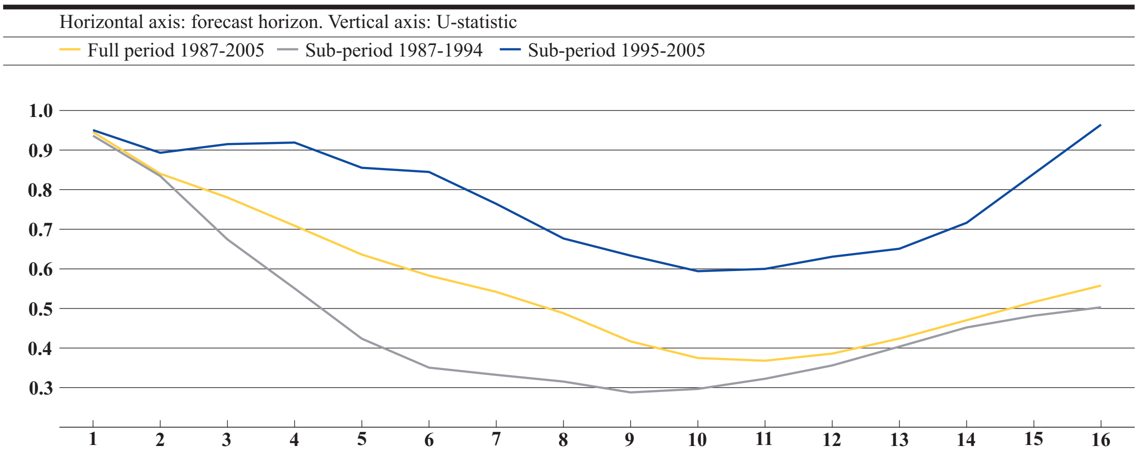
Graph 11
Bayesian estimation



3.9 Forecasts since 1987

As can be seen from Graph 2, the simulation period is divided into a high inflation period including a full inflationary cycle from 1987–1994 and a low inflation period since 1995. For a closer examination of the forecast performance in periods of high and low inflation, the simulation is run separately for the period from 1987 to 1994 and the period from 1995 to 2005 (Graph 12).

Graph 12
High and low inflation periods



As is evident from Graph 12, the VAR models perform very well, as compared with a random walk forecast, for the high-inflation environment from 1987 to 1994. For a horizon of 12 quarters, their RMSE is only one third of the RMSE of a random walk. However, in the subsequent period from 1995 to 2005, the U-statistic is much higher.

Graph 12 confirms that inflation forecasting in the low inflation environment that has prevailed since 1995 has been a difficult task. This probably holds true not only for VAR models, but for other types of models as well, given that VAR models can be viewed as a reduced form of a large variety of other linear macroeconomic models. Many of the insights gained from VAR modeling are therefore also valid for other models used for unconditional inflation forecasting.

The problem of inflation forecasting in a low inflation environment lies in the fact that, when inflation is low, inflation volatility is also low. The lower the inflation volatility, the higher the unpredictable share of volatility caused by idiosyncratic shocks to prices, and the lower the share of predictable volatility due to variations in macroeconomic fundamentals. Therefore, the lower the inflation rate is, the more difficult it becomes to outperform the random walk model.

The mediocre performance of unconditional inflation forecasting since 1995 is not least a consequence of the success of monetary policy in maintaining a low and stable inflation rate. Since 1994, Swiss consumer price inflation has been below 2%, which is consistent with the SNB definition of price stability. The SNB's success in the pursuit of price stability, however, comes at the cost of a deterioration in the quality of the unconditional inflation forecast, relative to a random walk model.

4. Summary and conclusions

By modifying a procedure for unconditional inflation forecasting used at the Swiss National Bank and simulating forecasts since 1987, several insights concerning the predictability of Swiss consumer price inflation, its determinants, and the quality of VAR inflation forecasts are obtained. The simulations show that by combining forecasts from different VAR models, a significant reduction in the root mean squared error (RMSE) can be obtained. The simulations also show that optimal forecasts are obtained from a pool of only seven monetary and price variables (the CPI, the real exchange rate, mortgage loans, M_2, M_3, the bond yield and an index of rents for new apartments), of which mortgage loans and M_3 are the most relevant. Adding more variables increases the forecast error, sometimes even significantly so. Surprisingly, neither real GDP nor any other indicator for real activity reduces the RMSE. Models estimated in levels of variables are superior to models estimated in first differences, especially at longer forecast horizons. Classic VAR models probably suffer from overparameterization, as Bayesian estimation with a Litterman prior can further reduce the RMSE at forecast horizons of up to three years.

The model selection procedure yields a considerably smaller RMSE than both a benchmark four-variable VAR model and a "naive" random walk forecast, for the period from 1987 to 2005. However, the reduction in the forecast error relates mainly to the period 1987–1994. From 1995 on, the VAR forecasts are only slightly better than a random walk "no change" forecast.

However, the worsening forecasting performance since 1995 must not be interpreted as a failure. It is rather a natural consequence of the success of monetary policy. If inflation is low and stable, it can reasonably well be characterized as a random walk, which makes it hard to forecast even with sophisticated models.

A low and stable inflation rate is no reason for a central bank to rest on its laurels and neglect economic analysis. On the contrary, low and stable inflation requires a special effort at careful and unbiased analysis of potential threats to price stability, because a low and stable inflation rate may tend to undermine critical analysis and the formation of rational expectations both inside and outside a central bank. There are three arguments that support this assertion. The first is the apparent failure of unconditional inflation forecasting. A poor forecasting performance will discourage serious efforts at model-based inflation forecasting. It may also lead to the replacement of certain models which constantly (but wrongly) predict changes in the inflation rate by other models which do not forecast changes easily. Second, not only the inflation forecasting models of central banks and other institutions, but also inflation expectations of individuals, adapt to the random walk-like behaviour of inflation. As rational expectations are more costly to build than adaptive expectations, people tend to abandon rational expectation formation and switch to cheaper adaptive expectations when inflation has been low and stable for a long enough period. Third, due to these changes in the formation of inflation expectations, the observed interactions between variables in a low-inflation environment may be different from the same interactions in a high-inflation environment. Parameter estimates based only on observations in a low-inflation environment may therefore be misleading when it comes to anticipating the consequences of large shocks, which might cause a switch in inflation expectations from adaptive back to rational expectations. All of these arguments could, in the worst case, lead to a situation where inflationary pressure builds up at a time when neither economic models nor inflation expectations are anticipating inflation. Central banks should therefore not be lulled into a false sense of security by low inflation expectations, but be open and self-critical when analyzing inflationary risks. For this purpose they need a thorough knowledge of the potential and limitations of their forecasting models. In this respect, we hope that our investigation will make a useful contribution.

References

Clemen, Robert T. 1989. Combining Forecasts: a Review and Annotated Bibliography. *International Journal of Forecasting* 5: 559–583.

Doan, Thomas, Robert Litterman and Christopher A. Sims. 1983. Forecasting and Conditional Projection Using Realistic Prior Distributions. NBER Working Paper No. 1202.

Jordan, Thomas J. and Michel Peytrignet. 2001. Die Inflationsprognose der Schweizerischen Nationalbank (The inflation forecast of the Swiss National Bank). Swiss National Bank *Quarterly Bulletin* 2: 54–61.

Jordan, Thomas J. and Marcel R. Savioz. 2003. Does it Make Sense to Combine Forecasts from VAR Models? An Empirical Analysis with Inflation Forecasts for Switzerland. Swiss National Bank *Quarterly Bulletin* 4: 80–93.

Jordan, Thomas J., Peter Kugler, Carlos Lenz, and Marcel R. Savioz. 2002. Inflationsprognosen mit vektorautoregressiven Modellen (Inflation forecasting with vector autoregressive models). Swiss National Bank *Quarterly Bulletin* 1: 40–66.

Makridakis, Spyros, Allen Andersen, Robert Carbone, Robert Fildes, Michele Hibon, Rudolf Lewandowski, Joseph Newton, Emanuel Parzen, and Robert Winkler. 1982. The Accuracy of Extrapolation (Time Series) Methods: Results of a Forecasting Competition. *Journal of Forecasting* 1: 111–153.

Makridakis, Spyros and Robert L. Winkler. 1983. Averages of Forecasts: Some Empirical Results. *Management Science* 29: 987–996.

Newbold, Paul and David I. Harvey. 2002. Forecast Combination and Ecompassing. In *A Companion to Economic Forecasting*, Michael P. Clements and David F. Hendry (eds), 268–283. Oxford: Blackwell Publishers.

Rich, Georg. 2000. Monetary Policy Without Central Bank Money: A Swiss Perspective. *International Finance* 3(3): 439–469.

Sims, Christopher A., James H. Stock, and Mark W. Watson. 1990. Inference in Linear Time Series Models with Some Unit Roots. *Econometrica* 58(1): 113–144.

Stalder, Peter. 2001. Ein ökonometrisches Makromodell für die Schweiz (An econometric macro-model for Switzerland). Swiss National Bank *Quarterly Bulletin* 2: 62–89.