Forecasting euro area inflation: Does aggregating price component forecasts improve forecast accuracy?[1] [2]

Kirstin Hubrich
European Central Bank[3]

January, 2002

Preliminary

Abstract

Monitoring and forecasting price developments in the euro area is essential in the light of the second pillar of the ECB's monetary policy strategy. This study systematically analyses whether the forecasting performance of euro area inflation models can be improved by aggregating forecasts of HICP subindices in comparison to forecasting total euro area HICP inflation directly. The comparison is carried out across different methodological approaches, including univariate and multivariate time series models, as well as across different forecast horizons and aggregates. The results indicate that under certain conditions it might be better to forecast euro area inflation directly instead of aggregating forecasts of HICP subcomponents.

JEL Codes: E31, E37, C53, C32
Keywords: euro area inflation, subindex forecast aggregation, linear time series models

[1]An earlier version of the paper has been circulated as Research Memorandum No. 661, July 2001, Dutch Central Bank.

[2]I thank Peter Vlaar for many helpful discussions and Peter van Els, Lutz Kilian, Helmut Luetkepohl as well as participants of a seminar at the Dutch Central Bank for useful comments. The views expressed in this paper are those of the author and do not necessarily reflect those of the European Central Bank.

[3]European Central Bank, Research Department, Kaiserstrasse 29, D-60311 Frankfurt am Main, Germany; e-mail: kirstin.hubrich@ecb.int

1 Introduction

The European Central Bank (ECB) is monitoring and forecasting prices with the objective to assess the developments regarding the second pillar of its monetary policy strategy. The ECB is publishing its inflation forecast for the euro area since December 2000. Therefore, further insights regarding the performance of different forecasting strategies for euro area inflation are of high importance.

In the context of forecasting euro area inflation the question arises to what extent the forecasting performance of different methodological approaches can be improved by aggregating forecasts of disaggregated subcomponents. (Dis-)aggregation can be considered in two directions: the aggregation of national forecasts for euro area countries and the aggregation of HICP subcomponent forecasts. The forecasting performance of aggregating country-specific forecasts in comparison with forecasts based on aggregated euro area wide data has been analysed in e.g. Marcellino, Stock & Watson (2000). Other studies focus on how to incorporate national information for forecasting euro area wide inflation as Angelini, Henry & Mestre (2001b) and Angelini, Henry & Mestre (2001a).

In contrast, the aim of this analysis is to compare the forecasting performance in terms of root mean square forecast error (RMSFE) of models forecasting HICP total directly with aggregating forecasts for HICP subcomponents based on (aggregated) euro area inflation data. Many studies in the literature of inflation forecasting focus on some specific aspects or methods of modelling and forecasting euro area or single country inflation. A systematic study of the forecasting properties of aggregated subindex forecasts of euro area HICP inflation in comparison with forecasting euro area inflation directly is - to the knowledge of the author - not available. The present paper aims to fill this gap.

The rationale behind aggregating forecasts of different HICP subindices instead of forecasting euro area inflation directly is that HICP subcomponents might be better modelled and forecasted than total HICP implying smaller forecast errors. Furthermore, forecast errors of subindices might cancel partly. On the other hand, the aggregated series might be more smooth than its subcomponents and therefore better to forecast.

To examine whether aggregating forecasts of HICP subindices is really better than forecasting total HICP directly, both approaches to forecasting euro area inflation are compared under three different aspects: 1. across different forecasting methods, 2. across different forecast horizons and 3. across different aggregates of the HICP subcomponents. More specifically, they are compared across three different forecasting methods, i.e. a univariate autoregressive model used as a benchmark, a vector autoregressive model and a vector error correction model. Univariate and multivariate linear time series models are chosen for the comparison since these are often used for forecasting inflation in Europe on a national or euro area wide level. Nonlinear time series models are not considered here due to the short time series. The three approaches are also compared across short-term and medium-term forecast horizons. And finally, they are compared for different aggregates of the HICP subcomponents, namely for total HICP, HICP excluding energy and unprocessed food prices ('core' inflation) and HICP excluding energy prices.

The paper is structured as follows: In the second section some asymptotic and small sample results from the literature regarding the relative forecasting performance of aggregated forecasts of time series subcomponents are discussed. The third section presents the data and the model selection procedure underlying the forecast comparison presented in this paper. The forecast accuracy of different (sub-)index models of HICP inflation and some aggregated HICP measures is discussed in section four, whereas section five draws some tentative conclusions from the analysis and points at aspects that need further research.

2 Forecasting contemporaneously aggregated time series: Some results from the literature

In empirical analysis the researcher often has to work with temporally or contemporaneously aggregated variables. Recently, there has been some new interest into the consequences of temporal aggregation for empirical analysis (see Marcellino (1999)). Furthermore, the effects of contemporaneous aggregation across national variables in the context of modelling euro area developments have received increasing interest. However, the focus of this study is on contemporaneous aggregation of subcomponents of time series variables and its empirical implications which has found rather limited attention.

The literature analysing contemporaneous aggregation of time series has focussed to a large extent on the effect on forecasting which will be also of interest in this study. Lütekpohl (1987) provides a broad study considering asymptotic results, Monte Carlo simulations and an empirical example. With respect to the asymptotic results, that study is based on Kohn (1982) and Lütekpohl (1984). Related issues and results are presented, for example, in Granger & Morris (1976), Rose (1977), Tiao & Guttman (1980) as well as in Wei & Abraham (1981). Furthermore, Granger (1990) provides a survey on aggregation of time series variables.

A contemporaneously aggregated variable can be defined as a variable consisting of the sum or the weighted sum of a number of different variables at time t. The contemporaneous aggregate can be written as

$$y_t = f_1 x_{1t} + f_2 x_{2t} + ... + f_n x_{nt}, \qquad t = 1, ..., T,$$

where x_{it} $(i = 1, ..., n)$ are the disaggregate variables, n is the number of disaggregate variables and f_i, $i = 1, ..., n$, are the aggregation weights. It is assumed that the aggregation weights are fixed and therefore do not change over time. Thus, y_t is assumed to be a linear transformation of the stochastic processes x_{it}. The forecasts will be indicated as $\hat{y}$ in the following.

The following discussion of the asymptotic results and Monte Carlo simulations draws to a large extent on Lütekpohl (1987) since he not only presents asymptotic results, but also focuses on a comparison of the forecasting performance of aggregated forecasts with forecasting the aggregate time series directly in small samples. He compares 3 possibilities for forecasting a time series variable y_t that is a weighted sum of a number of disaggregate variables: first, the aggregation of multivariate forecasts of the components x_{it} $(i = 1, ..., n)$ of the time series considered, referred to as $\hat{y}_m^{agg}$ in the following; second, the aggregation of univariate forecasts of the components x_{it} $(i = 1, ..., n)$, $\hat{y}_u^{agg}$; and, third, to forecast the aggregate time series directly using a univariate model, $\hat{y}_u^{total}$. The range of possible forecasting methods considered for forecasting euro area inflation in section 4 will be even broader. However, the three possibilities considered in Lütekpohl (1987) will serve as a benchmark for the evaluation of the findings.

Based on asymptotic theory he derives the following results. The first option, the aggregated multivariate forecasts of the component series, $\hat{y}_m^{agg}$, is optimal in the sense of generating a linear minimum mean squared error predictor, if the DGP is known. If the assumption of a known data generation process (DGP) is relaxed and it is assumed that the process order is finite and estimated using a consistent criterion, the aggregation of multivariate forecasts of the components, $\hat{y}_m^{agg}$, can actually be inferior to the other two options, i.e. aggregating the univariate forecasts of the subcomponents, $\hat{y}_u^{agg}$, or forecasting the aggregated time series directly using a univariate model, $\hat{y}_m^{total}$. The relative efficiency of the second and third possibility to forecast y_t depends on the generation process of the disaggregated components and the aggregation weights. In special cases, when the estimation variability is the major source of mean square error differences the aggregated multivariate forecast might be outperformed by aggregated univariate forecasts and forecasting the aggregated series directly.

Furthermore, Monte Carlo simulations presented in Lütekpohl (1987) to analyse the relative small sample properties of the aggregated time series largely confirm the asymptotic results. The univariate forecasts, $\hat{y}_u$, can outperform the multivariate forecasts, $\hat{y}_m$, possibly due to higher sampling variability in heavily parameterised multivariate systems. Furthermore, aggregating the univariate forecasts of the subcomponents $\hat{y}_u^{agg}$ can be better than forecasting the aggregate time series directly using a univariate approach, $\hat{y}_u^{total}$, if the residuals of the individual univariate component models are uncorrelated.

Overall, previous findings suggest that the relative performance of the two approaches to forecast aggregated time series, i.e. forecasting the aggregated series directly or aggregating the forecasts of its subcomponents, will depend on many different aspects of the analysis: the method used for forecasting, whether a univariate or a multivariate model is used, the forecast horizon, the dimension of the multivariate model, the degree of dependence of the component series, the order selection procedure employed for choosing the model on which the forecast is based and, finally, the sample size. Those aspects will be explored empirically for euro area inflation in the following.

3 Model choice for forecasting inflation based on euro area HICP (sub-)indices

In this study three different methods of forecasting euro area inflation are compared in terms of their forecasting performance: univariate autoregressive (AR) models, where the respective HICP (sub-)index is only explained by its past values, vector autoregressive (VAR) models and vector error correction models (VECMs). For each of those methods a model selection procedure is applied. The data used in the analysis for the variables possibly included in the models are presented in section 3.1. The model selection procedure is outlined in section 3.2.

3.1 Data

The sample period covered in this analysis is from 1990(1) to 2000(12). The data used are monthly, seasonally non-adjusted data.[4] The total HICP price index and the HICP subindices in logarithm as well as annual inflation of the respective index are presented in Figures 1 and 2. The breakdown of the HICP (sub-)indices is chosen in accordance with the data published in the ECB Monthly Bulletin. The notation will be the following: HICP unprocessed food will be denoted p^{uf}, HICP processed food p^{pf}, HICP industrial production p^{i}, HICP energy p^{e} and HICP services p^{s}. Furthermore, total HICP will be denoted p^{total}. Total HICP and HICP processed food display a relatively smooth upward trend. HICP industrial production and HICP services show a seasonal pattern that, at least for prices of industrial production is changing over time, whereas HICP unprocessed food and HICP energy exhibit a much more erratic development.

Further variables possibly included in the multivariate models are industrial production, y, and nominal money M3, m, as endogenous variables. Additionally, producer prices, p^{prod}, are included as a proxi for import prices, since import prices for the euro area are only available including intra-euro area trade. This variable is included as endogenous variable. Commodity prices (world market prices excl. energy) in euro, wmp^{exe}, oil prices in euro, p^{oil}, the exchange rate Euro/USD, e^{eusd}, and a short-term interest rate, i^{3m}, are imposed to enter the models as exogenous variables. This model set-up has been chosen because euro area national central banks incorporate a common set of assumptions regarding the future development of these variables in the context of their joint inflation projection. Nominal hourly wages, $wages$, are also treated as exogenous variable here.

All variables except for the interest rate are in logarithm. The graphs displaying these variables are presented in Appendix 5.

[4] The data are taken from the ECB and Eurostat, except for nominal hourly wages that are based on data from the BIS and DATASTREAM as well as on own calculations.

Certainly this variable list is not exhaustive. Other variables like capacity utilisation, a variable measuring unit labour costs or indicators displaying information on certain countries, like indicators from the US or large countries within the euro area, could be additionally included and might improve the forecasting performance of the models. However, the variables included in this study cover the most important factors influencing euro area inflation and, therefore, the main results of this study will most likely not be very sensitive towards the inclusion of further variables.

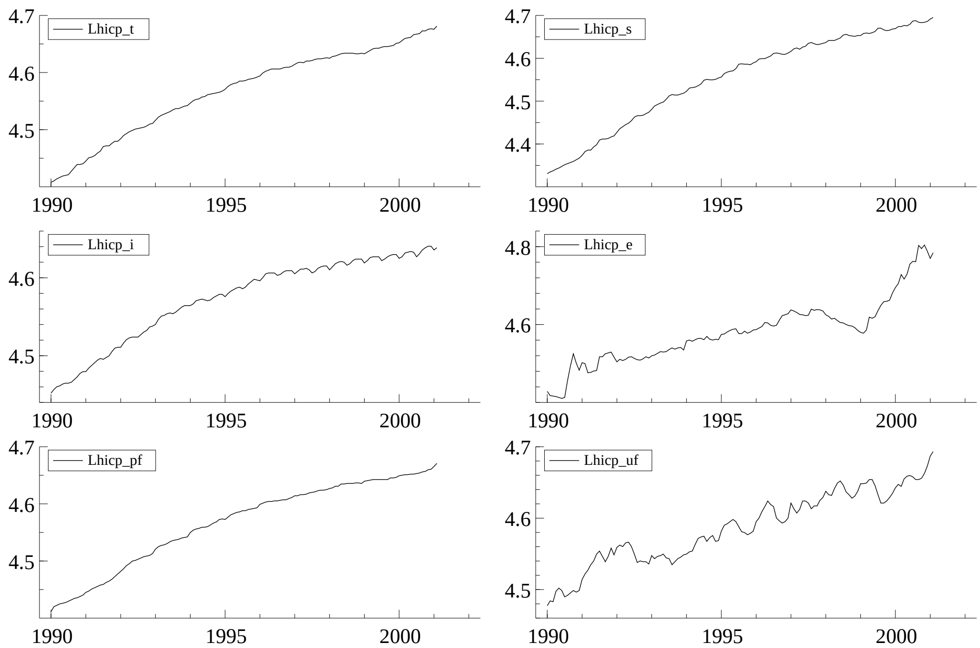

Figure 1: HICP (sub-)indices (in logarithm): HICP total, HICP industrial goods, HICP processed food, HICP services, HICP energy, HICP unprocessed food

3.2 Model selection

As mentioned above, three methodological approaches are compared in terms of their forecasting performance with respect to euro area inflation, including

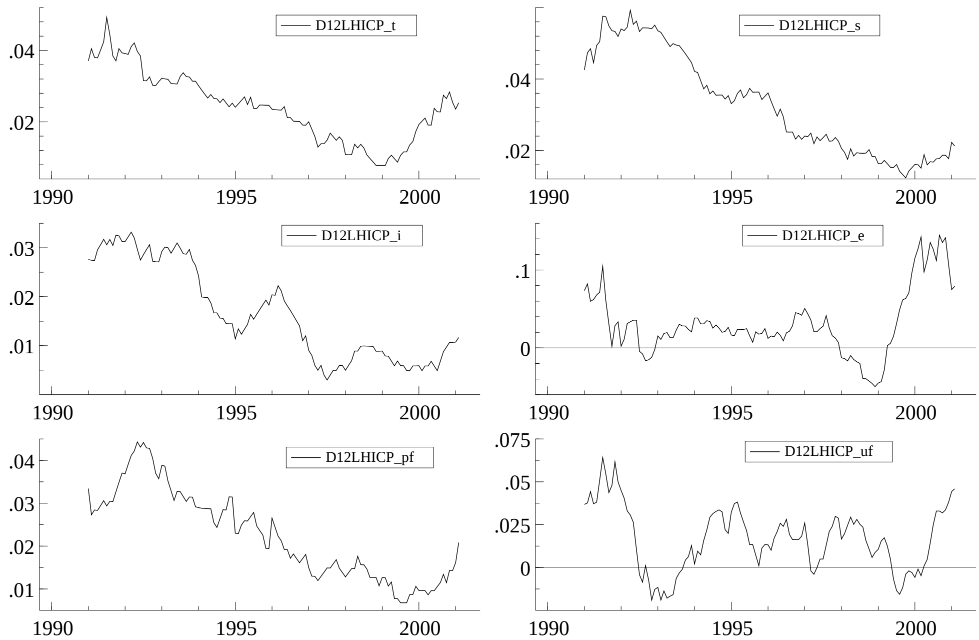

Figure 2: Annual HICP inflation, total and subindices

a univariate AR model, a VAR model and a VECM. The VECM model considered here allows only for one cointegration relation in each of the equations of the endogenous variables and the respective multivariate models are estimated as seemingly unrelated regressions, including the levels of the variables in each of the single equations.[5]

The model selection procedure is carried out on the basis of the out-of-sample RMSFE criterion for each of the three methodological approaches. This criterion is chosen because the in-sample fit of a model might not necessarily say much about the forecasting performance of a model. However, the choice of this criterion will be discussed later and the robustness of the results with respect to this choice will be analysed. For two of the subindices, HICP unprocessed food (p^{uf}) and HICP energy (p^{e}), a model in first differences is specified for each method considered since these variables do not exhibit a seasonal pattern. In contrast, the other subindices and HICP total are modelled in changes of the 12th differences, i.e. a seasonal unit root is imposed for these series due to a changing seasonal pattern. Each model includes seasonal dummies, even if a model in 12th differences is specified, to capture the remaining seasonal pattern. Furthermore, a constant is included in each of the models.

For each of the three methodological approaches - univariate AR, VAR and VECM - the best model is chosen based on the average out-of-sample performance 14 months ahead, where the estimation sample is starting in 1990(1) and ending in 1999(12). To provide some insight regarding the models chosen based on the out-of-sample RMSFE criterion, the resulting models based on the three different methods are presented for each of the subindices as well as total HICP for the estimation period 1990(1) to 1999(12) in Table 1. Regarding the models selected it is of interest to notice that money is included in most models - except for services HICP where money is not chosen for either of the multivariate models - indicating that money plays a role for inflation developments in the euro area.

4 Comparing the forecast performance

After choosing the best model within each methodological approach, the resulting forecasts are transformed to a comparable basis, annual inflation. The

[5]See Stock (1987) for the single equation approach.

Table 1: HICP (sub)indices: Model choice, sample 1990(1)-1999(12)

HICPindex	variables/ no. of lags/ Criterium	univariate AR	multivariate $(V)AR$	$VECM$
p^{uf}	exogenous	-	i^{3m}	wmp^{exe},
	endogenous	-	y, m, p^{prod}	m, p^{prod}
	lags	12	8	9
	EC term	-	-	p^{uf}, m, p^{prod}
p^{pf}	exogenous	-	i^{3m}, p^{oil}	wmp^{exe}, p^{oil}
	endogenous	-	y, m, p^{prod}	$y, m,$ p^{prod}
	lags	11	3	8
	EC term	-	-	p_{pf}, m
p^{i}	exogenous	-	i^{3m}	-
	endogenous	-	m, p^{prod}	$y,$ p^{prod}
	lags	11	4	8
	EC term	-	-	p^{i}, p^{prod}
p^{e}	exogenous	-	e^{eusd}, wages, i^{3m}	p^{oil}
	endogenous	-	y	$m,$ p^{prod}
	lags	12	8	7
	EC term	-	-	p^{e}, m, p^{oil}
p^{s}	exogenous	-	wmp^{exe}, i^{3m}	e^{eusd}, wmp^{exe}
	endogenous	-	-	$y,$ p^{prod}
	lags	10	12	10
	EC term	-	-	p^{s}, p^{prod}
p^{t}	exogenous	-	i^{3m}, p^{oil}	p^{oil},
	endogenous	-	m	$m,$ p^{prod}
	lags	12	10	8
	EC term	-	-	p^{t}, m, p^{oil}

Model choice: based on out-of-sample RMSFE, unprocessed food and energy prices: specified in first differences, other HICP subindices: specified as changes in annual inflation, EC term: error correction term forecasting performance of the different methods is then evaluated on the basis of the out-of-sample RMSFE for five subindices of HICP, that is annual inflation of unprocessed food ($\Delta_{12}p^{uf}$), of processed food ($\Delta_{12}p^{pf}$), of industrial production ($\Delta_{12}p^{i}$), of energy ($\Delta_{12}p^{e}$) and of services ($\Delta_{12}p^{s}$). The mean forecast error (MFE) is also presented and discussed for each model. Furthermore, the direct forecast of total annual HICP inflation based on aggregated data ($\Delta_{12}p^{total}$) is compared with the aggregated forecasts of the subindices ($\Delta_{12}p^{agg}$). Further aggregated component forecasts presented include core inflation ($\Delta_{12}p^{core}_{agg}$), i.e. total HICP inflation only including HICP of processed food, industrial production and services, and, furthermore, HICP inflation excluding energy ($\Delta_{12}p^{exe}_{agg}$). The weights used for the aggregation of the forecasts of the HICP subindices are the weights in the total of the index period 2000.

Since different forecast horizons might lead to different rankings of the forecasting methods, the comparison is carried out for short-term forecast horizons, 1 month and 6 months ahead, and a medium-term forecast horizon, 12 months ahead. The RMSFE evaluation is based on 'rolling' forecasts that involve an average of a number of 1-,6- and 12-horizon forecasts. The 'rolling' forecasts for 1-step ahead forecasts, starting with the forecast for 2000(1) based on the estimation sample 1990(1) to 1999(12), the second forecast is for 2000(2) based on the estimation sample up to 2000(1), etc., the 12th forecast for 2000(12) is then based on the estimation sample up to 2000(11). Therefore, the average RMSFE is calculated for 12 one-step-ahead forecasts based on 12 different estimation samples. Similarly, 6-period-ahead forecasts are carried out for 6 different forecast periods, first for 2000(1) to 2000(6), then for 2000(2) to 2000(7), etc. and 12 period ahead forecasts are carried out for 3 different forecast periods, i.e. for the periods 2000(1) to 2000(12), 2000(2) to 2001(1) and 2000(3) to 2001(2). Since only three 12-step-ahead forecasts are considered due to the short time series for euro area HICP, the respective results have to be interpreted with some caution. The results of the forecast performance comparison in terms of RMSFE and MFE are presented in Tables 2 to 7.

4.1 Relative forcast performance: Aggregation versus Disaggregation in forecasting euro area inflation

The RMSFEs results for the univariate AR model (see Table 2 to 4, first column) show that in this case forecasting total HICP p^{total} directly on average outperforms aggregating the forecasts of the subindices, p^{total}_{agg}, for a forecast horizon of $h = 1$. However, this result is not that unambiguous since p^{total} does only outperform p^{total}_{agg} in 6 out of 12 periods. For the longer forecast horizons 6 and 12, however, the aggregated forecast clearly performs better than forecasting total HICP directly.

The theoretical and small sample results reviewed in Section 2 suggest that the correlation between the subindices influences the relative forecasting performance between the aggregated and direct forecast. Therefore, the correlation matrix of the residuals of the univariate AR model for the different subcomponents of HICP for the sample period 1990(1) to 1999(12) is presented in Table 8. The correlations are relatively low, at most between 0.10 and 0.20 for the correlation between p^i and p^s with p^{uf} and p^{pf}. These low correlations might explain why aggregating the univariate forecasts performs better than forecasting total HICP directly with a univariate model for $h = 6$ and $h = 12$.

In contrast to the results for the univariate model, it is found for the VAR and the VECM across almost all forecast horizons considered that the forecast of total HICP p^{total} performs better on average than the aggregate forcast p^{total}_{agg} of the HICP subindices in terms of RMSFE (see Tables 2 to 4). The same is found for the mean forecast errors (see Tables 5 to 7). However, for the forecast horizon h=1 (Table 2) the forecast performance of p^{total}_{agg} and p^{total} is rather similar in terms of RMSFE, with p^{total}_{agg} performing slightly better for the VAR, whereas p^{total} performs slightly better for the VECM. If the different one-step-ahead forecasts are considered separately, p^{total} performs better for 7 out of 12 periods. This finding holds for both the VAR and the VECM.

Comparing the different forecasting methods, the VECM performs better for p^{total} and p^{total}_{agg} on average in terms of RMSFE than the VAR and the univariate AR model over all forecast horizons (see Tables 2 to 4) except for p^{total}_{agg} and forecast horizon $h = 12$, where the VAR performs better. The same can be found for the MFE (see Tables 5 to 7): The VECM also exhibits a smaller bias than the other methods for p^{total}_{agg} and p^{total}.

These results are in principle in line with Lütekpohl (1987) who finds on the basis of asymptotic results and Monte Carlo simulation as well as in an empirical example that the aggregated forecasts of the multivariate models outperform the aggregated forecasts of the univariate model which in turn outperform the univariate model forecasting total inflation directly. However, he does not consider forecasting total HICP inflation directly with a multivariate model. Furthermore, he does not consider cointegrated models but only differenced multivariate models.[6]

The finding of the good performance of forecasts of p^{total} in comparison with p^{total}_{agg} for multivariate models can have different reasons which are difficult to pinpoint from the empirical results. First of all, if high-dimensional multivariate models are used to forecast the subindices, aggregating these forecasts might worsen the forecast performance. In this context, it should be noted that the model selection procedure used here includes all lags up to the maximum lag chosen by the model selection criterion. One possility to improve the forecast performance might be parameter reduction by excluding insignificant lags and allowing for different lag lengths for different variables.

The results of this study regarding the relative forecast performance of forecasting euro area inflation directly and aggregating the subindex forecasts seem not to depend on the variables chosen by the model selection procedure, since the VAR and the VECM arrive at similar results in that respect although they represent different methodological approaches and different variables are chosen by the model selection procedure for the respective method.

It might be argued that the relatively bad performance of the aggregated multivariate forecast of euro area inflation is due to the fact that some subcomponents, namely p^{uf} and p^{e}, are not very well forecasted across all methods and forecast horizons. Then the better forecast performance for the other subindices can not compensate for the bad forecast performance, especially for p^{e}. However, the bad forecasting performance of the models for p^{e} and p^{uf} can also be observed for the univariate model. Therefore, this is not a plausible explanation for the results. An alternative explanation might be that total HICP resulting from aggregating the HICP subindices is a very smooth series that is easier to forecast than a very volatile series. However, since the results for the univariate model show that p^{total}_{agg} performs better this argument does not hold.

[6] Some results for the effects of contemporaneous aggregation on cointegration have been derived by Gonzalo (1992), but no general rules emerge.

The good performance of the VECM for p^{total} might be explained by allowing for a cointegration relation, especially between the HICP price indices and oil prices, producer prices and money. On the other hand, the good performance of the VAR for p^{total} is more difficult to explain. One reason can be found in the very low bias that both the VECM and VAR exhibit for p^{total}. Examining the MFE (Table 5 to 7) shows that for the VECM forecast error indeed seem to cancel in p^{total}_{agg}, especially for $h = 1$. However, for higher forecast horizons p^{total} nevertheless performs better.

Additionally, the forecast performance of p^{core}_{agg} is considered since it is often used as an indicator for price developments for policy analysis. It turns out that it always performs better than p^{total} and p^{total}_{agg} in terms of RMSFE except in case of the VECM and $h = 12$. In contrast to total aggregated HICP inflation, for p^{core}_{agg} the VAR is always better than the VECM in terms of RMSFE. However, the differences between the VAR and the VECM are not very large. p^{exe}_{agg} performs somewhat worse than core inflation, but better than total HICP inflation for forecast horizons $h = 1$ and $h = 6$.

Furthermore, the forecast performance for the HICP subindices is considered. The univariate AR model is almost outperformed by the VAR and the VECM in terms of RMSFE for all subindices. However, for some of the subindices p^{uf} and p^{pf}, the univariate model displays a lower bias than the other two methods, especially for lower forecast horizons. In general it can be seen that with increasing forecast horizon the RMSFE and the MFE increase as would have been expected.

4.2 Some robustness analyses

The robustness of the results can be analysed by modifying the model selection procedure. One might argue that more rigorous testing and modelling of possible structural breaks should be carried out at the stage of the model selection procedure. A broader range of misspecification and structural break tests might be included into the model selection procedure. In a first step, misspecification tests have been carried out for the univariate models for the period 1990(1) to 1999(12) indicating no serious problem with autocorrelation. Furthermore, structural breaks were mainly detected for energy HICP inflation models.

Further research should also focus on the impact of a different model selection criterion, for example the SC (Schwarz) criterion. First results indicate that the main findings regarding the relative forecast performance

Table 2: **Forecasts of HICP (sub-) indices: Forecasting performance, annual inflation, forecast horizons: 1**

	univariate AR	multivariate (V)AR	VECM
	RMSFE(1)	RMSFE(1)	RMSFE(1)
p^{uf}	0.0032	0.0032	0.0035
p^{pf}	0.0013	0.0007	0.0009
p^{i}	0.0008	0.0004	0.0003
p^{e}	0.0190	0.0127	0.0114
p^{s}	0.0012	0.0009	0.0010
p^{total}	0.00181	0.00108	0.00088
p^{total}_{agg}	0.00185	0.00106	0.00090
p^{core}_{agg}	0.0008	0.0004	0.0004
p^{exe}_{agg}	0.0010	0.0005	0.0006

Table 3: **Forecasts of HICP (sub-)indices: Forecasting performance, annual inflation, forecast horizons: 6**

	univariate AR	multivariate (V)AR	VECM
	RMSFE(6)	RMSFE(6)	RMSFE(6)
p^{uf}	0.0165	0.0070	0.0095
p^{pf}	0.0021	0.0018	0.0029
p^{i}	0.0016	0.0013	0.0014
p^{e}	0.0371	0.0269	0.0225
p^{s}	0.0023	0.0013	0.0018
p^{total}	0.0038	0.0024	0.0015
p^{total}_{agg}	0.0036	0.0026	0.0016
p^{core}_{agg}	0.0018	0.0006	0.0014
p^{exe}_{agg}	0.0029	0.0008	0.0010

Table 4: **Forecasts of HICP (sub)indices: Forecasting performance, annual inflation, forecast horizons: 12**

	univariate AR	multivariate (V)AR	VECM
	RMSFE(12)	RMSFE(12)	RMSFE(12)
p^{uf}	0.0314	0.0042	0.0164
p^{pf}	0.0046	0.0034	0.0030
p^{i}	0.0040	0.0016	0.0025
p^{e}	0.0437	0.0300	0.0316
p^{s}	0.0050	0.0018	0.0031
p^{total}	0,0054	0.0020	0.0016
p^{total}_{agg}	0,0038	0.0022	0.0029
p^{core}_{agg}	0.0044	0.0013	0.0018
p^{exe}_{agg}	0.0068	0.0012	0.0016

Table 5: **Forecasts of HICP (sub-)indices: Mean forecast errors, annual inflation, forecast horizons: 1**

	univariate AR	multivariate (V)AR	VECM
	MFE(1)	MFE(1)	MFE(1)
p^{uf}	0.0024	0.0010	0.0010
p^{pf}	0.0007	-0.0001	-0.0003
p^{i}	0.0004	0.0001	-0.00005
p^{e}	-0.0027	0.0055	0.0030
p^{s}	0.0006	0.0001	-0.0007
p^{total}	0.00044	0.0003	0.0002
p^{total}_{agg}	0.00040	0.0006	0.00001
p^{core}_{agg}	0.0005	0.0001	-0.0004
p^{exe}_{agg}	0.0007	0.0001	-0.0003

Table 6: **Forecasts of HICP (sub-)indices: Mean forecast errors, annual inflation, forecast horizons: 6**

	univariate AR	multivariate (V)AR	VECM
	MFE(6)	MFE(6)	MFE(6)
p^{uf}	0.0146	0.0042	0.0062
p^{pf}	0.0009	0.0009	-0.0021
p^{i}	0.0007	-0.0010	-0.0009
p^{e}	-0.0088	0.0223	0.0165
p^{s}	0.0022	0.0005	-0.0011
p^{total}	0.0027	0.0008	0.0003
p^{total}_{agg}	0.0016	0.0022	0.0010
p^{core}_{agg}	0.0014	-0.00002	-0.0012
p^{exe}_{agg}	0.0026	0.0004	-0.0005

Table 7: **Forecasts of HICP (sub-)indices: Mean forecast errors, annual inflation, forecast horizons: 12**

	univariate AR	multivariate (V)AR	VECM
	MFE(12)	MFE(12)	MFE(12)
p^{uf}	0.0277	0.0001	0.0126
p^{pf}	0.0030	0.0016	-0.0023
p^{i}	0.0028	-0.0008	-0.0021
p^{e}	-0.0308	0.0340	0.0282
p^{s}	0.0043	0.0007	-0.0004
p^{total}	0.0045	0.0007	0.0004
p^{total}_{agg}	0.0025	0.0031	0.0024
p^{core}_{agg}	0.0035	0.0002	-0.0014
p^{exe}_{agg}	0.0057	0.0002	-0.0001

Table 8: **Correlation Matrix of the residuals of the univariate AR of the different subcomponents of HICP, 1990(1) - 1999(12)**

	p_{uf}	p_{pf}	p_i	p_e	p_s
p_{uf}	1.0000				
p_{pf}	-0.0036102	1.0000			
p_i	0.11405	0.17502	1.0000		
p_e	-0.068229	-0.029844	-0.025862	1.0000	
p_s	0.14040	0.20147	0.075927	0.094155	1.0000

of the aggregated forecasts of the subindices and forecasting the aggregated series directly do not change with the choice of the model selection criterion. For an additional robustness check, a different model selection strategy will be employed, i.e. an automatic general-to-specific modelling procedure (Hendry & Krolzig, 2001).

5 Tentative conclusions and further research

The results presented in the previous section, although based on rather short samples of the 1990s, indicate that for the univariate AR model the forecasting accuracy of the aggregated forecasts of HICP subindices for the euro area is higher than the forecasting accuracy of a model for total HICP. Such a result can be explained by the relatively low correlation between the subcomponents. In the case of multivariate models in contrast, forecasting total euro area HICP directly tends to outperform aggregating the forecasts of the HICP subindices in terms of forecast accuracy.

However, this finding depends on the forecast horizon. For longer forecast horizons the finding is robust whereas for a one period ahead forecast the results are somewhat mixed. A possible explanation for the better predictability of total HICP inflation is the fact that the series is smooth leading to a lower bias of the forecast.

Regarding the performance across different aggregates of HICP subcomponents, 'core' inflation - that is HICP inflation excluding the most volatile HICP components unprocessed food and energy - is found to be better predictable than total HICP inflation and also better predictable than total HICP inflation excluding energy.

The finding that multivariate models outperform univariate models in forecasting euro area HICP and HICP subcomponents at various horizons implies that the inclusion of macroeconomic variables based on economic considerations seems to improve the forecasting performance. Furthermore, it is of interest to notice that money is included in the majority of models selected by the model selection procedure, indicating that money plays a role for inflation developments in the euro area.

Further research needs to be carried out in several directions. First, the robustness of the results towards a change in the model selection procedure has to be further examined. including other model selection criteria and procedures. Additionally, the conditions under which the VECMs outperform the other models considered in this study might be further explored. A further extension of the study would be a broader investigation into the performance of aggregating HICP subcomponent forecasts on the basis of country data.

However, the results presented here so far suggest that aggregated price component forecasts have to be considered with some caution when forecasting overall inflation in the euro area since they might be less accurate than directly forecasting total HICP inflation.

Appendix A Data

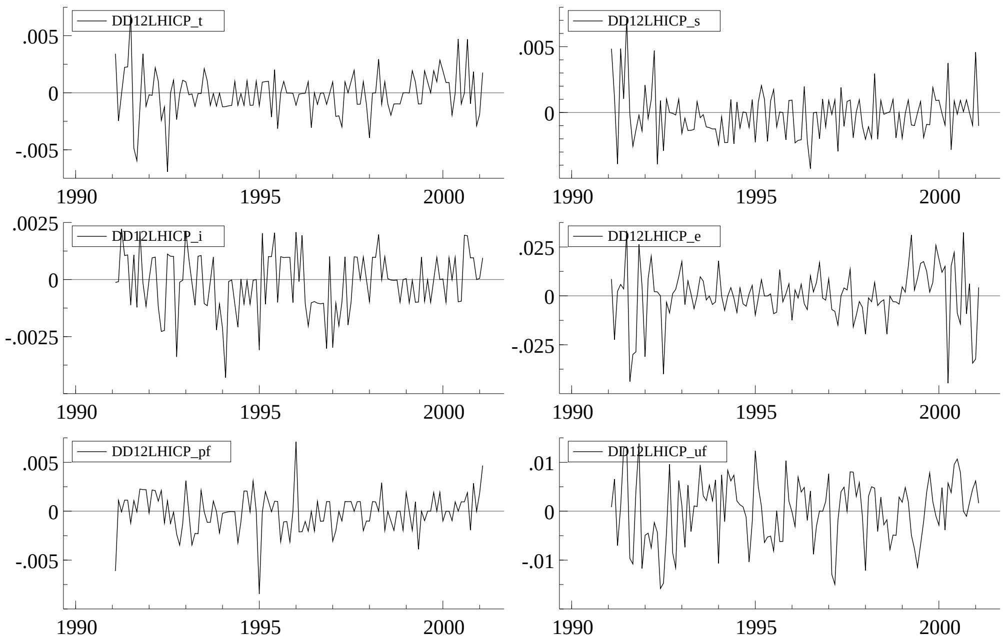

Figure 3: Changes in annual HICP inflation, total and subindices

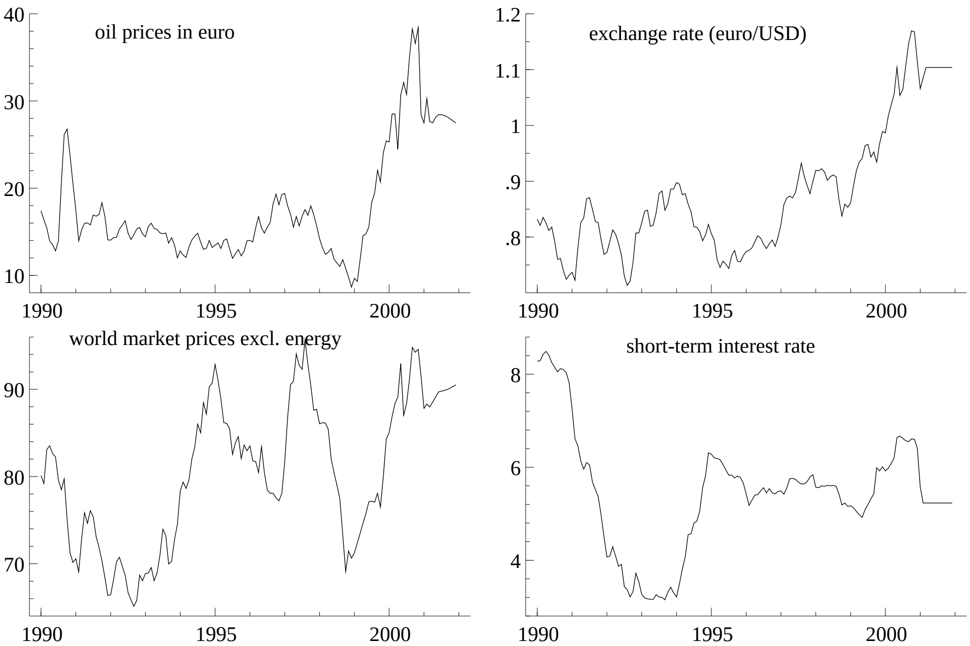

Figure 4: Oil prices, world market prices, euro/USD exchange rate, interest rate (in logarithm except for interest rates)

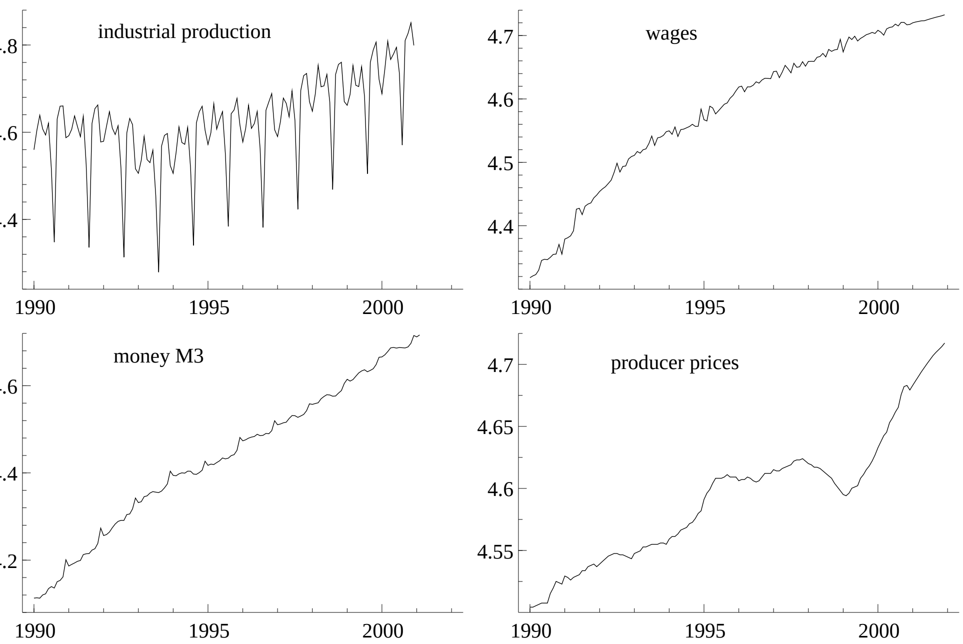

Figure 5: Industrial production, nominal hourly wages, money M3, producer prices (in logarithm except for interest rates)

References

Angelini, E., Henry, J. & Mestre, R. (2001a). Diffusion index-based inflation forecasts for the euro area, *Working Paper 61*, European Central Bank.

Angelini, E., Henry, J. & Mestre, R. (2001b). A multi-country trend indicator for euro area average inflation: Computation and properties, *Working Paper 60*, European Central Bank.

Gonzalo, J. (1992). Cointegration and aggregation, *Working Paper 89-30R*, University of California, San Diego Department of Economics.

Granger, C. W. J. (1990). Aggregation of time-series variables: A survey, *in* T. Barker & M. H. Pesaran (eds), *Disaggregation in econometric modlling*, Routledge, London and New York, pp. 17–34.

Granger, C. W. J. & Morris, M. J. (1976). Time series modlling and interpretation, *Journal of Royal Statistical Society* **139**: 246–257.

Hendry, D. F. & Krolzig, H.-M. (2001). New developments in automatic general-to-specific modelling, *in* B. P. Stigum (ed.), *Econometrics and the Philosophy of Economics, forthcoming.*

Kohn, R. (1982). When is an aggregate of a time series efficiently forecast by its past?, *Journal of Econometrics* (18): 337–349.

Lütekpohl, H. (1984). Linear transformations of vector ARMA processes, *Journal of Econometrics* (26): 283–293.

Lütekpohl, H. (1987). *Forecasting Aggregated Vector ARMA Processes*, Springer-Verlag.

Marcellino, M. (1999). Some consequences of temporal aggregation in empirical analysis, *American Statistical Association* **17**(1): 129–136.

Marcellino, M., Stock, J. H. & Watson, M. W. (2000). Macroeconomic forecasting in the euro area: Country specific versus area-wide information, mimeo.

Rose, D. E. (1977). Forecasting aggregates of independent ARIMA processes, *Journal of Econometrics* (5): 323–345.

Stock, J. H. (1987). Asymptotic properties of least-squares estimators of cointegrating vectors, *Econometrica* **55**: 1035–1056.

Tiao, G. C. & Guttman, I. (1980). Forecasting contemporal aggregates of multivple time series, *Journal of Econometrics* (12): 219–230.

Wei, W. W. S. & Abraham, B. (1981). Forecasting contemporal time series aggregates, *Communications in Statistics - Theory and Methods* (A10): 1335–1344.