Forecasting inflation through a bottom-up approach: How bottom is bottom?*

Cláudia Duarte — Banco de Portugal

António Rua† — Banco de Portugal

January 2007

Abstract

The aim of this paper is to assess inflation forecast acurracy over the short-term horizon, using Consumer Price Index (CPI) disaggregated data, through a bottom-up approach. That is, aggregating forecasts is compared with aggregate forecasting. A new dimension to the question of to bottom-up or not is introduced by considering different levels of data disaggregation, namely a higher disaggregation level than the one considered up to now. This raises modelling issues that one has to cope with. In particular, it is suggested the use of a new strand of models, the Factor-Augmented SARIMA models. Considering as case-study the Portuguese one, we find an inverse relationship between the forecast horizon and the amount of information underlying the forecast, when minimizing the RMSFE.

Keywords: Inflation forecasting; Bottom-up; Factor-Augmented SARIMA; Dynamic common factors.

JEL classification: C22, C32, C43, C53, E31, E37

---

*Forthcoming Economic Modelling

†Corresponding author: Economic Research Department, Banco de Portugal, Av. Almirante Reis no 71, 1150-165 Lisboa, Portugal. Email: antonio.rua@bportugal.pt

1 Introduction

In many countries, and especially in the European Union, the primary objective of monetary policy is price stability. For example, according to the European Central Bank (ECB), price stability is defined as "*a year-on-year increase in the Harmonised Index of Consumer Prices (HICP) for the euro area of below*" but "*close to 2% over the medium term*" (ECB (2003a)). Therefore, to ensure that this objective is attained, the monetary authority needs to be constantly monitoring and forecasting the evolution of prices. The existence of lags, caused by transmission mechanisms, and economic shocks, which endanger price stability, explains why inflation forecasting is regarded as a crucial tool for conducting monetary policy. Actually, Jean-Claude Trichet (ECB (2003b)) said that inflation forecasts are "*useful, even indispensable, ingredients of monetary policy strategy*".

Thus, forecasting inflation for the euro area as a whole is very important for monetary policy purposes. However, it is also relevant to forecast country level inflation. First of all, country level inflation forecasting contributes to a better understanding of the different transmission mechanisms in each country. Furthermore, Marcellino, Stock and Watson (2003) found evidence that forecasting inflation at the country level and then aggregating the forecasts increases accuracy against forecasting at the aggregate level. Finally, the usefulness of inflation forecasts is not restricted to monetary policy purposes. For example, improving short-term forecasting is also important because it allows to better anticipate near-term future developments, which enhances evaluation of current trends and improves the conjunctural assessment of central bankers. Moreover, inflation forecasts are also quite relevant in other areas, such as fiscal policy, wage bargaining and financial markets.

In this context, it is crucial to have more and more accurate inflation forecasts. One possible way of improving forecast accuracy is by considering more data, in particular, disaggregated one. Several studies have focused on whether using this kind of information increases forecasting accuracy. Actually, two strands of literature about the role of disaggregated information on improving forecast accuracy can be distinguished. One strand focus on using the forecasts of disaggregates to obtain the forecast for the aggregate – the bottom-up approach. Let y_t be the variable of interest for forecasting. Suppose that this variable can be decomposed in n subcomponents, y_{it} $(i = 1, ..., n)$, the disaggregated series. Then, $y_t = \sum_{i=1}^{n} \alpha_i y_{it}$, where α_i are the weights associated with each subcomponent. Instead of forecasting y_t by fitting a model directly to this variable, the bottom-up approach advocates that one can obtain forecasts for the aggregate series by aggregating the forecasts of the disaggregated ones, $\widehat{y}_t = \sum_{i=1}^{n} \alpha_i \widehat{y}_{it}$. According to this approach, using disaggregated information can contribute to increase forecast accuracy, which means that aggregating the forecasts of disaggregated series can be better than forecasting the aggregate directly.

For example, Lütkepohl (1984a, 1984b) argues that if the disaggregated data are generated by a known vector ARMA process then it is preferable to forecast the disaggregated variables first and then aggregate the forecasts, rather than forecast the aggregated time series directly. However, in practice, this is not always true, because of parameter and model uncertainty. Lütkepohl presents evidence that suggests that the forecasts from the aggregated process are superior to the aggregated forecasts from the disaggregated process for large lead times h if the orders of the processes are unknown. So, does aggregation of disaggregated forecasts improve forecasting accuracy? The answer to this question is not clear-cut. One advantage of the bottom-up approach is the possibility of capturing idiosyncratic characteristics of each variable by modelling each one individually. However, disaggregated forecast inaccuracy might increase if models are misspecified. Also, what happens with forecast errors is not unambiguous. Forecast errors of the disaggregated variables might cancel out or not.

The second strand explores the possibility of condensing the large information set into a small number of variables, which retains the main features of the original dataset. This means that, for forecasting purposes, a large number of predictors is replaced by a reduced number of variables – diffusion indices, or dynamic common factors – without a significant loss of information (see Stock and Watson (1998)).

This paper tries to merge both strands (see also Marcellino, Stock and Watson (2003)). On one hand, dynamic common factors are introduced in the forecasting models, while, on the other hand, this is done for both aggregate and disaggregated series, whose forecasts are subsequently aggregated. Therefore, two dimensions - the model and the disaggregation level - are simultaneously analysed.

In particular, we try to assess if forecasting consumer price index (CPI) subcomponents individually and then aggregating those forecasts is better than forecasting the aggregate index. Currently, there seems to be some evidence in favour of the bottom-up approach for short-term inflation forecasting. For example, Hubrich (2005) and Benalal *et al.* (2004) conclude that, for the euro area, the bottom-up approach is relevant in the very short-run while Fritzer, Moser and Scharler (2002) (for Austria) and Reijer and Vlaar (2003) (for the Netherlands) find that it is also important up to six and seven-months ahead, respectively.

Additionally, we consider different levels of CPI disaggregation for the bottom-up approach. In the literature, it is used a rather low level of disaggregation. In general, the aggregate index is divided in five components, namely unprocessed food, processed food, non-energy industrial goods, energy and services. However, the results may not remain unchanged if other levels of disaggregation are considered. This paper tries to provide further insight into this question, by considering three different CPI disaggregation levels: the lowest disaggregation level, given by the aggregate price index itself; an intermediate level, in which appear the traditional five components; and a higher disaggregation level, with almost sixty subcomponents.

Using such high level of disaggregation renders intractable structural and VAR modelling. Therefore, other kind of models should be considered. Among the univariate models, the random walk (RW) and a simple autoregressive model (AR) are standard benchmarks, while the SARIMA (Seasonal Autoregressive

Integrated Moving Average) is a quite general univariate model that tries to capture the variable's dynamics based on its past behaviour. Regarding the SARIMA model, one can augment it by including exogenous variables. In particular, the exogenous variables used are the common dynamic factors (see Stock and Watson (1998)). Following Bernanke, Boivin and Eliasz (2005) and Stock and Watson (2005) these models can be called Factor-Augmented SARIMA (FASARIMA) models. By introducing the dynamic common factors in the SARIMA models, the FASARIMA models get closer to the typical VAR models because they also account for potentially relevant information about the variables' comovement, through the dynamic common factors.

The forecasting performance of the different approaches and models is evaluated by an out-of-sample forecast exercise. The criterion used to compare the forecasting performance of the different methods is the root mean squared forecast error (RMSFE). To test whether the differences are statistically significant we use the modified version of Diebold and Mariano (1995) test proposed by Harvey, Leybourne and Newbold (1997).

Considering as case-study the Portuguese one, we find that, for very short-term inflation forecasting, it is better to pursue a bottom-up approach with a high disaggregation level and exploit the variables' comovement, while for longer horizons simpler models perform better and the required disaggregation level decreases over the forecast horizon.

The remainder of the paper is organised as follows. In section 2, a brief description of data is given and some preliminary issues are addressed. In section 3, modelling is discussed and, in section 4, inflation forecasts accuracy is evaluated. Finally, section 5 concludes.

2 Data

The monthly dataset refers to Portuguese CPI and covers the period from January 1988 to December 2004, comprising 204 observations, for the aggregate CPI, its partition in five components (unprocessed food, processed food, non-energy industrial goods, energy and services) and in 59 subcomponents (see Table 1). We exclude from our analysis administered and housing prices. In the first case, administered prices behaviour is hardly captured by an econometric model since these prices are adjusted according to specific national regulations[1] while, in the second case, housing prices, before 1997, were collected on an annual frequency only.

Prior to modelling, data are transformed and examined to account for possible factors that can distort future analysis. First, all series are transformed to logarithms. Second, following Marcellino, Stock and Watson (2003), it was not found evidence of the presence of large outliers.

As Diebold and Kilian (2000) point out, unit root pre-testing can be very useful for model selection purposes. We perform three different kinds of unit root tests. In first place, Dickey and Pantula (1987) tests are carried out. These tests indicate that price indices are not integrated of order two but are integrated of order one. The latter evidence is also supported by Augmented Dickey-Fuller (ADF) and Kwiatkowski, Phillips, Schmidt and Shin (KPSS) (1992) tests. Therefore, we first difference all series.

Following Hylleberg, Engle, Granger and Yoo (HEGY) (1990), seasonal unit root tests are also carried out. In particular, we use the corresponding test procedure for monthly data (see Beaulieu and Miron (1993)). In general, we reject the null hypothesis, which means that we reject unit roots at most frequencies. Based on this, there seems to be no reason for seasonal differencing.

3 Modelling

Regarding model selection, one can choose the models based on their fit within the sample excluding the last observations, which are reserved for an out-of-sample exercise, or one can pick the models based on their out-of-sample performance. Models with the best in-sample fit are not necessarily the best forecasting models. However, selecting a model based on its out-of-sample per-

[1] We also exclude fuel prices because they were subject to regulation until quite recently.

formance means that its selection will strongly rely on a short sample period. For example, Hendry and Clements (2001) argue that one should not focus on forecasting performance for model identification and Inoue and Kilian (2006) show that, under standard conditions, the in-sample method is more reliable than the out-of-sample one. Therefore, for model specification, we focus on in-sample analysis, excluding the last part of the sample, which is used to perform a true out-of-sample exercise.

In contrast with some literature (see, for example, Stock and Watson (1999)), models are not specified as a linear projection of the h-step ahead (h = 1,. . . ,12) interest variable onto t-dated and lagged regressors. The latter is called 'direct forecasting' while here we consider 'iterated forecasting'. 'Iterated forecasting' is done by using a one-period model iterated forward. In fact, Marcellino, Stock and Watson (2006) found that iterated forecasts typically outperform direct forecasts and iterated forecast accuracy increases with the forecast horizon.

3.1 Univariate models

In addition to the AR models, chosen using standard information criteria, SARIMA models are considered. One should note that, although SARIMA models are based on the series' past behaviour only, these models are able to capture rich dynamics, both seasonal and non-seasonal, where the former is typically important in price series.

The SARIMA modelling follows Box and Jenkins (1976) well-known methodology. Basically, this methodology comprises three stages: identification, estimation and diagnostic checking. This process is iterative, which means that when the model chosen is not satisfactory, a new cycle begins and the same steps are repeated until a suitable model is found. To account for deterministic seasonality, whenever it seems appropriate, seasonal dummy variables are added to the model. Hence, the model for series y_t has the following form:

$$\phi(L)\varphi(L^s)(\Delta y_t - \alpha - \beta_1 D_1 - ... - \beta_{11} D_{11}) = \theta(L)\delta(L^s)\varepsilon_t \tag{1}$$

where α is a constant, D_i is a seasonal dummy $(i = 1, ..., 11)$, β_i its corresponding coefficient $(i = 1, ..., 11)$ and ε_t is a white noise. The lag polynomials ($\phi(L)$ - autoregressive polynomial; $\varphi(L^s)$ - seasonal autoregressive polynomial; $\theta(L)$ - moving average polynomial; $\delta(L^s)$ - seasonal moving average polynomial) are defined as usual.

The SARIMA models are estimated by non-linear least squares and selected resorting to coefficient significance tests, Schwarz Bayesian Criterion, residual correlation plots and Ljung-Box tests.[2]

3.2 Factor-Augmented models

When one estimates SARIMA models for each series one is ignoring potentially relevant information about the variables' comovement. However, as mentioned before, when considering a large set of variables renders VAR models intractable. As Granger and Yoon (2001) put it, "*VAR models are the major tools for investigating linear relationships between* ***small*** *groups of variables*" (our emphasized). Therefore, one has to consider alternative models. For example, one can extend the SARIMA models by including exogenous variables that affect the dynamic behaviour of the dependent variable. The additional variables considered are the dynamic common factors, extracted from the large disaggregated price dataset. The key role of the common factors is to summarise large amounts of information in a few variables, which capture the main features of the original data. In fact, the idea behind the factor model is that variables have two components: the common component, which can be captured by a small number of variables – the common factors; and the idiosyncratic component, which reflects variable-specific features. Hence, the purpose of using common factors is to reduce the dimension of data, by pooling the most significant information from the initial series while excluding their idiosyncratic component. By introducing the dynamic common factors in the SARIMA models, we get Factor-Augmented SARIMA (FASARIMA) models, which allow to account for potentially relevant

[2] The chosen models and corresponding estimation results are available from the authors upon request.

information about the variables' comovement, like VAR models.

In particular, the common factors are extracted from the dataset comprising almost sixty CPI subcomponents. Nevertheless, one could have considered alternative dynamic common factors extracted from information sets that included other kind of series, such as activity related variables. However, the existing evidence seems to indicate that nominal variables (like the disaggregated price indices) are more useful for inflation forecasting purposes (see, for example, Angelini, Henry and Mestre (2001a)).[3]

For the dynamic common factors extraction, we follow Stock and Watson (1998). Accordingly, it is possible to estimate dynamic common factors consistently in a "static" version of the dynamic factor model, when both time series and cross-sectional dimensions are large.

Let X_t be a N-dimensional multiple time series of variables, observed for $t = 1, \ldots, T$. Assume that the dynamic factor model can be represented by

$$X_{it} = \lambda_i(L) f_t + e_{it} \tag{2}$$

for $i = 1, \ldots, N$, where e_{it} is the idiosyncratic disturbance, $\lambda_i(L)$ are lag polynomials in nonnegative powers of L and f_t is the r^* common dynamic factors vector. If we assume that $\lambda_i(L)$ have finite lags (for example, m lags), then it is possible to rewrite the dynamic factor model with r common "static" factors (F_t). Thus, the model can be redefined as

$$X_t = \Lambda F_t + e_t \tag{3}$$

where $F_t = (f_t', \ldots, f_{t-m}')'$ is a $r = r^*(m+1)$ vector of stacked vectors and Λ is a $(N \times r)$ parameter matrix. The main advantage of the "static" version of the dynamic factor model is that it can be consistently estimated by principal components.

[3] One should note that we also tried to include in FASARIMA models a composite indicator reflecting real activity developments but we found that it is not statistically significant for almost all models.

Due to different model representations, factors can be extracted alternatively from the contemporaneous values of X_t only, or from a stacked set of variables, including X_t and its lagged values (see Stock and Watson (1998)). Theoretically, adding more variables (lagged values of X_t) could lead to an improvement in the finite sample performance of the models. However, Stock and Watson (2002) conclude that, for US monthly price series, forecasts based on the stacked data perform less well than those based on the unstacked data. Therefore, as Angelini, Henry and Mestre (2001a, 2001b) and Marcellino, Stock and Watson (2003), we extract the common factors from the contemporaneous values of X_t only.[4]

Once obtained the common factors, the next step is to consider them in modelling. In particular, using common factors as exogenous variables in the model has some advantages. There are gains in terms of additional information that is brought into the analysis (especially, the one about variables' comovement) and the number of variables in the model does not increase substantially. Moreover, Stock and Watson (1998) show that the estimated factors can efficiently replace the true factors in forecasting models.

For the determination of the number of factors to include in the model, we use Bai and Ng (2002) criteria (in particular, IC_1, IC_2 and IC_3). These criteria are similar to the well-known information criteria (AIC and SBC, among others) but the penalty is also a function of the cross-sectional dimension (N). These criteria are valid for the "static" version of the dynamic factor model only. The optimal number of factors minimizes the information criteria.

The results suggest the relevance of one factor only (see Table 2). This evidence is also supported by the variance of the original series explained by each factor. As it can be seen in Figure 1, there is a major difference between the variance explained by the first principal component and by the other principal

[4] Before extracting the factors, the 59 subcomponents are transformed to logarithms and first differenced. Since we are not interested in capturing relations based only on common seasonal patterns, the series are also seasonally adjusted. Finally, all series are standardized.

components, which also favours the use of just one factor.[5]

The FASARIMA model considered for each series y_t can be written as

$$\phi(L)\varphi(L^s)(\Delta y_t - \alpha - \beta_1 D_1 - ... - \beta_{11} D_{11} - \upsilon(L)x_t) = \theta(L)\delta(L^s)\varepsilon_t \qquad (4)$$

where α is a constant, D_i is a seasonal dummy ($i = 1, ..., 11$), β_i its corresponding coefficient ($i = 1, ..., 11$) and ε_t is a white noise. The lag polynomials ($\phi(L)$ - autoregressive polynomial; $\varphi(L^s)$ - seasonal autoregressive polynomial; $\theta(L)$ - moving average polynomial; $\delta(L^s)$ - seasonal moving average polynomial; $\upsilon(L)$ - polynomial associated with the exogenous variable x_t) are defined as usual.

The identification of FASARIMA models comprises five stages (see, for example, Enders (2004)). The first one consists in fitting an ARMA model to the exogenous variable, that is, the common factor. Following the univariate modelling strategy, an AR(1) model was chosen. The corresponding residuals are the filtered values of the exogenous variable. By applying the same filter to the variable of interest, in the second step, we obtain the filtered values of the price series. In the next step, both filtered series are used to build a cross-correlogram. The pattern exhibited by the cross-correlations between the common factor and price series helps to determine the number of lags of both variables that should be introduced in the FASARIMA models.

The fourth step consists in estimating plausible models of the following form

$$\phi(L)\varphi(L^s)(\Delta y_t - \alpha - \beta_1 D_1 - ... - \beta_{11} D_{11} - \upsilon(L)x_t) = e_t \qquad (5)$$

and selecting the model with the best fit. The residuals of the resulting model (e_t) are not necessarily white noise. So, the examination of the residual autocorrelation should suggest plausible orders for the $\theta(L)$ and $\delta(L^s)$ polynomials. The last step consists in estimating altogether the FASARIMA model.[6]

[5] As a sensitivity analysis, we also allowed for more than one factor in FASARIMA models, but the forecasting results did not improve.

[6] The chosen models and corresponding estimation results are available from the authors upon request.

4 Out-of-sample forecast evaluation

Forecasting performance is evaluated through an out-of-sample exercise. For each series and model, a recursive estimation process is implemented. Starting from the estimation period (up to December 2000), each round a new observation is added to the sample. In each round of this recursive estimation process one to twelve step ahead forecasts are computed. Thus, for each forecast horizon, 37 forecasts are used.[7]

For each forecast horizon, the forecast series of all 59 subcomponents are aggregated, using the corresponding CPI weights. Then, from the index forecasts, year-on-year inflation rate forecasts are obtained. The same is done for the intermediate disaggregation level (5 components) and for the aggregate index.

Even though the forecast evaluation framework is similar for all models, factor-augmented forecasting requires some additional steps. The reason why this happens is quite obvious. Forecasting the dependent variable also requires forecasting the exogenous variable.

However, forecasting the common factor is not a straightforward issue. It is possible to obtain 'direct' or 'indirect' factor forecasts. The direct forecasts can be obtained by fitting a model directly to the common factor and using it to produce one to twelve months ahead forecasts, for each recursive sample. In particular, we use the model already fitted to the common factor in the previous section (an AR(1) model). Obviously, this brings up another drawback - the potential misspecification of the common factor model. The 'indirect' approach relies on the fact that the common factor is a weighted linear combination of the disaggregated price series. Therefore, factor forecasts can be obtained by weighting price series forecasts, resulting from SARIMA models.

The results obtained are quite interesting. Unsurprisingly, the RW models are the ones that present the worst performance in terms of RMSFE (see Table 3). Although simple autoregressive models outperform random walks, these

[7] We set the number of forecasts for each horizon to be the same so as to allow a fair comparison also across forecast horizons (in practice, we dropped the last $12 - h$ forecasts for the h-step ahead horizon).

models present a worse performance than more general univariate models.

Disregarding the benchmark models, whatever the model considered, the bottom-up approach improves on aggregate forecasting for forecast horizons up to five-months ahead. Furthermore, within these forecast horizons, the highest disaggregation level delivers better results than the one with five components (see Figure 2). That is, the gains of aggregating forecasts against aggregate forecasting increase with the disaggregation level. Moreover, the gains are rather statistically significant for a very short-term forecast horizon range (see Table 4). The test used for comparing forecast accuracy is based on the modified Diebold-Mariano test proposed by Harvey, Leybourne and Newbold (1997). This test compares, for each kind of model, the forecasts obtained through aggregating the disaggregated series' forecasts (5 or 59) against the forecasts obtained by fitting a model to the overall index. So, from the results obtained, it seems that, the benefit in terms of additional information stemming from disaggregated price data through a bottom-up approach is such that compensates the loss due to potential model misspecification and parameter uncertainty. As mentioned earlier, the advantage of the bottom-up approach is that the modelling is done at a more disaggregated level, which allows capturing the idiosyncratic characteristics of each component underlying the aggregate. However, the true data generating process is unknown and errors may cancel out or not. This means that the benefits from the bottom-up approach may overcome or not the losses resulting from the fact that the true model is unknown and has to be specified and estimated. Since the bottom up-approach dominates in the short-run, at least for such horizon, the above mentioned net balance is positive.

Concerning the models, for one and two-months ahead, the forecasting procedure that minimizes the RMSFE is the one that considers the highest disaggregation level (59 subcomponents) and models each subcomponent as a FASARIMA resorting to direct common factor forecasts. For three up to five-months ahead, the disaggregation level that provides the best results continues to be the highest one, but now each subcomponent is modelled as a SARIMA. Hence, as the forecast horizon increases, the gains resulting from the additional information on the variables' comovement, captured by the common factor, are outbalanced by the factor forecast errors. For even longer horizons, more accurate forecasts can be obtained by simply forecasting the aggregate index through a SARIMA model.

Therefore, we find an inverse relationship between the forecast horizon and the amount of information underlying the forecast when minimizing the RMSFE. That is, for very short-term inflation forecasting, it is worthwhile to follow a bottom-up approach with a high disaggregation level and exploit the variables' comovement, while for longer horizons simpler models appear to deliver better results and the required disaggregation level decreases over the forecast horizon.

5 Conclusion

The purpose of this paper is to assess if one can improve forecasting accuracy by considering disaggregated price data. Although the question of to bottom-up or not has also been addressed elsewhere, a new dimension to this discussion is introduced by considering different levels of data disaggregation, namely a higher disaggregation level than the one considered up to now. In particular, we consider three CPI disaggregation levels: the lowest disaggregation level, given by the aggregate price index itself; an intermediate level, in which appear the traditional five components; and a higher disaggregation level, with almost sixty subcomponents. However, such high disaggregation poses modelling difficulties that have to be dealt with. Hence, we resort to a new strand of models, the Factor-Augmented SARIMA models. Using as case-study the Portuguese one, the forecasting accuracy (up to twelve months ahead) of the bottom-up approach is evaluated through an out-of-sample forecast exercise. Our results point to an inverse relationship between the forecast horizon and the information set used in forecasting. That is, for very short-term inflation forecasting, it is better to pursue a bottom-up approach with a high disaggregation level and take into account the variables' comovement, while for longer horizons simpler models seem to perform better and the required disaggregation level decreases over the forecast horizon.

References

[1] Angelini, E., Henry, J., Mestre, R. (2001a) Diffusion index-based inflation forecasts for the Euro Area. ECB Working Paper no 61.

[2] Angelini, E., Henry, J., Mestre, R. (2001b) A Multi Country Trend Indicator for Euro Area Inflation: Computation and Properties. ECB Working Paper no 60.

[3] Bai, J., Ng, S. (2002) Determining the Number of Factors in Approximate Factor Models. Econometrica 70(1): 191-221.

[4] Beaulieu, J., Miron, J. (1993) Seasonal unit roots in aggregate US data. Journal of Econometrics 55: 305-328.

[5] Benalal, N., del Hoyo, J., Landau, B., Roma M., Skudelny, F. (2004) To aggregate or not to aggregate? Euro area inflation forecasting. ECB Working Paper no 374.

[6] Bernanke, B., Boivin, J. and Eliasz, P. (2005) Measuring the effects of monetary policy: a factor-augmented vector autoregressive (FAVAR) approach. Quarterly Journal of Economics 120: 387-422.

[7] Box, G., Jenkins, G. (1976) Time Series Analysis: forecasting and control. Holden-Day.

[8] Dickey, D., Pantula, S. (1987) Determining the order or differencing in autoregressive processes. Journal of Business and Economic Statistics 5(4): 18-24.

[9] Diebold, F., Kilian, L. (2000) Unit-Root Tests Are Useful for Selecting Forecasting Models. Journal of Business and Economic Statistics 18(3): 265-273.

[10] Diebold, F., Mariano, R. (1995) Comparing Predictive Accuracy. Journal of Business and Economic Statistics 13(3): 253-263.

[11] Enders, W. (2004) Applied Econometric Time Series, 2^{nd} edition. Wiley.

[12] European Central Bank (2003a) The ECB's monetary policy strategy. ECB Press Release, 8 May 2003.

[13] European Central Bank (2003b) The ECB's monetary policy strategy after the evaluation and clarification of May 2003. Speech by Jean-Claude Trichet, 20 November 2003.

[14] Fritzer, F., Moser, G., Scharler, J. (2002) Forecasting Austrian HICP and Its Components Using VAR and ARIMA Models. Oesterreichische Nationalbank Working Paper no 73.

[15] Granger, C., Yoon, G. (2001) Self-generating variables in a cointegrated VAR framework. University of California San Diego Discussion Paper no 4.

[16] Harvey, D. I., Leybourne, S. J. and Newbold, P. (1997) Testing the equality of prediction mean squared errors. International Journal of Forecasting, 13, 281-291.

[17] Hendry, D., Clements, M. (2001) Economic Forecasting: Some Lessons from Recent Research. ECB Working Paper no 82.

[18] Hubrich, K. (2005) Forecasting euro area inflation: Does aggregating forecasts by HICP component improve forecast accuracy? International Journal of Forecasting, 21: 119-136.

[19] Hylleberg, S., Engle, R, Granger, C, Yoo, B. (1990) Seasonal Integration and Cointegration. Journal of Econometrics 44: 215-238.

[20] Inoue, A., Kilian, L. (2006) On the selection of forecasting models. Journal of Econometrics 130: 273-306.

[21] Kwiatkowski, D., Phillips, P., Schmidt, P., Shin, Y. (1992) Testing the null hypothesis of stationarity against the alternative of a unit root. Journal of Econometrics 54: 159-178.

[22] Lütkepohl, H. (1984a) Forecasting Contemporaneously Aggregated Vector ARMA Processes. Journal of Business and Economic Statistics 2(3): 201-214.

[23] Lütkepohl, H. (1984b), Linear transformations of vector ARMA processes, Journal of Econometrics, 26: 283-293.

[24] Marcellino, M., Stock, J. H., Watson, M. W. (2003) Macroeconomic forecasting in the Euro area: Country specific versus area-wide information. European Economic Review 47: 1-18.

[25] Marcellino, M., Stock, J. H., Watson, M. W. (2006) A Comparison of direct and iterated multistep AR methods for forecasting macroeconomic time series. Journal of Econometrics 135: 499-526.

[26] Reijer, A., Vlaar, P. (2003) Forecasting inflation: An art as well as a science! De Nederlandsche Bank Staff Reports no 107.

[27] Stock, J. H., Watson, M. W. (1998) Diffusion Indexes. NBER Working Paper no 6702.

[28] Stock, J. H., Watson, M. W. (1999) Forecasting inflation. Journal of Monetary Economics 44: 293-335.

[29] Stock, J. H., Watson, M. W. (2002) Macroeconomic Forecasting Using Diffusion Indexes. Journal of Business and Economic Statistics 20(2): 147-162.

[30] Stock, J. H., Watson, M. W. (2005) Implications of Dynamic Factor Models for VAR Analysis. NBER Working Paper no 11467.

Table 1 - CPI breakdown

	Subcomponents	Components
Series 1	Rice	Pf
Series 2	Other cereal products	Pf
Series 3	Pasta products	Pf
Series 4	Bread and other bakery products	Pf
Series 5	Potatoes and other tubers	Unpf
Series 6	Dried vegetables	Unpf
Series 7	Fresh and preserved vegetables	Unpf
Series 8	Fruit	Unpf
Series 9	Meat of sheep and goat	Unpf
Series 10	Meat of swine	Unpf
Series 11	Meat of bovine animals	Unpf
Series 12	Other meat	Unpf
Series 13	Sausages and preserved meat	Unpf
Series 14	Poultry	Unpf
Series 15	Fresh, chilled or frozen fish	Unpf
Series 16	Fresh, chilled or frozen seafood	Unpf
Series 17	Other preserved or processed fish and seafood and fish and seafood preparations	Unpf
Series 18	Dried, smoked or salted fish and seafood	Unpf
Series 19	Eggs	Pf
Series 20	Milk	Pf
Series 21	Yoghurt, cheese and other milk-based products	Pf
Series 22	Edible oils	Pf
Series 23	Butter, margarine and other fats	Pf
Series 24	Sugar and confectionery	Pf
Series 25	Cocoa and powdered chocolate	Pf
Series 26	Coffee	Pf
Series 27	Tea	Pf
Series 28	Sauces, condiments and salt	Pf
Series 29	Baking powders, preparations and soups	Pf
Series 30	Catering	Serv
Series 31	Alcoholic beverages	Pf
Series 32	Mineral waters, soft drinks and juices	Pf
Series 33	Clothing materials and garments	Neig
Series 34	Dry-cleaning, repair and hire of clothing	Serv
Series 35	Footwear	Neig
Series 36	Repair and hire of footwear and shoe-cleaning services	Neig
Series 37	Gas	Enrgy
Series 38	Heating and cooking appliances, refrigerators, washing machines and similar major household appliances	Neig
Series 39	Furniture, furnishings and carpets	Neig
Series 40	Repair of furniture, furnishings and floor coverings	Serv
Series 41	Household textiles	Neig
Series 42	Glassware, tableware and household utensils and tools	Neig
Series 43	Non-durable household goods	Neig
Series 44	Repair of household appliances	Serv
Series 45	Therapeutic appliances and equipment	Neig
Series 46	Medical, paramedical and hospital services	Serv
Series 47	Motor cars, motor cycles and bicycles and spare parts and accessories	Neig
Series 48	Maintenance and repairs; other services in respect of personal transport equipment	Serv
Series 49	Telephone and telefax equipment	Neig
Series 50	Education	Serv
Series 51	Equipment for the reception, recording and reproduction of sound; other major durables for recreation and culture	Neig
Series 52	Repair of audio-visual, photographic and data processing equipment	Serv
Series 53	Receational and cultural services	Serv
Series 54	Newspapers, books and stationery	Neig
Series 55	Accommodation services	Serv
Series 56	Package holidays	Serv
Series 57	Electrical appliances and products for personal care	Neig
Series 58	Hairdressing salons and personal grooming establishments	Serv
Series 59	Insurance and banking services	Serv

Note: Unpf - Unprocessed food; Pf - Processed food; Neig - Non-energy industrial goods; Enrgy - Energy; Serv - Services.

Table 2 - Bai and Ng criteria

Number of factors	IC_1	IC_2	IC_3
r=1	9.005	9.013	8.987
r=2	9.026	9.041	8.988
r=3	9.049	9.071	8.993
r=4	9.072	9.102	8.998
r=5	9.102	9.140	9.009
r=6	9.133	9.178	9.021
r=7	9.162	9.215	9.032
r=8	9.190	9.250	9.041

Note: Shaded area denotes the minima.

Table 3 - Root Mean Squared Forecast Errors

	Forecast horizon											
	1	2	3	4	5	6	7	8	9	10	11	12
RW_1	0.437	0.780	1.012	1.138	1.241	1.421	1.703	2.034	2.329	2.554	2.751	2.990
RW_5	0.453	0.804	1.043	1.175	1.286	1.481	1.774	2.116	2.428	2.666	2.872	3.131
RW_59	0.439	0.793	0.945	1.077	1.225	1.555	1.880	2.225	2.495	2.699	2.859	3.096
AR_1	0.360	0.713	1.000	1.154	1.215	1.278	1.482	1.803	2.096	2.304	2.461	2.630
AR_5	0.256	0.455	0.658	0.803	0.933	1.070	1.264	1.503	1.721	1.910	2.089	2.298
AR_59	0.243	0.543	0.670	0.801	0.943	1.256	1.539	1.807	2.035	2.256	2.443	2.626
SARIMA_1	0.230	0.353	0.448	0.492	0.549	0.571	0.602	0.662	0.724	0.761	0.798	0.830
SARIMA_5	0.209	0.305	0.403	0.487	0.588	0.692	0.802	0.938	1.079	1.220	1.365	1.512
SARIMA_59	0.207	0.254	0.322	0.408	0.542	0.729	0.856	0.969	1.126	1.286	1.436	1.544
FASARIMA_1_dir	0.244	0.347	0.477	0.544	0.654	0.721	0.856	1.019	1.184	1.316	1.468	1.626
FASARIMA_5_dir	0.226	0.277	0.387	0.482	0.595	0.680	0.804	0.982	1.145	1.284	1.457	1.626
FASARIMA_59_dir	0.201	0.234	0.335	0.436	0.577	0.822	0.946	1.120	1.332	1.491	1.645	1.810
FASARIMA_1_ind	0.237	0.364	0.479	0.563	0.655	0.738	0.851	0.993	1.128	1.246	1.375	1.515
FASARIMA_5_ind	0.212	0.306	0.410	0.506	0.610	0.710	0.828	0.973	1.116	1.254	1.403	1.560
FASARIMA_59_ind	0.213	0.255	0.347	0.451	0.602	0.811	0.949	1.115	1.286	1.455	1.616	1.750

Note: RW stands for Random Walk, AR for Autoregressive, SARIMA for Seasonal Autoregressive Integrated Moving Average, FASARIMA for Factor-Augmented SARIMA; 1 for the aggregate index, 5 for the intermediate disaggregation level, 59 for the high disaggregation level; "dir" for direct common factor forecast, "ind" for indirect common factor forecast. Shaded area denotes the minima.

Table 4 - Gain of aggregating forecasts against aggregate forecasting

	Forecast horizon											
	1	2	3	4	5	6	7	8	9	10	11	12
RW_5	-3.6	-3.0	-2.9	-3.2	-3.5	-4.1	-4.0	-3.9	-4.1	-4.2	-4.2	-4.5
RW_59	-0.6	-1.6	7.1	5.6	1.3	-8.6	-9.4	-8.6	-6.6	-5.4	-3.8	-3.4
AR_5	40.6 ***	56.6 ***	52.1 ***	43.6 ***	30.3 ***	19.4 ***	17.2 ***	19.9 ***	21.8 ***	20.7 ***	17.8 ***	14.4 ***
AR_59	48.3 ***	31.4 ***	49.4 ***	44.0 ***	28.8 ***	1.7	-3.7	-0.2	3.0	2.1	0.7	0.1
SARIMA_5	10.1	15.6	11.3	1.1	-6.7	-17.5	-24.9	-29.4	-32.9	-37.6	-41.5	-45.1
SARIMA_59	11.3	39.1 **	39.3 **	20.8 *	1.2	-21.6	-29.6	-31.6	-35.7	-40.8	-44.4	-46.3
FASARIMA_5_dir	7.9	25.0 ***	23.4 ***	13.0 **	9.9 **	5.9	6.5 *	3.8	3.4	2.5	0.8	0.0
FASARIMA_59_dir	21.6 **	48.1 ***	42.4 ***	24.8 **	13.4 *	-12.3	-9.5	-9.0	-11.1	-11.8	-10.7	-10.2
FASARIMA_5_ind	11.9 *	18.8 **	16.8 **	11.3 **	7.5 *	3.9	2.8	2.1	1.1	-0.7	-2.0	-2.9
FASARIMA_59_ind	11.4	42.4 ***	38.0 ***	24.7 **	8.8	-9.0	-10.2	-11.0	-12.3	-14.4	-14.9	-13.4

Note: The gain is computed as ($RMSFE_{\text{aggregate forecasting}}/RMSFE_{\text{aggregating forecasts}}$ - 1)*100. The asterisks mean that the null hypothesis of the modified Diebold-Mariano test, for equal performance of aggregate forecasting and aggregating forecasts, is rejected. The *** stands for a significance level of 1%, ** stands for a significance level of 5% and * stands for a significance level of 10%.

Figure 1 - Variance of price series explained by each principal component

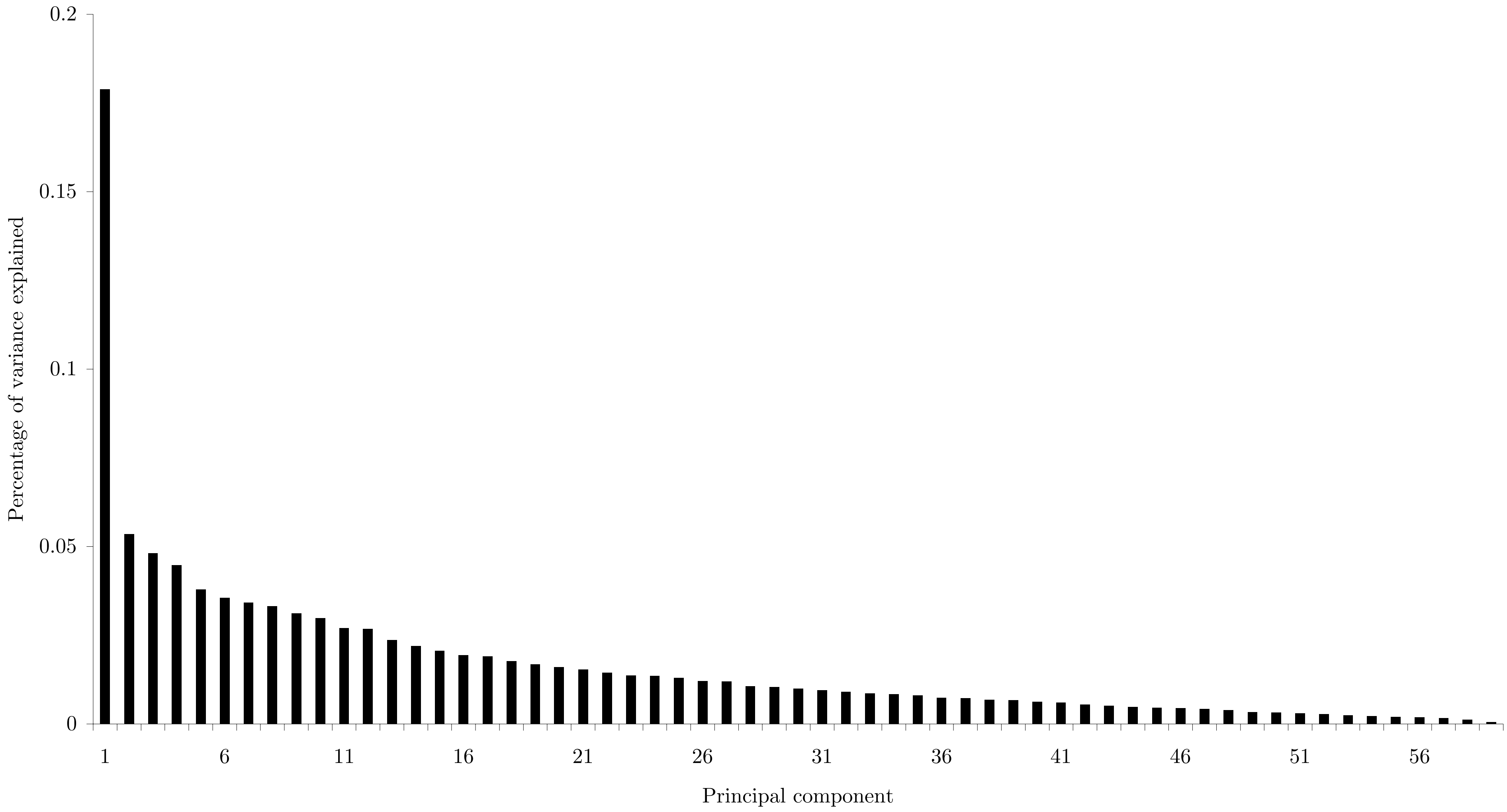

Figure 2 - Gain of aggregating forecasts against aggregate forecasting