ECONOMETRIC MODELLING FOR SHORT-TERM INFLATION FORECASTING IN THE EMU.

Antoni Espasa and Rebeca Albacete*

nº: 03-43 (9)

Abstract

This paper faces the problem of forecasting monthly inflation in the euro zone by breaking down the HICP in different sectors through the different countries, grouping the eight smallest countries all together for the purpose of this work. The paper shows relevant features about the HICP, as the inflation in the sample considered is I (0), that there are significant seasonal breaks in the data and that the component prices are related by firm cointegration restrictions. For forecasting purposes the disaggregation by countries and sectors with the use of VEqCM model with a block diagonal restriction separating the different sectors provides the best results. They are significantly different for horizons one to twelve with respect to forecasts from aggregate models and for long horizons (around 12 months) with respect to other simpler disaggregations. International crude oil prices are only a good leading indicator for horizons one and two.

Keywords: disaggregated econometric modelling, cointegration, core inflation, residual inflation, leading indicators, VEqCM, seasonal breaks.

First version: February 2004
Revised version: July 2004

*Espasa, Departamento de Estadística, Universidad Carlos III de Madrid, C/ Madrid 126, 28903 Getafe (Madrid), e-mail: espasa@est-econ.uc3m.es; Albacete, Departamento de Estadística, Universidad Carlos III de Madrid, e-mail:albacete@est-econ.uc3m.es

1. INTRODUCTION

Agents in financial markets and monetary authorities between others demand frequent updates of inflation expectations. Monthly updates seem to be a good compromise because they can use new data on prices. This can be done using consumer prices indices (CPI). Like other previous studies, Espasa et al. (1987), (2002a) and (2002b) and Lorenzo (1997), this paper focus on monthly CPI. Inflation in the Economic Monetary Union is directly measured by the Harmonised Indices of Consumer Prices (HICP). Other measures are available, such as the GDP or consumption deflators, but they are not based directly and exclusively on price data. Thus, even though they cover more of the economy than the HICP, they are not so widely used as relevant inflation measures. Besides, these alternative measures can only be broken down into a relatively small number of components, at least when they are originally published. Since an important aim of this paper is to study the question of whether prices in different markets follow a single common trend or not, a degree of detail on the disaggregation of the price indicator is important. Finally, macroeconomic deflators are only available on a quarterly basis and monthly updates are in high demand to forecast inflation, and recent price information turns to be more important for these updates than old information on explanatory variables. Also, the accuracy in forecasting monthly inflation is relatively high. For all these reasons, the paper centres on the HICP.

Inflation is usually presented as a big aggregate through markets and certainly through geographical regions. The aggregation poses the question if for forecasting and diagnosis purposes one should consider disaggregated data and in the affirmative case which are the limits of disaggregation.

Disaggregation is one way to increase information in the forecasting process. What matters is that this additional information must be relevant and with the possibility of an adequate econometric treatment. The economic theory suggests that this could be the case. In considering a market disaggregation, it can be noted that markets are not homogeneous. By the demand side, the consumer's preferences change differently through markets and there are different information levels (captive demand). And by the supply side, the degree of incorporating technological innovations differs through sectors and there are different possibilities for the entrance of new suppliers and for quality improvements. So, the economic theory supports that in the medium term prices through markets will have different trends. If these economic arguments matter different trends in empirical data will be found. The argument can also be extended to seasonality, cycles and volatility.

Inflation in the Euro-zone can be analysed by breaking down the aggregate HICP in different ways. Three of them are as follows. One refers to the breakdown into price indexes corresponding to large groups of markets, denoted as sectors in this paper, throughout the Economic Monetary Union as a whole; another considers the HICP by national aggregates, and the last approach takes into account both types of information and breaks down the HICP by components corresponding to different sectors in each country. The two first disaggregations are analysed in Espasa et al. (2002a) and this paper focus on the last one. In the first and last cases the determination of the sectors is not arbitrary, but constructed so that each price component is an index for a (sector) group of markets in which prices are relatively homogeneous. Therefore these breakdowns must be guided by economic theory, taking into consideration the demand and supply characteristics of each sector and confirming the suitability of the decomposition by appropriate data analysis. For national economies these three disaggregation alternatives also apply, considering regions (states) instead of countries.

The above mentioned disaggregations can be interesting for many countries and economic areas like the EMU because in general it turns out that prices show, between other things, different trends. This means that in a vector of *n* different price components they are restricted by some, but not by a possible maximum number *(n-1)* of cointegration relationships. In this respect, it could be said that the different price components are cointegrated, if there is at least one long-term restriction, but not fully cointegrated. The absence of full cointegration between the *n* elements of a vector time series implies that the *n* trends in the component time series are generated by more than one common (trend) factor, (see Escribano and Peña, 1994) and this indicates that there is no full convergence between the components, in this case between the different prices.

If in an aggregate of *n* components there is more than one common trend then the univariate models for components will have different trend formulations. Consequently, it can be expected that the model for the aggregate will have a certain complexity due to restrictions in the trend. These restrictions could be more important at some times than others and in those cases could not be easily detected, because the complexity is related to information on the components which is concealed in the aggregate. Thus, too simple univariate models are usually obtained for macroeconomic indicators. The experience certainly shows that the possible complexity and restrictions implied by the components are difficult to capture modelling directly the aggregate. By contrast, modelling the components with simple parsimonious models, a complex model with important restrictions could result for the aggregate. Consequently, the gain in efficiency can be large with disaggregation.

Evidence of this complexity can be seen comparing the forecasting functions for the aggregate derived from the disaggregated and aggregate models. If in both cases the models are linear, the forecasting functions sooner or later collapse to deterministic linear functions. In general, the forecasting function from an aggregate model collapses forecasting constant annual growth sooner that the forecasting function from the corresponding disaggregated model.

The lack of full cointegration between prices implies that the innovations in the aggregate will have different long-run effects depending on the common trend from which they mainly stem. In these circumstances disaggregation is interesting because is a way to increase the information about future price trends. Besides the disaggregation proposal is also convenient for similar arguments on the seasonality of price 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 data properly, then the disaggregated models could be wrong 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, for some components at least, the linear approximation in modelling would not be supported by data, requiring non-linear structures which could be quite difficult to construct.

In linear modelling outliers could correspond to aberrant observations or to very important points (VIP's). The last ones indicate that the linear model is wrong and some non-linear alternative is required.

Strong nonlinearities in some components have much reduced importance in the aggregate, but they could matter a lot at some local points in time. Therefore, at the aggregate level it could be difficult to detect and model nonlinearities, but they are present. Consequently, disaggregating could be an easier way to capture nonlinearities in the aggregate, and therefore a way to end up with a more reliable and efficient model.

The impulse response functions in non-linear models are not constant along time. They depend on the previous history, the sign and the magnitude of the innovation. Therefore, it matters to specify from which component the innovation comes from because is going to have different short and long-term effects.

Forecasting inflation can be approached in different ways: a) time series models; b) leading indicator models; and c) congruent models (Hendry 2001). The disaggregated analysis can be applied with each type of models. When using models with explanatory variables, the disaggregation allows specific effects of the common explanatory variables in each equation and the inclusion of specific explanatory variables in each equation.

Stock and Watson (1999) present a sophisticated application on USA data based on a leading indicator constructed using a large number of macroeconomic time series and following the methodology described in Stock and Watson (2002). But in this model the leading indicator is not cointegrated with inflation, which in USA is I(1). In other words, the innovations have persistent effects in inflation. Aggregate indicators very often are constructed with weights, which are very different from those used to construct CPI. The disaggregation approach also has the advantage that it is possible to consider leading indicators for the price components and the indicators can be general ones with different effects throughout components or specific indicators for each price component.

It is possible to establish different inflation modelling alternatives according to the economic theory employed. Some of the most relevant models proposed in the literature are commented below.

- Monetary models are based on monetary theories, such as the quantity theory of money or theories that relate inflation with interest differentials. Stock & Watson (1999) show that this kind of models generate worse forecasts than models employing real indicators.
- Mark-up models. The economic theory behind these models - Duesenberry (1950), Richards & Stevens (1987) and Franz & Gordon (1993) -, supposes that in the long-run, the level of consumer prices represents a mark-up on the total unit costs, including unit labour costs, import prices and energy prices. An equilibrium correction specification of this theory can be seeing in De Brouwer & Ericsson (1998) and Banerjee et al. (2001). It is difficult to construct these models at monthly level. In any case explanatory variables as labour costs and import prices are observed with delay.
- Models based on a Philips curve, which relate changes in inflation with past changes of inflation and past values of unemployment gap – i.e., the difference between the unemployment rate and the NAIRU. Stock & Watson (1999, 2002) employ a generalised Philips curve including real activity measures such as production capacity, commodities rate of growth, etc., instead of an unemployment rate. Stock & Watson (1999) show that predictions derived from this generalised Philips curve are more accurate than the forecasts obtained from the conventional Philips curve.

  Stock & Watson (1999) construct a leading indicator, which they propose that can be used as a proxy for unemployment gap in a Phillips curve model. This is intended to join leading indicator and economic-theory-based models. In general is more appropriate to see the leading indicator models as some sort of reduced form econometric model.

In any case, disaggregation is also of interest in this context, because allow different relationships for prices of different markets.

In monthly inflation modelling it is difficult to construct congruent models because several explanatory variables are not observed with this periodicity. In order to improve inflation forecasting results, the approach taken in this paper – originated in Espasa et al. (1987), (2002a), (2002b) and Lorenzo (1997) – is a disaggregated econometric modelling in order to make use of more information, starting by increasing the amount of information on prices themselves. The idea is that the behaviour of prices through different markets and countries is sufficiently diverse in trend, seasonality, short-term oscillations and erraticity, that forecasting results can be considerably improved if all this information is taken into consideration. Having established the interest in increasing information by disaggregating the HICP, a subsequent step of this methodology deals with the question of introducing general and specific leading indicators for each price component and finally formulating non-linear structures when required. Certainly there is evidence about non-linearity in the consumer energy prices.

In economic forecasting, congruent econometric models would be the most useful ones because they provide forecasts and an economic explanation of them. But there are drawbacks in the use of congruent models in short-term forecasting inflation. In order to incorporate properly the monthly information on prices, monthly models are needed. This means that some explanatory variables, which are not observed with this periodicity, must be replaced by some proxies (indicators) and the resulting monthly model could have less theoretical support. The quality of data for monthly prices is relatively high, but this is not the case for many leading indicators, which appear as explanatory variables in inflation models. A good forecasting performance of a congruent model requires accurate forecasts of the explanatory variables (or indicators) and this is not always the case.

Hendry 2001 shows that a single inflation theory does not explain inflation in the UK and an eclectic model, which includes all theories, demand-pull, cost-push, monetary, imported inflation, etc, is needed. In more precise terms, see Hendry (2001), what is required is a model which includes as explanatory variables the disequilibria in different markets: monetary and financial markets, labour markets, good and services markets, international markets, etc. If for a given sector a certain theory or certain disequilibria are more important than others, then the proposal made in this paper is in accordance with these results, capturing in part the causal diversity through the disaggregation.

On the basis of the above argument, having different inflation forecasts for the different sectors provide the analyst with solid clues to propose a causal explanation for the aggregated forecast inflation. So, disaggregation is also useful for economic policy. It can give more accurate forecasts for the aggregate, and since the forecasts include different trend projections for prices in different sectors, it shows which sectors are provoking undesirable forecasts for the aggregate. In this sense, if an expected drop in inflation in the EMU at a particular time is due to the fall in energy and unprocessed food markets, it could be said that the factors behind inflation fall in the EMU are mainly foreign.

Causal factors of inflation could be the output gap, money supply, salaries, productivity, Euro-dollar exchange rate, etc. Given the uncertainty of all these explanatory variables, although they could explain adequately the sample period, they could generate forecasts with high uncertainty if one considers that the relevant confidence intervals correspond to ex-ante predictions. A subsequent paper could deal with different alternatives for causal analysis with monthly models, formulating congruent monthly models or relating the monthly forecasts from time series or leading indicators models with forecasts from a quarterly congruent model.

The econometric models in the methodology proposed in this paper are not causal ones. They are reduced form models if including leading indicators or final form models if only past information on prices is used. In both cases they can incorporate long-run restrictions between price components. All models in this paper can be obtained from a structural Marshallian Macroeconomic Model with n-sectors, see Zellner (2000) and Zellner and Israilevich (2003) and references in them.

Aggregate linear models forecast constant growth relatively soon with increasing confidence intervals. Nevertheless, disaggregating and/or non-linear modelling can capture interesting oscillations and constant growth forecast appears on longer horizons.

This forecasting approach has been successfully applied to the USA, EMU, Spain and some Spanish regional economies. The advantages of this approach increase when price components are not fully seasonally cointegrated or important differences exist in the stationary behaviour of their cyclical and short-term oscillations.

The paper is organised as follows. In section 2 some considerations for building monthly forecasting models for Euro-zone inflation are made. Section 3 studies the number of country components required in the disaggregation. Section 4 presents a disaggregated econometric modelling of HICP by two sectors in five economic areas. Section 5 shows an approach based on block-diagonal vector models and develops the modelling of specific effects, such as the incorporation of Greece and the entrance of the Euro, and the introduction of international indicators in energy prices. Section 6 analyses the forecasting performance of the models proposed in previous sections and finally section 7 presents the conclusions.

2. CONSIDERATIONS FOR BUILDING MONTHLY FORECASTING MODELS FOR EURO-ZONE INFLATION: DISAGGREGATION AND ECONOMETRIC METHODOLOGY

After detecting the importance of the breaking down CPI’s by groups of markets, see Espasa et al. (1987), (2002a) and (2002b) and Lorenzo (1997), and that disaggregation by countries is also required, see Espasa et al. (2002a), one could use doubly indexed panel data and study aggregate inflation by considering a price index for each big group of markets in each country. The two previous breakdown alternatives have the interest of being simpler. The fact is that modelling a panel data for sectors and countries is not going to be easy, because the heterogeneous behaviour of each price index in the panel can not be reduced to a fixed or random effect. This heterogeneity includes different responses to the cointegrated restrictions and different transitory dynamics. In any case, the most complex question arises from the fact that, as it is shown in this paper, there are cointegrated relationships between sectors and countries. So, this paper focuses on vector models with difference equations.

For the Euro-area, a minimum of six sectors and twelve countries, as suggested in the above references, represent a large number of components, so the econometric model for the resulting vector is unfeasible. Instead of facing such an approach this paper considers an analysis that takes into account just four or five geographical components and two (aggregated) sectors, and develops two ways of analysing this information. Jointly in a vector composed by the HICP’s of the two sectors in each geographical component or separately in two vectors taking into account country-component vectors for each one of the two sectors.

A geographical breakdown of the EMU in a maximum number of four of five components, can be done, for instance, in one of the following two ways. One is to consider the set of four or five big countries in the EMU, assuming that they are quite representative of the whole set. This was done in Espasa et al. (2002a), where the analysis by countries only covers France, Germany, Italy and Spain, with a global weight in the Euro-zone inflation around 80%. The question is that the group of eight remaining countries could affect the cointegration relationships. The alternative taken in this paper is to consider an aggregate of the above mentioned remaining eight countries, denoted by Rest in what follows, and work with a five-component geographical disaggregation formed by the four big countries and the Rest. For the purpose of this paper the Rest will also be called country. In this five time series there is no full cointegration between them. The lack of full cointegration appears as an indicator of convergence problems within the EMU. As it will be shown below, this analysis with five geographical components is more robust than the previous study with four countries and estimates more cointegration relationships than previously.

The breakdown of HICP by sectors is approached in Espasa et al. (2002a) by taking into account theoretical considerations about differences in supply and demand, which could result in prices having different trends. Following Espasa et al. (1987) this led to consider the following price indexes corresponding to five sectors: (1) Processed food, (2) Non-energy industrial goods, (3) Services, (4) Unprocessed food and (5) Energy.

Given a minimum number of countries, four or five, and taking into account the dimension of the sample, this paper considers a less disaggregated scheme based in just two big sectors, one of them derives from the aggregation of the above first three and the other one results of the sum of the last two. So, the breakdown of HICP by sectors in each geographical area has been approached considering the following two price indexes corresponding to: (I) Overall Index excluding energy and unprocessed food, and (II) Energy and unprocessed food. With (I) a core inflation measure can be calculated, and through (II) a residual inflation rate is composed. This core measure proposed in Espasa et al. (1987) is also used by the European Central Bank, and this also give a rational for this disaggregation. This paper illustrates that innovations are more persistent in the price index from which core inflation is obtained than in the price index corresponding to residual inflation. In fact, this and no others, seem to be the main reason for using core inflation as an economic indicator. Thus the disaggregation used in this paper also faces the utility of underlying measures of the total inflation.

The weights of the HICP derive from Household Budget Surveys, so their revision could reflect a change in the consumer preferences. Therefore, the disaggregation approach seems quite convenient because it incorporates the information of changing weights of core HICP and residual HICP in total HICP.

Considering five countries and two sectors could result excessive, so it will be necessary to check if a vector model composed by ten equations could be approached through a block-diagonal vector model. So, the original model could be simplified by two vector models, each one with five endogenous variables.

This paper has been elaborated in a sequential way. Therefore, the models become more complex attending to the detection of problems in the data which require more sophisticated econometric techniques. Consequently, the samples used in estimating different models are different because of the sample size and, what is more important, for significant updates done by Eurostat to introduce the consideration of sales in the construction of the HICP. Besides, the larger samples have problems, like the effect of the incorporation of Greece to the EMU or the introduction of the Euro, which were not present in shorter ones. In all cases, the samples start in January 1996, but due to the stability problems mentioned below in all cases some observations are lost at the beginning.

This fact has a special relevance in the case of the introduction of sales prices[1] in the measure of the HICP. The inclusion of sales into the HICP has changed the seasonal pattern of the indexes, and therefore, of the month-on-month rates. Germany and France have included sales prices into their HICP since the start of the HICP, but Belgium incorporated them into its HICP in 2000 and Italy and Spain in 2001.

This section also shows the econometric methodology applied in all the subsequent analysis in order to analyse the Euro-zone inflation using monthly models. In what follows the different steps composing this methodology are described.

1. Integration analysis: Before modelling the HICP's, it is necessary to determine the orders of integration for the variables considered. The order augmented Dickey-Fuller (1981) statistics for the variables in logarithms and for their first differences point out that all variables considered in this study appear to be integrated of order 1 - I(1) - with the hypothesis of a second unit root being rejected in all cases. Therefore, the inflation variables in this paper are stationary. This integration analysis is applied assuming constant seasonality. In all cases the correlograms of the first difference for the variables in logarithms also support this evidence. The alternative of stochastic seasonality with the necessity of testing all the seasonal roots is not considered in this study due to the small sample size.

[1] The term sales prices also consider the price reductions derived from offers. For the introduction of sales prices in the HICP see "Compendium of HICP. Reference Documents". Eurostat. March 2001.

2. Time Dimension: Given the low number of observations, the analyses began with a VAR model of order 5 using the logarithmic transformation of the price indexes in levels as dependent variables and the models were specified with a constant term and seasonal dummies. Appropriate tests show that the fifth-order VAR model can be reduced in all cases to a first-order VAR.
3. Stability tests: There is some sort of instability at the beginning of the sample period and recursive estimations for the presence of the cointegration relationships, β vector of cointegration parameters, α adjustment parameter values and the eigenvalues, have been applied all along the paper[2]. Stability is not rejected from certain month of 1996, depending on the model considered.
4. Cointegration analysis: Cointegration analysis helps to clarify the long-run relationships between integrated variables. The Johansens' (1988, 1991) procedure for finite-order vector autoregressions (VAR's) is applied. The greatest eigenvalue and trace eigenvalue statistics, corresponding to the first-order VAR, reject in all cases the null of no cointegration relationship in favour of some number of cointegrating relationships which always is lower than *n-1*. So, this analysis indicates that there exists certain number of restrictions between the harmonised consumer prices and, at the same time, that there is not full cointegration between prices, pointing out the presence of more than one common factor driving the price trend in the Euro-zone.
5. Weak exogeneity test: The weak exogeneity test statistics for the variables in logarithms indicate that the speed of adjustment corresponding to some price

[2] These recursive estimations and other additional information, which do not appear in the paper, are available by request to the authors.

indexes can be taken as zero and these results are used to simplify the models.

6. Imposition of additional restrictions in order to obtain an exactly identified model.
7. Estimation of the restricted Vector Autoregression Model with Equilibrium-Correction Mechanism.
8. Comparison of the fits and forecasting performances of the vector models with the ones from corresponding univariate models.

3. STUDY OF A GEOGRAPHICAL DISAGGREGATION FOR THE INFLATION IN THE EMU

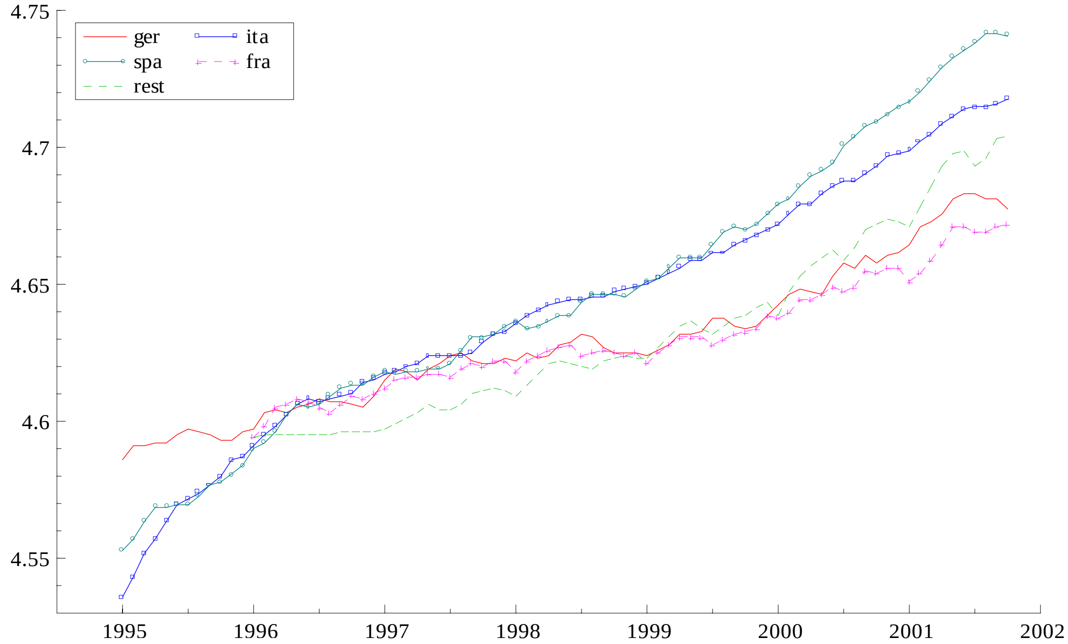
In this paper total HICP of the Euro-zone is broken down in the HICP's corresponding to the five geographical components mentioned in the previous section: Germany (GER), France (FRA), Italy (ITA), Spain (SPA) and Rest (RES).

The main purpose in this section is to study if Rest could be considered as exogenous or not.

Table A1 in the appendix gives the weights for different EMU countries in the calculation of HICP for the years 1996 to 2002. In 2001, Greece joins to the EMU. This table shows that the HICP's corresponding to Germany and France have lost weight along the sample period.

Graphs of the five indexes can be found in figure 1, which shows that the HICP's of Germany and France move together and also the HICP's of Italy and Spain. The HICP for Rest is in the middle.

Figure 1: ***Harmonised Indices of Consumer Prices in different countries in logarithms***



Source: Eurostat

The sample used in this section goes from January 1996 to March 2001[3]. Stability appears after September 1996 and this is the date that has been considered for the estimation of the HICP for these countries.

The greatest eigenvalue and trace eigenvalue statistics, corresponding to the Johansens' (1988, 1991) cointegration analysis, reject the null of at least one cointegration relationship in favour of at least two cointegrating relationships. This last hypothesis is not rejected in favour of a hypothesis with more than two cointegration relationships. So, this analysis indicates that there exist two long-run restrictions between the country consumer prices.

[3] Eurostat till March 2001only published HICP for France and Austria from January 1996 and this fact determined the beginning of the sample used

The estimated restricted cointegration relationships could be expressed as:

$$\text{(1) } \log(\text{HICP RES}) - \underset{(0.26)}{1.13} \log(\text{HICP FRA}) = \underset{(0.05)}{0.58} \log(\text{HICP SPA}) - \underset{(0.21)}{0.65} \log(\text{HICP GER});$$

$$\text{(2) } \log(\text{HICP ITA}) - \underset{(0.16)}{0.88} \log(\text{HICP SPA}) = \underset{(0.90)}{4.00} \log(\text{HICP FRA}) - \underset{(0.72)}{3.53} \log(\text{HICP GER});$$

where standard errors of estimates are in brackets.

The cointegration relationships show restrictions between relative prices. One, the relation number (2), is like the unique relationship found in Espasa et al. (2002a) which yields that relative prices between Italy and Spain, two countries characterised by a higher inflation but lower level prices, equals to relative prices between France and Germany to the power 4. In the other long-run restriction enters Rest and could be interpreted in the sense that relative prices between Rest and France equals to relative prices between Spain and Germany to the power 0.6.

The considered variables are price indexes, but price levels in the reference period are different among countries. The nominal convergence among countries will lead to a greater price growth for Italy, Spain and Rest, which are countries with lower price levels. The cointegration relationship could show that this convergence process is taken place.

Espasa et al. (2002a) make a bivariate cointegration analysis of original CPI's, for France, Germany, Italy and Spain. They found that there are no cointegration relationships between any other pair of countries other than those composed by Germany and France and Italy and Spain. The second cointegration relationships shown above points out that the joint evolution of the logarithmic trends in the prices of France and Germany is proportional to the corresponding joint evolution in Italy and Spain.

Binary cointegration analysis between each of the eight countries in the Rest group and each of the four big countries, show that, in a set of countries, say Rest 1, formed by Austria, Belgium, Finland, Luxembourg and Greece, which weight 56% in the Rest, each country is cointegrated with the four big countries. On the other hand for the set of countries, say Rest2, formed by the Netherlands, Portugal and Ireland, no element of the set shows a cointegration relationship with the big countries

The above results shows that Rest should be considered in the geographical disaggregation of EMU inflation, because enters in one cointegration relationships and it is not exogenous, see below. At the same time five geographical areas seems to be a minimum number, because a further breakdown of Rest in Rest1 and Rest2 could be of interest. Unfortunately the sample size available prevent us of going beyond five elements in this decomposition.

A Vector Autoregression Model with Equilibrium-Correction Mechanism for the five countries has been estimated and results are shown in table 1. The model also includes seasonal dummies and CI_1 represents the first cointegration relationship and CI_2 the second one.

Table 1: ***VEqCM model for countries.***

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0.19L & 1+0.31L & 0 & -0.29L \\ 0 & 0 & 0 & 1-0.28L & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \Delta\log(HICPGER)_t \\ \Delta\log(HICFRA)_t \\ \Delta\log(HICITA)_t \\ \Delta\log(HICPSPA)_t \\ \Delta\log(HICPRES)_t \end{pmatrix} = \begin{pmatrix} 0.0011 \\ 0.0010 \\ 0.0017 \\ 0.0020 \\ 0.0015 \end{pmatrix} +$$

$$\begin{pmatrix} -0.58 \\ 0 \\ -0.29 \\ -0.31 \\ -0.37 \end{pmatrix} (CI_{1t-1} + 0.53) + \begin{pmatrix} -0.14 \\ 0 \\ 0 \\ -0.13 \\ 0 \end{pmatrix} (CI_{2t-1} - 0.24) + \begin{pmatrix} a_{1t} \\ a_{2t} \\ a_{3t} \\ a_{4t} \\ a_{5t} \end{pmatrix} + Seasonal \quad dummies$$

This model shows that there is not much cross-dependence between the variables in the short run, except for the equation corresponding to Italy. The residual covariance matrix indicates that only the contemporaneous correlation between HICP in Germany and France seems important (0.74).

This VEqCM model shows that a disaggregating analysis of HICP by countries carried out by separate single-equation models would be inefficient. Nevertheless, for forecasting purposes, ARIMA models for the HICP of each country could be entertained. Univariate models for the variables in logs for these five indexes are summarised in table 2.

Table 2: ***Univariate ARIMA models for countries HICP in logarithms.***

	Difference order	Constant	Stationary structure	Seasonal Dummies
GER	1	0.0010	White noise	Yes
FRA	1	0.0009	White noise	Yes
ITA	1	0.0013	$\frac{1}{(1-0.26L^2)}a_t$	Yes
SPA	1	0.0020	$\frac{1}{(1-0.46L)}a_t$	Yes
RES	1	0.0016	White noise	Yes

Table 3 shows the standard residual deviations with degrees of freedom correction from the VEqCM and ARIMA models. The VeqCM gets a better fit than univariate one.

Table 3: ***Standard residual deviations for total inflation equations in different countries***

	VEqCM	Univariate ARIMA
GER	0.20%	0.21%
FRA	0.17%	0.17%
ITA	0.10%	0.11%
SPA	0.13%	0.14%
RES	0.16%	0.15%

The breakdown of HICP by sectors is approached in Espasa et al. (2002a). This analysis has been updated with a longer sample and it is obtained the same result of only one cointegration relationship between sectors.

4. A DISAGGREGATED ECONOMETRIC MODELLING OF HICP BY TWO SECTORS IN FIVE ECONOMIC AREAS

In order to disaggregate the HICP in the EMU in a vector of small dimension, a first intent to join the both types of information sets, sectors and countries, consisted in taking into account the two sectors – core and residual HICP's – in the four big countries. The sample used in this analysis went from January 1996 to July 2001.

For this vector the null of no cointegration is rejected in favour of at least three cointegration relationships, showing that there is not full cointegration between the harmonised core and residual indexes in different countries. The restricted estimated cointegration relationships include in all cases a mixture of core and residual prices.

Given the results of section 3 the previous analysis should be extended by incorporating the core and residual HICP of the Rest, concluding with a vector of ten components. For this study a longer sample till June 2002 was used.

The plots of the five core indexes and the five residual indexes are in section 5 in figures 9 and 11, respectively. In the same section figures 2 to 7 show the plots of core and residual indexes for the EMU and the five countries considered here.

The stability analysis shows stable results from March 1996 and the null of no cointegration is rejected in favour of at least three cointegration relationships. Therefore, the inclusion of Rest does not increase the number of long-run restrictions in this case.

The weak exogeneity test statistics indicate that the speed of adjustment corresponding to the core HICP of Germany and to the residual HICP of Germany, Spain and Rest could be zero. The core HICP for Italy, Spain and Rest and the residual HICP for France rejects the weak exogeneity test at 5%.

The restricted estimated cointegration relationships can be written as shown table 4:

Table 4: ***Restricted cointegration relationships considering two sectors HICP in five countries***

Variable	GERC	GERR	FRAC	FRAR	ITAC	ITAR	SPAC	SPAR	RESC	RESR
first	1.00	0.00	0.00	-0.24	-2.83	-0.92	2.69	0.22	-1.20	0.87
vector	-	-	-	(0.24)	(0.48)	(0.28)	(0.53)	(0.27)	(0.35)	(0.27)
second	0.00	1.00	0.00	2.28	9.88	4.41	-4.09	-1.39	-2.84	-6.58
vector	-	-	-	(1.44)	(2.93)	(1.71)	(3.21)	(1.62)	(2.10)	(1.65)
third	0.00	0.00	1.00	-0.01	-0.29	0.04	0.30	-0.02	-0.69	0.04
vector	-	-	-	(0.03)	(0.07)	(0.04)	(0.08)	(0.04)	(0.05)	(0.04)

(1) Series are taken in logarithms

(2) Standard errors reported in parentheses

The cointegration vectors can be interpreted as follows. The first cointegration relationship can be approached by,

$$[GER(CORE) - RES(CORE)] - [ITA(RESIDUAL) - RES(RESIDUAL)] = 2.8[ITA(CORE) - SPA(CORE)]$$

It mixes core and residual indexes and it can be interpreted saying that core relative prices between Italy and Spain to the power 2.8 equal the core relative prices between Germany and the Rest divided by the residual relative prices between Italy and the Rest.

The second vector mainly relates residual indexes and expresses the core index in Italy as a function of residual indexes in Germany, Italy and Rest. And finally, the third relation can be approached by,

$$[RES(CORE) - FRA(CORE)] = 0.3[SPA(CORE) - ITA(CORE)]$$

It only involves core indexes and faces the differential between core indexes in Rest and France and the differential between core indexes in Spain and Italy.

As in the previous section, the variables corresponding to Rest can not be considered as exogenous. This argument favours the extension to a vector of ten components.

A Vector Autoregression Model with Equilibrium-Correction Mechanism for the ten components has been estimated and results are shown in table 5. The model also includes seasonal dummies and CIa_t, CIb_t and CIc_t represents the cointegration relationships.

*Table 5**: VEqCM model for core and residual inflation in five countries***

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0.57L & 0 & 0 & 0 & -0.55L \\ 0 & 0 & 1+0.26L & 0 & -0.17L & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1+0.25L & 0 & 0 & 0 & 0 & 0 & -0.62L \\ 0 & 0 & 0 & 0 & 1+0.25L & 0 & -0.43L & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1+0.20L & 0 & 0 & 0 & -0.17L \\ 0 & 0 & 0 & 0 & 0 & 0 & 1-0.35L & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & -0.46L \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1+0.41L & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.25L & 0 & 1-0.36L \end{pmatrix} \begin{pmatrix} \Delta GERCt \\ \Delta GERRt \\ \Delta FRACt \\ \Delta FRARt \\ \Delta ITACt \\ \Delta ITARt \\ \Delta SPACt \\ \Delta SPARt \\ \Delta RESCt \\ \Delta RESRt \end{pmatrix} =$$

$$\begin{pmatrix} 0.0009 \\ 0.0018 \\ 0.0010 \\ 0.0018 \\ 0.0019 \\ 0.0018 \\ 0.0023 \\ 0.0012 \\ 0.0017 \\ 0.0027 \end{pmatrix} + \begin{pmatrix} -0.05 & -0.01 & 0.00 \\ 0.00 & 0.00 & 0.00 \\ 0.00 & -0.01 & -0.30 \\ 0.00 & 0.00 & -1.13 \\ 0.00 & 0.00 & 0.00 \\ 0.32 & 0.03 & -1.27 \\ -0.17 & -0.03 & 0.00 \\ 0.00 & 0.00 & 0.00 \\ 0.00 & -0.02 & 0.00 \\ 0.00 & 0.00 & -0.69 \end{pmatrix} \begin{pmatrix} CI_{at-1}+1.89 \\ CI_{bt-1}-12.25 \\ CI_{ct-1}-1.74 \end{pmatrix} + \begin{pmatrix} a1t \\ a2t \\ a3t \\ a4t \\ a5t \\ a6t \\ a7t \\ a8t \\ a9t \\ a10t \end{pmatrix} + Seasonal \quad dummies$$

Residual contemporaneous correlations greater than 0.25 are shown in table 6.

Table 6: ***Correlation matrix of residuals derived from the VEqCM model for core and residual inflation in five countries***

Variable	**ΔGERC**	**ΔGERR**	**ΔFRAC**	**ΔFRAR**	**ΔITAC**	**ΔITAR**	**ΔSPAC**	**ΔSPAR**	**ΔRESC**	**ΔRESR**
ΔGERC	1.00									
ΔGERR	-	1.00								
ΔFRAC	0.31	-	1.00							
ΔFRAR	-	0.74	-	1.00						
ΔITAC	-	-	-	-	1.00					
ΔITAR	-	0.48	-0.28	0.40	-	1.00				
ΔSPAC	-0.25	-	-0.30	-	0.51	-	1.00			
ΔSPAR	-	0.60	-	0.64	-	0.46	-	1.00		
ΔRESC	-	-	-	-	-	-	-	-	1.00	
ΔRESR	-	0.84	-	0.81	-	0.46	-	0.64	0.29	1.00

This model shows that the long run equilibrium equations enter in seven equations, concretely in four core inflation equations corresponding to Germany, France, Spain and Rest, and three residual inflation equations of France, Italy and Rest.

There is more contemporaneous correlation between these residuals than in the breakdown by sectors or by countries (see Espasa et al. 2002a). The correlations are greater between the residuals derived from the residual inflation equations than between the residuals coming from the core inflation equations. A possible interpretation could be that energy prices, concretely fuel prices, depend on the OPEC policy and in the sector of unprocessed food, prices are very influenced by the weather. All these constitute general conditions, which may be affect to different countries at the same time. The greatest contemporaneous correlation between the residuals derived from the residual inflation equations is the corresponding to Germany and Rest, and in the case of core inflation, between Italy and Spain. The cross-correlograms for the residuals do not show significant values.

Once again, there is not much dependency among the endogenous variables in the short-run. Nevertheless, this dependency is more extended than in the previous approaches. Almost all the variables depend on their own lag, and residual inflation equations for all variables depend on the lagged residual inflation in Rest, which turns to be a sort of leading indicator for other residual price indexes. It can be concluded that extending the disaggregation framework allows capturing more dynamic relationships.

The estimated univariate ARIMA models for each price index are summarised in table 7. Table 8 shows the standard residual deviations with degrees of freedom correction in both approaches. A drawback of these results is that for the residual price index of the Rest the fit in the VEqC model is worst than the univariate fit, and this price enters in all equations for residual prices of other countries.

Table 7: ***Univariate models for core and residual HICP in logarithms in five countries***

	Difference order	Constant	ARIMA structure	Seasonal Dummies
GERC	1	0.0009	$(1\text{-}0.26L^2)\ a_t$	Yes
GERR	1	0.0023	white noise	Yes
FRAC	1	0.0010	$1/(1+0.29L)\ a_t$	Yes
FRAR	1	0.0017	white noise	Yes
ITAC	1	0.0018	$(1\text{-}0.48L^2)\ a_t$	Yes
ITAR	1	0.0017	$1/(1\text{-}0.28L^2)\ a_t$	Yes
SPAC	1	0.0022	white noise	Yes
SPAR	1	0.0022	$1/(1\text{-}0.49L+0.27L^2\text{-}0.18L^3)\ a_t$	Yes
RESC	1	0.0012	white noise	Yes
RESR	1	0.0020	white noise	Yes

Table 8: ***Standard residual deviations for core and residual inflation equations in five countries***

	VEqCM	Univariate ARIMA
GERC	0.13%	0.12%
GERR	1.00%	0.95%
FRAC	0.10%	0.12%
FRAR	0.75%	0.76%
ITAC	0.19%	0.18%
ITAR	0.35%	0.38%
SPAC	0.24%	0.26%
SPAR	0.52%	0.51%
RESC	0.14%	0.17%
RESR	0.69%	0.45%

In all the equations, but the residual price index in Rest, the VEqCM fit is similar or better than the univariate one, as in the initial study with four countries and two sectors. Nevertheless now the residual standard deviations for Italy and Spain are worse than in the vector with eight components. A deeper analysis of these results points out that the extension of the sample carried out to estimate the 10-equation model is not neutral for the fit, because the new sample is affected by the consideration of sales prices in Italy and Spain since January 2001, but first published by Eurostat in January 2002 and substantially revised in April this year.

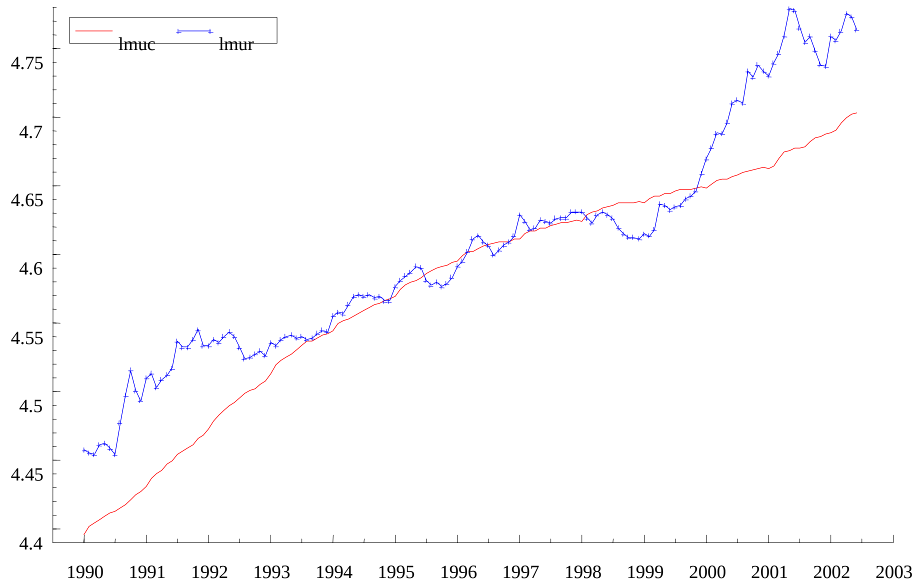
In table 8 it can be seen that the fits of core prices are better than the ones corresponding to residual prices. Being particularly bad for the case of Germany, France and Rest. Since in this simultaneous equation system the contemporaneous correlations are important and there are several cross-equation dynamic effects, a bad specification in one equation translates to others. The results in this section indicate that based on asymptotic tests, data in a given small sample reject a separate modelling of core and residual prices. At the same time the model obtained with the sample available can not be considered appropriate because the bad fit of some equations. In these circumstances the theoretically preferable 10 equations system should be simplified, but maintaining the disaggregation level of the 10 variables, which allows capturing more adequately the different trend factors in the HICP for the EMU. This is studied in the next section.

5. DISAGGREGATED ECONOMETRIC MODELLING OF HICP BY SECTORS AND BY COUNTRIES ASSUMING BLOCK-DIAGONAL VEqC MODELS

A first aim in this section is to study if pairs of core and residual prices share a common trend or not. This analysis is applied to the Euro-zone as a whole and to each one of the five countries, using a sample that ends in June 2002.

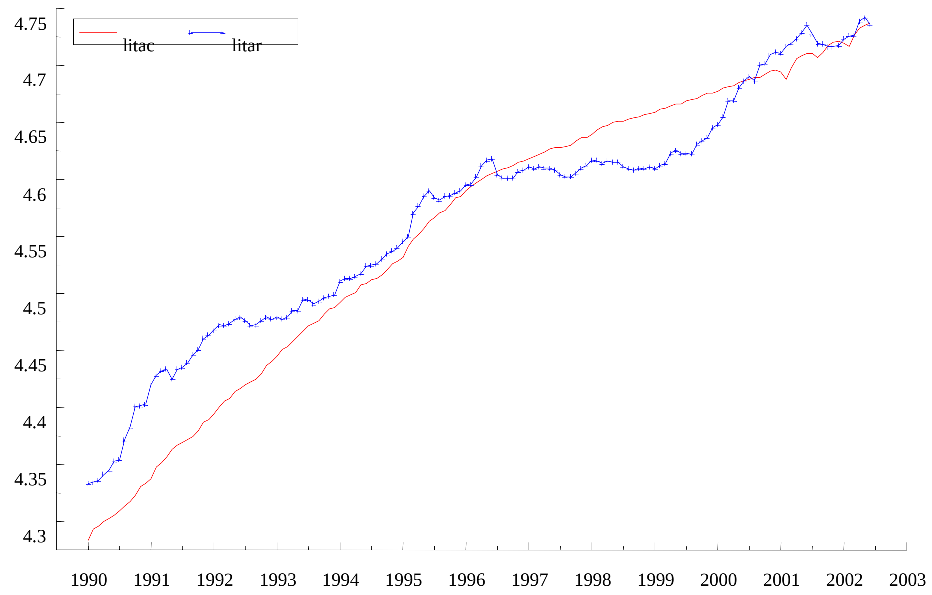
In order to calculate the core HICP and residual HICP corresponding to Rest, their weights in total HICP of each country are needed. Tables A2 and A3 in the appendix show these weights corresponding to the years 1996-2002 and tables A4 and A5 give the weights of country core HICP in the core HICP for EMU and similarly for country residual HICP. Figures 2 to 7 show the graph of core and residual HICP for each country and the EMU[4].

Figure 2 ***Core and residual Harmonised Indices of Consumer Prices in the EMU in logarithms***



Source: Eurostat Date: 17 July, 2002

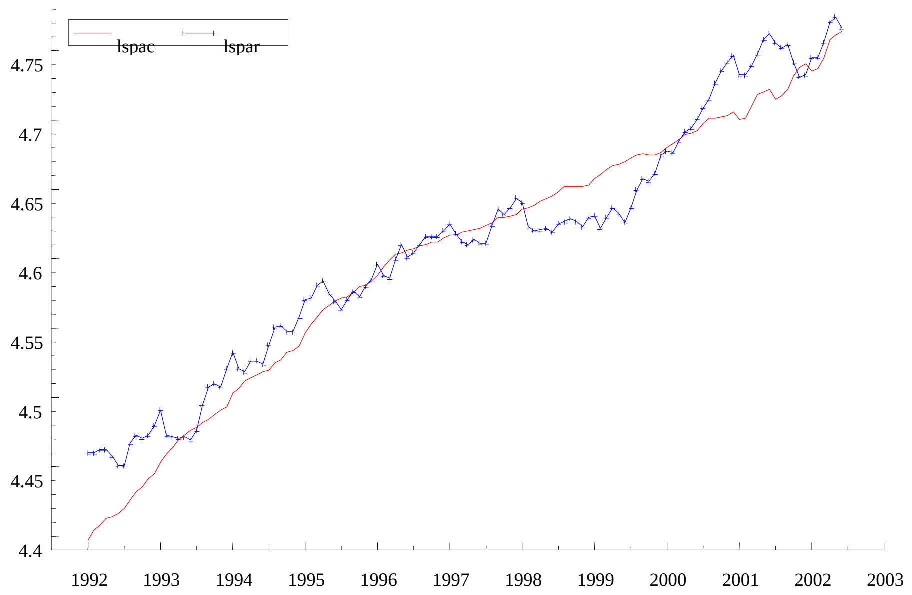
Figure 3: ***Core and resiual Harmonised Indices of Consumer Prices in Italy in logarithms***



Source: Eurostat

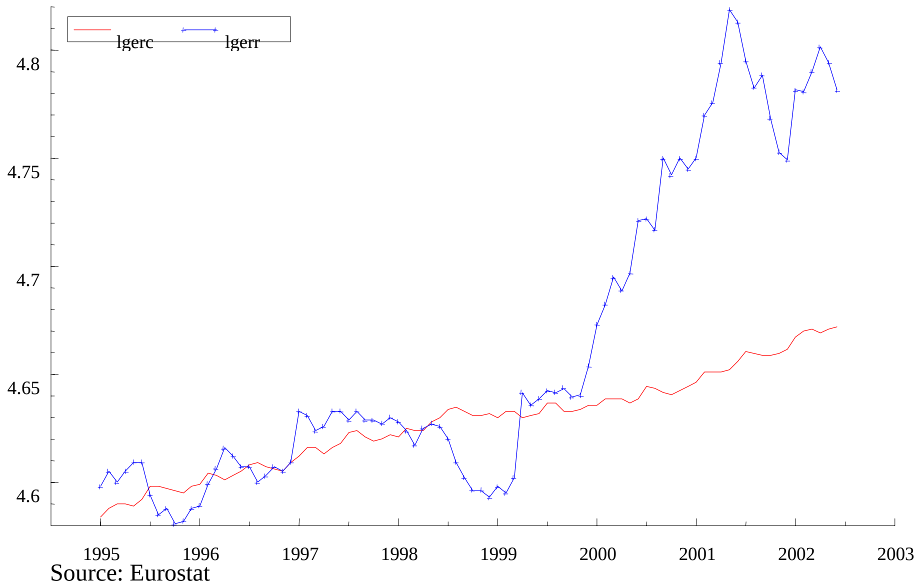
Date: 17 July, 2002

Figure 4: ***Core and residual Harmonised Indices of Consumer Prices in Spain in logarithms***



Source: Eurostat Date: 17 July, 2002

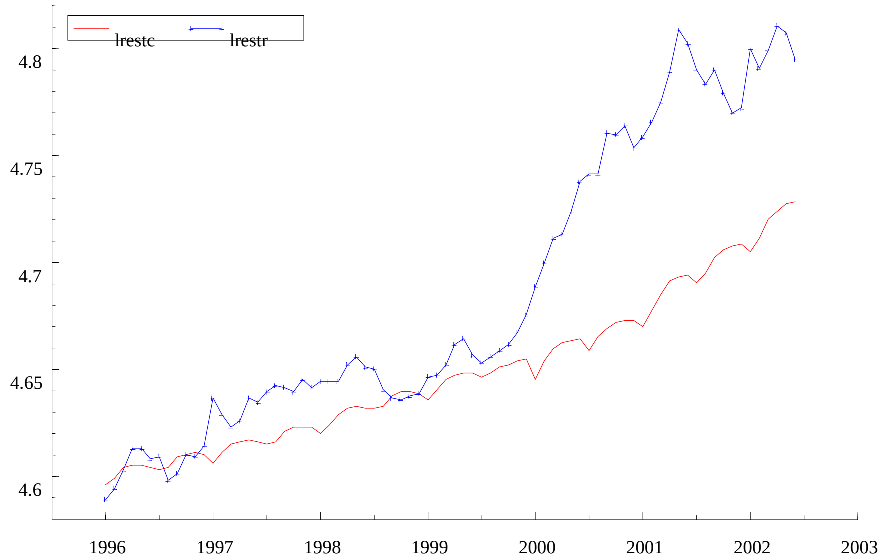
Figure 5: ***Core and residual Harmonised Indices of Consumer Prices in Germany in logarithms***



Source: Eurostat

Date: 17 July, 2002

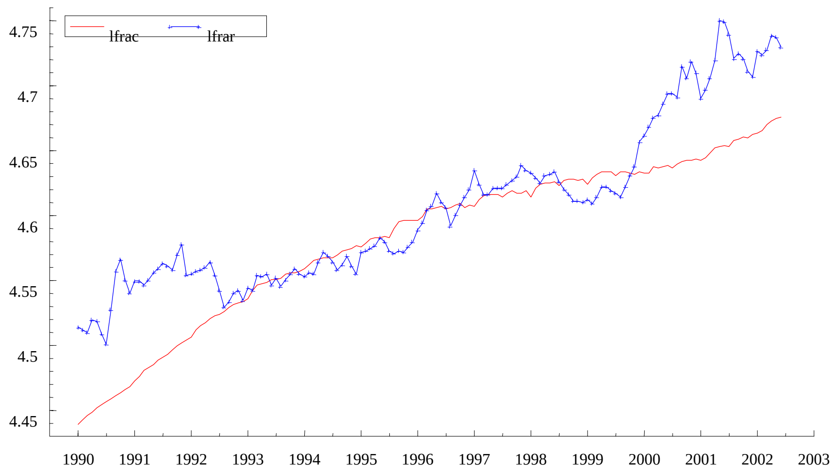
Figure 6: ***Core and residual Harmonised Indices of Consumer Prices in Rest in logarithms***



Source: Eurostat

Date: 17 July, 2002

Figure 7: ***Core and residual Harmonised Indices of Consumer Prices in France in logarithms***



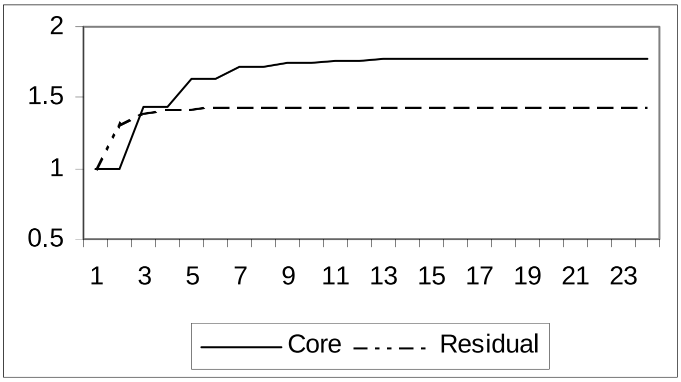
Source: Eurostat

Date: 17 July, 2002

The Johansens' (1988, 1991) cointegration analysis shows the absence of cointegration between the core HICP and the corresponding residual HICP in all cases - EMU, France, Germany, Italy, Spain and Rest - pointing out the presence of two common factors in each pair of core and residual price indexes.

An 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. In fact, the breakdown of a vector variable like HICP when there is not full cointegration between its components is in any case important for diagnosis purposes. For instance, an innovation from services prices, properly weighted, does not have the same implications in total CPI as one from non-processed food prices. In fact, institutions which perform monthly inflation analyses occasionally alert readers by claiming that unexpected inflation in a given month is particularly worrying because it comes from prices included in the core inflation index, like prices of services. In other cases, these institutions could refer to an innovation of the same magnitude in the CPI as not being particularly important because it comes from the set of prices corresponding to residual inflation, like unprocessed food prices. To illustrate this point, figure 8 shows the impulse response function to an innovation in core and residual consumer prices in the Euro zone. Their effects settle gradually to 1.8 and 1.4, respectively. This result support the use of core inflation as an economic indicator, because it separate from the global consumer price index, prices with lower long-run multiplier effects. Therefore the interest on core inflation does not relay much on being a good predictor of total inflation, but on being an additional economic indicator formed by prices on which innovations are persistent with a higher multiplicative effect.

Figure 8: ***Response function to an innovation in the core and residual price indexes***



To simplify the vector model with ten equations of the last section, the results from the mentioned bivariate cointegration analysis are indicative. Given the absence of cointegration relationships in all the pairs of core and residual indexes it could be considered to approximate the 10-component vector by two separate vectors, one composed by the five core indexes and another by the five residual indexes. This implies a block diagonal formulation for the original 10-component vector. This restriction is not truly correct because even when pairwise there are not cointegration relationships between core indexes and residual indexes, it was seen in section 4 that considering the ten components jointly the cointegration relationships include some mixture of core and residual indexes. But even in this situation, from the three cointegration relationships one was between core indexes only and another related almost exclusively residual indexes. So it can be concluded that even when the block-diagonal restriction is not fully supported by data it can be seen as 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.

5.1 ANALYSIS BY FIVE CORE HICP's

In this section a vector composed by the core-harmonised indexes in Germany, France, Italy, Spain and the Rest is considered. All this analysis has been elaborated taking into account the last available figures of HICP corresponding to December 2002. Graphs of the five indexes and their year-on-year rates of growth can be found in figures 9 and 10.

Figure 9: ***Core Harmonised Indices of Consumer Prices in different countries in logarithms***

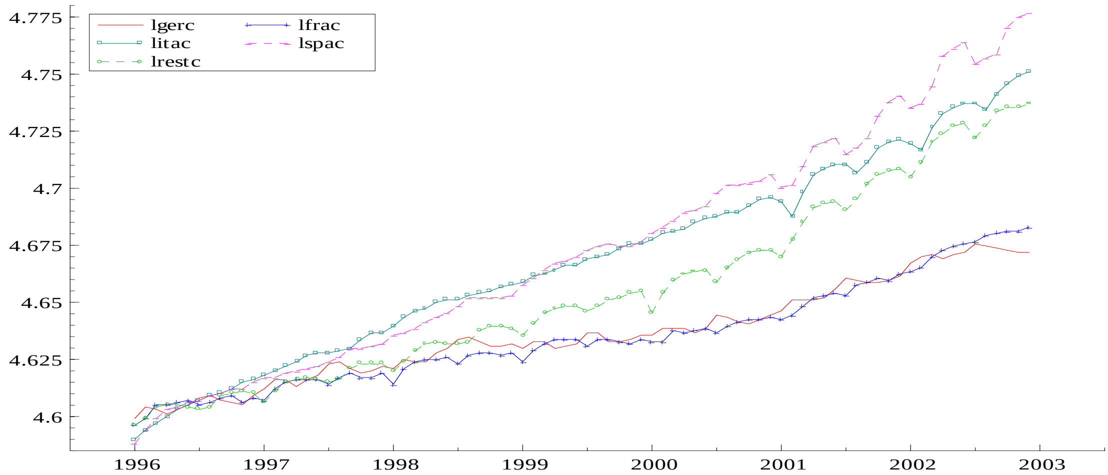
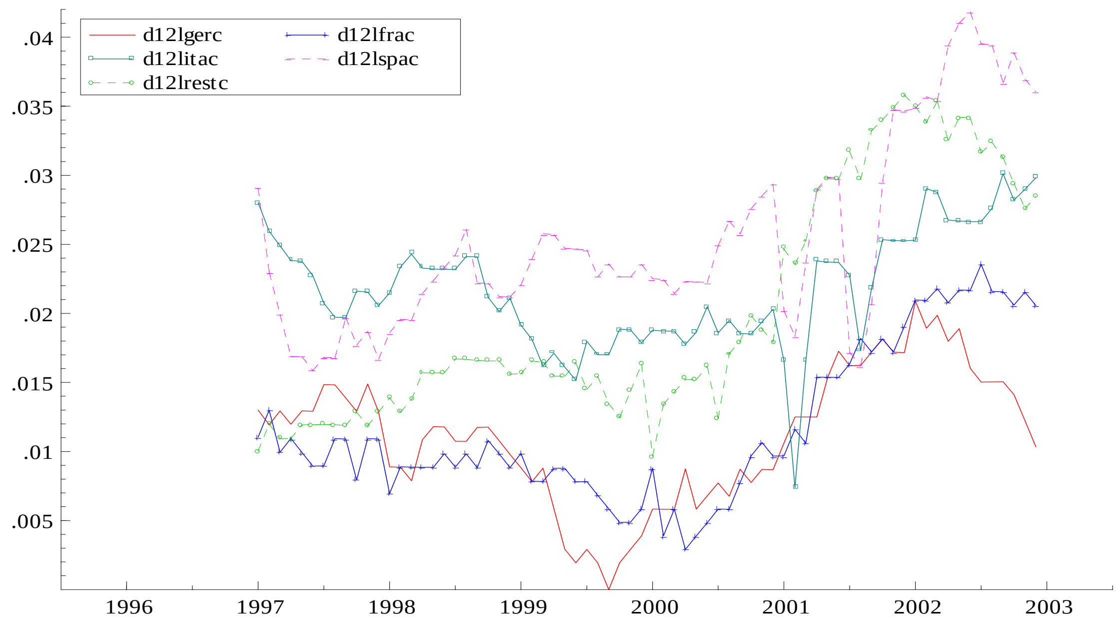


Figure 10: ***Year-on-year rates of core inflation in different countries.***



Source: Eurostat

As it happens in figure 1 for HICP, figures 9 and 10 show that the indexes of Germany and France seem to move together and also the indexes of Italy and Spain. The indexes for Rest are in the middle. Another feature revealed in figure 9 is the big impact of the introduction of sales prices in the calculation of HICP, in 2000 because of Rest and in 2001 because Italy and Spain.

The stability analyses show stable results from March 1996.

The standard statistics and estimates for Johansen's procedure reject the null of no cointegration in favour of at least one cointegration relationship, indicating the lack of full cointegration between the harmonised core indexes in different countries. But a change in the constant of the cointegration relationship appears in 2001, due probably to the introduction of Greece in the EMU in 2001.

In order to pick up the incorporation of Greece in 2001 and the introduction of the Euro in 2002 two step dummies, denoted as Euro and Greece, were included in the system. These two variables were also included inside the cointegration relationships with the objective of capturing the change in the constant term.

Now, the null of no cointegration is rejected in favour of at least two cointegration relationships. The stability analyses show stable values from September 1996.

The restricted estimated cointegration relationships can be written as:

$$\underset{}{\log(gerc)} + 14.71 \log(frac) - 5.53 \log(itac) + \underset{(0.99)}{4.63 \log(spac)} - \underset{(1.48)}{8.09 \log(restc)} - \underset{(0.02)}{0.08\ Euro} - \underset{(0.02)}{0.03\ Greece};$$

$$-0.80 \log(gerc) - 0.28 \log(frac) + \log(itac) - \underset{(0.12)}{0.68 \log(spac)} + \underset{(0.18)}{0.27 \log(restc)} + \underset{(0.002)}{0.01\ Euro} + \underset{(0.003)}{0.003\ Greece};$$

If more restrictions are imposed we can obtain:

$$\{\log(restc) - \log(frac)\} = 0.5 \{\log(spac) - \log(gerc)\} + 0.001\ Euro + 0.006\ Greece;$$

$$\{\log(itac) - \log(spac)\} = 5.7 \{\log(gerc) - \log(frac)\} + 0.01\ Euro + 0.05\ Greece;$$

These cointegration relationships are very similar to the long-run restrictions derived from the analysis of total HICP in five countries. The first cointegration vector could be interpreted in the sense that relative core prices between Rest and France equals to relative core prices between Spain and Germany to the power 0.5 and the second one relates relative core prices between Italy and Spain whit relative core prices between Germany and France.

A Vector Autoregression Model with Equilibrium-Correction Mechanism for the five components has been estimated and results are shown in table 9. The model also includes seasonal dummies in all equations, and dummies to pick up the effect of including sales prices in the Rest equation from 2000 and in the equations for Spain and Italy from 2001; $CI1_{t,}$ and $CI2_t$ represent the cointegration relationships.

Table 9: ***VEqCM model for core inflation in different countries***

$$\begin{pmatrix} 1-0.07L-0.03L^2-0.07L^3 & -0.28L^3 & 0 & 0 & 0 \\ 0 & 1+0.34L & -0.16L & 0 & 0 \\ 0 & 0 & 1-0.07L+0.26L^2-0.07L^3 & 0.20L^2 & 0 \\ 0.36L^2 & -0.36L-0.39L^2-0.26L^3 & 0.49L & 1-0.14L-0.23L^2 & 0 \\ 0 & -0.27L & 0 & -0.12L & 1+0.21L-0.02L^2-0.36L^3 \end{pmatrix} \begin{pmatrix} \Delta GERCt \\ \Delta FRACt \\ \Delta ITACt \\ \Delta SPACt \\ \Delta RESCt \end{pmatrix} =$$

$$\begin{pmatrix} 0.0009 \\ 0.0010 \\ 0.0019 \\ 0.0022 \\ 0.0017 \end{pmatrix} + \begin{pmatrix} 0.00 \\ -0.02 \\ 0.00 \\ -0.02 \\ 0.00 \end{pmatrix} (CI1_{t-1}-30.99) + \begin{pmatrix} 0.13 \\ -0.12 \\ -0.09 \\ 0.23 \\ 0.00 \end{pmatrix} (CI2_{t-1}+2.26) + \begin{pmatrix} 0.0056 \\ 0.0030 \\ 0.0081 \\ 0.0050 \\ 0.0071 \end{pmatrix} (\Delta Euro) + \begin{pmatrix} 0.0000 \\ 0.0000 \\ 0.0000 \\ 0.0000 \\ 0.0084 \end{pmatrix} (\Delta Greece) + \begin{pmatrix} a1t \\ a2t \\ a3t \\ a4t \\ a5t \end{pmatrix} + \textit{Seasonal dummies} + \begin{matrix} \textit{Dummies for the} \\ \textit{introduction of sales} \\ \textit{in ITAC, SPAC and RESC} \end{matrix}$$

The greatest contemporaneous correlation between the residuals derived from the core inflation equations is the corresponding to Germany and Rest (0.39). The cross-correlograms for the residuals do not show significant values.

The estimated univariate ARIMA models for each price component are summarised in table 10 and include the same dummies considered in the VEqC model.

Table 10: ***Univariate ARIMA models for core HICP in logarithms in five countries***

	Difference order	Constant	Stationary structure	Euro effect	Greece effect	Seasonal Dummies	Sales Dummies
GERC	1	0.0008	$(1\text{-}0.13\ L^2\text{+}0.21L^3)\ a_t$	0.0056	-	Yes	-
FRAC	1	0.0010	$1/(1\text{+}0.27\ L)\ a_t$	0.0032	-	Yes	-
ITAC	1	0.0018	$(1\text{-}0.40\ L^2\text{+}0.58L^3)\ a_t$	0.0060	-	Yes	Yes
SPAC	1	0.0021	$(1\text{+}0.20\ L\text{+}0.27\ L^2)\ a_t$	0.0103	-	Yes	Yes
RESC	1	0.0016	$1/(1\text{-}0.29\ L^{2\text{-}}0.32\ L^3)\ a_t$	0.0048	0.0044	Yes	Yes

Table 11 shows the standard residual deviations with degrees of freedom correction in both approaches, which are very similar in both cases.

Table 11: ***Standard residual deviations for core inflation equations in different countries***

	VEqCM	Univariate ARIMA
GERC	0.12%	0.11%
FRAC	0.10%	0.12%
ITAC	0.08%	0.07%
SPAC	0.10%	0.10%
RESC	0.09%	0.08%

5.2 ANALYSIS BY FIVE RESIDUAL HICP's

A vector composed by the residual-harmonised indexes in Germany, France, Italy, Spain and the Rest is considered. This analysis has been elaborated taking into account the last available figures of HICP corresponding to December 2002. Graphs of the five indexes and their year-on-year rates of growth can be found in figures 11 and 12.

*Figure 11**: Residual Harmonised Indices of Consumer Prices in different countries in logarithms***

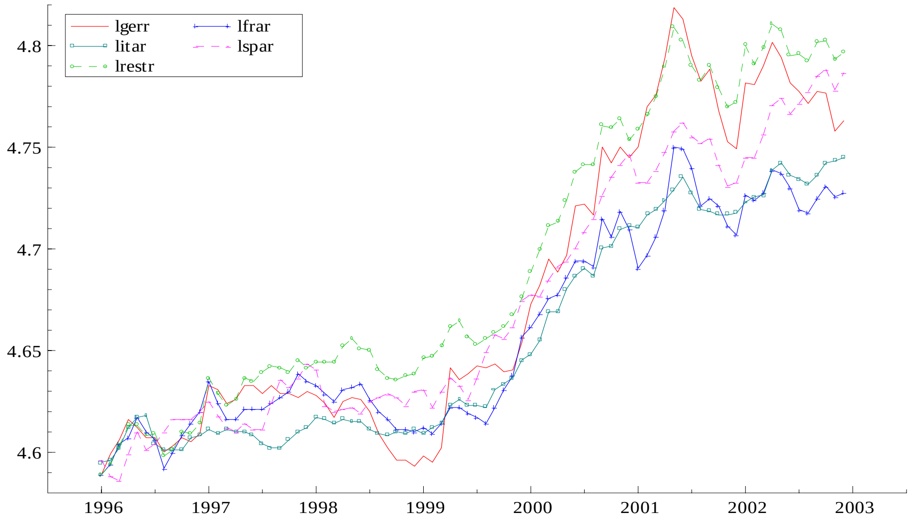
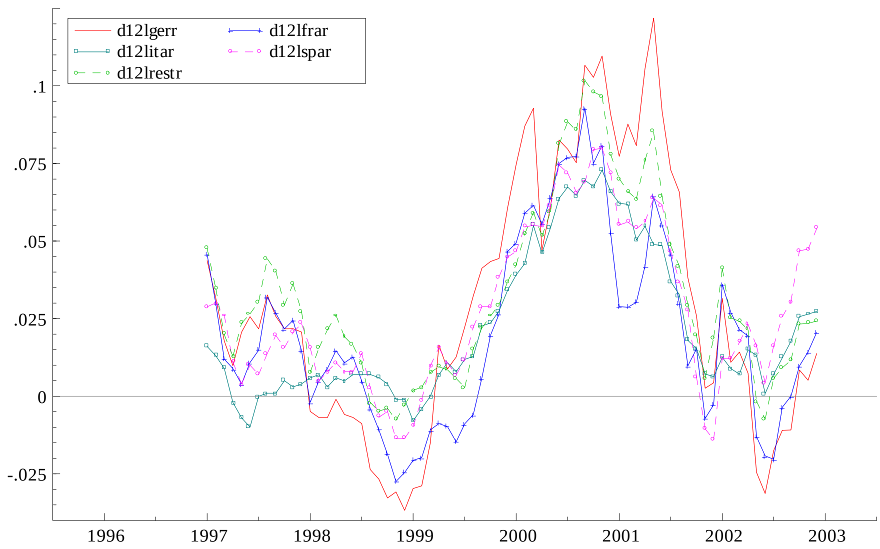


*Figure 12**: Year-on-year rates of residual inflation in different countries.***



Source: Eurostat

The stability analyses show stable results from July 1996 and the null of no cointegration is rejected in favour of at least one cointegration relationship.

The weak exogeneity test statistics for the variables in logarithms indicate that the speed of adjustment corresponding to the residual HICP of Germany, France, Spain and Rest could be zero.

The restricted estimated cointegration relationship can be written as:

$$\log(gerr) - \underset{(0.77)}{0.48} \log(frar) - \underset{(1.17)}{7.88} \log(itar) + \underset{(1.07)}{4.68} \log(spar) + \underset{(0.82)}{1.33} \log(restr);$$

if more restrictions are imposed one ends up with:

$$\log(gerr) = 13.18 \{\log(itar) - \log(spar)\}$$

So the cointegration vector relates residual index in Germany with the differential in residual indexes between Spain and Italy.

A Vector Autoregression Model with Equilibrium-Correction Mechanism for the five components has been estimated. The model also includes seasonal dummies in all equations, and the above dummies on sales; $euro_t$ represents a dummy to capture the introduction of Euro in 2002 and $CI_{t,}$ represents the cointegration relationship. In this case the entrance of Greece in the EMU has not had a significant effect.

As it was already mentioned, there is much more contemporaneous correlation between these residuals, than between the residuals from core inflation of the previous section. The cross correlograms only show significant contemporaneous correlations. The greatest contemporaneous correlation between the residuals derived from this model is between residual inflation in France and Rest and between Germany and Rest. All the variables depend on lagged residual inflation in Rest.

Comparing the standard residual deviations from the VEqC model with the corresponding ones from the univariate models in almost all the equations the VEqCM fit is worse than the univariate one. Also both approaches yield high residual standard deviations. As was aforementioned, the evolution of energy consumer prices depends in a great measure of the behaviour of international crude price. At the same time unprocessed food prices have experienced in the last part of the sample some movements in their evolutions due to very specific factors which must be included in the models. In what follows an analysis including all these factors is discussed. In order to pick up the effects derived from the unprocessed food crisis, as a consequence of the adverse weather conditions, a detailed study of these prices has been realised in each of the five countries and in the EMU as a whole. As a result of it, a new dummy variable it is considered - denoted as unpfint -, which captures the intervention analysis on unprocessed food prices in the EMU derived from the effects of several level shifts in France, Germany, Italy and Rest.

The following analysis also includes dummies to pick up the effect of the introduction of the ecological tax in Germany. This ecological tax reform affects motor fuel and electricity prices, started at April 1999 and continue in the following years. Finally the brent crude price in euros is also included in the analysis.

The stability tests point to stable results from February 1996.

The standard statistics and estimates for Johansen's procedure reject the null 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 price could be zero. Thus, in the following estimation of the Vector Equilibrium Correction model this variable is considered as exogenous.

The restricted estimated cointegration relationship can be written as:

$$\underset{}{\log(gerr)} + \underset{(0.42)}{1.59\log(frar)} - \underset{(0.52)}{0.63\log(itar)} + \underset{(0.56)}{2.40\log(spar)} - \underset{(0.43)}{2.89\log(restr)} - \underset{(0.02)}{0.17\log(brent)} - \underset{(0.42)}{1.09\,unpint};$$

If the restriction that the coefficient of log(itar) is not significant is imposed it can be obtained:

log(gerr) + 1.99 log(frar) + 2.35 log(spar) – 3.45 log(restr) -0.20 log(brent) – 1.22 unpint;

(0.50) (0.53) (0.51) (0.02) (0.50)

A Vector Autoregression Model with Equilibrium-Correction Mechanism for the five harmonised-residual indexes has been estimated and results are shown in table 12. The model also includes seasonal dummies in all equations, dummies to pick up the effect of sales prices in the specific equations mentioned above, and $CI_{t,}$ represents the cointegration relationship. The variables corresponding to the five harmonised-residual indexes and brent crude price are taken in logs. The model also includes dummies to pick up the effect of the ecological tax reform in the Germany equation.

Table 12: ***VEqCM model for residual inflation in different countries***

$$\begin{pmatrix} 1+0.23L & 0 & -0.28L^3 & 0 & -0.32L \\ 0.27L & 1-0.16L^2 & 0 & 0 & -0.46L \\ 0 & 0 & 1+0.27L-0.19L^2 & -0.37L & 0 \\ 0 & 0 & 0 & 1-0.27L+0.09L^2-0.10L^3 & 0 \\ 0.23L^3 & 0 & 0 & 0 & 1-0.27L^3 \end{pmatrix} \begin{pmatrix} \Delta GERRt \\ \Delta FRARt \\ \Delta ITARt \\ \Delta SPARt \\ \Delta RESRt \end{pmatrix} =$$

$$\begin{pmatrix} 0.0021 \\ 0.0017 \\ 0.0018 \\ 0.0023 \\ 0.0025 \end{pmatrix} + \begin{pmatrix} -0.09 \\ -0.06 \\ -0.05 \\ -0.07 \\ -0.08 \end{pmatrix} \left(CI_{t-1}-8.17\right) + \begin{pmatrix} 0.05+0.01L-0.02L^2 \\ 0.03+0.02L-0.02L^2 \\ 0.01+0.01L-0.01L^2 \\ 0.02-0.01L^2-0.01L^3 \\ 0.03-0.02L^2 \end{pmatrix} \left(\Delta BRENTt\right) + \begin{pmatrix} 0.69 \\ 0.86 \\ 0.00 \\ 0.38 \\ 0.67 \end{pmatrix} \left(\Delta unpint_t\right) + \begin{pmatrix} a1t \\ a2t \\ a3t \\ a4t \\ a5t \end{pmatrix} + \begin{matrix} \textit{Seasonal Dummies } + \\ \textit{Dummies for the} \\ \textit{introduction of sales} \\ \textit{in Italy, Spain and Rest} + \\ \textit{Dummies for the} \\ \textit{introduciton of ecological tax} \\ \textit{in Germany} \end{matrix}$$

The contemporaneous correlations for the residuals are shown in table 13.

Table 13: ***Correlation matrix of residuals derived from VEqCM model for residual inflation in different countries***

Variable	ΔGERR	Δ FRAR	Δ ITAR	ΔSPAR	ΔRESTR
ΔGERR	1.00				
Δ FRAR	0.62	1.00			
Δ ITAR	0.40	0.37	1.00		
Δ SPAR	0.15	-0.01	0.20	1.00	
Δ RESR	0.74	0.66	0.45	0.26	1.00

These contemporaneous correlations have lower values than those shown in the previous system without considering international indicators. The greatest contemporaneous correlations between the residuals derived from the residual inflation equations are the corresponding to Germany and Rest and France and Rest. The cross-correlograms for the residuals do not show significant values.

The estimated single-equation econometric models for each price component are summarised in table 14 and include the same dummies considered in the VEqC model.

Table 14: ***Single-equation econometric models for residual HICP in different countries***

	Difference order	Constant	Stationary structure	Transfer Function Brent	unpint effect	Seasonal Dummies	Sales Dummies
GERR[1]	1	0.0004	$1/(1-0.27L)\ a_t$	$(0.06+0.03L-0.02L^2)$	0. 84	Yes	-
FRAR	1	0.0004	$1/(1-0.26\ L^2)\ a_t$	(0.03+0.03L)	1.08	Yes	-
ITAR	1	0.0015	$1/(1-0.40\ L^2)\ a_t$	(0.01+0.02L)	-	Yes	Yes
SPAR[2]	1	0.0020	$1/(1-0.24\ L-0.27\ L^3)\ a_t$	(0.02+0.02L)	-	Yes	Yes
RESR[3]	1	0.0020	$1/(1-0.26\ L^{-}0.26\ L^3)\ a_t$	$(0.03+0.02L-0.01L^2)$	-	Yes	Yes

1. The model also includes the dummies that pick up the effect of the ecological tax reform.
2. As shown a previous analysis the euro effect is not significant either.
3. Although the unpint effect is not significant, the euro effect is considered whit a coefficient of 0.0176, as shown a previous analysis.

The contemporaneous coefficient in the transfer function with the brent in both approaches, univariate and multivariate, makes sense because this indicator is known immediately and the HICP appears one month after the reference month.

Table 15 shows the standard residual deviations with degrees of freedom correction. The adjustment is similar in both approaches.

Table 15: ***Standard residual deviations for residual inflation equations in different countries***

	VEqCM	**Single-equation**
GERR	0.60%	0.55%
FRAR	0.55%	0.50%
ITAR	0.31%	0.28%
SPAR	0.35%	0.31%
RESR	0.46%	0.40%

The standard deviations of the residuals of the 10-component vector in table 8 are higher than those from core and residual vectors in tables 11 and 15, but this is due to the inclusion of dummy variables in the latter case. In any case it seems that the block-diagonal restriction is a reasonable approximation.

6. FORECASTING INFLATION IN THE ECONOMIC MONETARY UNION

This section evaluates and compares the forecast performance of the ARIMA and VEqCM models proposed in previous sections. In order to attain this objective three forecast exercises are elaborated, which are limited to the models which have been estimated with longer samples – till July 2003. They are listed and denoted in table 16. All these models contain different dummy variables to take account of specific events on prices, like (a) the entrance of Greece in 2001, (b) the introduction of sales prices in Rest in 2000 and (c) the introduction of sales prices in Spain and Italy in 2001, (d) the incorporation of the Euro in 2002, (e) the unprocessed food crisis caused by animal diseases and poor harvests, among others.

Table 16: ***Models used in the forecasting exercise and terminology used***

Type of model		Brent prices (leading indicator, l. i.) and forecasting procedure			
		Without including l. i.	Including the l. i.[(a)] and forecasting it by		
			Future-market values	Univariate model	Using observed values
Univariate models	**Aggregated univariate model**	AU.1	AU.2	AU.3	AU.4
	Disaggregated univariate models for 5-sectors components of HICP	D5SU.1	D5SU.2	D5SU.3	D5SU.4
	Disaggregated univariate models for 5-countries components of HICP	D5CU.1	D5CU.2	D5CU.3	D5CU.4
	Disaggregated univariate models for 10 components of HICP (2 sectors by 5 countries)	D10U.1	D10U.2	D10U.3	D10U.4
VEqCM models	**VEqCM for 5-sectors components of HICP**	V5S.1	V5S.2	V5S.3	V5S.4
	VEqCM for 5-countries components of HICP	V5C.1	V5C.2	V5C.3	V5C.4
	VEqCM for 10 components of HICP (2 sectors by 5 countries)	V10.1	V10.2	V10.3	V10.4
	VEqCM for 10 components of HICP (2 sectors by 5 countries) with a block-diagonal restriction	V10BD.1	V10BD.2	V10BD.3	V10BD.4

(a) The leading indicator is included in the equation for the energy price index or in the equations of the price indexes which include energy prices

For all these events except (e) it was known beforehand when they were to happen 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. The first exercise takes into account the models that use a disaggregated analysis by sectors or by countries. They are models D5CU.1, V5C.1, D5SU.1, V5S.1 described in table 16. The second one considers the models that employ both information sets. These are models D10U.1, V10.1, V10BD.1 in table 16. Finally, the third one considers the models which include the Brent price as a leading indicator: models D5SU.3, V5S.3, D10U.3, V10DB.3, V5S.4 and V10DB.4 in table 16. In the first two cases the exercise compares the forecasting accuracy of the different models **with** a univariate aggregate model, AU.1; in the third case the models are compared with an aggregate univariate model enlarged with the crude Brent variable, model AU.3.

Table 17 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 AU.1 univariate model for the total HICP in the EMU. The models were re-estimated till December 1999 in order to consider a forecasting period of more than three years, January 2000 – July 2003, to evaluate the performance of the forecasts from one to twelve periods ahead.

Table 17: ***Forecasting errors for the year-on-year rates of total HICP in the EMU. January 2000- July 2003.***								
Horizon	**Statistics**	**Aggregate Univariate (a) AU.1**	**Disaggregated Analysis by 5 countries (b)**		**Disaggregated Analysis by 5 sectors (c)**		**Combination of forecasts (b) and (c)**	
			UNIV D5CU.1	VEqC V5C.1	UNIV D5SU.1	VEqC V5S.1	UNIV	VEqC
1	**RMSE**	0.1245	0.1151	0.1114	0.0997	0.0988	0.0999	0.0953
	$\frac{MSE(i)}{MSE(univ)}$	1	0.86	0.80	0.64	0.63	0.64	**0.59**
	DM	-	1.1675	1.1426	1.9808	1.9692	2.2583	1.8139
	p-value	-	0.2496	0.2597	0.0542	0.0555	0.0303	0.0783
3	**RMSE**	0.2278	0.2256	0.2157	0.1793	0.1722	0.1904	0.1739
	$\frac{MSE(i)}{MSE(univ)}$	1	0.98	0.90	0.62	**0.57**	0.70	**0.58**
	DM	-	1.3016	0.5862	2.1478	1.9273	1.7688	2.0102
	p-value	-	0.2007	0.5610	0.0380	0.0613	0.0856	0.0522
6	**RMSE**	0.2930	0.3131	0.2467	0.2738	0.2553	0.2843	0.2345
	$\frac{MSE(i)}{MSE(univ)}$	1	1.14	0.71	0.87	0.76	0.94	**0.64**
	DM	-	-0.6772	2.2644	0.7981	1.1721	0.1164	1.9151
	p-value	-	0.5026	0.0295	0.4300	0.2489	0.9080	0.0639
12	**RMSE**	0.3933	0.4866	0.3246	0.4521	0.3615	0.4567	0.3121
	$\frac{MSE(i)}{MSE(univ)}$	1	1.53	0.68	1.32	0.84	1.35	**0.63**
	DM	-	-1.6040	2.1817	-1.9926	0.3743	-1.4458	2.2547
	p-value	-	0.1195	0.0371	0.0555	0.7109	0.1577	0.0307

RMSE stands for root mean squared error.
MSE stands for mean squared error.
DM stands for the Diebold and Mariano test.
Numbers in bold type correspond to the lowest values per horizon.

The $\frac{MSE(i)}{MSE(univ)}$ 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 in each case.

Table 17 shows that for one and three periods ahead the best results are obtained when disaggregating by sectors, but for a medium run the geographic disaggregation yields the best forecasts. Therefore, further disaggregation by countries and sectors seems promising. Another important result shows that vector equilibrium correction models provide better forecasts than the corresponding univariate models in all cases.

This last result focuses on the relevance of considering cointegration relationships not only in estimation but also in forecasting, in order to pick up the long run restrictions between the variables. Regarding the Diebold and Mariano test, statistically significant differences can be found between the different approaches in favour of the disaggregated methodology for all periods.

Finally, both disaggregations are based on different information sets but the forecast errors for different periods do not differ in a relevant way between the two approaches, so the combination of the forecasts derived from both approaches, based on the average of the forecasts following Granger & Jeon (2004), makes sense and yields the best results, see last columns in table 17, suggesting that both disaggregations matter.

Table 18 shows the forecast errors for the year-on-year rates of total HICP derived from the two approaches that take into account both types of information – sectors and countries -: (1) in a model for the 10-elements vector composed by the HICP's of the two sectors – core and residual - in each geographical component (D10.1 and V10.1), (2) the previous model with the block-diagonal restriction (VBD10.1) and (3) the aggregate univariate model for the total HICP in the EMU (AU.1).

Table 18 shows that from the two-steps ahead the best results are obtained when disaggregating by countries and by sectors introducing the block-diagonal restriction and confirms the previous conclusions: (1) vector models forecast better than single equation models and (2) the Diebold and Mariano tests find statistically significant differences among the aggregated and disaggregated approaches in favour of the latter for all periods. Comparing the results from tables 17 and 18, we can observe the importance of disaggregating by sectorial blocks in different countries in order to improve the forecasts. The Diebold and Mariano tests show that the forecasting differences are significant in favour of this disaggregation for long horizons.

Table 18: ***Forecasting errors for the year-on-year rates of total HICP in the EMU. January 2000- July 2003.***

Horizon	Statistics	Aggregate Univariate AU.1	Disaggregated Analysis by sectors and countries		
				VEqCM	
			UNIV D10U.1	Vector of 10 components V10.1	Vector of 5 core HICP's and Vector of 5 residual HICP's V10BD.1
1	**RMSE**	0.1245	0.0856	0.0841	0.0937
	$\frac{MSE(i)}{MSE(univ)}$	1	**0.47**	**0.46**	0.57
	DM	-	2.6838	2.7706	2.0966
	p-value	-	0.0104	0.0083	0.0421
3	**RMSE**	0.2278	0.1675	0.1717	0.1561
	$\frac{MSE(i)}{MSE(univ)}$	1	0.54	0.57	**0.47**
	DM	-	1.9661	1.7012	2.8432
	p-value	-	0.0564	0.0969	0.0070
6	**RMSE**	0.2930	0.2482	0.2583	0.1949
	$\frac{MSE(i)}{MSE(univ)}$	1	0.72	0.78	**0.44**
	DM	-	1.4435	1.0339	2.9051
	p-value	-	0.1575	0.3081	0.0062
12	**RMSE**	0.3933	0.3955	0.4192	0.2468
	$\frac{MSE(i)}{MSE(univ)}$	1	1.01	1.14	**0.39**
	DM	-	-0.0858	-0.4862	2.0126
	p-value	-	0.9321	0.6305	0.0535

RMSE stands for root mean squared error.
MSE stands for mean squared error.
DM stands for the Diebold and Mariano test.
Numbers in bold type correspond to the lowest values per horizon.

Table 19 presents the forecast errors for the year-on-year rates of total HICP derived from the three approaches that take into account Brent prices in euros. The first approach considers a single equation leading-indicator model for the total HICP in the EMU (AU.3). The second one takes into account the sectorial disaggregation in five components in which the equation for energy price index includes the leading indicator

(D5S.3 and V5S.3). And the third one considers two sectorial blocks, core and residual, in five countries (D10U.3 and V10BD.3); the residual vector includes the Brent price in euros.

Table 19: ***Forecasting errors for the year-on-year rates of total HICP in the EMU, using international brent prices in euros as a leading indicator. January 2000- July 2003.***

Periods	**Statistics**	**Single-equation model AU.3**	**Disaggregated econometric modelling by 5 sectors**		**Disaggregated econometric modelling by sectors and countries**		**Disaggregated econometric modelling including observed brent prices**	
			UNI D5S.3	VEqC V5S.3	UNI D10U.3	VEqC V10BD.3	VEqC by sectors V5S.4	VEqC by sectors and countries V10BD.4
1	**RMSE**	0.1024	0.0969	0.0792	0.0858	0.0772	0.0792	0.0768
	$\frac{MSE(i)}{MSE(univ)}$	1	0.89	0.60	0.70	**0.57**	-	-
	DM	-	0.6192	1.2614	1.8498	2.1067	-	-
	p-value	-	0.5398	0.2155	0.0728	0.0424	-	-
3	**RMSE**	0.2393	0.2103	0.2110	0.1872	0.1797	0.1670	0.1565
	$\frac{MSE(i)}{MSE(univ)}$	1	0.77	0.78	0.61	**0.56**	-	-
	DM	-	2.0110	0.9854	2.0002	1.9054	-	-
	p-value	-	0.0521	0.3314	0.0535	0.0650	-	-
6	**RMSE**	0.3363	0.2761	0.2958	0.2493	0.2694	0.2614	0.2468
	$\frac{MSE(i)}{MSE(univ)}$	1	0.67	0.77	**0.55**	0.64	-	-
	DM	-	2.3192	1.6989	4.0369	2.2680	-	-
	p-value	-	0.0265	0.0985	0.0003	0.0296	-	-
12	**RMSE**	0.4207	0.2953	0.3525	0.2870	0.4014	0.4509	0.4320
	$\frac{MSE(i)}{MSE(univ)}$	1	0.49	0.70	**0.46**	0.91	-	-
	DM	-	2.7830	1.4410	2.9559	0.2747	-	-
	p-value	-	0.0087	0.1587	0.0056	0.7853	-	-

RMSE stands for root mean squared error.
MSE stands for mean squared error.
DM stands for the Diebold and Mariano test.
Numbers in bold type correspond to the least values.

It must be noted that this indicator, international Brent prices in euros, is known immediately while the HICP appears one month after the reference month. Therefore, the indicator is available for one-period ahead forecasts 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 of table 19 show that in one-period ahead forecasts the errors for the year-on-year rates of total HICP derived from the approaches including the Brent 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. From horizons two onwards the indicator must be forecast and the forecasting advantage of the models with indicator soon disappears.

Comparing all the previous forecasting results, it emerges that the best strategy to forecast inflation in the euro zone is to use the VEqCM model, constructed for a disaggregation of two sectors and five geographical areas, 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 including international Brent prices in euros as leading indicator.

This limited initial forecasting advantage of the models with the Brent prices and the subsequent deterioration of their forecasts deserves a further analysis. Thus, in this paper we have considered forecasting the energy price index component of the HICP, which has been denoted above as ENE. In forecasting ENE the following alternatives have been implemented: (a) it does not include international crude prices in euros as a leading indicator; (b) it includes this indicator and in forecasting ENE, univariate forecasts of the indicator are used; (c) like (b) but the indicator is forecasting using the future markets values; (d) like the two previous ones but taking future Brent prices as known.

These alternatives are applied to a single-equation model for ENE or to a VEqCM for the five sectors including the indicator in the equation of ENE. In all alternatives the VEqCM performs better than the corresponding univariate approach. For the first lags, in both types of models, forecasting the indicator by univariate model is better than the use of prices traded in future markets, but the opposite is true for longer horizons. The question is that for most forecasting origins the univariate forecasts for the crude price tend to keep increasing in medium and long horizons. In any case after two periods the best forecasts for ENE are obtained with the VEqCM which does not include Brent prices. For 12-period ahead horizon the RMSE is 2.17%. If we could know observed future Brent values when forecasting ENE, then the forecasts improve and the RMSE for the mentioned horizon falls to 1.62% with the VEqCM. This reduction is marginal, to 1.96%, when a univariate model for ENE is employed. In conclusion, it can be said that the use of Brent prices as an indicator when forecasting domestic consumer energy prices (ENE) is very useful for one-period ahead forecasts, but this advantage disappears soon and for horizons longer than two the ENE index is better forecast ignoring the international indicator. These conclusions are conditional to the linear models we are using and preliminary work done by the authors suggests that different conclusions could be obtained working with non-linear models.

The results on table 18 for the HICP for horizons longer than two just reflect this difficulty in forecasting crude oil prices.

A pairwise comparison of the forecasts of the different models used by means of the Diebold and Mariano tests shows that in forecasting inflation in the euro zone, disaggregation and the use of VEqCM models are important. The best disaggregation considers different sectors in different geographical areas. Since in this case the dimension of the vector becomes large and the samples available are short, then imposition of a block-diagonal restriction between sectors improves the forecasts. For short horizons, one and two, better forecasting accuracy can be obtained including international Brent prices in euros as a leading indicator. If the disaggregation is limited to sectors or geographical areas in the euro zone, then the combination of the forecasts from both sources improve the results.

7. CONCLUSIONS

From the results of the previous sections the following conclusions emerge:

1. For the sample period analysed, January 1990 – July 2003, the hypothesis that prices – HICP - in the euro-area are I(1) is not rejected. The annual inflation rate has been fluctuating around a mean of 1.9%. This can not 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.
2. The target adopted by the ECB that inflation in the medium run should not be above 2%, is compatible with the above result, interpreting the target in the sense that it refers to the mean in reasonable periods of time. For short periods of time, a few years, inflation could be systematically higher than 2%.
3. Disaggregating the HICP in five components, the four big countries and the geographical area which includes the remaining countries (REST), only two cointegrating relationships appear; one relating relative prices of Italy and Spain with those between France and Germany. The other restriction includes the HICP for the Rest. In fact this group of countries can not be considered as exogenous when analysing inflation in the euro-area.
4. Extending the sample before 1996 is not of much help, because the estimated models can only be considered stable from some point in 1996.
5. The disaggregation of the HICP in five sectors (energy, unprocessed food, processed food, other goods and services) was studied in Espasa et al. (2002a) and updating the study with data for an additional three-year period, the same results were

obtained. Only one cointegration relationship is found and price indexes for processed and unprocessed food can be considered as exogenous.

6. In modelling the HICP in the euro-area the break down of the aggregate by countries and sectors in n components matters. This is so because there are cointegration relationships between components and also several common trend factors. Thus disaggregation is a way to increase the information about the different trend factors affecting prices and a disaggregated model is a convenient framework to consider the long-run restrictions between the different price indexes. Disaggregating only by countries or only by sectors at the euro zone level is not enough to capture the different trend factors in prices.
7. Disaggregating the HICP by the five mentioned geographical components and two sectors, core and residual, in each case, three cointegrating relationships are found; one relating core indexes only, another relating residual indexes with the core index in Italy and a third one mixing core and residual indexes of different countries.
8. The above results indicate that core and residual indexes should be modelled jointly. But doing that with the sample available the estimation of the vector presents fit problems in some equations and some restrictions to accommodate the estimation to the existing number of degrees of freedom should be considered, as the one proposed in point 11.
9. Breaking down the national HICPs in each case in two components, one – core index – including the prices from which core inflation is computed, and another – residual index – including all other prices, it is shown that cointegration is not found in any pair of national prices. The same result is obtained for the euro-area as a whole.

10. The above result implies that core inflation is not a good leading indicator for total inflation. The interest of core inflation as a macroeconomic indicator lies **in** the fact that it is computed from prices in which innovations are more persistent.
11. Results in 9 show that the estimation of the 10-components vector could be carried out imposing a block diagonal structure with two blocks, one formed by the core indexes and other by the residual indexes.
12. The forecasting exercise shows that the forecasts from the VEqCM are better than the forecasts from univariate models.
13. A VEqCM 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-criteria disaggregation scheme.
14. Independently of the advantages of the disaggregation in forecasting, disaggregated forecasts are useful for policy because they inform which sectors have the higher expected inflation rates and how persistent are the shocks arriving to the different sectors.

ACKNOWLEDGEMENT

The authors acknowledge financial support from the Spanish Ministry of Science and Technology, project BEC2002-03720.

Appendix

Table A1: ***Country weights in the total EMU- HICP***

Country	**Weights**						
	1996	**1997**	**1998**	**1999**	**2000**	**2001**	**2002**
Germany	346.46	345.46	345.15	345.18	346.51	309.08	305.57
France	219.55	219.35	218.72	210.53	209.07	205.46	204.12
Italy	180.45	181.19	181.71	188.15	183.08	187	193.36
Spain	88.22	88.67	89.01	91.45	90.83	104.44	103.43
Austria	30.47	30.51	30.36	28.9	29.1	32.7	31.85
Belgium	38.15	38.23	38.01	39.89	39.9	33.5	33.97
Finland	15.57	15.51	15.53	14.83	15.07	15.9	15.94
Greece	0	0	0	0	0	24.28	24.68
Netherlands	53.21	53.1	53.47	51.27	56.54	52.52	52
Ireland	9.02	9.08	9.03	9.61	9.8	11.72	12.08
Luxembourg	2.17	2.17	2.17	1.98	1.99	2.46	2.56
Portugal	16.71	16.73	16.83	18.21	18.13	20.94	20.45
Rest	165.3	165.33	165.4	164.69	170.53	194.02	193.53
EMU	1000	1000	1000	1000	1000	1000	1000

Source: Eurostat / Date: 17th July 2002

Table A2: ***Weights of core HICP in the total HICP of each country (=1000)***

Country	**Weights**						
	1996	**1997**	**1998**	**1999**	**2000**	**2001**	**2002**
Germany	830.38	830.31	829.3	835.8	836.29	827.91	829.69
France	803.05	798.72	798.79	815.24	823.36	822.05	830.81
Italy	820.2	821.38	823.86	838.12	844.55	839.5	853.53
Spain	762.71	761.81	760.6	767.04	781.06	783.7	791.83
Rest	818.40	818.40	816.65	818.29	828.80	828.83	839.62
EMU	814.59	813.72	813.43	822.73	828.81	824.43	832.53

Source: Eurostat / Date: 17th July 2002

Table A3: ***Weights of residual HICP in the total HICP of each country (=1000)***							
Country	**Weights**						
	1996	**1997**	**1998**	**1999**	**2000**	**2001**	**2002**
Germany	169.62	169.69	170.7	164.2	163.71	172.09	170.31
France	196.95	201.28	201.21	184.76	176.64	177.95	169.19
Italy	179.8	178.62	176.15	161.88	155.45	160.5	146.47
Spain	237.29	238.24	239.4	232.96	218.94	216.2	208.17
Rest	181.61	181.63	183.41	181.71	171.20	171.17	160.38
EMU	185.41	186.29	186.58	177.27	171.19	175.56	167.47
Source: Eurostat / Date: 17th July 2002							

Table A4: ***Weights of core HICP of each country in the core inflation of the EMU***							
Country	**Weights**						
	1996	**1997**	**1998**	**1999**	**2000**	**2001**	**2002**
Germany	287.69	286.84	286.23	288.50	289.78	255.89	253.53
France	176.31	175.20	174.71	171.63	172.14	168.90	169.58
Italy	148.01	148.83	149.70	157.69	154.62	156.99	165.04
Spain	67.29	67.55	67.70	70.15	70.94	81.85	81.90
Rest	135.28	135.31	135.07	134.76	141.34	160.81	162.49
EMU	814.59	813.72	813.43	822.73	828.81	824.43	832.53
Source: Eurostat & IFL / Date: 17th July 2002							

Table A5: ***Weights of residual HICP of each country in the residual inflation of the EMU***							
Country	**Weights**						
	1996	**1997**	**1998**	**1999**	**2000**	**2001**	**2002**
Germany	58.77	58.62	58.92	56.68	56.73	53.19	52.04
France	43.24	44.15	44.01	38.90	36.93	36.56	34.54
Italy	32.44	32.36	32.01	30.46	28.46	30.01	28.32
Spain	20.93	21.12	21.31	21.30	19.89	22.58	21.53
Rest	30.02	30.03	30.34	29.93	29.19	33.21	31.04
EMU	185.41	186.29	186.58	177.27	171.19	175.56	167.47
Source: Eurostat & IFL /Date: 17th July 2002							

4. In order to elaborate bivariate cointegration analyses, the longest available sample in each case has been considered. Taking into account the EMU, the sample goes from January 1990 to June 2002, but data for the period January 1990 - December 1996 in the case of core HICP and for January 1990 - October 1996 in the case of residual HICP, are estimates constructed by Eurostat. For Germany, the sample goes from January 1995 to June 2002 for both indexes. In the case of France, the sample goes from January 1990 to June 2002, but data for the period January 1990 - December 1995 are estimates done by Eurostat for both indexes. The sample for Italy goes from January 1990 to June 2002, but data for the period January 1990 - December 1995 in core HICP and for January 1990 - December 1994 in residual HICP are estimates due to Eurostat. In the case of Spain, the sample goes from January 1992 to June 2002, but data for the period January 1992- December 1994 have been estimated by Eurostat for both indexes. For Rest, the sample goes from January 1996 to June 2002, so the reference base for HICP is 1996=100 and Rest is constructed through the aggregation of the indices corresponding to the countries that make it up. Data for 1996 in Belgium and Ireland are estimates made by Eurostat. There is some sort of instability at the beginning of the sample period and recursive estimations have been applied in all cases.

REFERENCES

Banerjee, A., Cockerell, L. and Russell, B. (2001), "An I(2) Analysis of Inflation and the Markup", *Journal of Applied Econometrics*, 16, pp. 221-240.

Bulletin EU & US Inflation and Macroeconomic Analysis. Instituto Flores de Lemus. Universidad Carlos III de Madrid.

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

De Brouwer, G. and Ericsson, N.R. (1998), "Modeling Inflation in Australia", *Journal of Business & Economic Statistics*, 16(4), pp. 433-449.

Dickey, D.A. and Fuller, W.A. (1981), "Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root", *Econometrica*, 49(4), pp. 1057-1072.

Diebold, F. X. and Mariano, R. S. (1995), "Comparing Predictive Accuracy", *Journal of Business & Economic Statistics*, 13, pp. 253-263.

Duesenberry, J. (1950), "The Mechanics of Inflation", *Review of Economics and Statistics*, 32, 144-149.

Escribano, A. and Peña, D. (1994), "Cointegration and Common Factors", *Journal of Time Series Analysis*, 15, 577-586.

Espasa, A., Matea, M. L., Manzano, M. C. & 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. and Albacete, R. (2002a), "Forecasting Inflation in the European Monetary Union: A Disaggregated Approach By Countries and By Sectors", The *European Journal of Finance,* 8, 402-421.

Espasa, A., Poncela, P. and Senra, E. (2002b), *Forecasting Monthly US Consumer Price Indexes Through a Disaggregated I(2) Analysis,* Working Paper, Universidad Carlos III.

Franz, W. and Gordon, R. J. (1993), “ German and American Wage and Price Dynamics: Differences and Common Themes” (with discussion), *European Economic Review,* 37, 719-762.

Granger, C.W.J and Jeon, Y. (2004). “Thick modeling”. *Economic Modelling,* 21, 323-343.

Hendry, D. (2001), “Modelling UK Inflation, 1875-1991”, *Journal of Applied Econometrics,* 16, 255-275.

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

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

Lorenzo, F. (1997), "Modelización de la inflación con fines de predicción y diagnóstico" Ph. D. Dissertation. September 1997. Universidad Carlos III de Madrid.

Richards, T. and Stevens, G. (1987), “Estimating the Inflationary Effects of Depreciation”, Research Discussion Paper 8713, Reserve Bank of Australia.

Stock, J. H. and Watson, M. W. (1999), “Forecasting Inflation”. *Journal of Monetary Economics,* 44, 293-335.

Stock, J. H. and Watson, M. W. (2002), “Macroeconomic forecasting using diffusion indexes”. *Journal of Business and Economic Statistics,* 20, 147-62.

Zellner, A. and Tobias, J. (2000), "A note on Aggregation, Disaggregation and Forecasting Performance", *Journal of Forecasting*, 19 (5), 457- 469.

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

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