Universidad Carlos III de Madrid

e-Archivo
Institutional Repository

This is a postprint version of the following published document:

Espasa, A., Mayo-Burgos, I. (2013). Forecasting aggregates and disaggregates with common features. International Journal of Forecasting, v. 29, pp. 718-732. Avalaible in: http://dx.doi.org/10.1016/j.ijforecast.2012.10.004

Forecasting aggregates and disaggregates with common features

Antoni Espasa [a,*], Iván Mayo-Burgos [b]

[a] Universidad Carlos III de Madrid, Dpt. Estadística and Instituto Flores de Lemus, Getafe, Madrid, Spain
[b] Universidad Carlos III de Madrid, Instituto Flores de Lemus, Getafe, Madrid, Spain

A R T I C L E   I N F O

Keywords:
Common trends
Common serial correlation
Inflation
Euro area
UK
US
Cointegration
Single-equation econometric models
Disaggregation maps

A B S T R A C T

This paper focuses on the provision of consistent forecasts for an aggregate economic indicator, such as a consumer price index and its components. The procedure developed is a disaggregated approach based on single-equation models for the components, which take into account the stable features that some components share, such as a common trend and common serial correlation. Our procedure starts by classifying a large number of components based on restrictions from common features. The result of this classification is a disaggregation map, which may also be useful in applying dynamic factors, defining intermediate aggregates or formulating models with unobserved components. We use the procedure to forecast inflation in the Euro area, the UK and the US. Our forecasts are significantly more accurate than either a direct forecast of the aggregate or various other indirect forecasts.

1. Introduction

The demand for macroeconomic forecasts has increased considerably over the last twenty years, and with it, the requests for quicker and more detailed releases of official data. In this context, one important phenomenon is the steadily growing flow of information available to forecasters; in particular, data are increasingly becoming available at higher degrees of disaggregation, at the regional, temporal and sector levels. Therefore, the traditional debate about whether to forecast an aggregate variable directly, or indirectly by aggregating the forecasts of its components, has recently received a considerable amount of attention. Usually, this discussion concentrates only on the forecasting accuracy of the aggregate. In contrast, the starting point of this paper is that all data—aggregate and components—are relevant, both for a full understanding of the aggregate, and for the formulation of useful economic policies. The focus of this paper is on inflation, but the question as to the usefulness of disaggregated information for econometric modelling and forecasting is relevant to many other macroeconomic variables, and the proposals in this paper could therefore be applied to these too.

Behind an aggregate lies a large amount of data that should not be ignored when generating the forecasting results that economic agents need for designing economic policy measures, making investment decisions, and related activities. For instance, when analysing the price components of a Consumer Price Index (CPI), a frequent observation is that several prices share features such as common trends or common serial correlation, whereas others do not, perhaps because they are affected by technological changes in a particular way, or because they are affected differently by changes in preferences. Similar remarks apply when considering the specific sectoral industrial production indexes of a national industrial production index, or the individual components of aggregates such as exports and imports. In examples such as these, a valid hypothesis is that a certain subset of components of the aggregate share a common feature, but others do not. Consequently, it seems convenient to use disaggregated information and exploit the restrictions existing between the components in econometric modelling in order to provide decision makers with forecasts that refer to the aggregate and its components. For example, a forecast of 2.2%

* Corresponding author. Tel.: +34 916249803, +34 916249847.
E-mail addresses: antoni.espasa@uc3m.es (A. Espasa), imaybur@gmail.com (I. Mayo-Burgos).

for headline inflation next year with a large percentage of the price components forecast to grow at around the same rate is quite different from the same forecast where the rate of growth of energy prices is forecast at 15% and many other prices are forecast to grow at very small percentages. We advocate the consideration of n components of an aggregate, which we call basic components. We aim to provide joint consistent forecasts of the aggregate and its basic components, as well as of useful intermediate aggregates. The validation of our proposal would involve showing that the indirect forecast of the aggregate is at least as accurate as the direct one, in which case disaggregation is clearly useful.

The literature in this area of research considers three main time series forecasting procedures: (F1) the direct approach, which works with a scalar model of the time series of the aggregate; (F2) the disaggregated procedure, based on univariate models for each of the basic components; and (F3) the multivariate disaggregated approach, which works with a vector model for the time series of all of the basic components. This paper develops a fourth alternative (F4), a disaggregated approach based on single-equation models of the basic components that include restrictions between them. In the applications in Section 5, we extend this alternative, considering three possibilities which are denoted by FP2, FP3 and FP4.

Theory shows that when the data generation process (DGP) is known, the forecasting accuracy of F3 is at least as good as that of the other procedures. Nevertheless, if the number of components is large, as is usually the case when working with basic components—numbering 160 for the US CPI in this study, for example—F3 is not feasible; in any case, it would be subject to a great deal of estimation uncertainty, on which we comment below. On the other hand, F2 can be either better or worse than F1, depending on the properties of the data. As will become clear, the existence of restrictions between the components is one of the main reasons why the disaggregated approach could be useful. In this paper, we develop an intermediate—relative to F1, F2 and F3—approach, called F4, which is based on single-equation models that take into account important restrictions between the components arising from the fact that some share common features. We keep this approach simple by using bivariate methods to identify a unique common feature in a subset of components and by using single-equation models to forecast each basic component. Furthermore, the basic components that do not share common features are aggregated into an intermediate aggregate, which is forecasted directly. Our procedure differs from that used in the dynamic factor literature because we consider the possibility of common features when analysing the relative behaviour of each of the variables—the basic components of an aggregate—and we only estimate common features between the basic components that truly share them—the estimation restriction. Then, each factor is used only in modelling and forecasting the basic components that have the corresponding common features—the forecasting restriction. At the same time, the procedure requires that the presence of common features be stable. In the dynamic factor literature, when large numbers of series are considered, as in our case, all elements are assumed to incorporate a common factor without the above estimation restriction, leaving the estimation process to determine which components enter with a zero weight. If the application of the estimation restriction is appropriate, the common factors (features) in our procedure could be estimated more precisely in small samples, and may also have a more direct economic interpretation.

Recently, Hendry and Hubrich (2006, 2010), hereafter HH, proposed a procedure for forecasting an aggregate by using a model for that aggregate that includes its own lags as well as lags of the components as regressors. They use Autometrics (see Doornik, 2009), and follow the general-to-specific approach when building the model. Our procedure differs from that of HH in two main respects. The first arises because our procedure incorporates specific identified and tested restrictions between the basic components when forecasting the aggregate. Because their model does not include all of the components in the equation for the aggregate, HH implicitly incorporate unknown restrictions between the components. However, as was shown by Clark (2000), specific restrictions, such as cointegration restrictions, should also be taken into account. The second difference is that our procedure naturally provides forecasts for the basic components, which are considered to be of interest because they could be necessary for policy decision makers. HH only provide results for the aggregate, because the forecasts at different horizons are made using horizon-specific estimated models, where the dependent variable is the multi-period-ahead value being forecasted, and therefore, they only need observed values of the independent variables.

The remainder of the paper is structured as follows. In Section 2, in relation to forecasting an aggregate, we comment on theoretical efficiency, estimation uncertainty and the relevant restrictions. In Section 3, we describe the data, the intermediate aggregation schemes with basic components, and the tests for positive and seasonal unit roots. In Section 4, we present our forecasting approach and the classification of the basic components using a disaggregation map which takes into account some of the common features between them. In Section 5, we use our procedure to forecast inflation in the US, the Euro area (EA) and the UK, and compare these results with those obtained from a direct forecast, an indirect forecast based on univariate models and an indirect forecast based on models with a stationary dynamic factor. In Section 6, we draw conclusions and propose extensions for future work. The applications in this paper include many results which cannot be reported here, but the interested reader will find more details of them on the first author's website.[1]

2. Theoretical efficiency, estimation uncertainty and relevant restrictions

Previous theoretical results for stationary variables—for details, see Kohn (1982) and Lütkepohl (1984), Rose

[1] Detailed results for all of these tests can be obtained from the first author's website: http://halweb.uc3m.es/esp/Personal/personas/espasa/esp/publications/ExtendedResults.html.

(1977), Tiao and Guttman (1980), and Wei and Abraham (1981), among others—have shown that, in general, procedure F3 will provide more accurate forecasts of the aggregate than the other procedures. It is only if the data satisfy special conditions—conditions for efficiency of the direct forecast (CEDFs)—that the direct approach is efficient; see Kohn (1982). In the case of one aggregate and n basic components, these conditions require that, when applying the vector of aggregating weights to the polynomial matrix of the vector moving average (VMA) representation of the components, one obtains a vector in which all of its n elements are simply the dynamic polynomial of the MA representation of the aggregate. Similarly, the condition can also be formulated for VAR processes.

A CEDF is a very restrictive condition, and when it is not satisfied, the use of the direct forecasting approach implies that invalid restrictions are imposed on the DGP, defined as the set of all of the basic components. To avoid imposing invalid restrictions in this sense, one can work from the basic components. This is because, if we break down the aggregate into a smaller number of components, which we term intermediate aggregates, these intermediate aggregates will be aggregates of the basic components; then, when modelling these intermediate aggregates, one could find that invalid restrictions are being imposed on the basic components included in them. If this is the case, we can use a wider disaggregation to improve the modelling and forecasting of these intermediate aggregates, and thereby forecast the overall aggregate more accurately. There is also another, perhaps more important, reason for considering the basic components. Assume that a subset of basic components share a common feature. Our procedure reduces the variance of the forecasting errors of the aggregate by taking these restrictions into account. However, intermediate aggregates based on official or ad hoc breakdowns generally include a subset of basic components which are cointegrated, as well as others which are not. Therefore, when testing pairs of intermediate aggregates for cointegration, it is often found that they are not cointegrated. For instance, Espasa and Albacete (2007) show that, in breakdowns of the CPIs of different Euro area countries into two components, Core CPI and the rest, these components are not cointegrated. In these cases, the cointegration present in the basic components cannot be exploited by working with intermediate aggregates.

Lütkepohl (1987) shows that CEDFs hold, for instance, when the components are uncorrelated and have identical stochastic structures. This can be taken as an indication that when components have different distributions—for instance when some have conditional heteroskedasticity or have a conditional mean with a nonlinear structure—or when there are cross-restrictions between them, disaggregation could be important. In this paper, we limit ourselves to considering the case in which there are restrictions between the components. This does not mean that distributional differences are unimportant; it merely allows us to study the problem in a way that is easier to solve in a general framework.

The case in which the components are nonstationary and cointegrated has recently been considered. Our approach is inspired by the results of Clark (2000), who showed that, when the model is known, the indirect forecast from a vector equilibrium correction model (VEqCM) for the components is more accurate than the direct forecast. Again, it is only under very specific conditions that the two forecasts are equivalent. These conditions include ones similar to those specified by Kohn (1982) for the transitory dynamics of the VEqCM, as well as the requirement that the aggregation of the matrix of equilibrium correction coefficients is a vector of zeros, in which case aggregation does not cause the loss of relevant information about the aggregate. Clark (2000) shows the importance, in general, of taking the cointegration restrictions into consideration when forecasting the aggregate, and proposes testing for cointegration, then testing the CEDFs. For the latter, the model for the aggregate needs to include lags of all but one of the components and the error correction terms, and the null hypothesis tested will be that the corresponding coefficients are zero. The problem with this is that when the number of components is large, one cannot perform even the initial cointegration tests. Thus, in this paper, we only consider what we call full cointegration, meaning that in a vector of n variables, there is only one common trend; that is, $(n - 1)$ cointegration restrictions. In this case, one can test for the presence of a unique common trend by using bivariate cointegration tests between all possible pairs of elements in the vector. The tests are implemented following the Engle and Granger (1987) approach. Thus, if a vector of n elements contains a subset of n_1 elements such that all possible pairs formed with its elements are cointegrated, then there is only one common trend in this subset. Therefore, in order to develop a simple procedure for capturing common trends, we restrict ourselves to finding subsets of basic elements that are fully cointegrated. Our application refers to inflation. In Section 3 we test for positive and seasonal unit roots in CPI components, and conclude that most of them have a positive unit root—meaning that they are I(1)—and that some of them have deterministic seasonality. Consequently, all of the cointegration tests in this paper are applied, including the appropriate seasonal dummies in the equation proposed by Engle and Granger (1987).

In this paper, by following an approach similar to that of Engle and Kozicki (1993), we also consider common serial correlation as another possible common feature in the data. The number of studies of comovements among stationary time series has increased considerably since the 1990s, and the different common features that have been defined and proposed include co-dependence (Gourieroux, Monfort, & Renault, 1991) and polynomial serial correlation (Cubadda & Hecq, 2001). Most of these features can be encompassed in the notion of the weak form of polynomial serial correlation proposed by Cubadda (2007). In this paper, we restrict our attention to the concept of common serial correlation, as defined by Engle and Kozicki (1993). That is, two stationary time series have common serial correlation if each series exhibits serial correlation and there is a linear combination of them that is white noise. The coefficients of the linear combination define the co-feature vector. For the general case of a vector of n stationary variables, y_t, the presence of common serial correlation implies a reduced rank in the matrix of coefficients, Γ, on the variables that are used to capture the common feature, lags of y_t. This matrix will have rank $(n - r)$ if there are r linear combinations that are white noise, and consequently, there will be $(n - r)$ common serial correlation factors. Thus, testing for common serial correlation involves testing the rank of Γ.

As has been proposed for common trends, we restrict ourselves to the case in which there is just one common serial correlation factor (CSCF) in a vector of n_2 components; this means that there are $(n_2 - 1)$ linear combinations that are white noise. This can be tested by applying the canonical correlation test proposed by Engle and Kozicki (1993) to all possible pairs of components in the vector. If, for each possible pair, we do not reject the hypothesis of one zero canonical correlation (one CSCF), each component will have one CSCF with any one of the other components, which will be common to all components of the vector. Suppose that n_2 is three and that there is a CSCF in all possible pairs of the three elements; each element can therefore be expressed by two different equations in terms of a CSCF plus a white noise. This implies that there is just one linearly independent CSCF.

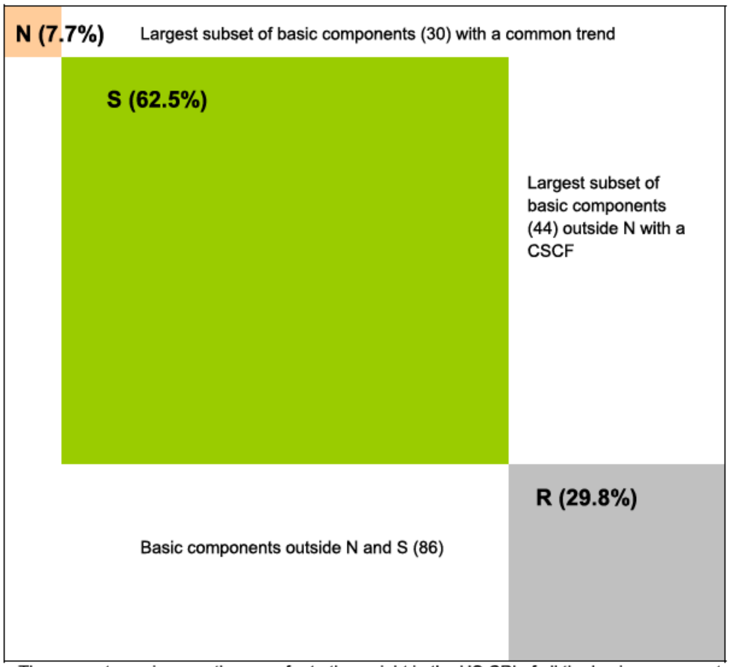
CPI components could have cointegration restrictions between them. In this case, as was shown by Vahid and Engle (1993), the test for common serial correlation in the stationary transformation of the original data should also consider the lags of the cointegration restrictions. This implies that all cointegration restrictions must be taken into consideration, not only those derived from full-cointegration; however, as was mentioned above, this is not feasible for vectors with a large number of basic components. Our procedure could incorporate full-cointegration restrictions when testing for CSCF in small dimension subsets, but this is not the case for subsets N in this paper. Consequently, we apply the Engle and Kozicki (1993) method in the following way. We look for non-overlapping subsets of basic components with a common trend and a CSCF, respectively. Thus, for a vector of n basic price index components, we first test for the largest subset of (n_1) basic components having just one common trend, subset N, then we test for common serial correlation in the first differences of the remaining $(n-n_1)$ basic components in which this common trend is absent. We also include appropriate seasonal dummies in these tests.

Other cointegration restrictions, such as those which are potentially present in the second largest subset of basic elements with only a single common trend—which could be identified by using bivariate methods—are very few in our applications, and have dimensions much smaller than those of subset N, as is shown on the website cited at the end of the introduction. In particular, for the US CPI, where subset N contains 30 elements, there is only one additional subset with a stable common trend, and it has only four elements; for UK and the EA we find one additional subset of dimension two for the former, and none for the latter. Thus, in this paper we consider only the largest subset of basic components with a common trend. Ignoring other subsets with a common trend means that we lose information, but that does not seem very important for the applications in this paper, since there are just a few subsets of this type and their dimensions are very small. In this first formulation of our procedure, we intend simply to show that it works. Succeeding while ignoring some potentially useful information only increases the procedure's interest. It could be widened to include the ignored information on other subsets with one common trend and the consideration of overlapping subsets of small dimensions with common features, but these are not the only possible extensions, or even the most important ones, and will be covered better by another paper defining a more general procedure.

Hence, our approach, that of first finding the largest subset of basic components with a common trend and then looking for CSCFs in the remaining basic components (based on the work of Engle & Kozicki, 1993), represents a simple and appropriate procedure for identifying relevant restrictions when working with the basic components of an aggregate. As we explain below, this approach is also consistent with that of Giacomini and Granger (2004).

When dealing with aggregated macroeconomic variables, in order to exploit the possible advantages of disaggregation, it is generally necessary to work with a large number of components, since these variables typically comprise many basic components. The theoretical results relating to the advantages of aggregating component forecasts from a multivariate model over forecasting the aggregate directly apply when the DGP is known. Because this is rarely the case in practice, the mean squared error (MSE) of the forecasts includes an additional factor, which is $1/T$ times a term that depends on the number of parameters to be estimated; see Giacomini and Granger (2004) and the references therein. Then, as is widely recognized in the literature, the question of which is the best procedure for forecasting the aggregate is mainly empirical. However, various results from the literature also shed light on this issue. Giacomini and Granger's (2004) results for space–time models suggest the existence of a trade-off between the efficiency gain achieved from specifying the fully disaggregated system and the loss in efficiency that arises from parameter estimation errors. In this context, Giacomini and Granger (2004) also consider four forecasting procedures, F1 to F4, which can be related to the time series procedures considered here. The F1 procedure is equivalent to a direct forecast of the aggregate; F2 is equivalent to an indirect forecast of the aggregate using ARIMA models for the components; F3 is equivalent to an indirect forecast based on a multivariate model for the components; and F4 is related to the forecasting procedure which we propose in this paper.

Giacomini and Granger (2004) show that imposing constraints in the fully disaggregated model improves the forecasts. One way to impose constraints is to use their F4 procedure instead of the theoretically optimal F3. Our proposed forecasting procedures, denoted FP2, FP3 and FP4 below, are also ways of imposing a large number of constraints in the vector model of the basic components. Note that the purpose of this paper is not to obtain new theoretical results but to formulate a procedure that is useful in practice, and one that deals with specification and estimation issues. Thus, the contributions



a The percentages in parentheses refer to the weight in the US CPI of all the basic components in the corresponding subset.

Fig. 1. Disaggregation map for the basic components of the US CPI.

of this paper are as follows. First, as has already been argued, the advantages of disaggregation must be explored from the most disaggregated level, in order to ensure that one is making a proper and efficient use of all of the available information and considering important restrictions between components. This approach facilitates the definition of possible useful intermediate aggregates and points out that their formulation is an endogenous question that must be investigated based on the properties of the basic components.

In this paper, we limit our attention to restrictions arising from the fact that non-overlapping subsets of basic components share a common trend (subset N) or a CSCF (subset S). Used in conjunction with bivariate methods, this procedure generates a disaggregation map in which the basic components are classified into subsets N, S and R. The disaggregation map obtained from our results can be represented as a squared $n \times n$ matrix, M, the elements of which are arranged in such a way that the upper left corner of the matrix collects the n_1 basic components with a common trend, followed by the n_2 basic components with a CSCF. In Fig. 1, we report the results for the US. Detailed results for the disaggregation maps of the EA and the UK are given in the mentioned website.

Our work represents a first attempt to build disaggregation maps for the basic components of an aggregate. The results could be useful for several purposes other than forecasting, such as the application of dynamic factors, the formulation of models with unobserved components and the design of economic policies. Indeed, in this paper we apply stationary dynamic factors to the elements in S, and the applications will show that we obtain a better forecasting accuracy for the aggregate in this way than by just applying stationary dynamic factors to the whole set of basic components, ignoring the results of the disaggregation map, as is standard practice. The disaggregation map can be extended to include overlapping subsets of basic components with common features, and by incorporating, for example, additional common-trend restrictions, the types of common cyclical features identified by Cubadda (2007), common seasonality, co-breaks, common non-linearity, common volatility, etc.

Our second and most important contribution is to develop a simple indirect forecasting procedure based on single-equation models, along the lines of Giacomini and Granger (2004), that departs from the use of vector models, imposes important restrictions on those models, and can produce improved forecasts in practice, as we will see in Section 5. At the same time, our procedure improves on direct forecasting by including additional relevant information, so that it more than compensates for the greater cost arising from estimation errors in more complex models. We expect this procedure to be widely applicable because it works with the basic components that share selected common features. Moreover, having tested for these restrictions, the basic components that do not share the common features specified in the procedure—subset R—are aggregated,[2] using the official

[2] The aggregation methodology is explained in Section 3.2.

weights, into an intermediate aggregate r_t, which is forecast using a scalar model.

We also show that the procedure works when forecasting inflation in three big economies, the US, the EA and the UK. In this context, our work is intended to provide not only better forecasts of an aggregate, but also forecasts of the basic components and of any intermediate aggregates which one could require. The basic elements of R can also be forecast by using ARI(p, 1) models under the restriction that the aggregation of those forecasts gives the direct forecast of the intermediate aggregate r_t. This could be done following Guerrero and Peña (2000).

3. The data

3.1. Data sets and aggregation procedure

We apply our procedure to the US CPI and the harmonized EA and UK CPIs. These economies were selected because they represent almost 50% of the global GDP, and because most econometric applications relate to at least one of these economies. For the US, we use monthly CPI data for all urban consumers, CPI-U, seasonally unadjusted, published by the Bureau of Labor Statistics. The sample goes from January 1999 to December 2010. The aggregate is broken down into 160 basic components. For the EA and the UK, we use the monthly Harmonised Index of Consumer Prices data (HICP), seasonally unadjusted, published by Eurostat. The samples used for the EA and the UK start in January 1995 and finish in December 2010, and the breakdowns of the aggregates have 79 and 70 elements, respectively.[3] One of the study's outputs is a disaggregation map based on estimated common features of the basic components. Some intermediate aggregates are formulated from the basic components in the paper, using the official weights and normalizing the sum of the weights of all of the basic components in an intermediate aggregate to 100 and applying the normalizing factor to the weight of each basic component in this intermediate aggregate.

3.2. Trend and seasonal factors

The basic components have trends, and some have seasonal oscillations; we therefore need to test for the presence of positive and seasonal unit roots in the data. To implement these tests with a large number of series, we have developed a standard procedure which could be applied to each series automatically. This almost prevents the possibility of considering the presence of outliers when performing these tests, and we have ignored the corrections for outliers in them. The tests were performed using the log transformation of the data, and their results can be found on the website cited in the introduction. We applied the Osborn, Chui, Smith, and Birchenhall (1988) tests, hereafter OCSB, and the Hylleberg, Engle, Granger, and Yoo (1990) test, as extended by Beaulieu and Miron (1993), hereafter HEGY. Using the terminology employed in the first paper, $I(r, s)$, where r and $s$can take values of one or zero, means that the data need r regular differences and s annual differences in order to be stationary. Following both references, we can test whether a particular series is I(1, 1), I(1, 0), I(0, 1) or I(0, 0), and in the second and fourth cases we can test whether the seasonal dummies are significant. All of the tests are performed at the 1% significance level, and the critical values are taken from Rodrigues and Osborn (1997). Following the OCSB test, hypothesis I(1, 1) is rejected in all cases, except for 4, 9 and 3 basic elements in the EA, US and UK, respectively. Hypothesis I(0, 1) is also rejected in most cases, with just 1, 8 and 2 exceptions in the above economic areas, respectively. Finally, the null hypothesis I(1, 0) is only rejected in two cases, one each in the UK and the US. In the latter hypothesis, the set of seasonal dummies can be appropriate, and this can be tested by an F test. In many cases—24 in the EA, 17 in US and 14 in UK—the presence of seasonal dummies is not rejected. Thus, based on these results, for the purposes of this paper, we consider that all of the basic components are integrated of order one, and some of them exhibit deterministic seasonality. To corroborate this conclusion, we apply the HEGY test. This test refers to the twelve π_i coefficients, following the notation of Beaulieu and Miron (1993), and the critical values were also taken from Rodrigues and Osborn (1997). At the above-mentioned significance level, we get results similar to those obtained with OCSB: the need for seasonal differencing is strongly rejected (by an $F_{1,12}$ test on the null: $\pi_i, i = 1, \ldots, 12$, are zero), but regular differencing is required in all cases (by a t-test on the null that π_1 is zero). In particular, the null I(0, 1) is rejected for all series except for one in the EA. The null of a positive unit root is not rejected in any cases, while the null of eleven seasonal unit roots (by an $F_{2,12}$ test on the null that $\pi_i, i = 2, \ldots, 12$, are zero) is rejected in all cases but five, four in the basic components of the US and one in the EA data. Since I(1, 0) has not been rejected, these last cases form a contradiction with the results for the $F_{1,12;}$ however, this is something that can occur in finite samples. In summary, the I(1, 0) hypothesis with possible deterministic seasonality seems quite acceptable for the data.

In addition, when applying the ADF test to the differences of the basic components, the null of I(2) for the basic components is rejected in all cases at the 1% significance level. The critical values are taken from MacKinnon (1991, Chap. 13) for the case in which a constant is included. This is the result expected, because otherwise innovations in the distant past would have a greater impact on the contemporaneous value of a price index than recent innovations.

In the EA, seasonality in the harmonized index of consumer prices (HICP) has a break at the beginning of 2001 because of a change in Eurostat's data collection methodology. Thus, in all of the tests and models for the EA, following Espasa and Albacete (2007), we always include two sets of seasonal dummies, one of which applies up to December 2000 and the other operating from January 2001. Given the initial 1995–2003 sample, seasonal change is estimated with few degrees of freedom, so this could be seen as a necessary correction for outliers. With the use of recursive samples—samples in which the

[3] In all cases, the data correspond to the existing published versions as of 15th March 2011.

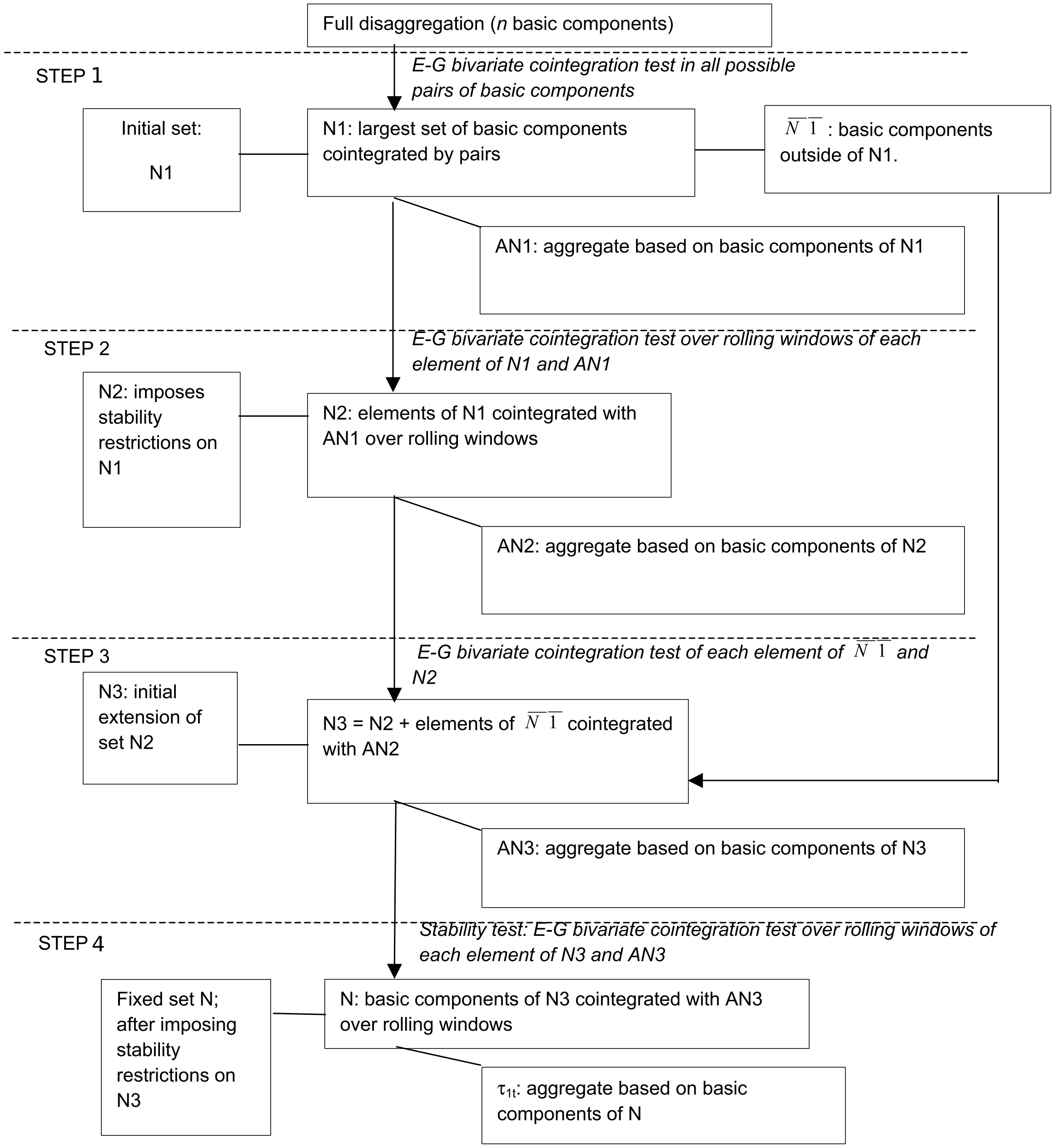


Fig. 2. Identification of the largest subset of basic components with just one common trend.[4]

initial observations remain fixed but which are extended each time that the base of the forecast moves forward—in the forecasting process, seasonal change is ultimately estimated more precisely.

All of the models estimated in the next section include the appropriate sets of seasonal dummies when required, even when denoted as ARI.

4. Our procedure

In our procedure, we distinguish between the following three phases: (1) the selection of the relevant common features, which in our case are a stable single common trend and a stable single CSCF; (2) the construction of a disaggregation map, with the largest non-overlapping subsets of basic components sharing one of the above common features; and (3) the construction of single-equation forecasting models for the elements of the disaggregation map.

4.1. Construction of the disaggregation maps

Fig. 2 summarizes the process followed to identify the elements in N, the largest subset of basic components with

[4] This procedure can also be used to look for the second largest subset with a common trend, applying it to the basic components outside N .

just one common trend. This is done using cointegration tests based on the Engle–Granger (EG) procedure, although the Johansen test could also be used. Since we have rejected the I(2) hypothesis for the data, the EG tests are performed including a constant and seasonal dummies in the models, and the critical values are obtained by simulation, following MacKinnon (1991, Chap. 13), for sample sizes of 60 and 108, depending on the economic region, at the 10% significance level.[5] In the initial step, we apply the cointegration test to all possible pairs of basic components and select the largest subset, say N1, of n_0 basic components in which all pairs are cointegrated. The tests are performed using a restrictive approach in order to end up with the presence of bivariate cointegration. Thus, we conclude that two basic components are cointegrated when the hypothesis is not rejected after applying the EG test in both directions.

The second step involves testing whether the bivariate cointegration relationships found in the previous step are stable over time, in the sense that they are evident in shorter subsamples. For this purpose, an intermediate aggregate, AN1, including all of the elements of N1, is constructed. Each element of N1 must be cointegrated with AN1, and the stability of this restriction is investigated by estimating and testing for cointegration across the sample by using a rolling window. The elements of N1 that do not pass this "stability test" are removed from N1, and the resulting subset is denoted by N2.

A third step is used to check whether it is possible to enlarge N2. Thus, we consider the basic elements outside N1 as potential candidates and perform a bivariate cointegration test between each of them and the intermediate aggregate AN2. Any elements that are cointegrated with this AN2 are added to N2 to form a new subset at the end of step 3, termed N3, and the corresponding intermediate aggregate AN3 is constructed.

The final step tests for stability in the bivariate cointegration relationships of the elements of N3, proceeding as in the second step, but relating each element of N3 to AN3. Removing the basic components that do not pass the test from N3 results in the final subset N, which is taken as the largest subset of basic components with only a single (stable) common trend. With the elements of N, the intermediate aggregate τ_{1t} is formed as is described in Fig. 2, and τ_{1t} can be seen as a proxy for the common trend in the basic components of N.

To apply the procedure proposed by Engle and Kozicki (1993), we look for the largest subset of basic components outside N with just a single CSCF, subset S. The elements of S can be identified by using a four-step procedure similar to that used to identify a common trend, but now testing for a CSCF. The intermediate aggregate τ_{2t} is then formed with the elements of S. In this case, the CSCF can be approximated by the univariate fit of $\Delta\tau_{2t}$, as we did in this paper for the purpose of the disaggregation map, or by applying the dynamic factor analysis to the components of S.

The procedure could also be extended to identify other subsets of basic components with other types of common trends or CSCFs. For example, one could consider the subset of basic components outside N in which all of the elements share two common trends with the elements of N. In addition, the disaggregation map could consider the type of cyclical features identified by Cubadda (2007), as well as other common features such as seasonality, cobreaks, common non-linearities and volatility.

4.2. The final disaggregation maps

The US data used correspond to a breakdown of the US CPI into 160 basic components (listed on the website referred to above). A useful sectorial breakdown of the CPI includes the following sectors: energy (ENE); non-processed food (NPF); processed food (PF); non-energy industrial goods (MAN); and services (SERV). We use these "broad CPI categories" to present the disaggregation maps for the basic CPI components. Note, however, that the correspondence is not perfect, because a basic component could include prices belonging to two broad categories.

According to Table 1, for the US, the subset N contains 30 basic components that account for 7.66% of the CPI, and belong mainly to MAN (2.90 percentage points (pp)) and SERV (2.15 pp); see Table 2. The number of basic components in subset S—basic components with a CSCF—is 44, and they account for 62.5% of the CPI. The elements of S are more widely distributed among the broad CPI categories; see Table 2. This subset of the disaggregation map has the most weight in the CPI, and includes prices of food, fuels, heat energy, transport and tourism services, nondurable household goods, sporting equipment, and goods related to new technologies. The subset R has 86 elements and they account for 29.81% of the CPI.

Table 1 also presents results for the EA and the UK. Although they differ from the US results, an important source of this difference is that the US CPI includes (in S) the "owner's equivalent rent of primary residence" (with a weight in the CPI of around 24%), which the HICPs in the EA and the UK do not include. Nevertheless, correcting for this divergence in CPI composition methods, the basic components with a common trend carry less weight in the US than in the EA or the UK, whereas the basic components with a CSCF carry relatively more weight. In any case, one result that emerges from these applications is that the characteristics of the intermediate aggregates τ_{1t}, τ_{2t} and r_t differ greatly between countries. This is illustrated in Fig. 3. In our sample period, τ_{1t} and τ_{2t} have been diverging considerably in the US and the UK, but not so much in the EA.

5. Forecasting results for inflation in the US, the EA and the UK

In this section, we use our procedure to forecast inflation in the US, the EA and the UK.[6]

[5] The simulated critical values are −3.15 for US, −3.13 for UK and EA, with a unique set of dummies, and −3.33 for EA with two sets of dummies. All critical values are at the 0.1 significance level.

[6] Additional results for all three regions following the procedure FP3, estimating the CSCF by the fit of $\Delta\tau_{2t}$, can be obtained from the above-mentioned website.

Table 1
Composition of the largest subsets of basic components sharing a common trend (N) or a CSCF (S), and the subset of the remaining basic components (R) in the US, the EA and the UK.

	Subset N	Subset S	Subset R	Total
US CPI				
Number of basic components	30	44	86	160
Weight in the CPI(%)	7.66	62.53	29.81	100
EA HICP				
Number of basic components	26	23	30	79
Weight in the HICP(%)	39.51	22.08	38.41	100
UK HICP				
Number of basic components	26	19	25	70
Weight in the HICP(%)	39.30	30.40	30.30	100

Table 2
Classification by broad categories in the US CPI of the basic components belonging to the subsets N, S and R.[a]

	Number of basic components	Weights of basic components in subset N (%)	Weights of basic components in the corresponding broad category (%)	Weights of basic components in CPI (%)
Basic components in subset N (one common trend)				
NPF	5	6.96	16.21	0.53
ENE	2	17.56	18.12	1.35
PF	4	9.49	10.88	0.73
MAN	15	37.91	12.67	2.90
SERV	4	28.08	3.60	2.15
TOTAL	30	100.00		7.70
Basic components in subset S (one CSCF)				
NPF	5	1.40	26.70	0.88
ENE	3	9.72	81.88	6.08
PF	11	4.15	38.89	2.60
MAN	13	18.80	51.31	11.76
SERV	12	65.92	69.03	41.23
TOTAL	44	100.00		62.50
Basic components in subset R				
NPF	15	6.29	57.10	1.88
ENE	0	0.00	0.00	0.00
PF	16	11.25	50.23	3.36
MAN	32	27.67	36.02	8.25
SERV	23	54.79	27.37	16.35
TOTAL	86	100.00		29.80

[a] The broad categories are: non-processed food (NPF); energy (ENE); processed food (PF); other goods (MAN); and other services (SERV).

5.1. Forecasting procedures

We examine the forecasting results for the year-on-year inflation rates, approximated by the annual differences in the logarithmic transformation of the price indexes. The data are monthly and the samples range from 1995:01 or 1999:01 to 2010:12. The data till December 2003 are used for the specification and initial estimation of the models, and the remaining data are employed for evaluating the forecasts of the different methods. To do this, we replicate real-time forecasting by using recursive windows for the different forecasting procedures and employing all of the available disaggregated information at each step, and forecasting up to 12 periods ahead. This means that, starting with information up to December 2003, we forecast up to December 2004. Then, extending the sample to January 2004, we again test for the number of lags, check for outliers, re-estimate the models, and forecast from February 2004 up to January 2005, and so on. In the forecasting procedures, we include dummies for additive outliers.

The forecasting accuracy of each formulation is evaluated using the root mean squared forecast errors (RMSFEs) at any given forecast horizon. The Diebold and Mariano (1995) test is implemented to test for significant differences between pairs of RMSFEs; and in addition, we also use Capistran's (2006) version of the Diebold–Mariano test based on a multivariate loss function to test the forecast accuracies of two procedures jointly over the 12 horizons.

The following forecasting exercise compares the performances of alternative formulations of the indirect procedures proposed in the paper, and of an indirect procedure that uses stationary dynamic factors, against the performance of the direct procedure using an ARI(p, 1) for the aggregate variable, called Y_t. The direct procedure models the aggregate variable as a constant, with the corresponding deterministic dummies and the own past. All of the models include dummies for additive outliers (AO). This is our benchmark model. Stock and Watson (2004, 2007)

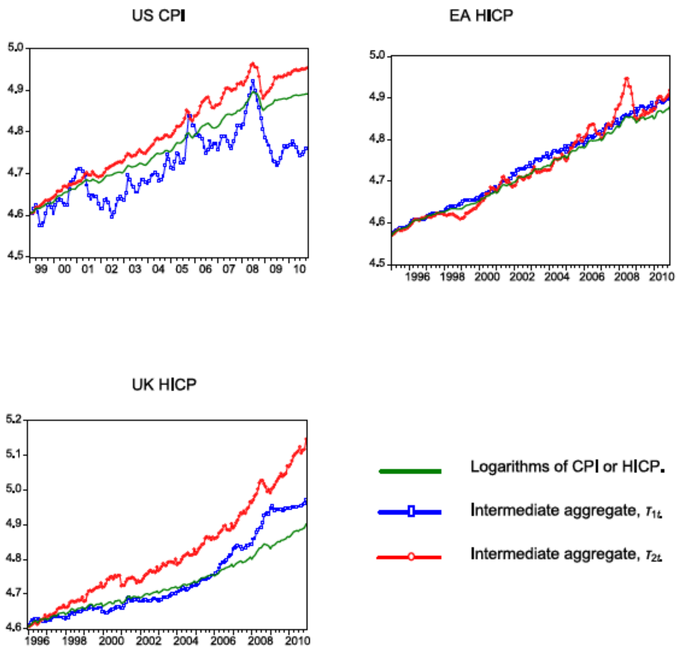


Fig. 3. Logarithms of CPI and HICP, and logarithms of the intermediate aggregates τ_{1t} and τ_{2t}, for the US, the EA and the UK.

Table 3
Summary of forecasting approaches.

The models refer to the first differences of the log transformation of the variables.
(I) Direct approach
(FP1) Forecast the aggregate data by an AR(p) model.
(II) Indirect approaches based on the disaggregation map proposed in this paper
(FP2) Procedure using a common-trend restriction.
Forecast each basic component of N and the intermediate aggregate formed with the remaining basic components.
(FP3) Procedure using a CSCF restriction.
Forecast each basic component of S and the intermediate aggregate formed with the remaining basic components.
(FP4) Procedure using a common-trend and a CSCF restrictions.
Forecast each basic component of N and S and the intermediate aggregate formed with the remaining basic components.
(III) Indirect approach based on factor-augmented models for the full disaggregation
(FP5) Procedure using a stationary dynamic factor.
Forecast each basic component by including a stationary dynamic factor.
(IV) Indirect approaches based on AR(p) models
(FP6) Forecast each basic component by an AR(p) model.

argue that it is difficult to improve on the results from a simple univariate model with gradually evolving relevant parameters.

$$\Delta \log Y_t = \alpha + \sum_{j=1}^{m} \delta_j D_j^{AO} + \sum_{j=1}^{11} \gamma_j D_j \quad \text{(I): constant, dummies for additive outliers (AO), and seasonal dummies} \qquad (1)$$
$$+ \sum_{j=1}^{s} \omega_j \Delta \log Y_{t-j} \quad \text{(II): own lags}$$
$$+ \varepsilon_t \quad \text{(III): residual term.}$$

The forecasting approaches applied in this paper, denoted as FP, could be classified into four groups (see Table 3). The first, approach FP1, is the direct procedure, which uses the simplest information set, and therefore, a univariate model for the first differences of the aggregate price index. It corresponds to approach F1 in the introduction. The model used is given above in Eq. (1). The second group, comprising approaches FP2 to FP4, includes indirect procedures based on the disaggregation map proposed in this paper. They represent different alternatives of approach F4 in the introduction. Each of these procedures uses a different subset of restricted basic components—N in approach FP2, S in approach FP3, and N and S in approach FP4—and in each case, a specific residual intermediate aggregate is formed from the remaining basic components. Consequently, the residual intermediate aggregates are different under procedures FP2, FP3 and FP4; specifically, it is only under approach FP4 that the residual subset is the subset R, as illustrated in Fig. 1. Then, all of the basic components selected in each approach are forecast using the appropriate specification of the general model presented below in Eq. (3), and the corresponding residual intermediate aggregate is forecast by using an ARI(p, 1) model like Eq. (1). In the final step, these forecasts are aggregated.

Approach FP3 can be run with two different options, corresponding to the two approaches for estimating CSCF. One is the fit of $\Delta\tau_{2t}$, and the other involves applying the dynamic factor analysis to the components of S. Thus, the results with FP3 could be compared with those from an application of dynamic factors to all basic components, i.e., ignoring the results from the disaggregation map, as is done in FP5 below.

The third group, approach FP5, collects indirect procedures based on factor-augmented models, as proposed by Bernanke, Boivin, and Eliasz (2005) and Stock and Watson (2005), for each basic component. Each model is an ARI(p, 1) model with stationary dynamic factors as regressors. The dynamic factor is estimated over all basic components by applying the procedure described by Stock and Watson (1998, 2002). We obtained the best forecasting results with just one dynamic factor, as did Duarte and Rua (2007). We denote the dynamic factor by F_t. We also found that, when using the dynamic factors, a better forecast of the aggregate is obtained by aggregating the forecasts of the basic components than by forecasting it directly. The general forecasting model used in FP5 for each basic component, $x_{i,t}$, is as follows:

$$\begin{aligned}
\Delta \log x_{i,t} &= \alpha_i + \sum_{j=1}^{m} \delta_j D_j^{AO} + \sum_{j=1}^{11} \gamma_{i,j} D_j \quad \text{(I): constant, dummies for additive outliers (AO), and seasonal dummies} \\
&+ \sum_{j=1}^{k} \lambda_j F_{t-j} \quad \text{(IIa): lags of the stationary factor } F_t \\
&+ \sum_{j=1}^{s} \omega_{i,j} \Delta \log x_{i,t-j} \quad \text{(III): own lags} \\
&+ \varepsilon_{i,t} \quad \text{(IV): residual term.}
\end{aligned} \tag{2}$$

The last group, approach FP6, is an indirect procedure based on a univariate ARI(p, 1) model for each basic component, similar to Eq. (1) for the direct procedure. It corresponds to approach F2 in the introduction.

We are interested in comparing the procedure FP3 proposed in this paper with an indirect procedure that also uses a common factor but that is extracted automatically from the whole set of basic components without having to test for restrictions between these basic components; this is approach FP5. This is the approach that would be followed by those working with dynamic factors from the set of basic components. Thus, it is interesting to compare it with FP3, where we also use a dynamic factor, but one estimated only from the elements of S.

This is important, because, as was mentioned in the introduction, our procedure incorporates an estimation restriction when calculating the common factors and a forecasting restriction when forecasting the basic components. In addition, because our procedure allows the isolation of the basic components that do not share the common features identified by the analysis, it can forecast the intermediate aggregate formed from those (residual) basic components directly.

The indirect forecasting approaches, FP2 to FP5, require some additional steps in the forecasting process. This is because forecasting the dependent variables requires forecasts of the explanatory variables: common trends, the CSCF, and the dynamic factor. For this forecasting exercise, at each forecast horizon h (>1) of a given base period n, we need forecasts of those explanatory variables for the time periods $n+h-1$. These are calculated by weighting the forecasts of the corresponding basic components obtained for previous horizons. For the common trend and the CSCF, we use official weights, as explained above. We use the loading vector for the dynamic factor. Approach FP3 was used in the two different options mentioned above. The option that estimates CSCF by applying dynamic factors to the elements of S gives better results, and is the only FP3 option for which we publish results here. The comparison of procedures FP3 and FP5 shows the usefulness of applying dynamic factors in the context of a disaggregation map such as that proposed in this paper.

5.2. Single-equation forecasting models

From the disaggregation map, we need to build single-equation forecasting models for the basic components in N and S, and for the intermediate aggregate r_t. Then, by aggregating these forecasts using the normalized official weights of the corresponding CPI, as explained in Section 3.1, we obtain the headline inflation rate forecast.

The general structure of the forecasting model of the $x_{i,t}$ basic component in N or S is as follows:

$$\begin{aligned}
\Delta \log x_{i,t} &= \alpha_i + \sum_{j=1}^{m} \delta_j D_j^{AO} + \sum_{j=1}^{11} \gamma_{i,j} D_j \quad \text{(I): } \Delta \log x_{i,t} \text{ denotes the first differences of the log of the } i\text{th basic component. The model includes a constant, dummies for additive outliers (AO), and seasonal dummies} \\
&+ \delta_i(\log x_{i,t} - \beta_{1,i} \log \tau_{1,t}) \quad \text{(IIa): cointegration relationship between the basic component } i \text{ of N and the intermediate aggregate } \tau_{1,t} \\
&+ \theta_i \left(\sum_{j=1}^{r} \lambda_j CSCF_{t-j} \right) \quad \text{(IIb): the estimate of the CSCF in S} \\
&+ \sum_{j=1}^{k} \delta_{i,j} \Delta \log \tau_{1,t-j} \quad \text{(IIIa): lags of } \Delta \log \tau_{1,t} \\
&+ \sum_{j=1}^{q} \varphi_{i,j} \Delta \log r_{t-j} \quad \text{(IIIb): lags of } \Delta \log r_t \\
&+ \sum_{j=1}^{s} \omega_{i,j} \Delta \log x_{i,t-j} \quad \text{(IIIc): own lags} \\
&+ \varepsilon_{i,t} \quad \text{(IV): residual term.}
\end{aligned} \tag{3}$$

The number of lags is selected based on the Akaike information criterion (AIC). Because we have not tested whether the basic components in N have the same CSCF

Table 4
RMSFEs (in percentage terms) of the direct approach FP1, and the RMSFE ratio for each approach to FP1.

US, year-on-year inflation rate						
Prediction horizon (months)	Direct procedure	Indirect procedures based on intermediate disaggregations considered in the paper			Indirect procedure based on factor-augmented models	Indirect procedure based on AR models
	FP1	FP2	FP3	FP4	FP5	FP6
1	0.42	1.04	0.96	0.95	1.05	0.98
2	0.83	0.93*	0.92*	0.90*	0.93	1.00
3	1.11	0.94	0.93	0.91*	0.91*	1.01
4	1.32	0.94	0.93	0.90*	0.91*	1.01
5	1.48	0.94*	0.93	0.90*	0.91*	1.01
6	1.59	0.93*	0.93	0.89*	0.92*	1.01
7	1.67	0.93	0.91	0.88*	0.93	1.01
8	1.76	0.93	0.87*	0.85*	0.92	1.01
9	1.85	0.92	0.84*	0.81**	0.91	1.00
10	1.94	0.91	0.81**	0.78**	0.89*	1.00
11	2.04	0.90	0.78**	0.76**	0.87*	1.00
12	2.15	0.91*	0.78**	0.75**	0.86**	1.01

Forecast sample: 2004/01–2010/12.
* Significantly different at the 5% significance level using the Diebold and Mariano test.
** Significantly different at the 1% significance level using the Diebold and Mariano test.
The base periods of the forecasts go from 2003/12 to 2010/11. For horizons 1 and 12, we have 84 and 72 forecasting errors, respectively.

Table 5
Diebold–Mariano test results based on a multivariate loss function for the path forecast between two approaches (Capistran, 2006).

US results	Direct procedure	Indirect procedures based on intermediate disaggregations considered in the paper			Indirect procedure based on factor-augmented models	Indirect procedure based on AR models
	FP1	FP2	FP3	FP4	FP5	FP6
FP1		**	**	**	**	
FP2			*	**		
FP3				*		
FP4						
FP5				*		
FP6		**	**	**	**	

* (**) indicates that the procedure appearing in the column performs significantly better than the procedure appearing in the row at the 5% (1%) significance level.

as that found for the basic components in S, we can now include the estimated CSCF for S in the models for the basic components in N, if this is significant. This is an indirect way of identifying the basic components that share not only the common trend of N but also the CSCF of S. Thus, the models for the basic components in N could include all of the terms in the above general structure.

However, because we rejected the hypothesis that the basic components in S have the common trend that is present in N, the models for these components can include all of the terms of the general structure in Eq. (3) except for (IIa).

For the basic elements of R, we proceed, as mentioned above, by forecasting the intermediate aggregate r_t directly using a model like Eq. (1). We do this because, in all of the applications described in Section 5.3, forecasting r_t directly generates a greater accuracy than does aggregating the individual forecasts of the basic components in R.

5.3. *Forecasting exercise*

The forecasting results are summarized in Table 4, with column FP1 reporting the root mean squared forecast error (RMSFE) of the direct procedure, which is our benchmark. The other columns report the ratios of the RMSFEs of the corresponding specific forecasting approaches to the RMSFE of FP1. Values below unity indicate an improvement in forecast accuracy with respect to the benchmark model. In this table, a single asterisk (*) indicates that the difference in RMSFEs is significant at the 5% significance level based on the Diebold–Mariano test, and ** indicates that the difference is significant at the 1% level. In Table 5, we report the Diebold–Mariano test results based on the multivariate loss function proposed by Capistran (2006), to test jointly the forecast accuracies of two procedures over 12 horizons. Tables 6–9 report similar results for the EA and the UK.

6. Conclusions and proposed extensions

6.1. *Conclusions*

These results, based on CPI data for economic regions that cover about 50% of the world's GDP, were obtained using a sample of seven years of forecasting errors. The results are therefore informative and generate interesting conclusions. The indirect forecast based on using ARI(p, 1) models for all of the basic components (FP6) is not better than the direct forecast. In fact, for the US, the forecasting performance of FP6 is similar to that of the direct approach for all horizons. In the EA and the UK, FP6 performs relatively poorly for most horizons. Therefore,

Table 6
RMSFEs (in percentage terms) of the direct approach FP1, and the RMSFE ratio for each approach to FP1.

EA, year-on-year inflation rate						
Prediction horizon (months)	Direct procedure	Indirect procedures based on intermediate disaggregations considered in the paper			Indirect procedure based on factor-augmented models	Indirect procedure based on AR models
	FP1	FP2	FP3	FP4	FP5	FP6
1	0.21	0.94*	0.88**	0.85**	0.90*	0.90
2	0.31	0.96*	0.91*	0.91*	0.94	1.27
3	0.41	0.97*	0.92*	0.91*	0.96	1.34
4	0.49	0.99	0.90**	0.89**	0.95	1.32
5	0.58	0.99	0.89**	0.88**	0.93	1.29
6	0.65	1.01	0.89**	0.89**	0.94	1.29
7	0.73	1.01	0.87**	0.87**	0.92	1.21
8	0.81	1.01	0.86**	0.86**	0.91	1.14
9	0.88	1.01	0.87**	0.87**	0.91*	1.09
10	0.94	1.01	0.90**	0.90**	0.92*	1.07
11	1.00	1.01	0.91**	0.90**	0.93*	1.05
12	1.05	1.02	0.93**	0.92**	0.95	1.04

Forecast sample: 2004/01–2010/12.
* Significantly different at the 5% significance level using the Diebold and Mariano test.
** Significantly different at the 1% significance level using the Diebold and Mariano test.
The base periods of the forecasts go from 2003/12 to 2010/11. For horizons 1 and 12, we have 84 and 72 forecasting errors, respectively.

Table 7
Diebold–Mariano test results based on a multivariate loss function for the path forecast between two approaches (Capistran, 2006): Euro area results.

EA results	Direct procedure	Indirect procedures based on intermediate disaggregations considered in the paper			Indirect procedure based on factor-augmented models	Indirect procedure based on AR models
	FP1	FP2	FP3	FP4	FP5	FP6
FP1			**	**	**	
FP2			*	**	**	
FP3						
FP4						
FP5			*	**		
FP6			**	**	**	

* (**) indicates that the procedure appearing in the column performs significantly better than the procedure appearing in the row at the 5% (1%) significance level.

Table 8
RMSFEs (in percentage terms) of the direct approach FP1 and RMSFE ratio for each approach to FP1.

UK, year-on-year inflation rate						
Prediction horizon (months)	Direct procedure	Indirect procedures based on intermediate disaggregations considered in the paper			Indirect procedure based on factor-augmented models	Indirect procedure based on AR models
	FP1	FP2	FP3	FP4	FP5	FP6
1	0.27	0.99	0.98	0.98	1.05	0.92
2	0.39	0.97	0.93	0.99	0.98	1.07
3	0.51	0.96	0.91**	0.90*	0.97	1.53
4	0.63	0.94	0.88**	0.84**	0.96	1.38
5	0.75	0.91*	0.86**	0.78**	0.92*	1.24
6	0.86	0.90**	0.85**	0.75**	0.93**	1.19
7	0.97	0.91*	0.84**	0.72**	0.93**	1.16
8	1.08	0.91**	0.82**	0.70**	0.93**	1.14
9	1.19	0.90**	0.80**	0.69**	0.92**	1.12
10	1.30	0.89**	0.79**	0.68**	0.92**	1.11
11	1.39	0.89**	0.79**	0.67**	0.92**	1.11
12	1.49	0.88**	0.79**	0.66**	0.92**	1.11

Forecast sample: 2004/01–2010/12.
* Significantly different at the 5% significance level using the Diebold and Mariano test.
** Significantly different at the 1% significance level using the Diebold and Mariano test.
The base periods of the forecasts go from 2003/12 to 2010/11. For horizons 1 and 12, we have 84 and 72 forecasting errors, respectively.

disaggregation in itself does not improve the aggregate forecast in these cases unless the relationships between components are also taken into consideration.

In contrast, for the US, the indirect approaches that incorporate information about common features (FP2 to FP4) or about stationary dynamic factors (FP5) perform significantly better than the direct approach for several horizons—all horizons but the first in FP4—(Table 4), and as a whole for the entire forecasting path (Table 5). Similar results are obtained for the EA, except for FP2. For the

Table 9
Diebold–Mariano test results based on a multivariate loss function for the path forecast between two approaches (Capistran, 2006).

UK results	Direct procedure	Indirect procedures based on intermediate disaggregations considered in the paper			Indirect procedure based on factor-augmented models	Indirect procedure based on AR models
	FP1	FP2	FP3	FP4	FP5	FP6
FP1			**	**		
FP2			*	**		
FP3				**		
FP4						
FP5			*	**		
FP6	**	**	**	**	**	

* (**) indicates that the procedure appearing in the column performs significantly better than the procedure appearing in the row at the 5% (1%) significance level.

UK, only approaches FP3 and FP4, which respectively incorporate information about a CSCF, and CSCF and the common trend based on the disaggregation map, significantly outperform the direct approach for the entire forecasting path. Moreover, for all three areas and for all horizons, all of the indirect approaches, FP2 to FP5 (except for FP2 in the EA), have RMSEs below that of the direct method, except in some cases for the one-period-ahead horizon. In addition, in the cases of the US and the UK, the relatively superior performances of these indirect methods improve with the length of the horizon.

Procedures FP3 and FP5 are similar, in the sense that both include only one stationary dynamic factor extracted from the basic components, with the former making use of the disaggregation map and the latter ignoring it. FP3 is significantly better than FP5 for the whole forecasting path for the EA and the UK, but not for the US. Therefore, in this experiment, FP3 performs better than FP5 overall, showing that dynamic factor analysis is more important when it is used in connection with a disaggregation map.

The relevance of our approach becomes more apparent when we consider the indirect procedure, FP4, which exploits the full disaggregation map by incorporating a subset of basic components that share a common trend or a CSCF. For all three areas, this approach gives the best forecasting results, in the sense that they are significantly better than those from all other approaches for the forecasting path as a whole, except for FP3 in the EA. Clearly, therefore, this is the preferred approach. In this exercise, for 12-period-ahead forecasts, using this approach reduces the RMSE by 8% for the EA, 25% for the US and 34% for the UK, relative to the direct method.

The above comments show that there is evidence that distinguishing basic components with a common trend or a CSCF from the rest matters. Also, Tables 5, 7 and 9 show that CSCF (FP3) is significantly more useful than a subset with a common trend (FP2). Our results show that, in a comparison between a direct forecasting procedure and the theoretically efficient one based on a vector model for all of the basic components—which is usually not feasible and often unreliably estimated—the indirect procedure based on single-equation models for the basic elements that share some common features is an intermediate alternative that can forecast inflation successfully in three different economies. The key point seems to be that the procedure incorporates important restrictions between the basic components. This suggests that, when using disaggregated information to forecast an aggregate, one should consider any relevant restrictions which are present in the disaggregated information.

6.2. *Proposed extensions*

To apply our procedure, one must classify a large number of basic components based on their shared features, conveyed by our disaggregation map, which we have shown can be obtained simply by using bivariate methods. The disaggregation map may also be useful in other areas, such as in the application of dynamic factors (as was illustrated in the comparison between FP3 and FP5 above), the definition of useful intermediate aggregates, and the formulation of models with unobserved components. In this paper, we have concentrated on exploiting the restrictions in non-overlapping subsets, the largest subset of basic components with a common trend (N), and the largest subset with a common serial correlation factor (S), but the disaggregation map could be made more sophisticated, as described in Section 2. An extension of the forecasting procedure presented here using a disaggregation map which considers common trends, CSCFs, common seasonality, common non-linearity, co-breaks and common volatility seems promising.

Because the current version of our indirect forecasting approach does not incorporate variables outside the whole information set of consumer prices, it cannot be used to explain the economic determinants of inflation. Nevertheless, the forecasts of the basic components generated by our procedure may shed light on what economic factors constitute the main drivers of inflation. In any case, our procedure could easily be extended to include exogenous variables in the models for the basic components or in the model for the intermediate residual aggregate, r_t, which could be done by applying Autometrics (see Doornik, 2009).

Acknowledgments

We appreciate the comments and suggestions made by Katerina Juselius and by an anonymous referee. We are also grateful for the comments of the participants at the Econometric Seminars of the Nuffield College and the European University Institute. Antoni Espasa acknowledges financial support from Ministerio de Educacin y Ciencia, proyect ECO 2009-08100. Any opinions expressed are solely those of the authors.

This paper was originally distributed as "Forecasting A Macroeconomic Aggregate Which Includes Components With Common Trends".

References

Beaulieu, J., & Miron, J. (1993). Seasonal unit roots in aggregate US data. *Journal of Econometrics*, *55*, 305–328.

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

Capistran, C. (2006). On comparing multi-horizon forecasts. *Economics Letters*, *93*(2), 176–181.

Clark, T. (2000). Forecasting an aggregate of cointegrated disaggregates. *Journal of Forecasting*, *19*(1), 1–21.

Cubadda, G. (2007). A unifying framework for analysing common cyclical features in cointegrated time series. *Computational Statistics and Data Analysis*, *52*(2), 896–906.

Cubadda, G., & Hecq, A. (2001). On non-contemporaneous short-run co-movements. *Economics Letters*, *73*(3), 389–397.

Diebold, F. X., & Mariano, R. (1995). Comparing predictive accuracy. *Journal of Business and Economic Statistics*, *13*, 253–265.

Doornik, J. A. (2009). Autometrics. In J. L. Castle, & N. Shephard (Eds.), *The methodology and practice of econometrics*. Oxford: Oxford University Press.

Duarte, C., & Rua, A. (2007). Forecasting inflation through a bottom-up approach: how bottom is bottom? *Economic Modelling*, *24*(6), 941–953.

Engle, R., & Granger, C. W. J. (1987). Cointegration and error-correction: representation, estimation and testing. *Econometrica*, *55*(2), 251–276.

Engle, R. F., & Kozicki, S. (1993). Testing for common features. *Journal of Business and Economic Statistics*, *11*(4), 369–395.

Espasa, A., & Albacete, R. (2007). Econometric modelling for short-term inflation forecasting in the EMU. *Journal of Forecasting*, *26*, 303–316.

Giacomini, R., & Granger, C. W. J. (2004). Aggregation of space–time processes. *Journal of Econometrics*, *118*(1–2), 7–26.

Gourieroux, C., Monfort, A., & Renault, E. (1991). *A general framework for factor models*. Document de Travail 9107, INSEE.

Guerrero, V. M., & Peña, D. (2000). Linear combination of restrictions and forecasts in time series analysis. *Journal of Forecasting*, *19*, 103–122.

Hendry, D. F., & Hubrich, K. (2006). *Forecasting aggregates by disaggregates*. European Central Bank Working Paper 589.

Hendry, D. F., & Hubrich, K. (2010). *Combining disaggregate forecasts or combining disaggregate information to forecast an aggregate*. European Central Bank, Working paper series, 1155.

Hylleberg, S., Engle, R., Granger, C., & Yoo, B. (1990). Seasonal integration and cointegration. *Journal of Econometrics*, *44*, 215–238.

Kohn, R. (1982). When is an aggregate of a time series efficiently forecast by its past? *Journal of Econometrics*, *18*(3), 337–349.

Lütkepohl, H. (1984). Linear transformations vector ARMA processes. *Journal of Econometrics*, *26*(3), 283–293.

Lütkepohl, H. (1987). *Forecasting aggregated vector ARMA processes*. Berlin: Springer-Verlag.

MacKinnon, J. G. (1991). Critical values for cointegration tests. In R. F. Engle & C. W. J. Granger (Eds.), *Long-run economic relationships: readings in cointegration*. Oxford: Oxford University Press.

Osborn, D. R., Chui, A. P. L., Smith, J. P., & Birchenhall, C. R. (1988). Seasonality and the order of integration for consumption. *Oxford Bulletin of Economics and Statistics*, *50*, 361–377.

Rodrigues, P. M. M., & Osborn, D. R. (1997). *Evidence on testing for unit roots in seasonal time series*. Discussion Paper 9728, School of Economic Studies, University of Manchester.

Rose, D. E. (1977). Forecasting aggregates of independent ARIMA processes. *Journal of Econometrics*, *5*(3), 323–345.

Stock, J. H., & Watson, M. W. (1998). *Diffusion indexes*. NBER Working Paper, 6702.

Stock, J. H., & Watson, M. W. (2002). Macroeconomic forecasting using diffusion indexes. *Journal of Business and Economic Statistics*, *20*(2), 147–162.

Stock, J. H., & Watson, M. W. (2004). Combination forecasts of output growth in a seven country dataset. *Journal of Forecasting*, *23*(6), 405–430.

Stock, J. H., & Watson, M. W. (2005). *Implications of dynamic factor models for VAR analysis*. NBER Working Paper, no. 11467.

Stock, J. H., & Watson, M. W. (2007). Why has U.S. inflation become harder to forecast? *Journal of Money, Credit, and Banking*, *39*(7), 3–34.

Tiao, G. C., & Guttman, I. (1980). Forecasting contemporal aggregates of multiple time series. *Journal of Econometrics*, *12*(2), 219–230.

Vahid, F., & Engle, R. F. (1993). Common trends and common cycles. *Journal of Applied Econometrics*, *8*(4), 341–360.

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

Antoni Espasa is a Fellow of the EEA. He won the Rey Jaime I Prize for Economics, and has a Ph.D. (LSE). He has been a senior economist at the Research Department of the Bank of Spain. At present, he has a chair in Econometrics at the UC3M and is the director of the Flores de Lemus Research Institute and of the publication, *Bulletin of EU and US Inflation and Macroeconomic Analysis*. He has published a book called 'The Spectral Maximum Likelihood Estimation of Econometric Models with Stationary Errors', as well as articles in the *Journal of the American Statistical Association*, *International Economic Review*, *International Journal of Forecasting*, *Journal of Forecasting*, and *Econometric Theory*, among others.

Iván Mayo-Burgos is fund manager in Welzia Management SGIIC S.A. He worked as coordinator of the publication, *Bulletin of EU and US Inflation and Macroeconomic Analysis*, at the Flores the Lemus Research Institute. He received his DEA in economics at Universidad Carlos III de Madrid while he was teaching assistant. Before, he had graduated as Economics at same university, and obtained a Diploma in Business Administration. His main research interests are time series models and forecasting.