Oesterreichische Nationalbank

Eurosystem

Workshops

Proceedings of OeNB Workshops

Macroeconomic Models and Forecasts for Austria

November 11 to 12, 2004

No. 5

Comment on "Forecasting Austrian Inflation"

Gerhard Rünstler[1]

European Central Bank

At least since Stock and Watson (1999), the use of composite indicators as derived from factor analysis has become a widely used approach to forecasting. Indeed, the empirical evidence tends to favour factor models as compared, for instance, to the standard VAR or ARIMA models. In this sense, the findings of the paper, which is competently done, stand largely in line with a number of recent studies on inflation forecasting conducted in the Eurosystem central banks.

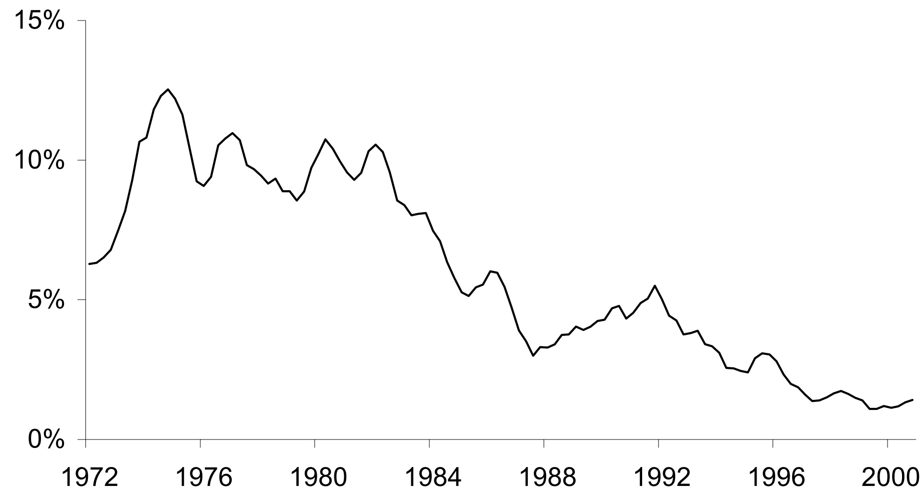
Still, a closer look at tables 2 and 4 reveals that the forecasts of the factor model still fall probably short of what we would like to achieve. According to the forecastability measure presented in table 4 it is only 2 among the 5 subcomponents of HICP, i.e. "Services" and "Processed food", for which the models produce informative 1-year ahead forecasts. As regards overall annual HICP inflation, the reduction in the standard error of the forecasts amounts to $1-\sqrt{1-0.31} \approx 17\%$.

However, these findings stand in line with the previous studies also in this respect. Inflation data have some undesirable properties and are simply very difficult to forecast. I shall demonstrate some of the difficulties with data for the euro area. I should however say beforehand that these issues may be somewhat less relevant for inflation in Austria, as the historical swings in the 1970s and 1980s are smaller compared to euro area data. However, apparently they apply to a number of euro area countries as well.

Chart 1 shows quarterly data for euro area quarterly and annual inflation, as measured by the consumer price index. The data range from 1970:1 to 2001:4.

[1] The opinions expressed in this contribution are those of the author and do not necessarily reflect the views of the European Central Bank.

Chart 1: Euro Area Annual CPI Inflation



Source: Rünstler (2002).

First of all, is inflation stationary? The graph does not suggest so. The data exhibit a clear downward trend and are far from fluctuating around a constant mean. Augmented Dickey-Fuller tests do not reject non-stationarity. This may stem from insufficient power of the tests. But there exist also unit root tests against stationarity as the null, such as the test due to Leybourne and McCabe (1984). This test clearly rejects stationarity.[2]

Note that a rejection of stationarity does not necessarily imply a unit root in the series. Arguably, such can hardly be a feature of inflation dynamics, as it would imply unbounded variance. However, a deterministic downward trend can hardly be regarded as an admissible model either, which becomes evident if one considers longer-term forecasts from such a model. Instead, some other approach to modelling the non-stationarity of inflation seems required. It has been proposed, for instance, to allow for infrequent jumps in the unconditional mean in time series models for inflation (e.g. Corvoisier and Mojon, 2005). However such infrequent deterministic jumps can be at best an approximation.

[2] The Leybourne-McCabe test statistics for quarterly inflation rates amounts to .278, which exceeds the 1% critical value of .216. The Augmented Dickey-Fuller test statistics, in turn, is found with -1.20, insignificant at the 10% level. Both tests exclude the possibility of a deterministic trend in inflation.

In any case, the level shifts in inflation may generate a good deal of parameter instability in time series models and thereby hamper their forecasting performance. Table 1 aims at demonstrating this for the euro area data (see also Rünstler, 2002). The table shows statistics of 4 and 8 quarters-ahead forecasts for inflation, based on various real activity and monetary indicators x_t. I consider two types of forecasts, i.e. conditional and leading-indicator forecasts.

The *conditional forecasts* for inflation π_{t+h} use the future values of the indicator $x_t, \ldots, x_{t+h}$. They are based on the ARIMAX equation

$$\Delta\pi_{t+1} = \mu + \theta_x(L)x_{t+1} + c(L)\Delta\pi_t + e_t \tag{1}$$

where L denotes the lag operator, $Lx_t = x_{t-1}$, and $\Delta = 1 - L$ is the difference operator. The forecasts for π_{t+h} are obtained from iterating this equation for $i = 1, \ldots, h$.

The *leading indicator (LI) forecasts* (Stock and Watson, 1999) predict π_{t+h} directly from the equation

$$\pi_{t+h} - \pi_t = \mu + \theta_x(L)x_t + c(L)\Delta\pi_t + e_t \tag{2}$$

The major difference between these two is that the LI forecasts use only the current (and past) values $\{x_s\}_{s=1}^{t}$ of the indicator, whereas the conditional forecasts also use future values, $\{x_s\}_{s=1}^{t+h}$. The latter should hence be more precise. While the LI forecast are more relevant in practice, conditional forecasts are of interest as a diagnostic instrument.

Crucially, the forecasts shown in Table 1 are either based on *full-sample* or on *recursive* parameter estimates.[3] In the latter case, parameters are re-estimated at each single point time. They hence, use only the information available to a forecaster in 'real time' and are therefore the relevant ones in practice. Naturally, recursive forecasts are more sensitive to parameter instability.

The table shows the root mean squared error (RMSE) of 4 and 8 quarters-ahead forecasts relative to the RMSE of a naive forecast.[4] A small value indicates good

[3] For full-sample forecasts, in turn, the parameters are estimated once over the entire sample. Lengths of lag polynomials c(L) and θ(L) have been determined from the Schwartz information criterion. Versions of equations (2) and (3) that use inflation in levels instead of first differences yield very similar results.

[4] The naive forecast for π_{t+h} amounts simply to the last observed value π_t.

forecasting performance, while a value of larger than one indicates that the forecast is uninformative (i.e. worse then the naive forecast). In addition, the table contains a test for Granger-causality of the indicator to inflation together with Andrews' (1993) test for stability of constant μ (with unknown breakpoint).

Table 1: Inflation Forecasts from Various Indicators

	GC test	Stability test (μ)	RRMSE Conditional (2) full-sample		RRMSE Conditional (2) recursive		RRMSE LI forecasts (3) recursive	
Forecast horizon			4	8	4	8	4	8
None							0.96	0.84
Capacity utilisation	**17.31	2.33	0.63	0.46	0.75	0.75	0.82	0.92
UR (level)		**16.15	0.92	1.02	1.39	1.81	0.89	0.81
UR (change)	2.20	3.35	0.86	0.91	1.30	1.67	0.83	0.76
GDP (growth)	**8.71	3.26	0.85	0.89	1.36	1.74	0.81	0.86
Long-term rate	**14.49	**8.07	0.72	0.67	1.31	1.76	0.72	0.79
Short-term rate	*7.88	**8.96	0.75	0.64	1.34	1.84	0.88	0.89
Money M3 growth	*8.59	**10.90	1.02	1.31	0.98	1.00	1.10	0.83

*Note: RRMSE denotes the root mean squared error of the forecasts relative to the one of the naive forecast. GC denotes the test for Granger causality of the indicator to inflation. Critical values for Andrews' (1993) stability test for constant μ are 6.05 and 7.51 for 10% and 5% significance levels, respectively * and ** denote significance at 10% and 5% levels, respectively. Estimation period starts in 1973Q1 with the exception of money and interest rates which start in 1981Q1. The forecast evaluation period ranges from 1991Q1 to 2000Q4.*

Source: Rünstler (2002).

The results of table 1 contain a few interesting features. First, a number of indicators, e.g. capacity utilisation, interest rates and to a lesser extent, GDP growth provide good full-sample forecasts.. When it comes to recursive forecasts however, most of the indicators perform worse than the naïve forecast. This strongly suggests parameter instability, which, in some cases, can be attributed to the constant μ as indicated by Andrews' (1993) test.

Second, and somewhat surprisingly, the recursive LI forecasts perform better than the conditional forecasts. This holds despite the smaller information set and the presence of parameter instabilities. Long-term interest rates and money M3 growth are perhaps the most striking examples. The test for Granger-causality of M3 is only significant at the 10% level and for both series, the instability of constant μ leads to values of the RMSE statistics of well above one. However, the LI forecasts show the indicators as those with the highest information content for future inflation, thereby turning the results from conditional forecasts on their head.

Overall, it seems difficult to find leading indicators for inflation that exhibit a stable relationship with the latter and this seems to stem from the pronounced low-frequency shifts in the level of inflation over most of the available data period. Appropriate ways to accounting for these shifts seems a precondition for obtaining reliable forecasts. The results of table 1 also suggest that leading indicator forecasts may be a rather unreliable guide to model selection, given a few technical issues that so far have not been investigated in detail.[5]

References

Andrews, D. W. K. (1993), Tests for parameter instability and structural change with unknown change point, Econometrica 6(4), 821–56.

Leybourne, S. J. and P. B. McCabe (1994), A consistent test for a unit root, Journal of Business and Economics Statistics 12(2), 157–66.

[5] One important issue is the overlapping nature of the forecasts obtained from LI regressions. With using π_{t+h} as the dependent variable, the forecasts are necessarily subject to an MA(h) structure. This hampers the precision of parameter estimates and perhaps also of the RMSE statistcs. Such effect may become particularly severe in case of highly persistent series.

Corvoisier, S. and B. Mojon, B. (2005), Breaks in the mean of inflation: how they happen and what to do with them, ECB Working Paper 451.

Rünstler, G. (2002), The information content of real-time output gap estimates, ECB Working Paper 184.

Stock, J. and M. Watson (1999), Forecasting inflation, Journal of Monetary Economics 44(2), 293–335.