ECONOMETRIC MODELLING FOR SHORT-TERM INFLATION FORECASTING IN THE EURO AREA

Antoni Espasa and Rebeca Albacete*

Abstract

This paper examines the problem of forecasting macro-variables which are observed monthly (or quarterly) and result from geographical and sectorial aggregation. The aim is to formulate a methodology whereby all relevant information gathered in this context could provide more accurate forecasts, be frequently updated, and include a disaggregated explanation as useful information for decision-making. The appropriate treatment of the resulting disaggregated data set requires vector modelling, which captures the long-run restrictions between the different time series and the short-term correlations existing between their stationary transformations. Frequently, due to a lack of degrees of freedom, the vector model must be restricted to a block-diagonal vector model. This methodology is applied in this paper to inflation in the euro area, and shows that disaggregated models with cointegration restrictions improve accuracy in forecasting aggregate macro-variables.

Keywords: sectorial and geographical disaggregation, VEqCM, cointegration, core inflation, combination of forecasts

*Antoni Espasa, Departamento de Estadística, Universidad Carlos III de Madrid, C/ Madrid 126, 28903 Getafe, Madrid, (Spain); e-mail: espasa@est-econ.uc3m.es; telephone: +34 91 624 98 03; fax: +34 91 624 98 49; Rebeca Albacete, Departamento de Estadística, Universidad Carlos III de Madrid, e-mail:albacete@est-econ.uc3m.es

Authors' biographies:

Antoni Espasa is Professor of Econometrics. He obtained his PhD from the London School of Economics, has been employed as Chief Economist in the Banco de España and is a Fellow of the European Economic Association. In 1991 he was awarded the Rey Jaime I Prize in Economics. His published research includes spectral econometrics, dynamic models, applied econometrics, forecasting, and signal extraction in time series

Rebeca Albacete obtained her PhD in Economics from the Universidad Carlos III, Madrid, in 2004. She has collaborated with the European Forecasting Network cofinanced by the European Commission and coordinated the monthly publication, Bulletin of EU and USA Inflation and Macroeconomics. Currently, she is working at Fundación de las Cajas de Ahorro.

1. INTRODUCTION

This paper is concerned with the problem of forecasting macro-variables of interest for economic policy and for investment decisions in financial markets. The paper concentrates on variables which are observed monthly (or quarterly) and result from geographical and sectorial aggregation. The aim is to formulate a methodology which, when using relevant disaggregated information, could generate more accurate forecasts than models constructed for aggregate data and which could, by including a disaggregated explanation for the forecasts, provide useful information for decision-making.

Forecasting accuracy itself depends on the nature of the economic phenomenon considered (over which the analyst has no control), the data set used, and the econometric model employed. The smallest possible data set includes no more than the present and the past values of the macro-variable in question. This is a basic univariate set. This set can be enlarged in different non-exclusive directions as follows: (a) frequency enlargement, integrating more frequent data; (b) enlargement by means of the functional and geographical disaggregation of the macro-variable; and (c) enlargement with other related variables. In case (a) we continue to observe a single aggregate phenomenon and the resulting data set remains univariate. In (b) and (c) we have a multivariate data set in each case. In case (b) we denote it as an internal multivariate data set in the sense that it includes only information (disaggregated) corresponding to the phenomenon under study. Consequently, the data set in (c) is denoted as an external multivariate data set. This paper basically deals with information sets (a) and (b).

The development of the proposed methodology consists of detecting in which direction one should enlarge the information relevant for the forecast, for which the

underlying theory about the phenomenon of interest is decisive. The complexity of the appropriate econometric model will depend on the properties of the data set contemplated. In the process of increasing the amount of data used, and thus the complexity of the econometric models, we discern that certain lines of data enlargement are important for the forecast such that evaluation of the forecasting accuracy of the alternative econometric models at the different levels of data is essential.

For the main macro-variables in an economy, agents and authorities require frequent updates of trace forecasts covering at least the current and following years. A good compromise is to update forecasts each time that a new observation of the macro-variable in question is published. This implies monthly updates for inflation, industrial production, foreign trade, etc., but quarterly updates for macro-variables published in the national accounts. The analysis of an aggregate variable gives rise to the question of whether the best results are obtained by directly modelling the aggregate or, in contrast, by disaggregating and obtaining the aggregate forecast from the components.

Disaggregation is one way to increase information in the forecasting process. What matters is that this additional information must be relevant with the possibility of an adequate econometric treatment. Economic theory suggests that this could be the case. In considering market disaggregation, it can be noted that markets are not homogeneous. Furthermore, variables such as production, prices, employment, etc. can have different trends through markets, even though in general they are not independent. In particular, economic theory suggests that, in the medium term, prices through markets will have different but restricted trends. The argument can also be extended to seasonality, cycles, and volatility to show that, in general, disaggregation can be relevant.

The appropriate econometric treatment of the resulting internal multivariate data set requires vector modelling which captures the long-run restrictions between the different time series and the short-term correlations between their stationary transformations. Without this simultaneous modelling, there is no guarantee that one could obtain better forecasting results with the enlarged data set. On the contrary, proper modelling of the multivariate set ensures that the forecasting results derived from this model cannot be less accurate than those obtained with the corresponding univariate data set.

Macro-variables in an economic zone such as the euro area can be analysed by being broken down in different ways. Three such ways are as follows. One refers to the breakdown into components corresponding to large groups of markets, denoted as sectors in this paper, throughout the area as a whole; another considers the macro-variable by national aggregates or by regions. The final approach takes into account both types of information and breaks down the macro-variable by components corresponding to different sectors in each region.

These disaggregations can be interesting for many countries and economic areas because very often it turns out that components show, among other things, different trends. At the same time, in a vector of n components there are fewer than n independent trends, but more than one. This means that in a vector of n different components they are restricted by some, but not by the possible maximum number $(n-1)$ of cointegration relationships. In this paper, the case with $(n-1)$ restrictions is referred to as full cointegration. In this respect, it could be said that the different components in many macro-variables are cointegrated, but are not fully cointegrated. The presence of a certain degree of cointegration could make the disaggregation particularly relevant. In

this case, modelling the components with simple parsimonious models could result in a complex model with important restrictions for the aggregate. Consequently, the gain in efficiency can be large with disaggregation and cointegration.

The lack of full cointegration between components implies that innovations in aggregate will have different long-run effects depending on the common trend from which they mainly derive. In these circumstances, disaggregation is interesting because it is one way to increase the information about future trends. The disaggregation proposal is also convenient for similar arguments on the seasonality of components or for important differences in the stationary behaviour of their cyclical and short-term fluctuations.

Certainly, the practice of disaggregation has limits (see Zellner and Tobias, 2000). In particular, if the quality of data deteriorates when disaggregating, or the analyst does not succeed in modelling the data properly, then the disaggregated models could be faulty, and the forecasts derived from them for the aggregate could be much worse than the forecasts from an aggregate model. Modelling the vector of components becomes more complex than modelling the aggregate, not only because of the obvious question of dimensionality, but also because it is much more probable that, at least for some components, the linear approximation in modelling would not be supported by data, requiring non-linear structures which could be quite difficult to construct.

The disaggregate analysis proposed and developed in this paper is based on considering simultaneous modelling of the components as a way of capturing long-term interrelationships and the time-dependence between their stationary transformations. These simultaneous models also include appropriate processing by means of dummy variables for the special events with an impact on the observations of the components.

Espasa *et al.* (2002), for instance, consider disaggregation by sectors or by countries. In the methodology presented in this paper for forecasting a macro-variable, the disaggregation plays an important role and it is proposed that, when possible, it should be performed applying the double criterion of sectors and regions. The number of resulting components could then be too great to be processed by a vector model and restrictions are required to make it feasible. We are not able to produce a set of precise rules to specify those restrictions, but we point out some useful orientations in this respect. The aim is to end up with a model of a small dimension at the same time that an important amount of relevant information is maintained. Two types of restrictions appear to be of practical importance. One refers to breaking down the aggregate into a small number of components and the other to simplifying the model by imposing a structure of subvectors of variables which are treated as independent. In the first case, economic theory could be useful for specifying a reduced number of components with different but restricted trends and with cross-transitory dynamics. In the second case, where independence between subvectors in the long run cannot be fully justified by the data, the requirement is to formulate subvectors which only partially violate this independence hypothesis. Cointegration analysis for pairs of components could be useful for designing the restrictions to be imposed.

The disaggregated models enable the introduction of individual indicators in the component equations or general indicators which may have different effects by component. This is another of the major advantages of the models proposed. In this case we are dealing with an external multivariate data set. The econometric models in the methodology proposed in this paper are models in reduced form or final form models, which can be obtained from a structural Marshallian Macroeconomic Model with *n*-

sectors (see Zellner, 2000, and Zellner and Israilevich, 2003). This forecasting approach has been successfully applied to the USA, euro area, Spain, and some Spanish regional economies in the Bulletin of EU and US Inflation and Macroeconomic Analysis.

In this paper the macro-variable chosen is inflation in the European Monetary Union, hereafter referred to as the 'euro area'. The analysis is on the Harmonized Index of Consumer Prices (HICP) published by Eurostat. The first step in our procedure is the determination of the order of integration of the components of the macro-variable. In the case of HICP, the order-augmented Dickey–Fuller test statistics for the variables in logarithms and for their first differences indicate that all variables considered appear to be integrated of order 1 or I(1) with the hypothesis of a second unit root rejected in all cases. That is, the inflation variables are stationary. This integration analysis has been performed assuming constant seasonality. The alternative of stochastic seasonality with the necessity of testing all the seasonal roots is not considered because of the small sample size.

This paper studies the effects of special events with particular attention, as in the case of HICP they have given rise to a considerable change in price seasonality, occurring in different years in several countries. Finally, with inflation we perform an in-depth analysis of the inclusion of international crude oil prices in the different models considered and show that they improve sample fit, after other considerations are taken into account in forecasting.

The remainder of this paper is organized as follows. Section 2 examines the number of country components required in the disaggregation of the HICP and derives models for disaggregated data considering only regions or only sectors. Section 3 presents a disaggregated econometric modelling of HICP by two sectors in five regions

and an approach based on block-diagonal vector models, and develops the modelling of specific effects (such as the incorporation of Greece, the entrance of the euro, the incorporation of sales prices) and the introduction of international indicators in energy prices. Section 4 analyses the forecasting performance of the proposed models. Section 5 provides some concluding remarks.

2. GEOGRAPHICAL AND SECTORIAL INDEPENDENT DISAGGREGATIONS FOR INFLATION IN THE EURO AREA

In this section, the total HICP in the euro area is broken down corresponding to the following five geographical components: Germany (GER), France (FRA), Italy (ITA), Spain (SPA), and the remainder (RES), where RES denotes an aggregate of the remaining eight countries that comprise the euro area. The main purpose in this section is to examine whether the remaining economies could be considered as exogenous. A graphical analysis shows that the HICPs of Germany and France move together as do the HICPs of Italy and Spain.

The sample used is from January 1996 to March 2001. Stability analysis shows stable results from September 1996 and the null hypothesis of no cointegration is rejected in favour of at least two long-run restrictions. The cointegration relationships show restrictions between relative prices. One relationship is similar to the unique relationship found in Espasa *et al*. (2002), who only consider the vector formed by the four largest countries. In the other long-run restriction enters the other countries (RES). These results show that the other countries (RES) should be considered in the geographical disaggregation of euro area inflation, because it enters one cointegration relationship and it is not exogenous.

The variables analysed in this study are price indexes, but price levels in the reference period also differ among countries. The nominal convergence among countries leads to greater price growth for Italy, Spain, and the remainder (RES), which are countries with lower price levels. The cointegration relationships show that this convergence process is taking place. A Vector Autoregression Model with Equilibrium-Correction Mechanism (VEqCM) for the five regions, including seasonal dummies, is estimated. This model shows that there is little cross-dependence between the variables in the short-run. This model is used in Section 4 for the forecasting evaluation exercise.

The breakdown of HICP by sectors in the euro area as a whole is approached in Espasa *et al*. (2002) by taking into account theoretical considerations about differences in supply and demand through sectors, which could result in prices having different trends. Following Espasa *et al*. (1987), we consider the following sector price indexes: (1) Processed food; (2) Non-energy industrial goods; (3) Services; (4) Unprocessed food; and (5) Energy. The European Central Bank (ECB) later adopted this same breakdown. This paper updates the analysis in Espasa *et al*. (2002) with a longer sample and obtains the same result of only one cointegration relationship between sectors. This model for the five sectors is also used in Section 4 for evaluating the inflation forecasts.

3. A DISAGGREGATED ECONOMETRIC MODELLING OF HICP BY SECTORS AND BY COUNTRIES

Given a minimum number of five countries, and taking into account the dimensions of the sample, this paper considers a less disaggregated sector scheme based on just two large sectors, one of which derives from the aggregation of the first three enumerated in Section 2 and the other results from the sum of the last two. Therefore,

the breakdown of HICP by sectors in each geographical region has been approached considering the following two price indexes corresponding to: (I) overall index excluding energy and unprocessed food, denoted as core HICP; and (II) energy and unprocessed food, denoted as residual HICP. This core measure proposed in Espasa *et al.* (1987) is also used by the ECB, which provides a further rationale for this disaggregation.

To disaggregate the total HICP in the euro area, taking into account both criteria, sectors and countries, the HICPs corresponding to the two sectors – core and residual – in each of the five countries are considered. The sample used in this analysis is from January 1996 to June 2002. Stability analysis shows stable results from March 1996 and the null hypothesis of no cointegration is rejected in favour of at least three long-run restrictions. The result of having (seven) less than n (10) independent trends is what could be expected for the HICP on the basis of monetary theory. The first cointegration relation mixes core and residual indexes. The second mainly relates to residual indexes. The third and final relation only involves core indexes. As in the previous section, the variables corresponding to the group formed with the eight small countries (RES) cannot be regarded as exogenous. A Vector Autoregression Model with Equilibrium-Correction Mechanism for the 10 components is estimated. This model shows that there is greater contemporaneous correlation between these residuals than in the breakdown by sectors or by countries. The correlations are greater between the residuals derived from the residual inflation equations than between the residuals coming from the core inflation equations. One possible interpretation is that energy prices, more specifically fuel prices, depend on OPEC policy, while in the unprocessed food sector prices are

influenced by the weather. All of these constitute general conditions, which may affect different countries at the same time.

Once again, there are few transitory dynamics in the model. Nevertheless, this dependency is extended further than in the previous approaches discussed in Section 2. Almost all the variables depend on their own lags, and the residual inflation equations for all countries depend on the lagged residual inflation in RES, which turns out to be a leading indicator for other residual price indexes. It is concluded that extending the disaggregation framework allows more dynamic relationships to be captured.

The estimation results provide better fits for core prices than for residual prices; this is particularly relevant for the cases of Germany, France, and the remainder (RES). As the contemporaneous correlations are important in this multi-equation system, and there are several cross-equation dynamic effects, a poor specification in one equation translates to the others. The results in this section indicate that, based on asymptotic tests, data in the given small sample reject a separate modelling of core and residual prices. At the same time, the model obtained with the sample available cannot be considered appropriate because of the poor fit of some equations. In these circumstances, the theoretically preferable 10-equation system should be simplified, but while maintaining the disaggregation level of the 10 variables, which allows capturing more adequately the different trend factors in the HICP.

For simplification, the question arises whether or not pairs of core and residual prices share a common trend. This analysis is applied to the euro area as a whole and to each of the five countries. Johansen's (1988, 1991) cointegration analysis shows the absence of cointegration between the core HICP and the corresponding residual HICP in all cases – euro area, France, Germany, Italy, Spain, and the group of remaining

countries (RES) – pointing out the presence of two common factors in each pair of core and residual price indexes.

One implication of this result is that the core consumer price index in each of the above cases is not a sufficiently good indicator for forecasting total inflation. Core inflation is an interesting indicator for a different reason, which could be the higher persistence of shocks in core HICP than in residual HICP. In fact, the breakdown of a vector variable such as HICP when there is not full cointegration between its components is, in any case, important for diagnostic purposes. For example, an innovation from services prices, properly weighted, does not have the same implications in total CPI as one from non-processed food prices.

To simplify the vector model with 10 equations, the results from the bivariate cointegration analysis are indicative. Given the absence of cointegration relationships in all the pairs of core and residual indexes, we could consider approximating the 10-component vector model with two separate vector models, one referring to the five core indexes and the another to the five residual indexes. This implies a block-diagonal formulation for the original 10-component vector model. This restriction is not really acceptable because, even pairwise, there are no cointegration relationships between core indexes and residual indexes; it was seen above that considering the 10 components jointly the cointegration relationships include some mixture of core and residual indexes. However, even in this situation, of the three cointegration relationships one was between core indexes and another related almost exclusively to residual indexes. Accordingly, it can be concluded that, even when the block-diagonal restriction is not fully supported by the data, it can be seen as being close to it. In any case, some

approximation is required given that the fit of the 10-component vector to the small sample available is not good.

3.1 ANALYSIS BY FIVE CORE HICPs

In this section, a vector composed of the core-harmonized indexes in Germany, France, Italy, Spain, and the remainder (RES) is considered. All this analysis has been elaborated taking into account the last available HICP figures corresponding to December 2002.

The graphical analysis shows that the indexes for Germany and France appear to move together as do the indexes for Italy and Spain. Another feature revealed in the graphical analysis is the relevant impact of the introduction of sales prices in the calculation of HICP, in 2000 because of the group of other countries (RES), and in 2001 because of Italy and Spain.

To pick up the incorporation of Greece in 2001 and the introduction of the euro in 2002, two step dummies, denoted Euro and Greece, are included in the system. These two variables were also included in the cointegration relationships with the objective of capturing the change in the constant term. The stability analyses show stable values from September 1996 and the null of no cointegration is rejected in favour of at least two long-run restrictions. The restricted estimated cointegration relationships are very similar to the long-run restrictions derived from the analysis of total HICP in the five countries.

A Vector Autoregression Model with Equilibrium-Correction Mechanism for the five components is also estimated and the results shown in Table 1. The model includes seasonal dummies in all equations, as well as dummies to pick up the effect of including

sales prices in the equation for other countries (RES) from 2000 and in the equations for Spain and Italy from 2001; $CI1_t$ and $CI2_t$ represent the cointegration relationships.

Insert Table 1

ARIMA models with the same dummy variables of the model in Table 1 are also constructed for the five core price indexes. These are used in Section 4 for comparing the forecasting performance of the different models.

3.2 ANALYSIS BY FIVE RESIDUAL HICPs

This section deals with the vector comprised of the residual-harmonized indexes in Germany, France, Italy, Spain, and other countries (RES); the analysis is elaborated taking into account the latest available HICP figures corresponding to December 2002.

As discussed, the evolution of energy consumer prices depends to a great extent on the behaviour of international crude oil prices. At the same time, unprocessed food prices in the latter part of the sample experienced some movements due to very specific factors which must also be included in the models. In what follows, an analysis including all these factors is discussed.

To pick up the effects derived from the unprocessed food crisis, as a consequence of animal diseases and poor harvests due to adverse weather conditions, a detailed study of these prices has been realized in each of the five countries and in the euro area as a whole. As a result, a new dummy variable is considered – denoted as unpfint – which captures the intervention analysis on unprocessed food prices in the euro area derived from the effects of several level shifts in France, Germany, Italy, and the group of other countries (RES).

The analysis also includes dummies to pick up the effect of the introduction of the ecological tax in Germany. Commenced in April 1999, the ecological tax reform

affects motor fuel and electricity prices. Finally, the Brent Crude Oil price in euros is also included. The stability tests point to stable results from February 1996. The standard statistics and estimates for Johansen's procedure reject the null hypothesis of no cointegration in favour of at least one cointegration relationship.

The weak exogeneity test statistics indicate that the speed of adjustment corresponding to the Brent Crude Oil price could be zero. Thus, in the following estimation of the Vector Equilibrium Correction model, this variable is considered as exogenous. The model also includes seasonal dummies in all equations and dummies to pick up the effect of sales prices in the specific equations mentioned above; CI_t represents the cointegration relationship. The variables corresponding to the five harmonized-residual indexes and Brent Crude Oil price are taken in logs. The estimation results are in Table 2.

Insert Table 2

The contemporaneous coefficient for Brent in all equations makes sense because this indicator is known immediately and the HICP appears one month after the reference month. As in the previous section, ARIMA models with dummy variables for the five residual price indexes have also been estimated.

4. FORECASTING INFLATION IN THE EURO AREA

This section evaluates and compares the forecast performance of the ARIMA and VEqCM models proposed in the previous sections. To attain this objective, three forecast exercises are made. All of these models contain different dummy variables to take account of specific events on prices, including: (a) the entrance of Greece in 2001; (b) the introduction of sales prices in the group of other countries (RES) in 2000; (c) the

introduction of sales prices in Spain and Italy in 2001; (d) the incorporation of the euro in 2002; and (e) the unprocessed food crisis caused by animal diseases and poor harvests, among others.

For all of these events, with the exception of (e), it was known beforehand when they were to occur and that they would cause a structural break in the evolution of prices. Therefore, the forecast exercises are performed assuming that these effects are known by the inclusion of appropriate dummies in each model with coefficients estimated using the whole sample.

In this section the models are denoted by acronyms with three terms (x,y,z). The term x can take the letter U or V. The letter U refers to univariate models, possibly enlarged with dummy variables or leading indicators, and the letter V to vector models. The term y can take the symbols, A, 5S, 5C, 10 or 10BD. The first symbol corresponds to a single-equation model for the aggregate variable and the remaining symbols to vector models for the components: 5S when the components correspond to five sectors, 5C when corresponds to five countries, and 10 and 10BD when the components correspond to two sectors in five countries, with BD indicating that the model is restricted to be block-diagonal. Finally, the term Z takes a value of one when the models do not include leading indicators, a value greater than one when they include the Brent Crude Oil prices, a value of two when the Brent Crude Oil prices are forecast using future market price values, a value of three when the Brent Crude Oil prices are forecast by means of a univariate model, and a value of four when it is assumed that Brent Crude Oil prices are forecast without error. Future market prices forecast spot prices of Brent crude oil less accurately than univariate models and the following tables report only

results based on univariate forecasts. Altogether we consider eight different aggregated or disaggregated models with four alternatives for the leading indicators.

The first exercise takes into account the models of Section 2 which use a disaggregated analysis by sectors or by countries. The second considers the models which use data disaggregated by sectors within countries. Finally, the third includes the models which incorporate the Brent Crude Oil price as a leading indicator. In the first two cases, the exercise compares the forecasting accuracy of the different models with a univariate aggregate model, U.A.1; in the third case the models are compared with an aggregate univariate model enlarged with the Brent Crude Oil variable, model U.A.3.

Table 3 shows the forecast errors for the year-on-year rates of total HICP derived from the disaggregated analysis by countries, by sectors, a combination of both approaches, and the U.A.1 univariate model for the total HICP in the euro area.

The models were re-estimated until December 1999 to consider a forecasting period of more than three years, January 2000–July 2003, so as to evaluate the performance of the forecasts from one to 12 periods ahead.

The $\dfrac{MSE(i)}{MSE(univ)}$ ratio compares the forecast accuracy of the different models with respect to the aggregate univariate formulation. A less than unit value indicates an improvement with respect to the aggregate univariate model. The Diebold–Mariano statistic with respect to the baseline model is also reported.

Insert Table 3

The first columns in table 3 shows that, for one and three periods ahead, the best results are obtained when disaggregating by sectors, but for a medium run the geographical disaggregation yields the best forecasts. Therefore, further disaggregation

18

by countries and sectors seems promising. Another important result shows that vector equilibrium correction models provide more accurate forecasts than the corresponding univariate models in all cases. This final result focuses on the relevance of considering cointegration relationships to pick up the long-run restrictions between the variables. Regarding the Diebold–Mariano test, statistically significant differences can be found in favour of the disaggregated methodology for all periods.

Finally, although disaggregations are based on different information sets, the forecast errors for different periods do not differ in a relevant way between the two approaches; therefore, the combination of the forecasts derived from both approaches, based on the average of the forecasts following Granger and Jeon (2004), makes sense and yields the best results (see the last column in Table 3), suggesting that both disaggregations matter.

Table 4 shows the forecast errors for the year-on-year rates of total HICP derived from the two approaches that take into account the disaggregation by the combined criterion of sectors and countries.

Insert Table 4

Table 4 shows that, from two steps ahead, the best results are obtained when disaggregating by countries and by sectors introducing the block-diagonal restriction; this confirms the previous conclusions: (1) vector models forecast more accurately than single equation models; and (2) the Diebold–Mariano tests find statistically significant differences among the aggregated and disaggregated approaches in favour of the latter for all periods. A comparison of the results from Tables 3 and 4 reveals the importance of disaggregating by sectorial blocks in different countries to improve the forecasts. The

Diebold–Mariano tests show that the forecasting differences are significant in favour of this disaggregation for long horizons.

Table 5 presents the forecast errors for the year-on-year rates of total HICP derived from the three approaches that take into account the Brent Crude Oil prices in euros. The first approach considers a single equation leading-indicator model for the total HICP in the euro area (U.A.3). The second one takes into account the sectorial disaggregation of the euro area in five components – processed food, non-energy industrial goods, services, energy, and unprocessed food – in which the equation for the energy price index includes the leading indicator. In this case, we have a univariate model, U.5S.3, and a vector model, V.5S.3. The third considers two sectorial blocks, core and residual, in five countries, where the residual vector includes the Brent Crude Oil price in euros. The models entertained are a univariate model, U.10.3, and a vector model V.10BD.3.

It must be noted that this indicator, international Brent Crude Oil prices in euros, is known immediately while the HICP appears one month after the reference month. Therefore, the indicator is available for forecasts one period ahead, and its introduction in the system reduces the RMSE. For longer horizons, forecasts derived from a univariate model for the indicator are used.

The results in Table 5 show that, in forecasts made one period ahead, the errors for the year-on-year rates of total HICP derived from the approaches including the Brent Crude Oil price have smaller RMSE than the errors made with models without the leading indicator because former models can use observed values of the indicator for these forecasts.

Insert Table 5

From a two-period horizon onwards, the indicator must be forecast and the forecasting advantage of the models with the indicator soon disappears. These results reflect the difficulty in forecasting crude oil prices and are conditional to the linear models used. Preliminary work done by the authors suggests that different conclusions could be obtained by switching regimes models for the indicator.

Comparing all of the previous forecasting results, it emerges that the best strategy to forecast inflation in the euro area is to use the VEqCM model, constructed for a disaggregation of two sectors and five geographical regions, incorporating the block-diagonal restriction between equations for core and residual prices. For horizons one and two only, it is convenient to enlarge the model to include international Brent Crude Oil prices in euros as a leading indicator.

5. CONCLUSIONS

This paper shows that, using inflation data for the euro area, the breakdown of a macro-variable of an economic area by sectors in different countries or in groups of countries can be useful in forecasting the aggregate, mainly if a vector model with cointegration restrictions is used. Models built on data obtained by disaggregating only by sectors or by countries produce less accurate forecasts. Even when the combination of forecasts from both sources of data improve accuracy, it is still lower than those obtained after considering the double criterion of sectors and countries. This result could also apply to a disaggregation by sectors within geographical, political or administrative regions in a national economy.

In general, the n-components vector model is better estimated by imposing restrictions. Two types appear to be of practical relevance. One reduces the number of

sectors by defining them more broadly and groups the regions so as to obtain a more moderate number of components. The other imposes a block-diagonal structure to the vector model. Economic theory and cointegration analysis between all possible pairs of components could be useful for specifying the restricted model. These disaggregated block-diagonal vector models could also benefit from the inclusion of specific leading indicators for particular components and by using dummy variables to capture the effects of specific events affecting individual components. Finally, in addition to the advantages of disaggregation in forecasting, disaggregated forecasts are useful for policy because they inform which components have higher expected rates of growth and how persistent shocks are in the different sectors and regions.

With inflation in the euro area, this paper shows that for the sample period analysed, 1990–2003, the hypothesis that prices – HICP – in the euro area are I(1) is not rejected. The annual inflation rate fluctuates around a mean of 1.9%. This cannot be considered as an equilibrium value in the long run, because special future shocks could change the observed historical mean, but it shows that the inflation mean is not perturbed with the innovations arriving every month, as the I(2) hypothesis for HICP would imply. The target adopted by the ECB that inflation in the medium run should not be above 2% is compatible with this result.

In modelling the HICP in the euro area, the breakdown of the aggregate by countries and sectors in n components is important. Disaggregating only by countries or only by sectors at the euro area level is not enough to capture the different trend factors in prices. The paper proposes a disaggregation in five geographical regions and in two sectors, core and residual. For each region the prices in these sectors are not cointegrated and a block-diagonal structure by sectors is proposed to simplify the model

specification. This implies that core inflation is not a good leading indicator for total inflation. The interest in core inflation as a macroeconomic indicator lies in the fact that it is computed from prices in which innovations are more persistent. A VEqCM model on the disaggregation by countries and sectors with the block-diagonal restriction gives more accurate forecasts than aggregate models, or disaggregations based only on countries or sectors. The forecasting difference turns out to be more important with the length of the forecasting horizon in favour of the double-criterion disaggregation scheme.

REFERENCES

Bulletin of EU and US Inflation and Macroeconomic Analysis, Universidad Carlos III, Madrid, Spain. www.uc3m.es/boletin.

Compendium of HICP. Reference Documents. Eurostat. March 2001.

Diebold, FX, Mariano, RS. 1995. Comparing predictive accuracy. *Journal of Business & Economic Statistics* **13**: 253–263.

Dickey DA, Fuller WA. 1981. Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root. *Econometrica* **49**(4): 1057–1072.

Espasa, A, Matea, ML, Manzano, MC, Catasus, V. 1987. La inflación subyacente en la economía española: estimación y metodología. *Boletín Económico del Banco de España*, marzo: 32–51.

Espasa, A, Senra, E, Albacete, R. 2002. Forecasting Inflation in the European Monetary Union: A Disaggregated Approach By Countries and By Sectors. *The European Journal of Finance* **8**: 402–421.

Granger, CWJ, Jeon, Y. 2004. Thick modelling. *Economic Modelling* **21**: 323–343.

Johansen, S. 1988. Statistical Analysis of Cointegration Vectors. *Journal of Economic Dynamics and Control* **12(2/3):** 231–254.

Johansen, S. 1991. Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models. *Econometrica* **59(6)**: 1551–1580.

Zellner, A. 2000. The Marshallian Macroeconomic Model. Professor Ryuzo Sato Research Conference Paper. September 2000.

Zellner, A, Israilevich G. 2003. The Marshallian Macroeconomic Model: A Progress Report. University of Chicago. Paper 0302. May 2003.

Zellner, A, Tobias, J. 2000. A Note on Aggregation, Disaggregation and Forecasting Performance. *Journal of Forecasting* **19(5)**: 457–469.