EUROPEAN ECONOMY

EUROPEAN COMMISSION
DIRECTORATE-GENERAL FOR ECONOMIC AND FINANCIAL AFFAIRS

ECONOMIC PAPERS

ISSN 1725-3187
http://europa.eu.int/comm/economy_finance

Number 240 January 2006

Using Factor Models to Construct Composite Indicators from BCS Data - A Comparison with European Commission Confidence Indicators

by

Christian Gayer* and Julien Genet**

*Directorate-General for Economic and Financial Affairs

**Hendyplan, Brussels

Economic Papers are written by the Staff of the Directorate-General for Economic and Financial Affairs, or by experts working in association with them. The "Papers" are intended to increase awareness of the technical work being done by the staff and to seek comments and suggestions for further analyses. Views expressed represent exclusively the positions of the author and do not necessarily correspond to those of the European Commission. Comments and enquiries should be addressed to the:

European Commission
Directorate-General for Economic and Financial Affairs
Publications
BU1 - -1/13
B - 1049 Brussels, Belgium

ECFIN/A/3/REP/56011-EN

ISBN 92-79-01181-2

KC-AI-06-240-EN-C

Using Factor Models to Construct Composite Indicators from BCS Data

A Comparison with European Commission Confidence Indicators

Christian Gayer* and Julien Genet**

Abstract

In the framework of the Joint Harmonised EU Programme of Business and Consumer Surveys (BCS) the European Commission publishes two different kinds of monthly composite indicators: simple-average based so-called “Confidence Indicators” and the factor-analytic “Business Climate Indicator”. The Confidence Indicators are calculated for the five different sectors (industry, services, retail trade, building and consumers) covered by the BCS for all Member States, the euro area and the EU as a whole. The Business Climate Indicator is only calculated for the euro area on the basis of data from the industry survey. Starting from this duality and taking into account the recent rise in popularity of factor models in applied economics, the aim of this paper is twofold: First, we explore the option of extending the use of factor-based approaches to the non-industry sectors and to individual Member States. Second, we investigate possible performance gains of factor models over the Confidence Indicators currently computed by the European Commission in tracking the business cycle. For this purpose, we compare the Confidence Indicators with four different factor-analytic methods to extract composite indicators from BCS data.

Key Words: Business cycle, Confidence indicators, Factor models, Principal components

JEL Classification: C43, E32

---

* European Commission, Directorate General for Economic and Financial Affairs, Business Surveys Unit; corresponding author: christian.gayer@cec.eu.int

** Hendyplan, Brussels. The paper was written while J. Genet was working as a consultant for the European Commission; julien.genet@hendyplan.com

The views expressed in this paper exclusively represent the positions of the authors and do not necessarily correspond to those of the European Commission. The paper was presented at the EU Workshop on Recent Developments in Business and Consumer Surveys in Brussels, 16 November 2005. The authors would like to thank P. Weiss for valuable comments and suggestions. Thanks also to D. Grenouilleau for general discussion on factor models. We are grateful to M. Marcellino for providing the computer programmes to run the Forni et al. (2002) generalised dynamic factor model. Special thanks also to A. Carriero for help with the implementation of the Stock/Watson (1989, 1991) dynamic factor model.

1 Introduction

This paper compares different approaches to constructing composite business cycle indicators based on input series from the Joint Harmonised EU Programme of Business and Consumer Surveys (BCS). Currently, so-called *confidence indicators* are calculated for all of the five sectors covered by the BCS, i.e. industry, services, retail trade, building as well as households (consumers).[1] Based on simple averages of selected sector-specific results (balance series), these confidence indicators are calculated in a uniform way for all Member States, the euro area and the EU as a whole.[2]

Since 2001, the European Commission has also been calculating an alternative type of composite indicator, the *Business Climate Indicator* (BCI). This indicator is based on classical static factor analysis and is calculated exclusively for the industry sector of the euro area.

The aim of this research paper is to compare the traditional confidence indicator approach with BCI-like factor-based approaches and, by extending the analysis to all BCS sectors, to fill the gap in the availability of more advanced factor-based alternative indicators for the services, consumers, retail trade and construction surveys. In doing so, we explore the possibilities of improving the tracking performance of the currently published sector indicators relative to reference variables such as industrial production, private consumption, etc. On the methodological side, we employ different factor-analytic model specifications that have been proposed in the literature and compare their performance to that of the existing sector indicators. Our analysis of the five BCS sectors refers to both the euro area aggregate and individual Member States.

We structure our work as follows: first, we shortly present and compare the two existing alternative industry related indicators for the euro area, i.e. the Industrial Confidence Indicator (ICI) and the BCI. We then extend the comparison of the underlying index construction methodologies by computing factor-model-based indicators for all five BCS sectors and assessing their performance against the corresponding benchmark confidence indicators for the euro area.

Within the general class of factor models, several alternative modelling and estimation approaches can be used to extract coincident and leading indicators from a set of component series. In recent years, the focus has been particularly directed to dynamic factor models and their application in economic forecasting. We briefly review the main available approaches and, using exclusively sectoral data from the BCS data base, compare four different factor-extraction methods in terms of the correlation and turning point properties of the derived composite indicators. Moreover, we use the case of the euro area to explore the question of how many and possibly which sectoral input series should be taken into account in applying factor approaches to the available BCS data.

Based on the euro area experience, the analysis is then extended to the comparison of the competing composition approaches at the level of individual Member States, again referring to all five BCS sectors. Unlike the non-model-based confidence indicators, where the selection and (equal) weighting of input series was uniformly determined across countries with particular attention to the performance of the European aggregates, the weighting of individual input series implied by the factor approach is data-driven. This potential to take

---

1 For an overview of the BCS Programme, see European Commission (1997) or the user guide on http://europa.eu.int/comm/economy_finance/indicators/business_consumer_surveys/userguide_en.pdf

2 I.e. the selection of input series and the (equal) weighting thereof is identical across countries and EU aggregates.

into account country-specific patterns of the component series suggests that the gains from factor-model-based indicators could be particularly discernible at the country level.

The main results can be summarised as follows: comparing the factor-analytic BCI with the Industrial Confidence Indicator for the euro area, we find very little gain in using the more sophisticated method of indicator calculation. This is rationalised by the high degree of concordance of the underlying euro area balance series in the case of the industry survey, which renders their composition and weighting rather secondary. Moreover, the already remarkable tracking performance of the simple average-based Industrial Confidence Indicator for the euro area leaves very little room for further improvement from the outset.

A more differentiated picture is obtained in the subsequent systematic comparison of the traditional confidence indicators with different factor-analytic alternatives across all five survey sectors. Focussing on the *euro area* first, the performance differences between the alternative approaches of composite indicator construction are generally rather small. Using the same input data sets that underlie the current confidence indicators for the five sectors, none of the factor models consistently improves over the benchmark indicators. However, some improvements in terms of correlation with sector reference series are found for all sectors when the input data sets are extended to use more or all available balance series relating to the sector in question. These improvements range from a mere one percentage point in the above-mentioned case of the manufacturing industry to around four percentage points in the building and retail sectors. In line with a priori reasoning, we conclude that if a factor-based approach is used to calculate composite indices from (highly correlated) BCS data, it should preferably be based on all available sectoral balance series.

At the *individual country level*, we find stronger evidence that the model-based approaches exploiting the extended data sets lead to significant performance gains over the benchmark confidence indicators. These gains are especially discernible in the consumer and services sectors. We conclude that, from the point of view purely of tracking performance (correlation, mean leading properties), the employed factor model approaches, in combination with the use of expanded input data sets, provide a means of improving over the existing sectoral confidence indicators.

To distinguish between the performances of the different factor models is less easy, both in the case of the euro area and in the case of individual countries. The more advanced dynamic modelling approaches are apparently not able to fully exploit their methodological advantages in our restricted BCS set-up. Given occasional numerical problems with the classical static factor model, therefore, the widely used and methodologically simple principal components method may seem preferable for the regular production of factor-model-based BCS composite indicators of the type presented in this paper.

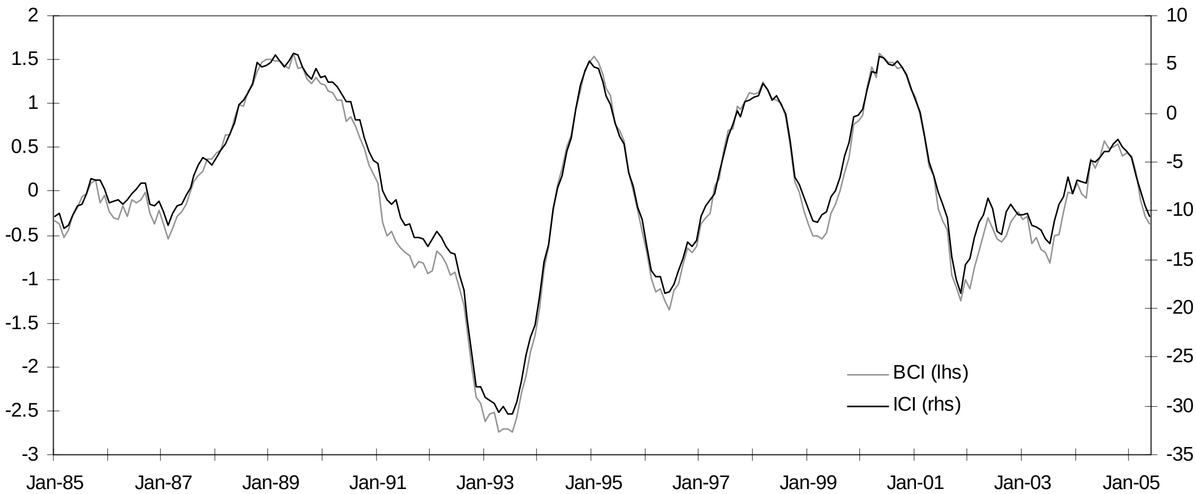
However, the improvements that can be achieved when using factor models instead of simple averages must be weighed carefully against some disadvantages. In particular, the problem of revisions and other practicability aspects linked to the use of any of the factor analytic approaches must be taken into account when gauging their merits compared to the existing confidence indicators. Moreover, using factor models also means that the uniform approach of calculating confidence indicators across all Member States and the euro area and EU aggregates will be lost. In view of not all-too-marked performance gains we stop short of actually recommending the use of factor-based models instead of simple average-based indicators. However, this assessment might change if some methodological refinements of the approaches investigated here are taken into account. In particular, we propose the use of disaggregate BCS input series, the extraction and combination of several factors as well as the implementation of a systematic data pre-selection step in order to potentially further increase the performance margin over the traditional confidence indicators.

2 Industrial Confidence Indicator versus the Business Climate Indicator for the euro area

The general idea underlying a composite business cycle index is to combine several indicators in order to invigorate the cyclical signal and smooth and clean it from idiosyncratic noise. The result can be seen as a proxy for the unobserved common components in the fluctuations of observable economic time series. For the euro area industry sector, the European Commission currently computes and publishes two different composite indicators: the *Industrial Confidence Indicator* (ICI) and the *Business Climate Indicator* (BCI), both of which relate to the manufacturing industry sector of overall business. The BCI (see European Commission (2000)) is based on a static factor analysis of five input questions (standardised seasonally adjusted balances) from the industry survey: *recent production trends, stocks, order books, export order books,* and *production expectations*.

The ICI, on the other hand, is the simple arithmetic average of three of these questions, namely *order books, stocks* (with inverted sign) and *production expectations*, where the aggregation is performed without prior standardisation of the variables. In both cases, the respective input series relate to the euro area as a whole, i.e. they result from a prior weighted aggregation of country survey results, based on value added in the manufacturing industry. Despite the marked differences in methodology and composition of the two composite indicators, they are very similar in practical terms, as shown by Graph 1.

GRAPH 1: BCI and ICI from 1985:01 to 2005:05



The cross-correlation between the two series assumes its maximum for a simultaneous relationship, indicating contemporaneous behaviour, and amounts to over 0.99. In terms of correlation with the reference series, industrial production, both series show an almost identical remarkable performance with values as high as 0.91 or 0.92.[3] Also the turning points of the two series, according to either mere visual inspection or the more formal Bry-Boschan NBER-algorithm, are nearly identical, showing a median lead/lag of zero (see Table 1).[4]

[3] More detailed correlation results will be presented later in section 4.2.

[4] For an earlier comparison of the two indicators see Bengoechea/Martinez (2002).

TABLE 1: BCI and ICI turning points

	P	T	P	T	P	T	P	T	P	T	P	T	T
BCI	10-85	2-87	12-88	4-93	1-95	6-96	3-98	4-99	5-00	11-01	12-02	7-03	7-04
ICI	9-85	2-87	6-89	6-93	12-94	5-96	3-98	3-99	5-00	11-01	10-02	7-03	10-04

Turning points due to Bry-Boschan (1971) algorithm

In terms of short-term volatility, there is hardly any visible difference between the series; both have evolved very smoothly between 1993 and 2001 but tend to reveal a higher degree of volatility during the last three years of the sample. The more erratic direction changes from 2002 onwards, previously only observed in the vicinity of turning points, are thus attributable to the characteristics of the input data rather than to either one of the two methodologies.

Despite this striking similarity of the indicators in general, there are cases where the two move into different directions for one particular month. This was last observed in December 2004 when the BCI increased but the ICI decreased, thus yielding contradictory cyclical signals. These divergences can be explained by the different composition of the BCI and the ICI. The first is driven by two more series than the latter. If, for example, these two series would rise markedly while the remaining three (making up the ICI) would decrease mildly, the ICI would point downwards, whereas the BCI would tend upwards. As can be seen from the joint history of the two series, however, such constellations seem to occur only very rarely, thus indicating a very high degree of cyclical conformity (correlation) of the movements of the five individual input series from the industry survey.

Another principal difference with practical implications between the two indicators stems from their different methodological backgrounds: whereas the ICI, directly following from the definite nature of its qualitative input series and the fixed weighting scheme, is not subject to revisions from one month to the other, the BCI is.[5] Since the method of factor analysis uses all available data points of the underlying input series each month to extract the common business cycle component, the recalculated values of the common component of all previous months are in general affected by the inclusion of the latest data point. Due to the negligible impact of one additional observation on the outcome of the method being applied to monthly data from 1985 onwards, the monthly revisions are hardly of any practical concern (for a closer evaluation, see Deroose et al. (2001)). Technically, the slight revisions from month to month are the cost of applying a weighting scheme that constantly adjusts to the available data so as to extract a factor that best describes the common variation of the input series.

To conclude the comparison of the two existing indicators for the euro area industry sector, we note that both are very close to each other. This observation can be rationalised by a high degree of homogeneity of cyclical information conveyed by the underlying input balance series from the industry survey. Relatedly, the two indicators are almost indistinguishable in terms of correlation with the sectoral reference series. Considering that the simple average-based confidence indicator already offers a remarkably high level of correlation at above 90%, it is obvious that the room for further improvement is extremely limited. These results are specific to the euro area industry sector and do not necessarily carry over to the other four BCS sectors. Indeed, we suspect that for some of the other sectors where the tracking performance of the traditional confidence indicators is less compelling, the benefits from the use of alternative, factor model-based approaches might be more substantial. The same argument applies for composite indicators calculated for individual Member States.

[5] The method used by DG ECFIN for seasonally adjusting the balance input series ("Dainties") is an asymmetric filter which does not give rise to any retrospective revisions. For a recent comparative evaluation of the method, see the study by Franses/Paap/Fok (2005).

3 Factor models: Theoretical background

Factor analysis is a class of statistical methods used to summarise a set of variables by constructing a few "common factors" related to all of the variables and "specific factors" related to each individual variable only. It therefore represents a formalisation of Burns and Mitchell's (1946) notion that business cycles represent co-movements in a set of economic time series. In the case of the BCI, being a common factor extracted from five monthly industry questions, the index is supposed to move contemporaneously with the global industrial activity in the euro area. The specific factors represent the specificities of each question.

The methodology is based on very classical principles. A set of p observed variables (balances of opinion) shall be summarised by a small number k<p of latent variables called common factors. The underlying model supposes that each of the p observed variables results from the combination of both a small number of *common factors* F_j and an idiosyncratic component u_i

$$z_{it} = l_{i1}F_{1t} + \ldots + l_{ik}F_{kt} + u_{it} \qquad , i = 1,\ldots,p, \tag{1}$$

where z denotes the value of the i^{th} standardised observed variable at time t. The term

$$\chi_{it} \equiv z_{it} - u_{it} = l_{i1}F_{1t} + \ldots + l_{ik}F_{kt}$$

defines the *common component* of variable i, and is itself driven by the k common factors.

The factors are supposed to be uncorrelated with each other and with the idiosyncratic factors. The latter are also uncorrelated with each other. The *loadings* l_{ij}, j=1,…,k, give the individual correlations of the k common factors with the ith variable. The squared loadings l_{ij}^2 provide the share of variance of variable i explained by factor j. Summing them over the k factors results in the *communality* of variable i, representing the share of variance explained by all of the extracted factors.

Assuming that the factors and idiosyncratic components are identically and independently distributed (i.i.d.) over time, the static model (1) can be estimated by Maximum Likelihood (ML). The derived common factors can be expressed as linear combinations of the original balance series in order to identify their individual contributions to the composition of the common factors.

Considering the auto-correlated nature of the BCS input series, the factors will, in general, not be i.i.d., but also correlated over time. This suggests formulating a different, i.e. dynamic parametric model for ML estimation. An example is provided by the analysis in Stock and Watson (1989, 1991), where both the factor and the idiosyncratic components are assumed to follow autoregressive processes. To allow for cyclical behaviour of the (single) factor, an initial specification would be to let the level follow an AR(2) process and assume an AR(1) process for the idiosyncratic components, as in

$$z_{it} = \lambda_i F_t + u_{it} \tag{2}$$

$$F_t = \varphi_1 F_{t-1} + \varphi_2 F_{t-2} + \varepsilon_t \tag{3}$$

$$u_{it} = \rho\, u_{it-1} + \varepsilon_{it} \qquad , i = 1,\ldots,p, \tag{4}$$

where ε_t and ε_{it}, the innovations of the common factor and the specific components, should follow white noise processes. The dynamics of the models is similar to that adopted by Doz and Lenglart (1999) for the case of French Business Survey data. The model (2)-(4) can be written in state space form allowing ML estimators of the factors to be obtained via the Kalman filter. A possible drawback of the Stock/Watson (SW henceforth) approach is that it cannot be directly extended to analyse large datasets due to computational problems.

Another means to extract common factors from a given set of data series is *principal component analysis* (PC), which can be viewed as a special method within the broader category of factor analysis (like e.g. ML-factor analysis or principal factor analysis). The objective is to explain the bulk of the variance of the observed data using only a few linear combinations of the original data. Unlike the general factor model (1), which explicitly assumes that the data consists of the underlying factors and the idiosyncratic residual components u_i, the latter specific factors are not taken into account in PC analysis, i.e. no assumptions are made concerning their correlations with each other or the common factors. The principal components are found as linear combinations of the p observed input variables satisfying the following conditions:[6]

i) they are uncorrelated,

ii) the first PC accounts for the maximum possible proportion of the total data variance, the second PC accounts for the maximum of the remaining variance and so on, until the complete variance is absorbed by the p-th principal component.

Schneeweiss (1997) shows that the principal components of the variables coincide under mild conditions with the ML-factor estimates derived under the i.i.d. assumption.

The PC approach has some practical advantages over more general factor estimation methods: As the resulting PCs are invariant to the total number of PCs extracted, one does not have to specify the parameter k (number of factors to be extracted) beforehand. In FA, on the other hand, the estimated factors generally depend on how many other factors are extracted. Moreover, and relatedly, PC offers a means to determine the number of relevant factors to be taken into account by simply looking at the magnitude of the ordered eigenvalues of the input data correlation matrix. The sum of the first eigenvalues to be taken into account divided by the number of input series gives the share of the total data variance explained by these first factors. In this context, PC can be used to gauge how many factors to extract using another estimation method (e.g. ML). Common thresholds for the share of data variance to be explained are 80% or 90%.[7]

Moreover, Stock and Watson (2002a, 2002b) have shown that, for large data sets, the static principal components provide reliable estimators for the factors also when the latter (and possibly the idiosyncratic components) are correlated over time. Given the closeness of the ML factor estimator and principal components as reported by Schneeweiss (1997), the static ML-estimator of model (1) should therefore also provide a reliable estimator when the factors are correlated,[8] and the estimates should be fairly similar to those obtained using the principal components of the variables.

Forni, Hallin, Lippi and Reichlin (2000) criticised the static PC approach on the grounds that it does not exploit the dynamic relationships across the variables. They suggested the use of dynamic principal components, as defined by Brillinger (1981). Basically, the idea is to compute the eigenvectors of the spectral density matrix of the variables at different frequencies and then combine them. However, since the estimation of the spectrum requires

---

6 Technically, the determination of the principal components corresponds to the solution of an eigenvalue problem for the data covariance matrix.

7 For other rules to determine the number of factors to be extracted (Kaiser criterion, screen plot), see e.g. Lawley/Maxwell (1971) or the discussion in OECD (2005). The decision of when to stop extracting factors basically depends on when there is only very little random variability left, and is rather arbitrary.

8 Doz and Lenglart (1999) confirm for French Business Survey data that the factor derived from classic static factor analysis is extremely close to the result of the Kalman filter. They conclude that classic static factor analysis seems to remain useful in the time domain when the data set shows a high degree of contemporaneous correlation.

the use of leads of the variables, this approach cannot be implemented in real time. A modified version of this procedure allowing application in real time was introduced by Altissimo et al. (2001). It is currently adopted to produce the CEPR's coincident Eurocoin indicator (see www.cepr.org). The underlying theory for this refined dynamic principal components approach is presented in Forni, Hallin, Lippi and Reichlin (2002). The main attraction of the dynamic model over static factor models is that the lead/lag structure of the dataset is taken into due account, i.e. the input series are no longer jointly restricted to be contemporaneously affected by the common factors only. Formally, model (1) is generalised to a specification with dynamic factor loadings:

$$z_{it} = l_{i1}(L)F_{1t} + \ldots + l_{ik}(L)F_{kt} + u_{it} \qquad , i = 1,\ldots,p. \tag{5}$$

The dynamic properties of leading and lagging variables are thus adequately exploited, without the need of prior ex-ante classification of the dataset into leading, contemporaneous and lagging series. As it is an important characteristic also of the BCS data that not all variables (questions) move exactly in phase, the explicitly dynamic nature of the Forni et al. approach (FHLR henceforth) might be an essential advantage in an attempt to extract a coincident index from the survey data.[9]

An alternative way to account for potential leads and lags of some of the input series is to introduce the variables not only contemporaneously but also at different lags and leads and then let a static factor model extract an implicitly dynamic common factor from that enriched data set.[10] Of course, the consideration of appropriately shifted lagging series in the data set is only feasible in-sample. It is not an option in a real-time application of such an implicitly dynamic approach.

The current practise of both the five sectoral confidence indicators and the BCI is to summarise the respective data sets within one single composite index, i.e. to extract one sole common factor in the BCI case. In this latter case, as the following results will show, the single factor is sufficient to explain more than 90% of the euro area data variation. At the same time, the extracted factor shows a very high degree of correlation with the obvious industrial activity reference series, the industrial production index (IP). In the case of less inter-correlated input data sets, however, it is likely that two or more factors have to be extracted to "sufficiently" summarize the characteristics of the data set.

In that case, the question emerges of how to best combine these extracted factors so as to achieve the basic goal of creating a one-dimensional composite index capable of tracing a broad activity measure for the given sector. Similar to the approach taken by Forni, Hallin, Lippi and Reichlin in the Eurocoin case, a pragmatic solution is to project the *n* input series on the two or more extracted factors to receive estimates of the *n* individual *common components* χ_i and to compute the desired one-dimensional composite index as the simple average thereof.[11]

---

9 It is worth noting that also the Stock/Watson (1989, 1991) single index model can easily be generalised to account for lagging rather than coincident input variables simply by allowing λ_i to be a lag polynomial in equation (2).

10 For such implicitly dynamic factor specifications see Grenouilleau (2004) or the "stacked version" of Stock/Watson's (1988) model.

11 If an explicitly defined reference series is part of the input data set, the best factor weighting would obviously be based on the respective loadings of that series, i.e. the logical candidate for the cyclical index would be the common component of the reference series (e.g. GDP in the case of Eurocoin). In the current BCS indicator practise, however, there is no explicit consideration of reference series in the stadium of factor extraction, or more generally, indicator combination. Alternatively, one could use *bridge equations* based on appended regressions of given reference series on a set of extracted factors to determine their "optimal" weights. However, all of these approaches rely on the availability of the last observation of the

However, depending on the quality of the survey data, including the adequacy of the questions with a view to the targeted phenomenon to be measured, it is theoretically possible that the second and following factors may explain important shared patterns of the input data while actually not being related to cyclical fluctuations and thus not helpful for tracing the reference series. The requirement of uncorrelated idiosyncratic components would force such survey-specific measurement errors to show up in (one of) the extracted factors.

A further important consideration in this area is that for both the currently calculated confidence indicators and the BCI, but also comparable indices calculated by other institutions, like e.g. INSEE's *synthetic indicator*,[12] the focus is not on producing "optimal" point forecasts (or nowcasts) for specific reference series, but rather on distilling sector-specific "climate variables", exclusively based on the survey series and thus disregarding other potentially helpful information. While such a climate variable should, depending again on the suitability of the survey questions, bear a rather close resemblance to an observable activity variable for that sector, it cannot represent more than the condensed latest developments regarded as characteristic or common to all of the input series.[13]

4 Empirical results for the euro area

In view of the above considerations, as a first approach to the study of factor-based indicators for all of the five BCS sectors and in line with the BCI practise, a single factor has been extracted from all of the sectoral survey data sets (for the shares of variance explained by this factor in the individual sectors and related measures, see Tables 4 and 5 below).

Eight different factor estimates have been generated. These eight estimates result from applying one of the four factor extraction methods at a time (classic factor analysis (FA), principal components (PC), dynamic principal components (FHLR), dynamic factor analysis (SW)) to either all of the monthly questions per survey as input series or just the subset of series currently used for the calculation of the five benchmark sector confidence indicators.

Table A1 at annex presents the individual data series per survey that were used as input for the different factor models. Concerning the Consumer Survey, question Q12 (current financial situation of the household) had to be excluded from the data set due to major structural breaks resulting from questionnaire changes e.g. in Germany and France in 2003/2004. Moreover, on the basis of a priori considerations of the nature of the respective series, two consumer questions relating to prices (Q5 and Q6) and two questions from the Industry Survey relating to price expectations (Q6) and employment expectations (Q7) were excluded from the available data base.[14]

---

reference variable and are thus not suited to exploit the fundamental advantage of the BCS indicators' publication lead over hard statistical data.

12 See http://www.insee.fr and for a theoretic background paper Doz/Lenglart (1999).

13 In principle, if one wants to tie the composite index as close as possible to a specific reference series y, a regularly updated regression of y on the balance input series could be envisaged for determining "optimal" weights. Such an index is, however, very unlikely to be of any empirical use.

14 Graphical analysis of all input series suggests that price questions generally appear to be less directly linked to the underlying "business or consumer climate". Furthermore, from a statistical point of view, studies such as Forni et al. (2001) show that price variables have low degrees of communality with the real economy. Due to the low number of available series for that sector, a price-related question was only retained in the building survey data set (Q5). The exclusion of the employment expectations question from the industry data set is motivated by its distinct lagging behaviour with respect to the other industry questions and the reference series. At the same time, the remaining input data set consisting of questions Q1 to Q5 corresponds to the set used for the calculation of the BCI.

4.1 Correlations with sector confidence indicators

Table 2 shows the correlations between the described eight factor-based indices and the confidence indicators for each of the sectors. Except for the services sector, where survey data for the euro area is only available from 1997 onwards, the samples cover the period from 1985 to May 2005.

TABLE 2: Correlations of factor estimates with confidence indicators

INDU	*FA_3*	*FA_5**	*PC_3*	*PC_5*	*FHLR_3*	*FLHR_5*	*SW_3*	*SW_5*
	0.99	1.0	1.0	0.99	1.0	1.0	0.97	0.97
CONS	*FA_4*	*FA_9*	*PC_4*	*PC_9*	*FHLR_4*	*FLHR_9*	*SW_4*	*SW_9*
	1.0	0.90	0.99	0.93	0.99	0.93	0.95	0.97
SERV[#]	*FA_3*	*FA_5*	*PC_3*	*PC_5*	*FHLR_3*	*FLHR_5*	*SW_3*	*SW_5*
	0.98	0.99	1.0	0.99	1.0	0.99	0.98	0.99
BUIL	*FA_2*	*FA_4*	*PC_2*	*PC_4*	*FHLR_2*	*FLHR_4*	*SW_2*	*SW_4*
	1.0	0.98	1.0	0.97	1.0	0.98	0.99	0.98
RETA	*FA_3*	*FA_5*	*PC_3*	*PC_5*	*FHLR_3*	*FLHR_5*	*SW_3*	*SW_5*
	0.94	0.96	0.98	0.96	0.99	0.96	0.95	0.94

*corresponds to the BCI. [#] Series starting in 1997 only. The number of input series used in each case is given by the integer following the underscore, e.g. PC_5 denotes static principal components based on five input series from the respective sector.

All of the reported contemporaneous correlations represent at the same time maximum cross-correlations. This implies that the relation between any of the factor estimates and the respective sector confidence indicator cannot be improved by leading or lagging one of the series. Therefore, all of the extracted factors can be regarded as coincident with the existing confidence indicators.

Moreover, the reported correlations are overall very high. In the case of industry (INDU), services (SERV) and building (BUIL), all of the eight new factor based indices are nearly perfectly correlated with the existing confidence indicators, at levels above 97%. The differences between the eight factors, and therefore both the influence of applying either one of the factor extraction methods and of using either all or only a subset of the input series, are obviously very limited for these three sectors. The entry for FA_5 in the INDU row gives the already mentioned 99% correlation between the ICI and the BCI. For retail trade (RETA), the level of correlation with the existing confidence indicator is overall very high, too, at levels around 94-99%.

In the consumer survey case (CONS), the correlations are more responsive to the set of input series: using only the four series currently establishing the consumer confidence indicator results in high correlations of above 99% for three out of the four factor methods, while using the whole set of nine questions markedly lowers the level of correlation to 90-93% for these models.[15] As the relative difference between the two input data sets (nine compared to four series) is more marked for CONS than for the other sectors, this seems a plausible initial result.

As to the four different factor extraction methods, there is no case where one of the methods would consistently (i.e. using either all or only part of the data set) show higher correlations with the benchmark series than all of the other methods. However, averaging over sectors and

[15] The exception is the SW model, where the factor based on nine series is slightly closer related to the confidence index than the more parsimonious factor.

data input sets, the factor estimates resulting from the dynamic SW approach seem to be slightly less related to the current confidence indicators than the estimates due to the other factor approaches.

Nevertheless, as a first general result, the different indicators generated by means of factor analysis bear a remarkable similarity with the simple unweighted average-based confidence indicators for all of the five sectors considered. This is corroborated by visual inspection of the graphs given in Annex C.

4.2 Correlations with sector reference series

In a second step, all of the generated indices were assessed with a view to their correlations with sector-specific reference series. For INDU, the industrial production index (annual growth rates) is the obvious choice. For CONS and RETA, private consumption was chosen (annual growth rates). In the latter case, annual growth of retail turnover served as a second reference series. For BUIL, the smooth trend-cycle component of the production volume index in construction was selected (in annual growth rates) and for SERV annual growth of gross value added (GVA) in services mirrors the sector activity. Reference series that are only available on a quarterly basis from the national accounts (private consumption, GVA in services) were transformed to monthly frequency by linear interpolation.

To gauge the level of correlation of the factor-based indicators with the selected reference series, the existing confidence indicators were also included in the exercise as benchmarks. Table 3 summarises the results, where the first row for each sector gives the *contemporaneous* correlations with the respective reference series and the second row reports the maximum correlations taking into account possible leading (+) or lagging (-) behaviour of the indicators. The central part of the table refers to indicators based on the same data sets underlying the sector confidence indicators, while the rightmost columns refer to indicators based on the extended data sets. The number of series taken into account in each case is given by the integer following the underscore in the indicators' identifiers.

TABLE 3: Correlations of factor estimates and confidence indicators with reference series

		Restricted data-sets					*Extended data-sets*			
INDU	IP growth	Confidence_3	FA_3	PC_3	FHLR_3	SW_3	FA_5	PC_5	FHLR_5	SW_5
	correl_0	0.91	0.91	0.91	0.91	0.89	0.92	0.92	0.92	0.93
	correl_max(*l*)	0.91 (-1)	0.91 (0)	0.91 (0)	0.91 (-1)	0.89 (0)	0.92 (-1)	0.92 (-1)	0.92 (-1)	0.94 (-1)
CONS	Priv. Cons. growth	Confidence_4	FA_4	PC_4	FHLR_4	SW_4	FA_9	PC_9	FHLR_9	SW_9
	correl_0	0.82	0.82	0.82	0.82	0.80	0.85	0.84	0.84	0.83
	correl_max(*l*)	0.82 (0)	0.82 (0)	0.82 (0)	0.82 (0)	0.81 (2)	0.86 (-2)	0.85 (-2)	0.85 (-2)	0.83 (-1)
SERV	GVA growth	Confidence_3	FA_3	PC_3	FHLR_3	SW_3	FA_5	PC_5	FHLR_5	SW_5
	correl_0	0.86	0.89	0.86	0.87	0.89	0.89	0.89	0.89	0.90
	correl_max(*l*)	0.89 (3)	0.89 (1)	0.88 (3)	0.88 (2)	0.89 (1)	0.90 (2)	0.91 (2)	0.90 (2)	0.91 (1)
BUIL	Production growth	Confidence_2	FA_2	PC_2	FHLR_2	SW_2	FA_4 *	PC_4	FHLR_4	SW_4
	correl_0	0.43	0.43	0.43	0.45	0.47	0.48	0.46	0.47	0.48
	correl_max(*l*)	0.45 (-2)	0.46 (-2)	0.46 (-2)	0.46 (-2)	0.52 (-3)	0.51 (-2)	0.49 (-2)	0.50 (-2)	0.52 (-3)
RETA	Turnover growth	Confidence_3	FA_3 *	PC_3	FHLR_3	SW_3	FA_5	PC_5	FHLR_5	SW_5
	correl_0	0.27	0.29	0.25	0.27	0.36	0.36	0.36	0.37	0.39
	correl_max(*l*)	0.58 (-1)	0.70 (-1)	0.56 (-1)	0.60 (-1)	0.60 (-1)	0.61 (-1)	0.58 (-1)	0.61 (-1)	0.60 (-1)
	Priv. Cons. growth	Confidence_3	FA_3 *	PC_3	FHLR_3	SW_3	FA_5	PC_5	FHLR_5	SW_5
	correl_0	0.65	0.52	0.59	0.61	0.56	0.68	0.70	0.69	0.62
	correl_max(*l*)	0.66 (-1)	0.57 (-2)	0.60 (-1)	0.62 (-1)	0.64 (-3)	0.71 (-2)	0.71 (-1)	0.72 (-2)	0.69 (-3)

Notes: See Table 2. Negative *l* marks a lag of the respective survey index. *Heywood cases: BUIL FA_4 equals Q4, RETA FA_3 equals Q1 of respective survey.

Focusing on INDU first, in line with the very high level of correlation of the different factor estimates with the benchmark confidence indicator, the performance of all indicators against IP growth is very similar. Importantly, no improvement over the confidence indicator can be achieved by applying the factor models to the same input data set of three series. Also comparing the factor estimates using the expanded data set with the confidence indicator, we see a negligible average correlation improvement of one percentage point only.[16] Obviously, the insensitivity of the composite index to the input of either three or five series is due to the high level of pair-wise correlation of these input series. As Table A2 in Annex A shows, correlations range from 85% to 94% in absolute terms, rendering the weighting issue (including zero weights) secondary.

Comparing the factor models using either the restricted or the extended data set, correlation is mostly a mere one percentage point (pp) higher when using all five instead of three questions only (uniformly 92% instead of 91% for FA, PC, FHLR). Only the dynamic SW approach shows a more substantial difference of 4 pps between the two input data sets, with the complete set producing better and the small set worse correlation results than the three rival models. In terms of mean temporal relationships with the reference series, three of the factor models seem to indicate a slight lagging effect of extending the data set to all five input series. In these cases, the maximum correlation is reached for a one month lag instead of a contemporaneous relation. However, the differences in correlation between the zero and the one month lag are negligibly small everywhere, not showing up in the rounded figures in Table 3. Moreover, the benchmark confidence indicator, although based on three series only, formally reveals a one-month lag itself.

Also in the consumer survey case (CONS), no improvements result from the different factor approaches when using the same input data set underlying the simple Consumer Confidence Indicator (CCI). However, the correlation with the reference series can be enhanced by around two percentage points by basing the index on all nine survey questions. The maximum correlation can be increased by on average three percentage points. On the other hand, the broader based indices show signs of a two months lag, while the benchmark CCI moves contemporaneously with private consumption growth. This observation can be rationalised by the fact that the restricted set of questions making up the CCI entirely refers to households' expectations (in theory 12 months ahead), whereas the remaining series ask for households' perception of the current and past situation, thus introducing a relative lag. In terms of the different modelling approaches, both indicators extracted by the SW model show a slight lead of one to two months with respect to the other factor approaches. At the same time, the level of correlation is somewhat lower for the SW indices.

For the services sector (SERV), a similar improvement of correlation as with CONS can be observed when including the full set of balance series. The contemporaneous correlation is improving by around three percentage points over the traditional confidence indicator, while the maximum correlation edges upwards by one to two points. Comparing the two different input data sets for each of the factor models, the slight improvement of correlation due to the two more series is not at the cost of a systematic adverse effect on the leading behaviour. Based on the original data set of three series, the factor estimates can not outperform the confidence index, especially when looking at maximum correlations.

Analogous results are obtained for the building sector (BUIL): The level of correlation with the production index can be enhanced by 2-5 pps using FA, PC or FHLR by including the two questions that are not part of the confidence index (past activity and price expectations),

[16] The case of the "FA_5" index, representing the BCI, replicates the observation from section 2 that the tracking performance of the BCI is practically undistinguishable from that of the ICI.

without any systematic effect on the lead/lag properties of the resulting index. The average gain over the confidence indicator is around 4 pps. Based on the two currently used balance series only, the results are extremely close again between the different factor estimates and the benchmark indicator. The only exception is the dynamic SW approach. Here, both factor estimates (based on two or four series) are extremely alike, as shown by a correlation coefficient of practically one. Essentially, the two factors turn out to be smoothed versions of one of the input series, namely Q4. This also explains their close similarity with the estimate derived from classic static factor analysis (FA_4). The latter is subject to a so called Heywood case, i.e. it coincides with one of the input variables (precisely Q4, expected employment change).[17] This might point to either too many or too few common factors, to a lack of data to provide stable estimates or generally to the fact that the common factor model is not appropriate for the four building survey series.

Concerning RETA, the level of contemporaneous correlation with the first reference series, turnover growth, can be enhanced considerably with respect to the traditional confidence indicator by using any of the factor approaches based on the extended set of questions. Looking at maximum correlations, however, the improvements vary from zero to three percentage points only.[18]

Choosing private consumption as reference series, the factor estimates using all five survey series consistently generate considerably higher correlations with the reference series than the factors based on the restricted data set. In fact, the latter perform consistently worse than the benchmark confidence indicator, consisting of the same input data set. Comparing the factor estimates based on the whole information set with the benchmark indicator, the correlation gain is generally between sizeable 3 to 5 pps in the case of contemporaneous correlation and 5 to 6 pps in the case of maximum correlation.[19] Overall, the level of correlation with the different indicators is significantly higher for the consumption-related reference series than for the highly volatile turnover series.[20] Consumption growth, therefore, seems to be the more relevant reference series for the composite retail indices.

To sum up, applying factor-analysis as opposed to simple averaging of an identical set of series does not generally improve the correlation results. Only in a few cases could the weighting implied by the factor models improve over the simple averaging implied by the confidence indicators; in other cases (RETA vs. consumption growth) the effect was actually negative. Also on average across the five sectors the effect is slightly negative, albeit insignificantly, at -1 percentage point.

On the other hand, an important advantage of the factor approach is that it allows for the use of a larger set of input data. Except for INDU, where the gain amounted to merely one percentage point, some noticeable improvements could be achieved by including more input series in the data sets than just the restricted number making up the individual confidence indicators. Averaging over sectors and factor models, and opting for consumption growth as the reference variable for the retail indicator, the average gain in contemporaneous correlation

---

17 Technically, the respective communality is set to one in the course of the ML-iterations. See e.g. Lawley and Maxwell (1971) for the problem of Heywood cases in the iterative ML estimation process.

18 The result of FA_3 is due to a Heywood case as was just discussed for FA_4 in the case of the building survey. Here, the extracted "common factor" equals retail question Q1 (past business activity). The difference compared to the building sector is, however, that for retail the Heywood problem only occurs in the case of the restricted input data set, but not for the whole data set of (here) five input series. Therefore, it does not invalidate the factor results using the complete data set.

19 Only the SW model produces somewhat different results.

20 Looking at the same shorter sample for consumption as for the turnover series (starting in 1996 only instead of 1992 in the case of consumption growth), differences are even more marked.

from the inclusion of the additional BCS series is between three and four percentage points. The average gain of the factors based on the expanded data set with respect to the traditional confidence indicators is around two and a half percentage points. These gains have to be gauged against a mean correlation level of the existing confidence indicators of close to 75%.

Lastly, comparing the different approaches to extract a common factor from the euro area BCS data, no clear conclusions arise as to which of the employed methods is preferable in terms of correlation with the selected reference series. Averaging over sectors and input data sets, all of the four approaches generate a more or less unique mean contemporaneous correlation of around 75%. Looking behind these average values, the dynamic factor model due to Stock/Watson appears to show slightly more heterogeneous results when compared to the always particularly close results of the FA, PC and FHLR approaches. This finding points to a high degree of robustness of the factor estimates generated by the latter three methods, while, on the other hand, it may point to a lower degree of computational stability of the dynamic ML-estimator using the Kalman Filter.

Moreover, the closeness of the FHLR results to the static PC and FA set-up does not support the idea that the explicitly dynamic nature of the factor loadings could allow for a beneficial exploitation of the dynamic properties of the input series in an attempt to extract a coincident index. Not only the correlations themselves but also the leading/lagging properties deduced from the maximum correlations do not show any systematic differences between the three approaches. Understandably, the FHLR approach, developed for very large input data sets, is more likely to reveal its methodical strengths in an environment where many more series with heterogeneous dynamics are available to the model than are in the present BCS case.

4.3 Turning point behaviour

To shed some more light on the lead/lag properties especially at cyclical turnarounds, the Bry-Boschan routine was used to compute the turning points of the different indicators. Tables A3 to A7 in the Annex provide the results, relative to the turning points (TPs) of the sector reference series. The sample starts are mostly determined by the latter series' availability and range from 1991:1 to 1997:1.

In the *industry* case, Table A3 shows that all of the indicators perform rather similar up to the trough in IP growth in late 2001. Importantly, all indicators capture all of the reference series' turning points. From 2002 onwards, however, only the models based on five input series reliably signal all turning points. On the other hand, they show a slightly higher mean lag of around half a month, which, albeit insignificant, refines the blurred picture of the correlation analysis. Again all factor models based on the restricted data set show very similar characteristics to the confidence indicator (ICI). Among the factor models, the SW approach reveals a slightly worse (i.e. more lagging) TP chronology. In terms of dispersion of turning points around the mean, as measured by the standard deviation of the record of individual leads and lags, the indicators show a very similar performance (last column of Table A3). With standard deviations of around two, all industry indicators reveal a rather stable timing of turning points with respect to the reference variable.

In the case of the *consumer* indices, Table A4 shows a rather homogeneous performance of the extracted factors, too. Notably, all factors based on nine input series and the three methods FA, PC and FHLR based on four series produce TPs that are very close to each other. Interestingly, while correlation analysis pointed to a lagging effect of the inclusion of the additional five consumer series, the turning point analysis indicates a systematically higher mean lead for those indicators based on all nine series. Therefore, when focussing on TP behaviour instead of mean lead/lag relations over the whole sample, there is no trade-off between correlation and timeliness due to the use of the expanded input data set.

The factor based on SW using the restricted data set shows more peculiar TPs. While it correctly anticipates the two TPs in 2002/03, missed by all other estimates, it suggests two additional TPs in 1999 that are not discernible in the reference series (false alarm). From visual inspection of the different indicator series, it is somewhat surprising that the Bry-Boschan algorithm does not identify the above mentioned mini-cycle in 2002/03 in the other indicators based on the restricted data set, including the confidence indicator. Unlike all factors based on the unrestricted set, completely missing out on the TPs, there are signs of a short-lived recovery and subsequent continued deterioration of consumer confidence in all five indices in the relevant period.

The most striking result of TP analysis of *service*-related indices is the finding of a mean lag at cyclical turnarounds between 1 and 3.5 months for all indices, where correlation analysis had pointed to mean leads between one and three months using the whole data sample. It is important to note, however, that the mean TP results are based on short records of mostly four TPs only. Generally, all factors and the service confidence index perform very similar. While all indices reveal long lags at the two TPs in 1998, most of them show a leading behaviour at the subsequent TPs in 2000 and 2003. Only the FHLR and the SW factors based on the restricted set of three series indicate the latest peak in Mai 2004.

The *retail* TP results reveal a rather high degree of heterogeneity between indicators. Most of the extracted factors show a mean lag at TPs, most pronounced for the two SW factors and the factor based on FA using three series only. Moreover, these indices reveal a marked dispersion of turning points around the historic reference dates, rendering their TP signals rather useless. While all factors reveal (at least) two additional TPs around 1998/99, the benchmark confidence indicator seems to trace the TP chronology of the reference series rather reliably. However, as most of the other indices, it missed the last mini-cycle of consumption growth in 2002/03. All in all, owing to the rather high volatility of all retail related indicators, the TP analysis suffers from a larger degree of indeterminacy and arbitrariness.

Overall, the performance of the *building* indices in tracking the cyclical TPs of the reference series (building production growth) is relatively weak. All of them miss several TPs and reveal rather long lags at the turnarounds they share with the activity variable. Except for SW, all factor estimates based on just two input series (FA, FHLR, PC) provide exactly the same TP record as the benchmark confidence indicator. This underscores the previous general finding that applying factor models to the restricted input data sets underlying the confidence indicators does not help to significantly improve the indicators' properties. Comparing the different input data sets, the mean lag is always smaller for those factors based on the full set of four questions.

Summarising the results of the turning point analysis, we find further evidence for the previous observation that, based on an identical data set, the characteristics of the composite indicators are largely unaffected by the use of either simple averages or any of the factor models. Also when including the additional sector time series, the turning point behaviour is in general not significantly altered. Among the four different factor approaches, the SW-based indicators display more peculiar TP characteristics in several cases. Lastly, we encounter some contradictions with the correlation analysis in terms of the mean leading or lagging behaviour of some of the composite sector indicators. However, considering the sometimes loose or arbitrary mapping of indicator turning points with those of the reference series, and taking into account the generally low number of available turning point pairs in the case of

CONS, SERV, RETA and BUIL, the presented turning point results should be interpreted with some caution.

4.4 Variance shares and loadings

In terms of the overall correlation level with the reference series, Table 3 showed that the indices derived from the industry survey lead the range with a value of around 92% with respect to IP. The service survey indicators come second with a still very high correlation level of around 89% with respect to GVA growth, followed by the consumer indices (82-86%). Using consumption as the reference variable for the retail indices results in correlation levels of around 60-70%, while using turnover produces comparable levels only when the indices are lagged by one month (56-70%); otherwise correlations are below 40%. The building indicators achieve a similar weaker correlation level with their activity series of around 45-50%.

Based on these correlation results, it is tempting to establish a ranking in terms of quality or usefulness of the individual surveys, where the industry survey would range first place due to the high level of correspondence with IP, followed by services, consumers and retail and, lastly, building. However, a low level of correlation between survey based indices and a sectoral reference series could, in principle, also mean that the latter series is the (statistical) problem, not the survey series. In that vein, the IP index is clearly a more reliable and cyclically relevant reference series than is for example the more volatile retail turnover.

Instead of comparing the indices with (possibly noisy) reference series, another way to measure their quality is to focus on the extent to which they capture common information in the individual qualitative judgements and expectations. In that sense, a high degree of representation of the total set of sectoral questions may be interpreted as a sign of congruency and reliability of a sector index. The following table gives the shares of variance explained by the extracted static principal components with respect to either the restricted or unrestricted input data sets.

TABLE 4: Shares of variance explained by the first PC

	Restricted set (#)	*Full set (#)*
INDU	0.928 (*3*)	0.916 (*5*)
CONS	0.842 (*4*)	0.718 (*9*)
SERV	0.893 (*3*)	0.885 (*5*)
BUIL	0.960 (*2*)	0.878 (*4*)
RETA	0.638 (*3*)	0.648 (*5*)

(#): Number of series of input data set

The table shows that for INDU, SERV and BUIL the one-dimensional PC-based indices account for up to 90% and more of the variance of the different data input sets. This, again, is of course the logical consequence of a very high degree of correlation within the sectoral data sets. (Absolute) bivariate correlation range from 85-94% for INDU, 78-92% for SERV and 77-92% for BUIL (92% in the case of the restricted set of two series), see Table A2 in Annex A.

In the case of RETA, the lower level of explained variance is directly attributable to the specific behaviour of question 2 concerning the *volume of stocks currently hold*, which is part of both the restricted and unrestricted data set. While the remaining four questions are interrelated at levels between 62 and 81%, the correlation of the stocks questions with any of these series is, as expected, negative but between 15 and 43% in magnitude only.

In the case of CONS it is question 10, asking *whether it is a moment to save*, that is markedly lower correlated with the remaining series than these are among each other (see table A.2). While the latter correlations never fall below 55%, Q10 shows correlations with the remaining variables of generally less than 20%. This idiosyncratic variation explains part of the lower share of explained variance for the full data set.

Due to the high level of correspondence with the extracted principal components, the reported results should also be indicative as to the degree of overall data explanation by the static factor models. Yet, some more details are available from the respective factor model output. Input series belonging to specific sectors have been looked at as rather homogeneous groups so far, without regard to their specific meaning in relation to the targeted reference variable, measured either directly or indirectly using the extracted factors. The loadings (or their squares, the communalities) derived from the static factor model estimations provide a means to have a closer look at individual series' relations with the estimated factors (Table 5). To save space, only the results for the unrestricted data sets are presented.

TABLE 5: Loadings and communalities due to static factor analysis

INDU	loadings	commun.
Q1	0.95	0.91
Q2	0.96	0.93
Q3	0.94	0.89
Q4	-0.94	0.88
Q5	0.93	0.87
average		*0.90*

SERV	loadings	commun.
Q1	0.97	0.94
Q2	0.89	0.79
Q3	0.94	0.88
Q4	0.95	0.90
Q5	0.88	0.78
average		*0.86*

CONS	loadings	commun.
Q1	0.93	0.87
Q2	0.94	0.88
Q3	0.99	0.99
Q4	0.86	0.74
Q7	-0.86	0.75
Q8	0.90	0.81
Q9	0.86	0.73
Q10	0.15	0.02
Q11	0.65	0.43
average		*0.69*

BUIL	loadings	commun.
Q1	0.84	0.70
Q3	0.92	0.85
Q4	1.00	1.00
Q5	0.92	0.84
average		*0.85*

RETA	loadings	commun.
Q1	0.89	0.80
Q2	-0.43	0.19
Q3	0.90	0.81
Q4	0.75	0.56
Q5	0.74	0.55
average		*0.58*

Bold entries: Components of confidence indicators

The results mirror the previous findings: The *industrial* input series are all highly correlated with the extracted factor, with *stocks* (Q4) showing the expected negative correlation with the activity proxy. All of the five *service* questions mirror different facets of activity in the services sector (business situation, demand, employment) and are thus highly and positively related to the extracted factor.

Concerning the consumer surveys, the results show the expected negative correlation of Q7 (*unemployment expectation*) with the extracted factor. As expected from the input data correlations (Table A2), Q10 *(is it a good moment to save?*) is nearly uncorrelated with the extracted consumption factor. Put the other way round, the respective communality of practically zero shows that the extracted factor does not at all contribute to the explanation of Q10. This points to the known problem of an imperfect harmonisation of this question's wording in the national questionnaires, ranging from *"is it a good moment to save*" to "*is it* *convenient to save*". As a consequence, the euro area aggregate, by averaging different concepts, is of very restricted informative value.[21]

The fact that the likewise savings-related Q11, which is part of the confidence indicator, is also comparably less correlated with the consumption factor than are questions Q1 to Q9 might point to specific determinants of households' saving behaviour. To get a clearer signal of "consumption propensity", it may be beneficial to exclude savings-related questions like Q11. As to the earlier finding of a markedly higher correlation with consumption growth when conditioning the factor on all nine questions instead of only four (Table 3), the results in Table 5 suggest that this is due to the inclusion of Q1 and especially Q3 (*general economic situation*), dominating the extracted factor together with Q2. As mentioned earlier, the gain in correlation with consumption is, however, at the cost of a deterioration in the indicator's leading behaviour vis-à-vis the reference series over the whole sample,[22] as only the four currently included series refer to households' *expectations* about the future.

Looking at the retail survey results, it is confirmed that the stocks question (Q2) is rightly negatively correlated with the retail activity factor, but also at a low absolute level. As above for the consumer index, to get a clearer activity related signal, one could consider to exclude this variable from the calculation of a sectoral index. However, while correlations might be improved by the exclusion of specific input series, the respective questions may add important information in special phases of the business cycle, cautioning against their elimination. More work would be needed here.

As was already mentioned above, the results for the building sector are due a maximisation problem (Heywood case), the loading of exactly one for Q4 (employment expectations) meaning that the extracted factor is identical to this particular input series. A variation of the sample for the factor extraction does not alter this finding.[23]

4.5 Scoring coefficients

A different view on the relationships between input variables and extracted factors is provided by the scoring coefficients. These are regression-based estimates of the weights given to individual input series in the linear combination forming the extracted factors. Results are presented for the static factor models based on the restricted data set, i.e. using the same input series as the current confidence indicators. This provides the possibility to compare the factor-model based "optimal" weights with the equal weighting implied by the simple average based confidence indicators.

The common factor in industry derived from the three components of the ICI is equal to

$F_{IND} = 0.26*Q2-0.51*Q4+0.25*Q5$.

The industry factor is thus slightly dominated by the inversed behaviour of the standardised Q4, i.e. stocks of finished products. Before comparing these weights with the equal weighting scheme of the ICI, however, one has to take the standardisation effects into account. While the given scoring coefficients apply to the standardised input series, the confidence indicators aggregate the raw balance series, such that more volatile components dominate the appearance of the index. For a fair comparison of individual series' contribution to the aggregate's profile, therefore, the scoring coefficients have to be divided by the standard deviations (SD) of the raw balance series. In the case of INDU, Q4 has a markedly lower SD,

---

21 The country-specific results (see section 5.2) show, however, that the communalities of Q10 are generally comparably low also at individual country level.

22 Focussing on the timing of the series at cyclical turning points only, such a trade-off was not confirmed.

23 A variation of the starting values for the maximisation process did not help either.

especially compared to Q2. The following table shows the original scoring coefficients along with SDs and correspondingly corrected weights ("cor. score"). The next to last column gives the appropriately rescaled weights so as to add up to one in absolute terms. These weights would apply to the non-standardised balance series and can thus be compared to the equal weights of one third of the confidence index weights given in the last column.

TABLE 6: Industry weights

	Score[1]	SD	Cor. Score	Weight	ICI Weight
Q2	0.26	13.91	0.02	**0.13**	**0.33**
Q4	-0.51	5.18	-0.10	**-0.66**	**-0.33**
Q5	0.25	7.75	0.03	**0.22**	**0.33**
Abs. sum			0.15	1	1

[1]The scores do not sum to up to one but are chosen such that the resulting factor has unit variance

Now the dominance of the stocks question becomes even more pronounced, with a double weight as compared to the ICI weighting. Notwithstanding these significant differences in weighting, the preceding comparisons in Table 2 and 3 showed that, due to the high level of correlation between the input balance series, the resulting composite indices are very similar.

The common factor derived from the four components making up the consumer confidence indicator (CCI) is

$F_{CONS} = 0.26*Q2+0.22*Q4-0.50*Q7+0.04*Q11$,

and is thus dominated by inverse Q7 movements (unemployment expectations). The following table shows the weights rendered comparable to those used in the consumer confidence indicator, thus referring to the non-standardised components.

TABLE 7: Consumption weights

	Score[1]	SD	Cor. Score	Weight	CCI Weight
Q2	0.26	3.68	0.07	**0.50**	**0.25**
Q4	0.22	8.88	0.03	**0.18**	**0.25**
Q7	-0.50	13.64	-0.04	**-0.26**	**-0.25**
Q11	0.04	4.67	0.01	**0.06**	**0.25**
Abs. sum			0.14	1	1

[1] see Table 6

Now, Q2 (*Financial position expectations*) receives the highest weight. Again, the differences in weighting compared to the confidence indicator are in marked contrast to the finding of a correlation of 99% between the two indices FA_4 and CCI (Table 2).

The common factor extracted from the three standardised components of the services confidence indicator (SCI) is

$F_{SERV} = 0.86*Q1+0.05*Q2+0.09*Q3$,

and is dominated by the influence of Q1 (*recent business situation*).

Taking the standardisation effects into account, the weights as presented in the next-to-last column of Table 8 can be compared to those of the confidence indicator (last column).

TABLE 8: Services weights

	Score[1]	SD	Cor. Score	Weight	SCI Weight
Q1	0.86	16.89	0.05	**0.81**	**0.33**
Q2	0.05	14.23	0.00	**0.06**	**0.33**
Q3	0.09	10.84	0.01	**0.13**	**0.33**
Abs. sum			0.06	1	1

[1] see Table 6

Here again, due to the extreme interrelations between the balance series involved, the significantly higher weight of Q1 has nearly no effect when compared to the equally weighted SCI, as was demonstrated by a correlation of 98% (Table 2).

The static factor extracted from the three standardised retail series equals question Q1 from the retail trade survey (see above) and is therefore not a valid factor solution. Recalling the correlation results from Table 3, series Q1 alone is apparently better suited to trace turnover growth but less suited to match consumption growth when compared to the confidence indicator giving equal weight to the three components Q1, Q2 and Q4.

For the building sector, the factor approach results in the following index

$F_{BUIL} = 0.5*Q3+0.5*Q4$.

It thus reproduces the equal weighting as imposed by the simple averaging approach of the confidence indicator. Only slightly changed weights apply to the non-standardised components, as shown by Table 9.

TABLE 9: Building weights

	Score[1]	SD	Cor. Score	Weight	Buil CI Weight
Q3	0.50	13.90	0.04	**0.47**	**0.50**
Q4	0.50	12.13	0.04	**0.53**	**0.50**
Abs. sum			0.08	1	1

[1] see Table 6

4.6 Robustness of euro area results

All of the presented factor results for the euro area are based on models where a single factor is extracted from the data. The techniques are used rather heuristically to arrive at data-driven weighting schemes for one-dimensional indices. As indicated by a non-negligible correlation between the idiosyncratic components of several input series in the case of the classic static factor model (1), the models may not be valid from a statistical point of view.[24] However, experiments with two-factor models showed that the extracted first factors are practically unchanged compared to the presented results; in the case of the static principal components method (PC), results are per se invariant to the number of extracted factors.

Moreover, the current set-up of both the confidence indicators and the BCI is that of "stand-alone indices", i.e. they combine BCS series without regard to specific reference series at the stage of construction. The underlying idea is thus that the indices distilled from the BCS information set intrinsically relate to, or even represent, the cycles of the various sectors. Consequently, with the current approach it is an open question of how one would have to combine two or more extracted factors. As was argued earlier, it is probable that additional factors extracted from the homogeneous BCS input data capture survey typical measurement errors rather than supplementary cyclically informative information. All in all, the limitation to single factor models seems justifiable, especially for the sectors INDU, SERV and BUIL where around 90% of the data variance were explained by the first factors.

24 Similar caveats apply to the dynamic SW factor model. Here, the residuals were checked for homoscedasticity, serial independency, and normality. While the first property was never rejected, the latter two tests pointed to problems in several cases. A more detailed specification search has shown that increasing the number of lags or tailoring the specification for each sector does not improve substantially the outcome of the diagnostic tests, but it also does not alter significantly the estimated factor. In line with the results from the other factor models, this suggests that the resulting indicator is rather robust to model specification.

4.7 Conclusions from the euro area case

Overall, performance differences between the nine compared approaches of composite index construction (four factor models applied to two different data sets plus the benchmark confidence indicators) are rather small. Especially for INDU and SERV, the high inter-correlation of balance input series explains the very low sensitivity of the output index to composition effects. A key result of the euro area related analysis is that none of the factor models consistently improves over the benchmark confidence indicators when using the same restricted input data set.

When the input data set is expanded to use more or all available balance series relating to the sector in question, some improvements in terms of correlation with sector reference series are found for all sectors. In the industry case, where the level of correlation of the confidence indicator is already very high at above 90%, these gains from exploiting the whole available data base are negligible at one percentage point. On average across all five sectors, the gain over the existing confidence indicators is some two and a half percentage points, which, compared to an average correlation level of these confidence indicators of already close to 75%, still appears rather limited.

Nevertheless, as a first general result in line with a priori reasoning, using a factor based approach to calculate composite indices from (highly correlated) BCS data, the procedure should apparently be based on all available sectoral balance series. At the same time, the consumer survey case showed that trade-offs may be encountered with regard to the leading properties of the extracted factors. Since only the restricted input data set underlying the confidence indicator refers to consumers' expectations, factors extracted from the broader set including backward-looking questions show a comparative lag when focussing on the mean behaviour over the whole sample. Though, in this case, the trade-off is not confirmed when focussing instead on turning points only, the case is still important. It illustrates the fact that the properties of the extracted factor do, of course, depend crucially on the characteristics of the input data set. If the latter shows systematic heterogeneities, e.g. with respect to the mean timing, a corresponding input data pre-selection step will lead to factor estimates with different properties. In the analyses carried out so far, we have not yet systematically considered the effects of screening the input data beforehand; we rather looked at the sector balance series as a fairly homogeneous information pool.

Discriminating between the four different model-based approaches to estimate the unobserved factors underlying the BCS data is not straightforward. The FA, PC and FHLR indicators are very close, while the indicators following from the SW approach show some peculiarities. This might point to a higher degree of robustness of estimates derived from the first three approaches.

5 Empirical results for individual Member States

The choice of questions used for the calculation of the confidence indicators was guided by the aim of achieving an as high as possible correlation of the confidence indicator with a reference variable at the *aggregate* (EU or euro area) level. While this selection of questions might be "optimal" at the aggregate level it is often less appropriate to achieve a high correlation between a national confidence indicator and the corresponding national reference series.

An important advantage of the factor approach over fixed-weight confidence indicators stems from its flexibility with respect to individual country characteristics, where some countries may show particularly poor results for some questions or sectors or may have specific questions that are more important in the national context than for European aggregates. Under these circumstances, the data-driven factor approach can be expected to produce better tailored, and thus more useful composite indicators.

Therefore, the next step is to extend the analysis to compare the traditional confidence indicators with their factor-based rivals on individual country level. At the same time, this should broaden the empirical basis for drawing conclusions as to the most appropriate method to calculate BCS-based indices.

Considering the euro area result that the application of the factor models to the restricted input data sets does not lead to improvements over the benchmark confidence indicators, the different models are run on the basis of individual countries' whole input data set only. The confidence indicators continue to serve as sectoral benchmark indices. Moreover, considering the above euro area experience, the factor approaches under study are restricted to static FA and PC plus the dynamic FHLR approach.

5.1 Industry: Only small improvements across euro area countries due to factor model based indicators

Coincident correlation with the IP reference series in euro area member states is only moderately affected by using either the current confidence indicator or one of the factor-based approaches, the latter exploiting two more input series. Looking across the twelve countries, average correlation edges up from 0.59 for the benchmark indicator to 0.60 (FA) or 0.61 (PC, FHLR) (see last row of Table 10). The two exceptions from this modest overall picture are Austria and Portugal. For the latter country, each of the three factor models generates factor estimates that enhance correlation with IP growth by around 5 percentage points. In Austria, on the other hand, only the static factor model (FA) produces a significant improvement of 6 pps, while the other models show more moderate improvements of merely one or three pps.

TABLE 10: Correlations and lead/lag properties of industry indicators vis-à-vis IP growth

	Correlation				Maximum of correlation				Lead (+) or Lag (-)			
	Conf indic.	FA	FHLR	PC	Conf indic.	FA	FHLR	PC	Conf indic.	FA	FHLR	PC
AT	0.65	0.72	0.68	0.66	0.69	0.75	0.72	0.70	-1	-1	-1	-1
BE	0.63	0.62	0.63	0.64	0.63	0.62	0.63	0.64	1	0	0	0
DE	0.84	0.83	0.85	0.85	0.84	0.83	0.85	0.85	0	0	0	0
EL	0.40	0.39	0.39	0.38	0.42	0.45	0.43	0.42	-1	-1	-1	-1
ES	0.68	0.68	0.70	0.71	0.72	0.73	0.75	0.76	-2	-3	-2	-2
FI	0.69	0.69	0.67	0.67	0.69	0.69	0.68	0.68	1	0	1	1
FR	0.79	0.81	0.81	0.81	0.82	0.85	0.85	0.85	-2	-2	-2	-2
IE	0.48	0.51	0.50	0.49	0.48	0.51	0.50	0.49	0	0	0	0
IT	0.68	0.67	0.69	0.71	0.70	0.72	0.74	0.73	-2	-2	-2	-2
LU	0.48	0.47	0.50	0.50	0.48	0.48	0.51	0.51	-2	-2	-1	-1
NL	0.44	0.45	0.47	0.47	0.46	0.46	0.47	0.47	-2	-2	0	0
PT	0.37	0.42	0.42	0.41	0.43	0.47	0.47	0.47	3	3	3	3
Average	0.59	0.60	0.61	0.61	0.61	0.63	0.63	0.63	-0.6	-0.8	-0.4	-0.4

Looking at the four largest countries, average improvements due to the factor approaches are negligible in the range of 1-2 pps. For Germany, Italy and Spain, the FA approach even leads to slight (insignificant) deteriorations, while the PC and FHLR indices record slight improvements between one and three pps. For France, all factor-based indicators consistently show an improvement of two pps. The results for the Netherlands and Belgium are similar, reflecting the general picture of no or minor positive changes due to the use of the factor models. Greece and Finland are the only countries where contemporaneous correlations due to the factor approaches are consistently below the benchmark confidence value, though by an insignificant average of 1 pp only.

The level of contemporaneous correlation with the reference series is overall rather high at around 60% across all countries and indicators. While comparably low correlations around 40% are observed for the smaller countries Greece and Portugal, the economic heavyweights display values around 70% (Italy, Spain) or even above 80% (Germany, France).

Taking leads and lags into consideration in the computation of correlation coefficients and focusing on the achieved maximum correlations, the average link between survey based indices and the reference series is increased to 63%. Italy and Spain now show correlations around 75%, Germany and France around 85%.

The results as to improvements due to the three factor models over the benchmark confidence indicator are hardly changed when looking at maximum correlations (mid panel of Table 10). Averaging over all countries, all three factor models show a moderate improvement of 2 pps over the confidence indicator, now including also Greece. Only Finland remains showing a tiny minus of 1 pp for two of the factor models in comparison to the benchmark confidence indicator. In three of the larger countries, France, Italy and Spain, we observe improvements around three pps. Again, the PC and FHLR approaches are somewhat ahead of the FA results for most of the larger countries.

Looking at the average leads or lags implied by the maximum cross correlations, we observe very few differences between the competing indicator approaches. The deterioration by one month implied by all factor approaches compared to the confidence indicator for Belgium is the only case where a consistent pattern can be identified. Taking into account the smooth evolution of cross-correlations as a function of the time lag, a difference of one month is, however, hardly relevant. For the Netherlands, the approaches PC and FHLR imply indicators with an improved timing of two months, i.e. the timing vis-à-vis IP growth is changing from a two-month lag in the case of the confidence indicator to coincident behaviour. Lastly, on average, we see a very small temporal advantage of the latter two factor approaches over the static factor approach (FA), which is, however, hardly significant.

To sum up the findings for the industry sector, we observe first of all that a change to factor-based approaches would induce no costs in the sense that neither correlation nor timing properties for any given country would deteriorate. On the other hand, the improvements that we can observe over the confidence indicator are rather small, both on average as well as for most individual countries. Portugal and France are the only countries that, while not seeing the leading/lagging behaviour of their indicators affected, could benefit from a correlation improvement of around three pps by using any of the three factor approaches. For the other countries, improvements are less visible and depend critically on which factor extraction method is employed.

Distinguishing between the three factor approaches is not straightforward: While the correlation averages based on all euro area countries point to more or less identical results, PC and FHLR dominate the static FA approach in the economically more important countries Germany, Spain, Italy, Netherlands and Belgium by 2-3 pps in terms of correlation. Moreover, FA produces slightly worse results in terms of average leading behaviour, as a result of somewhat higher lags (or lower leads) for Spain, Finland, Luxembourg and the Netherlands.

5.2 Consumption: Factor model based indicators are more highly correlated with consumption growth than the confidence indicator but suffer from a slight lag

For almost all euro area countries, the factor models under focus yield indicators that are better correlated with consumption growth than the current confidence indicator is. Their average contemporaneous correlations with consumption growth are, depending on the factor model approach, between 0.59 (FA) and 0.64 (FHLR, PC) compared to 0.55 for the confidence indicator (see leftmost panel of Table 11). The highest gain is obtained for Austria, where the improvement is between a marked 30 and 40 percentage points of correlation. Of course, here, the simple average-based benchmark indicator is very weakly correlated with consumption growth at 0.15, making it relatively easy for the factor models to improve upon by a better combination of the survey input variables but especially by exploiting information from the five additional survey questions.[25]

The gain in terms of correlation is moreover significant at around 10-12 pps for Germany and Spain and at around 5 to 7 points for Portugal and the Netherlands. The two factor approaches PC and FHLR provide also much higher correlations in the case of Finland and Italy, but not the static factor method (FA), which leads to Heywood cases. Therefore, the Finnish and Italian factors are simply equal to one of the input data series, namely the second and first question, respectively, of the consumption survey. Finally, the gain is still positive but comparably small at three pps in the case of France. Belgium is the only case where, depending on the employed factor model, there is no effect (FA) or even a negative effect (PC, FHLR) on contemporaneous correlation with consumption growth.

TABLE 11: Correlations and lead/lag properties of consumer indicators vis-à-vis consumption growth

	Correlation				Maximum of correlation				Lead (+) or Lag (-)			
	Conf indic.	FA	FHLR	PC	Conf indic.	FA	FHLR	PC	Conf indic.	FA	FHLR	PC
AT	0.15	0.56	0.52	0.45	0.17	0.61	0.52	0.47	3	3	2	3
BE	0.57	0.57	0.52	0.51	0.59	0.66	0.60	0.59	-3	-3	-3	-3
DE	0.44	0.58	0.56	0.55	0.47	0.59	0.56	0.55	3	1	1	1
ES	0.74	0.84	0.85	0.83	0.77	0.84	0.85	0.83	3	0	0	1
FI	0.25	*0.11	0.35	0.44	0.26	*0.15	0.38	0.45	2	*3	-3	-3
FR	0.75	0.78	0.78	0.78	0.75	0.79	0.79	0.79	1	-2	-1	-1
IT	0.56	*0.48	0.63	0.65	0.57	*0.55	0.68	0.68	-1	*-3	-3	-3
NL	0.75	0.81	0.80	0.80	0.78	0.82	0.81	0.81	3	3	2	3
PT	0.72	0.81	0.80	0.75	0.75	0.82	0.81	0.77	-3	-3	-3	-3
Average	0.55	0.59	0.64	0.64	0.57	0.61	0.66	0.66	0.9	-0.1	-0.9	-0.6

* Heywood cases. EL, IE, LU not included due to reference series unavailability

Looking at maximum correlations, i.e. taking possible leads or lags of the indicators into account, we observe improvements for all countries when comparing factor based indicators with the confidence indicator. Also for Belgium, the maximum correlation with consumption growth is now improved by the use of factor models. The only two cases showing lower maximum correlations than the confidence indicator are the mentioned Heywood cases of static factor analysis (FA) for Italy and Finland.

While there is thus a clear improvement in terms of correlation when using factor based indicators, the latter appear to be, on average across countries, suffering from the

25 In the Austrian case, all additional five survey questions used for the factor models are individually much higher correlated with the reference series than the four series constituting the confidence indicator. Therefore, their inclusion automatically raises the correlation with consumption growth for any combination of the enriched data set.

disadvantage of a minor lag with respect to consumption growth, while the confidence indicator is displaying a slight average lead (see rightmost panel of Table 11). For the small countries except Finland, the lag or lead is rather unaffected by the employed indicator approach, while for the four big euro area countries the lag or lead is turning for the worse by 2 to 3 months when using the factor models: While the German, French and Spanish confidence indicators display a lead of 3, 1 and 3 months, respectively, the factor estimates are leading for Germany by only 1 month, are lagging by 1 month for France and are contemporaneous for Spain. While the Italian confidence indicator is already lagging by 1 month, all factor model based indicators are lagging by a marked 3 months.

Among the different factor models, the FHLR and principal component methods seem preferable to the static factor analysis from a practical point of view, considering that this method leads to two Heywood cases for Italy and Finland. The latter result does not only render the FA estimates for these countries invalid from a theoretical perspective, but also implies inferior practical correlation results for the FA vis-à-vis the PC and FHLR approaches. On the other hand, excluding Italy and Ireland from the comparison, the FA approach displays a somewhat higher average correlation for the remaining seven euro area countries. In terms of leading behaviour, the three approaches are hardly distinguishable once the deceptive Heywood cases are excluded.

Looking at individual questions' contributions to the static factors extracted by Maximum Likelihood, Table B1 (Annex B) shows that, as in the euro area case, Question 10 (*is it a good moment to save?*) is only relatively loosely related to the factors for all countries but Portugal. Averaging across all countries, Q10 has a mean communality of 0.17 compared to between 0.34 (Q9) and 0.91 (Q3) for the remaining questions. The likewise savings-related Q11 plays a comparably minor role in explaining the common consumption factor, too. As in the euro area case, the latter appears dominated by questions Q3, Q2 and Q1 in the majority of countries.

To conclude, the correlation with consumption growth can be substantially improved for almost all analysed euro area countries by using factor models instead of simple average based confidence indicators. No country would see its consumer indicator deteriorate in terms of maximum correlation. Part of the explanation of the remarkable gains is certainly the larger flexibility provided by the factor model approaches in determining "optimal" weights to construct indicators capturing the country-specific common data trends underlying the various consumption survey series. The simple average based confidence indicator, on the other hand, simply keeps the same fixed weights for all countries. Yet, probably more important is the fact that the factor models are based on a larger dataset of nine questions from the consumption survey, whereas the confidence indicator uses only four of these questions. This data restriction is obviously too restrictive.

However, the apparent strong gain in correlation from using the enriched data set is, just as in the euro area case, at the cost of a lower lead (or higher lag) of the factor estimates relative to the confidence indicators for the four big euro area countries and Finland. On average, the higher correlation of the additional input series is thus partly offset by a relative lag with respect to the currently used four questions of the confidence indicator. As elaborated for the euro area, such a result seems plausible since the latter four questions are all asking for consumers' expectations, whereas the additional questions investigate assessments of the past or the current situation.

All in all, the choice between the traditional confidence indicators and the factor based approaches seems to be a question of whether one gives more weight to mean correlation behaviour or to the mean timing of the relationship with the reference variable as indicated by the correlogram. However, the detailed analysis for the euro area showed that while maximum correlations also pointed to a lagging effect of the inclusion of the additional five consumer series, the turning point analysis gave the opposite result of a systematically higher mean lead for indices based on all nine series. As can be seen in Tables B2-B6 (Annex B), this is partly also true for the individual country results. Comparing the turning points of the confidence indicators with those found in the factor-based indicators, we see that for France and Spain there is indeed evidence of a one to two months lag due to the use of the expanded data set. However, while there is no or no significant impact on turning points in Italy and Finland,[26] the turning points of the German cycle are actually indicated by the new indexes with a slight mean lead compared to the existing confidence indicator. The trade-off between correlation and timing is thus apparently less important when focusing on turning point behaviour. With this in mind, the factor based consumption indicators using the richer data set may seem preferable.

5.3 Services: Looking at correlations and timing, factor model based indicators outperform confidence indicators for all countries

For the services sector, we observe improvements in terms of both contemporaneous and maximum correlations for all countries but Spain. On average, correlations can be enhanced by sizeable 7 pps, up from 0.66 to 0.73 in the case of contemporaneous and from 0.71 to 0.78 in the case of maximum correlations. The largest improvements are recorded for the Netherlands, where correlation rises by 25 pps from 0.62 to around 0.87, for Finland (around plus 10 pps) and for Austria and Italy (plus 6 pps). For the remaining countries, the improvement is at around 3 pps, except for Spain, where contemporaneous correlation actually deteriorates to the same extent. Looking at maximum correlations, i.e. those achieved after leading or lagging the indicators, we see that for Italy the factor based indicators lead to a remarkable increase of around 15 pps over the confidence indicator.

TABLE 12: Correlations and lead/lag properties of services indicators vis-à-vis growth in value added in services

	Correlation				Maximum of correlation				Lead (+) or Lag (-)			
	Conf indic.	FA	FHLR	PC	Conf indic.	FA	FHLR	PC	Conf indic.	FA	FHLR	PC
AT	0.76	0.82	0.81	0.81	0.79	0.88	0.87	0.86	-2	-3	-3	-3
BE	0.62	0.64	0.64	0.64	0.63	0.68	0.66	0.67	-1	-2	-1	-1
DE	0.75	0.78	0.78	0.77	0.85	0.86	0.86	0.86	3	3	3	3
ES	0.74	0.72	0.70	0.71	0.75	0.72	0.70	0.71	-3	0	0	-1
FI	0.42	0.55	0.52	0.48	0.56	0.68	0.66	0.63	-3	-3	-3	-3
FR	0.84	0.87	0.86	0.87	0.85	0.87	0.87	0.88	-1	-1	-1	-1
IT	0.55	0.61	0.63	0.60	0.55	0.70	0.71	0.68	0	3	3	3
NL	0.62	*0.88	0.87	0.86	0.67	*0.88	0.87	0.87	-3	*-1	-1	-1
Average	0.66	0.73	0.73	0.72	0.71	0.78	0.78	0.77	-1.3	-0.5	-0.4	-0.5

* Heywood case. EL, IE, PT not included due to reference series unavailability. No services survey in LU

Looking across the three different factor extractions, we see that they perform very similar on the services survey series. Only for Finland and Italy there is some evidence of a slight underperformance of the PC approach compared to FA and FHLR. On average, however, differences are insignificant.

The favourable effects on correlation of using factor approaches to extract a services indicator from the full set of survey data are accompanied by positive effects on the average timing of

[26] Neglecting the invalid Heywood cases of the static factor model.

the indicators: While for five countries there are no or negligible effects on the lead or lag that maximises correlation over the sample, for three countries the temporal relationship with the reference series changes for the better. For Spain and the Netherlands, the relationship improves from a three-month lag of the confidence indicators to a coincident behaviour (Spain) or to a reduced lag of one month only (Netherlands). In Italy, the above mentioned distinct improvement of maximum correlation due to the factor based approaches is paralleled by a move from a coincident to a leading behaviour of three months.

To sum up, in the services sector, we see clear benefits on the country level from moving from confidence indicators based on three input series to factor based approaches using the full data set of five series. While this conclusion is straightforward for all countries but Spain, we assume that the 3-month improvement in terms of the timing of the Spanish services indicators at least compensates for the slight deterioration of mean correlation of 3 pps.

5.4 Building: Factor-based indicators are higher correlated with production growth for Ireland, Finland, Belgium and France, but lower especially for Germany and Italy

Generally, using factor models improves the correlation with building production growth for a majority of countries. The total unweighted average gain over the eleven countries in terms of contemporaneous correlation is around 2-3 points. This last result is, however, mainly attributable to strong improvements for two smaller countries, namely Finland and Ireland, where correlations rise from around 0.55 to around 0.65 and 0.75, respectively. Excluding these two countries from the calculation, the gain is on average zero. Apart from Finland and Ireland, Belgium and France also see their correlations with production growth significantly enhanced due to the use of the factor approaches. On the other hand, Germany and Italy see significantly lower correlations with a minus of around 5pps.

TABLE 13: Correlations and lead/lag properties of building indicators vis-à-vis production growth

	Correlation				Maximum of correlation				Lead (+) or Lag (-)			
	Conf indic.	FA	FHLR	PC	Conf indic.	FA	FHLR	PC	Conf indic.	FA	FHLR	PC
AT	0.12	0.13	0.14	0.13	0.14	0.18	0.21	0.19	-1	-1	-1	-1
BE	0.19	0.25	0.27	0.26	0.23	0.29	0.29	0.28	3	3	2	2
DE	0.85	*0.82	0.80	0.79	0.86	*0.82	0.80	0.79	2	*2	1	1
ES	0.76	0.76	0.74	0.74	0.78	0.78	0.76	0.76	-2	-2	-3	-3
FI	0.53	0.67	0.66	0.64	0.62	0.72	0.72	0.72	-3	-3	-3	-3
FR	0.51	0.54	0.55	0.56	0.53	0.56	0.57	0.58	-3	-2	-2	-2
IE	0.55	0.72	0.77	0.77	0.64	0.75	0.78	0.79	3	3	-2	-2
IT	0.41	0.35	0.34	0.35	0.49	0.46	0.46	0.47	-3	-3	-3	-3
LU	0.52	0.56	0.51	0.52	0.52	0.55	0.51	0.52	0	0	0	0
NL	0.64	0.65	0.68	0.65	0.69	0.70	0.71	0.69	-3	-2	-2	-2
PT	0.92	0.88	0.90	0.91	0.95	0.92	0.92	0.92	2	2	2	1
Average	0.55	0.57	0.58	0.57	0.59	0.61	0.61	0.61	-0.5	-0.3	-1.0	-1.1

* Heywood case. EL not included due to reference series unavailability

Correlations with production growth deteriorate to a lesser extent for Spain and Portugal, and slightly improve for Austria and the Netherlands. The results are less conclusive for Luxembourg where the change of correlation is either positive, zero or negative according to the factor model used. Finally, concentrating on the four big euro area countries (Germany, France, Italy and Spain) the factor model indicators are on average somewhat less correlated with building production growth than the current confidence indicators is (0.61 compared to 0.63).

The analysis of maximum correlations taking into account possible leads or lags yields similar conclusions except for some minor changes (mid panel of Table 13): the gain for Ireland and Finland is somewhat lower but still high, the deterioration is lower for Italy while the improvement for Austria is now higher than previously. However, considering the extremely low level of correlation for Austria, the changes are hardly of practical relevance.

The key to the correlation differences between the factor models and the simple average based confidence indicators lies in recalling the different datasets used for each method. While the factor model indicators are based on four questions of the building survey (Questions 1, 3, 4 and 5), the confidence indicator is only based on two of these (Questions 3 and 4). For Italy, Spain and Portugal, cases where the simple average outperforms the factor models indicators, the two additional questions used by the factor models are actually less correlated than the two questions used for the simple average. Conversely, for Belgium, Finland, France and Ireland, either the additional first or the fifth question of the building survey has a better correlation with production growth than the two currently used questions.

Comparing the three competing factor methods, we observe that they provide more or less the same positive average gain over the confidence indicator. Considering the four big countries, static factor analysis provides indicators which are slightly better in terms of correlation compared to the other two factor methods. However, in the case of Germany, static factor analysis leads to a "Heywood case", implying that the "factor" is actually mirroring the fourth question of the building survey only. While this indicator alone is apparently better than the two combinations provided by FHLR and PC, Table 13 shows that all three results can easily be beaten simply by averaging the third and fourth survey question, as done in the confidence indicator.

Comparing the indicators' lead or lag vis-à-vis building production growth, the only remarkable result is for Ireland where the confidence indicator and the FA-based indicator are leading by three months while the FHLR and the static principal component approaches show a lag of two months. For all other countries, the lead or lag remains either unaffected or the difference of plus or minus one month cannot be considered significant.[27]

To sum up, in the building sector, the factor model based indicators do not systematically outperform the simple confidence indicator in terms of correlation. While there is a marked improvement for Ireland and Finland and a significant improvement for Belgium and France, we observe a small deterioration for Spain and Portugal and a significant decline for Germany and Italy. Considering the four big euro area countries together, the simple average based confidence indicator displays a slightly higher correlation with production growth than the factor models, despite their clear advantage in the case of France. Finally, with one exception (Ireland), the mean timing of the indicators is not affected by the construction method. All in all, the building results are therefore inconclusive.

5.5 Retail trade: Correlation improvements over confidence indicators for all countries but Austria

In the retail sector, the effect of changing from confidence indicators based on three series to factors extracted from five series is predominantly positive, improving on average by a modest 3 percentage points. Except for Austria, where correlation decreases by around 6 pps for all factor model based indicators compared to the confidence indicator, and Italy, where exclusively the FA results show a deterioration, correlations with consumption growth can be seen to increase for all remaining countries.

[27] As already mentioned, the difference between correlations calculated with a lag or a lead of one month is generally very small.

These improvements are most impressive for France, where contemporaneous correlation rises by 18 pps. While the improvement is also significant for Finland with around plus 8 pps, the still very low level of correlation with the Finish reference series renders this results less interesting from a practical point of view. The bigger countries not mentioned so far, Germany and Spain, show enhancements of 4 and 2 pps, respectively. However, the level of correlation of German indicators is disappointingly low at around 0.45.

TABLE 14: Correlations and lead/lag properties of retail indicators vis-à-vis consumption growth

	Correlation				Maximum of correlation				Lead (+) or Lag (-)			
	Conf indic.	FA	FHLR	PC	Conf indic.	FA	FHLR	PC	Conf indic.	FA	FHLR	PC
AT	0.44	0.39	0.38	0.37	0.47	0.43	0.39	0.38	-2	-2	-2	-2
BE	0.56	0.62	0.60	0.58	0.56	0.62	0.60	0.58	0	0	0	0
DE	0.42	0.45	0.46	0.46	0.42	0.46	0.47	0.46	-1	-3	-1	0
ES	0.79	0.82	0.81	0.81	0.83	0.85	0.84	0.84	3	3	3	3
FI	0.06	0.17	0.12	0.11	0.14	0.27	0.23	0.21	-3	-3	-3	-3
FR	0.52	0.69	0.70	0.71	0.53	0.69	0.70	0.71	-1	0	-1	-1
IT	0.51	0.43	0.51	0.52	0.57	0.54	0.57	0.57	-3	-3	-3	-3
NL	0.73	0.77	0.75	0.74	0.75	0.77	0.76	0.76	-3	-1	-2	-2
PT	0.81	0.83	0.82	0.81	0.81	0.83	0.82	0.82	1	-1	-1	-2
Average	0.55	0.58	0.58	0.58	0.58	0.62	0.61	0.60	-1.0	-1.2	-1.2	-1.1

Note: EL, IE not included due to reference series unavailability. No retail survey in LU

The conclusions based on maximum correlations, i.e. taking leads and lags into account, are broadly unchanged. The average improvement due to the factor models is still at around 3 pps. Only the case of Italy shows that the above mentioned underperformance of the FA approach is much less pronounced, now pointing to a similar performance of all three factor estimates and the confidence indicator. Also on average across all countries, the three factor approaches offer similar correlation results.

Comparing the indicators' lead or lag vis-à-vis the reference series, the mean timing of the indicators is unaffected by the use of different construction methods for the majority of euro area countries. Only for Portugal, we see a deterioration of the correlation-maximising phase shift from a lead of one month for the confidence indicator to a lag of one or two months for the factor approaches.

To conclude, the results as to the use of factor models instead of the current confidence indicator in the retail sector are predominantly positive. The deterioration of correlation in the case of Austria and the slight lagging impact on the Portuguese indicator are the only downers that have to be contrasted especially with the considerable improvements for France.

5.6 Conclusions from individual country analyses

Based on country experience from all five BCS sectors, it emerges that the currently computed confidence indicators can be beaten in terms of correlation in almost all cases by indicators generated from factor models using an expanded data set. Numerous significant improvements are accompanied by only few deteriorations for individual countries and where this happens, the decrease in correlation is mostly negligible. Only in very few cases, deteriorations are observed for economically important countries, like e.g. for Germany in the building sector case.

Improvements are most pronounced in the case of the consumer and services confidence indicators with correlations increasing by around seven percentage points on average across countries. The result for the services sector is somewhat surprising as the euro area related analysis pointed to rather limited gains from moving to the factor-based approaches. The average improvements across countries in the industry, retail and building sectors are less discernible at around two to three percentage points.

Looking across countries and sectors, the documented correlation improvements amount to a moderate total average of around four percentage points. However, this has to be seen in conjunction with an average correlation level of the existing sector confidence indicators with their respective target series of only around 60%. A small caveat is that, due to reference data availability problems, not all countries could be covered in all sector comparisons.

As to the comparison of the different factor models, no clear-cut ranking in terms of correlation or timing emerges. However, the occurrence of four Heywood cases in the static classic factor model (FA) must be seen as a drawback of this particular estimation method. Considering its virtue of relative methodological simplicity, the fact that the static principal components analysis (PC) performs as well as the dynamic FHLR approach speaks in favour of using this particular method for a possible regular production of factor based composite BCS indicators. Apart from its predominant empirical application in the development of composite indicators,[28] PC analysis offers the additional practical advantage that it immediately provides the weights representing the contribution of individual input series to the composite indicator.

Finally, confronting the country results with those received for the euro area, we can confirm our hypothesis that improvements due to the application of factor models, implying a data-driven selection and weighting scheme, are more visible in the individual country case than in the case of the euro area aggregate.

6 Overall conclusions and outlook

The main results of our study can be summarised as follows: with respect to the euro area aggregate, the differences in performance between the compared approaches of composite index construction (simple averages, static factor analysis (FA), principal components (PC), two dynamic factor models (SW, FHLR)) are rather small. For the industry and services sectors in particular, a high inter-correlation of the input series is responsible for the observed low sensitivity of the derived indicators to composition effects. One essential result of the euro area-related analysis is that none of the factor models gives a consistent improvement over the benchmark confidence indicators when using the same (restricted) input data set.

However, for all sectors some improvements in terms of correlation with sector reference series are found when the input data set is expanded to use more or all available balance series relating to the sector in question. The average gain over the euro area-sector confidence indicators is two and a half percentage points, which must be gauged against a mean correlation level of the existing confidence indicators of around 75%.

In line with a priori reasoning, we conclude that if a factor-based approach is used to calculate composite indices from BCS data, it should preferably be based on all available sectoral balance series. An important qualification of this finding is, of course, that each individual input series actually needs to be driven to a certain extent by the underlying unobserved common variable that is believed to mirror the sector cycle. Adding variables that relate intrinsically to different phenomena or suffer from statistical problems such as structural breaks does not improve the factor estimates. Therefore, a certain degree of a priori data screening seems desirable. In our case we excluded three series from the consumer survey and two from the industry survey on the basis of a priori considerations. We did not, however, try

[28] See the OECD (2005) Handbook on Constructing Composite Indicators.

to intentionally tie the extracted factors to the selected reference series from the outset by systematically excluding input variables with lower correlations.

A systematic data pre-screening can also help to better understand and exploit the temporal structure of the input data set. The consumer survey case showed that including variables referring to the past or present economic situation in a set of variables that measure expectations can increase the correlation with the common statistical reference series. At the same time, as one would expect, there seems to be a dilution of the leading properties of the extracted factor. This trade-off illustrates the fact that the properties of the extracted factor, of course, depend crucially on the characteristics of the input data set. Depending on the purpose of the desired indicator, prior selection of the input data according to some well-specified criteria, e.g. a certain minimum leading behaviour, will produce factors with different properties.

In our study, as is the case with the currently published confidence indicators, the focus was on the correlation performance of approximately coincident composite indicators. No systematic attempts were made at this stage to screen the input data beforehand with a view to possibly sharpening the resulting indicators' timing properties. However, first tentative analyses point to the plausible finding that, since leading individual input series can be lagged in a real-time setting but not the other way round, the positive effects of shifting input series on synchronisation are accompanied by an inherent loss of leading properties.

An important general advantage of the factor approach over fixed-weight confidence indicators is its flexibility with respect to individual country developments, which allows a tailor-made, data-driven weighting of input series. Comparing three of the above-mentioned factor models (FA, PC, FHLR) to the benchmark confidence indicators in a cross-country and cross-sector study, we indeed find clearly more significant improvements than in the aggregate euro area case. With very few exceptions, the model-based approaches using the expanded data sets outperform the confidence indicators, on grand average by four percentage points. These improvements should be seen against the background of a benchmark average correlation level of the confidence indicators with their respective target series of around 60%. Improvements are most pronounced in the case of the consumer and services confidence indicators with correlations increasing by around seven percentage points on average across countries.

Therefore, if the focus is solely on performance aspects of the derived indicators (i.e. correlation, and controlling for mean timing properties), it can be concluded that the employed factor model approaches, in combination with the use of expanded input data sets, can generate some notable improvements over the existing sectoral confidence indicators.

Distinguishing between the performances of the different factor models was less straightforward, in the cases of both the euro area and individual countries. However, the following conclusions can be drawn: first, a both theoretical and practical problem with the classic ML-factor estimator is that it occasionally led to defective Heywood cases, rendering its possible regular application for the production of composite indicators rather hazardous. Secondly, a potentially important differentiation is the treatment of data dynamics. As not all input series of a given sector move perfectly in sync, taking into account their dynamic structure is expected to sharpen the composite's signal. However, the results set out above show that the methodological generalisations to dynamic modelling approaches (SW, FHLR) do not yield any practically relevant improvements.

The FHLR approach, developed for very large input data sets, is more likely to reveal its methodical strengths in an environment where many more series with heterogeneous dynamics are available to the model than are in the present BCS case. The SW model appeared to sometimes produce rather peculiar results. Moreover, and relatedly, its estimation via the Kalman filter is comparatively cumbersome and time-consuming. Based on the results derived so far, it therefore seems questionable whether the application of the methodologically more advanced dynamic factor approaches can be justified in an attempt to generate composite indicators from BCS balance data.

In view of the above considerations, the straightforward but apparently effective principal components method would appear preferable for the regular production of model-based BCS composite indicators of the type presented in this paper.

In gauging the relative merits over the existing confidence indicators, aspects other than performance differences must of course also be taken into account (e.g. the problem of continuous subsequent revisions of published indicators in the case of all model-based approaches, practicability aspects linked to possible estimation problems). With this in mind, the improvements of tracking behaviour derived so far, based on balance series referring to total sectors and restricting the number of extracted factors to one, might appear limited. This applies mainly to the euro area but partly also to individual country results.

Of course, one principal advantage of the indicators derived from factor analysis is that the existence of an underlying model makes it possible to prove that this indicator really is the one which, statistically, best summarises the information available from the surveys. To potentially further increase the performance margin over the traditional confidence indicators, several generalisations of the approach followed so far can be envisaged. These generalisations, applied to the four different factor models, could include in particular

- the use of the percentage series of positive, unchanged or negative answers instead of the aggregated balance series
- the breakdown to branch-specific survey results rather than using the aggregate sector-specific results
- the extraction of more than just one factor from the data (where appropriate) and subsequent projections to arrive at one-dimensional indicators.

The first two generalisations should provide a richer set of cyclical information and improve the data basis for a fruitful application of the factor models, and here especially of the more advanced dynamic approaches. The third generalisation may become indispensable especially when the input data set gets larger and more diverse.

Moreover, as already mentioned above in the context of the consumer survey results regarding a possible trade-off between correlation and leading properties, the performance of the factor models may be further improved by a systematic pre-screening of the input data sets. Two possible ways of doing so could be

- shifting the input series according to their individual lead/lag structure with respect to the reference series,
- imposing general correlation thresholds with respect to the reference series and excluding those survey series with low correlation.

While the first approach would be expected to sharpen the composite indicator's cyclical signals, the second should enforce a closer mechanical link between the survey-based composite indicator and the particular statistical reference series it is supposed to refer to. Both approaches would appear particularly promising when applied to the broader set of disaggregate (percentage or branch-level) input series mentioned above.

References

Altissimo, F., Bassanetti, A., Cristadoro, R., Forni, M., Lippi, M., Reichlin L. and Veronese, G., (2001), "EuroCoin: A real time coincident indicator of the Euro area business cycle", CEPR Working Paper 3108

Bengoechea, Pilar and Martinez, Carlos (2002): A Proposal for an Integral Exploitation of Industry Surveys, Paper presented at the EU Workshop on Business and Consumer Surveys, Brussels, November 2002.

Brillinger, David R. (1981), Time series data analysis and theory, Holt, Rinehart and Winston, New York.

Bry, Gerhard and Boschan, Charlotte (1971): Cyclical Analysis of Time Series: Selected Procedures and Computer Programs, NBER Technical Paper 20, New York

Burns, A. F. and W. C. Mitchell (1946), Measuring business cycles, NBER Studies in Business Cycles no. 2, New York.

Deroose, Servaas, Philippe Mills and Bertrand Saint Aubin (2001), Business Climate Indicator for the Euro Area, Paper prepared for CEPR / Banca d'Italia Conference on "Monitoring the euro area business cycle", Rome, 6-7 September 2001.

Doz, C. and Lenglart, F. (1999): Analyse factorielle dynamique: test du nombre de facteurs, estimation et application à l'enquête de conjoncture dans l'industrie, Annales d'Economie et Statistique, n°54.

European Commission (2000): Business Climate Indicator for the Euro area (Presentation Paper), November 2000, available at:
http://europa.eu.int/comm/economy_finance/indicators/business_climate/2001/presentation_climate.pdf

European Commission (1997): The joint harmonised EU programme of business and consumer surveys, European Economy, No 6.

Forni, M., Hallin, M., Lippi, M. and L. Reichlin (2000), The generalized factor model: identification and estimation, The Review of Economics and Statistics 82(4): 540-554.

Forni, M., Hallin, M., Lippi, M. and L. Reichlin (2001): Coincident and Leading Indicators for the Euro Area, The Economic Journal, Vol. 111, No. 471.

Forni, M., Hallin, M., Lippi, M. and L. Reichlin (2002): The generalized dynamic factor model: one-sided estimation and forecasting, CEPR Discussion Paper No. 3432, June.

Franses, P. H., Paap, R. and Fok, D. (2005): Performance of Seasonal Adjustment Procedures: Simulation and Empirical Results, Final Report (May), available at: http://europa.eu.int/comm/economy_finance/indicators/business_consumer_surveys/studies_en.htm.

Grenouilleau, Daniel (2004): A sorted leading indicators dynamic (SLID) factor model for short-run euro-area GDP forecasting, European Economy, Economic Papers No. 219, December 2004.

Lawley D. N. and Maxwell, A. E. (1971): Factor Analysis as a Statistical Method, Macmillan, New York.

OECD (2005): Handbook on Constructing Composite Indicators: Methodology and User Guide, OECD Statistics Working Paper, August 2005

Schneeweiss, Hans (1997): Factors and Principal Components in the Near Spherical Case, Multivariate Behavioural Research, 32 (4), 375-401.

Stock, J. H. and M. W. Watson (1989), New indexes of coincident and leading economic indicators, in: Blanchard, O., and S. Fischer, eds., NBER Macroeconomics Annual, MIT Press (Cambridge, MA): 351-394.

Stock, J. H. and Watson, M. W. (1991), A probability model of the coincident indicators", in Lahiri, K., and G. H. Moore, eds., Leading Economic Indicators: New approaches and forecasting records, Cambridge University Press (Cambridge, UK), 63-89.

Stock, J. H. and Watson, M. W. (2002a), Macroeconomic Forecasting Using Diffusion Indexes, Journal of Business and Economic Statistics, 20, 147-62.

Stock, J. H. and Watson, M. W. (2002b), Forecasting Using Principal Components from a Large Number of Predictors, Journal of the American Statistical Association, 97, 1167-1179.

Annex A

TABLE A1: Survey questions

INDU	
Q1	Production trends over past 3 months
Q2	Order books
Q3	Export order books
Q4	Current stock of finished products
Q5	Production expectations over next 12 months
CONS	
Q1	Financial situation of household over last 12 months
Q2	Financial position expectations of household over next 12 months
Q3	General economic situation over past 12 months
Q4	General economic expectations over next 12 months
Q7	Unemployment expectations
Q8	Moment to make major purchases?
Q9	Expenditure expectations on major purchases over next 12 months
Q10	Moment to save?
Q11	Save expectations over next 12 months
SERV	
Q1	Business Situation development over past 3 months
Q2	Evolution of demand (turnover) over past 3 months
Q3	Expected demand (turnover) change over next 12 months
Q4	Evolution of employment over past 3 months
Q5	Expected employment change over next 3 months
RETA	
Q1	Business activity (sales) development over past 3 months
Q2	Volume of stocks currently hold
Q3	Expected orders change over next 3 months
Q4	Expected business activity (sales) change over next 3 months
Q5	Expected employment change over next 3 months
BUIL	
Q1	Development of activity over past 3 months
Q3	Evolution of order books
Q4	Expected employment change over next 3 months
Q5	Expected charged price changes over next 3 months

Bold: questions included in confidence indicators

TABLE A2: Input series correlations

INDU	Q1	Q2	Q3	Q4	Q5
Q1	1.00				
Q2	0.91	**1.00**			
Q3	0.88	0.94	1.00		
Q4	-0.91	**-0.90**	-0.85	**1.00**	
Q5	0.90	**0.87**	0.88	**-0.90**	**1.00**

CONS	Q1	Q2	Q3	Q4	Q7	Q8	Q9	Q10	Q11
Q1	1.00								
Q2	0.85	**1.00**							
Q3	0.93	0.93	1.00						
Q4	0.68	**0.88**	0.86	**1.00**					
Q7	-0.69	**-0.91**	-0.86	**-0.92**	**1.00**				
Q8	0.89	0.82	0.90	0.68	-0.71	1.00			
Q9	0.92	0.77	0.85	0.59	-0.59	0.83	1.00		
Q10	0.19	0.21	0.13	0.12	-0.20	0.05	0.13	1.00	
Q11	0.66	**0.78**	0.63	**0.54**	**-0.67**	0.59	0.55	0.54	**1.00**

SERV	Q1	Q2	Q3	Q4	Q5
Q1	**1.00**				
Q2	**0.86**	**1.00**			
Q3	**0.92**	**0.81**	**1.00**		
Q4	0.91	0.87	0.88	1.00	
Q5	0.85	0.78	0.83	0.85	1.00

RETA	Q1	Q2	Q3	Q4	Q5
Q1	**1.00**				
Q2	**-0.43**	**1.00**			
Q3	0.81	-0.40	1.00		
Q4	**0.66**	**-0.15**	0.68	**1.00**	
Q5	0.65	-0.35	0.66	0.62	1.00

BUIL	Q1	Q3	Q4	Q5
Q1	1.00			
Q3	0.77	**1.00**		
Q4	0.84	**0.92**	**1.00**	
Q5	0.77	0.80	0.92	1.00

TABLE A3: Industry sector turning points

	T	P	T	P	T	P	T	P	T	P	missed	mean lag (-)	stdev
IP	93M02	94M12	96M04	97M10	99M02	00M05	01M12	02M11	03M06	04M05			
FA3	-2	0	-1	-5	-1	0	1				3	-1.1	2.0
FA5	-2	-1	-2	-5	-2	0	1	-1	-1	-2		-1.5	1.6
FHLR3	-2	0	-1	-5	-1	0	1	1	-1		1	-0.9	1.8
FHLR5	-2	-1	-2	-5	-2	0	1	-1	-1	-2		-1.5	1.6
PC3	-2	0	-1	-5	-1	0	1	1	-1		1	-0.9	1.8
PC5	-2	-1	-2	-5	-2	0	1	-1	-1	-2		-1.5	1.6
SW3	-4	0	-1	-5	-1	0	1				3	-1.4	2.2
SW5	-4	-1	-2	-5	-3	-1	1			-2	2	-2.1	1.9
ICI	-4	0	-1	-5	-1	0	1	1	-1		1	-1.1	2.1

TABLE A4: Consumer sector turning points

	T	P	T	P	T	P	T	missed TP	add. TP	mean lag (-)	stdev
Consum.	93M02	95M05	97M02	00M05	02M05	03M02	03M11				
FA4	-5	0	7	0			8	2		2.0	5.4
FA9	-5	6	5	0			8	2		2.8	5.3
FHLR4	-5	-1	7	0			8	2		1.8	5.5
FHLR9	-5	7	7	0			8	2		3.4	5.7
PC4	-5	-1	7	0			8	2		1.8	5.5
PC9	-5	7	7	0			8	2		3.4	5.7
SW4	-1	4	7	0	6	9	8		2	4.7	3.9
SW9	-5	7	7	0			8	2		3.4	5.7
CCI	-5	0	7	1			8	2		2.2	5.4

TABLE A5: Services sector turning points

	P	T	P	T	P	missed	mean	stdev
GVA Serv	98M02	98M11	00M05	03M05	04M05	TP	lag (-)	
FA3	-6	-9	-1	2		1	-3.5	4.9
FA5	-4	-9	1	2		1	-2.5	5.1
FHLR3	-4	-9	1	5	3		-0.8	5.7
FHLR5	-4	-9	1	2		1	-2.5	5.1
PC3	-6	-8	1	2		1	-2.8	5.0
PC5	-4	-8	1	2		1	-2.3	4.6
SW3	-6	-9	-4	2	3		-2.8	5.2
SW5	-4	-10	-1	2		1	-3.3	5.1
SCI	-6	-8	1	2		1	-2.8	5.0

TABLE A6: Retail trade sector turning points

	T	P	T	P	T	P	T	missed	add.	mean	stdev
Consum.	93M02	95M05	97M02	00M05	02M05	03M02	03M11	TP	TP	lag (-)	
FA3	-22	-11	1	-1	3	-14		1	2	-7.3	9.9
FA5	-1	7	1	-1	-4	-18		1	4	-2.7	8.4
FHLR3	2	7	1	-1	0			2	2	1.8	3.1
FHLR5	-1	-10	1	-1	-4			2	2	-3.0	4.3
PC3	2	7	4	-1	2			2	2	2.8	2.9
PC5	2	-10	1	-1	0			2	2	-1.6	4.8
SW3	-22			-1	-2	-18		3	2	-10.8	10.8
SW5	-23	-13	1	-1	-4	-18		1	2	-9.7	9.8
RCI	-1	7	1	-1	0			2		1.2	3.3

TABLE A7: Building sector turning points

		T	P	T	P	T	P	T	P	missed	mean	stdev
Prod.	92M02	93M03	94M11	96M02	99M11	01M03	02M01	03M01	04M01	TP	lag (-)	
FA2		-7	1	-2	-9	-17				4	-6.8	6.9
FA4		0	1	-2	-7			-8		4	-3.2	4.1
FHLR2		-7	1	-2	-9	-17				4	-6.8	6.9
FHLR4			1	-2	-4	-20				5	-6.3	9.4
PC2		-7	1	-2	-9	-17				4	-6.8	6.9
PC4			1	-2	-4	-17				5	-5.5	7.9
SW2		0	0	-2	-9	-20				4	-6.2	8.6
SW4			1	-2	-5			-8	-7	4	-4.2	3.7
BCI		-7	1	-2	-9	-17				4	-6.8	6.9

Annex B

TABLE B1: Communalities due to static FA (based on consumer survey questions)

	Q1	Q2	Q3	Q4	Q7	Q8	Q9	Q10	Q11
BE	0.79	0.68	0.98	0.75	0.87	0.52	0.12	0.02	0.09
DE	0.84	0.77	0.99	0.80	0.62	0.64	0.63	0.05	0.78
EL	0.79	0.91	0.58	0.61	0.33	0.28	0.04	0.24	0.54
ES	0.93	0.81	0.98	0.67	0.72	0.87	0.77	0.00	0.29
FR	0.81	0.69	0.99	0.65	0.72	0.85	0.03	0.05	0.28
IE	0.92	0.98	0.84	0.53	0.76	0.86	0.27	0.19	0.38
NL	0.70	0.79	0.94	0.54	0.88	0.84	0.44	0.00	0.30
PT	0.88	0.94	0.96	0.89	0.59	0.82	0.40	0.82	0.73
Average	**0.83**	**0.82**	**0.91**	**0.68**	**0.69**	**0.71**	**0.34**	**0.17**	**0.42**

TABLE B2: Turning points of alternative consumption related indicators: Germany

DE	T	P	T	P	T	P	T	P	missed	add.	mean	stdev
Confidence	93M10	95M03	96M06	99M02	99M10	00M05	03M05	04M02	TP	TP	lag (-)	
FA9	-3	0	0	0	-1	0	5		1		0.1	2.4
FHLR9	0	0	0	0	-1	0	5		1		0.6	2.0
PC9	0	5	0	0	1	0	5		1		1.6	2.4

TABLE B3: Turning points of alternative consumption related indicators: Spain

ES	P	T	P	T	P	T	P	T	P	missed	add.	mean	stdev
Confidence	89M12	90M09	91M05	92M10	94M12	95M07	00M03	03M03	04M04	TP	TP	lag (-)	
FA9	0			-10	0	0	0	0	0	2		-1.4	3.8
FHLR9	0			-10	0	0	0	0	0	2		-1.4	3.8
PC9	0			-10	0	0	0	0	0	2		-1.4	3.8

TABLE B4: Turning points of alternative consumption related indicators: Finland

FI	P	T	P	T	P	T	missed	add.	mean	stdev
Confidence	97M10	98M11	00M02	01M11	02M04	03M06	TP	TP	lag (-)	
FA9*			0	-1	-5		3		-2.0	2.6
FHLR9	0	-3	0	0	0	0			-0.5	1.2
PC9		0	0	0	0	0	1		0.0	0.0

*Heywood case

TABLE B5: Turning points of alternative consumption related indicators: France

FR	P	T	P	T	P	T	P	T	P	T	missed	add.	mean	stdev
Confidence	89M08	91M12	92M10	93M08	95M06	96M09	01M01	01M09	02M05	03M03	TP	TP	lag (-)	
FA9	-7	-1	0	0	-2	0	0			-9	2	1	-2.4	3.6
FHLR9	-7	0	0	0	-2	0	0			-9	2		-2.3	3.7
PC9	0	0	0	0	0	5	0			-9	2		-0.5	3.9

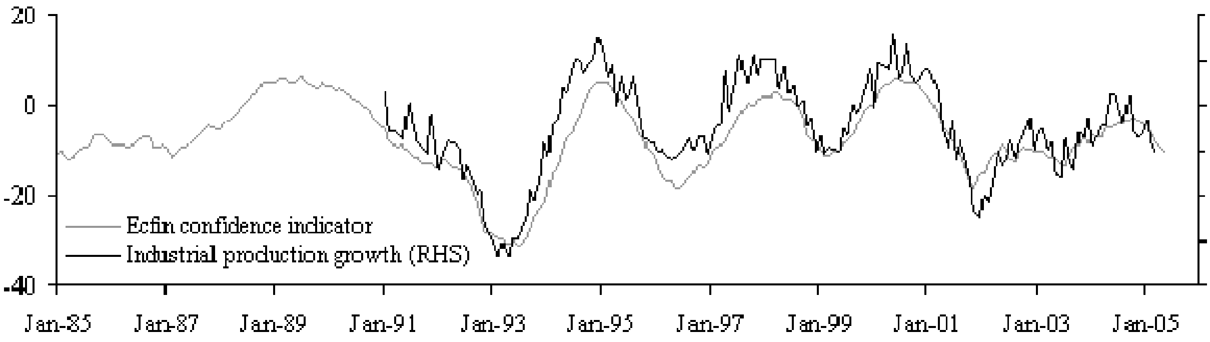
TABLE B6: Turning points of alternative consumption related indicators: Italy

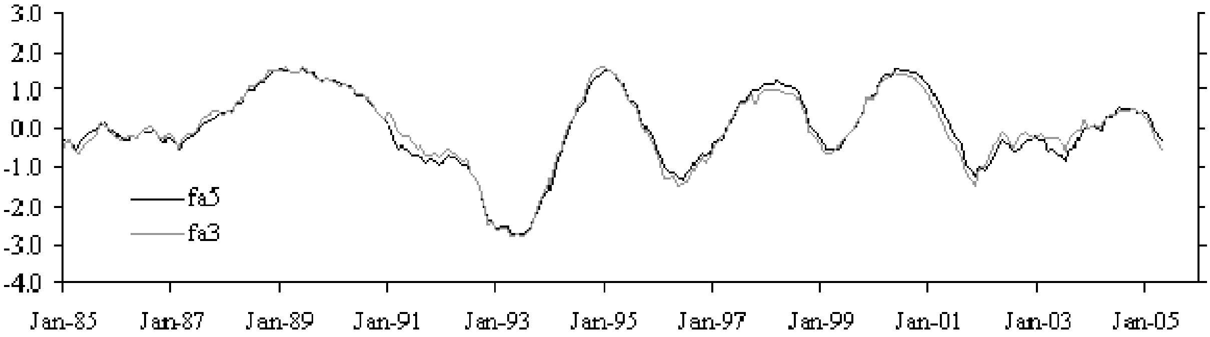
IT	P	T	P	T	P	T	P	missed	add.	mean	stdev
Confidence	90M04	93M04	95M02	97M06	98M11	99M05	01M06	TP	TP	lag (-)	
FA9*	-1	-9					-8	4	2	-6.0	4.4
FHLR9	-1	0	0	7	-2	-4	-4		1	-0.6	3.7
PC9	-1	0	0	1	-2	-4	1		1	-0.7	1.8

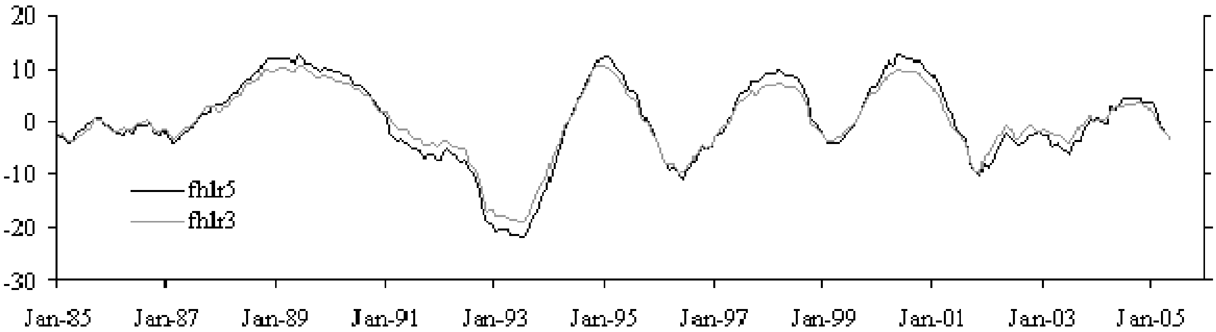
*Heywood case

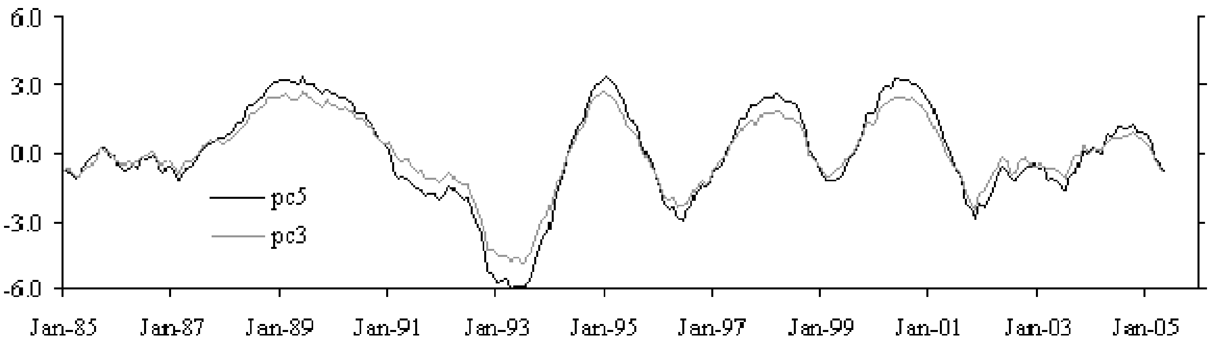
Annex C: Reference series, confidence indicators and factor based indicators for the euro area

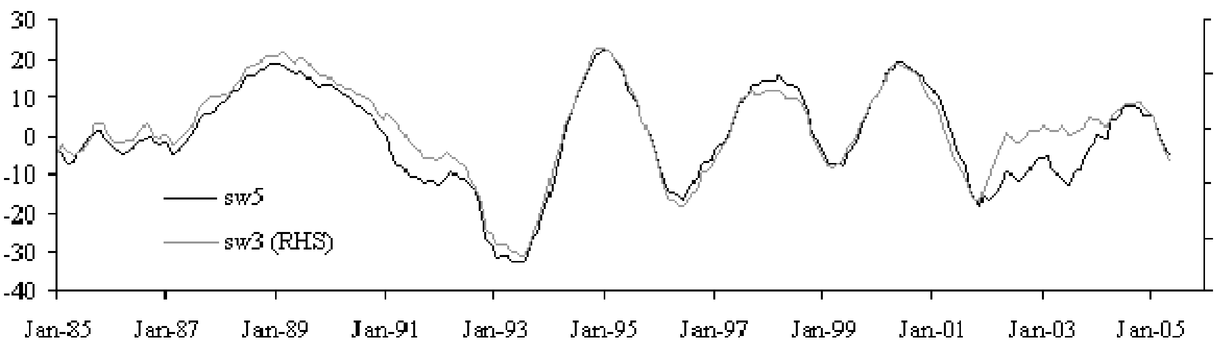
GRAPH C1: Industry



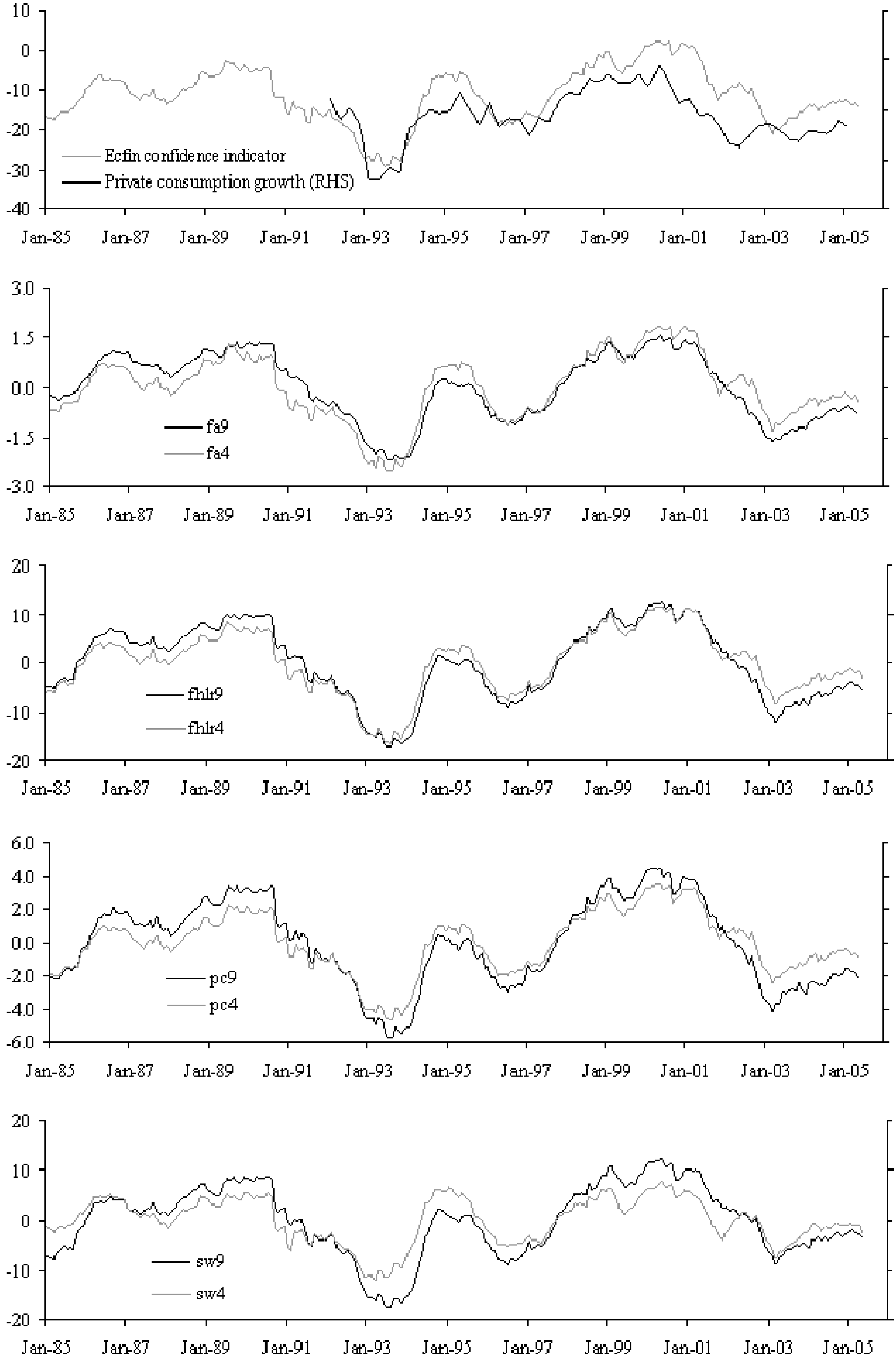




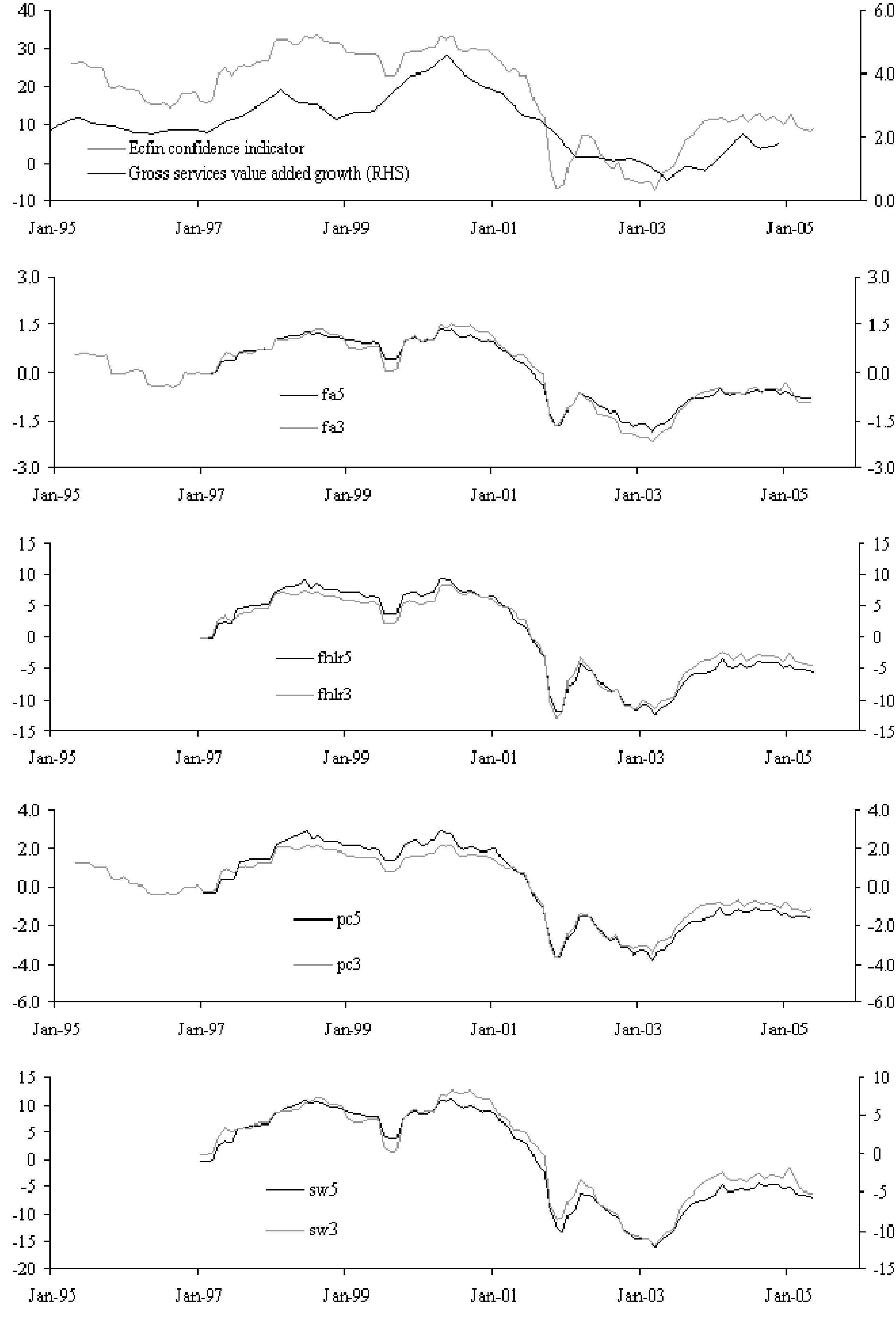




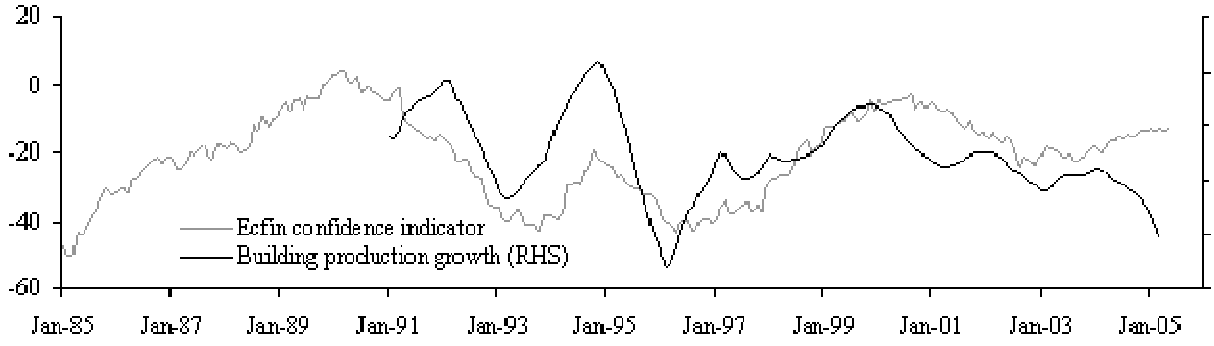
GRAPH C2: Consumption

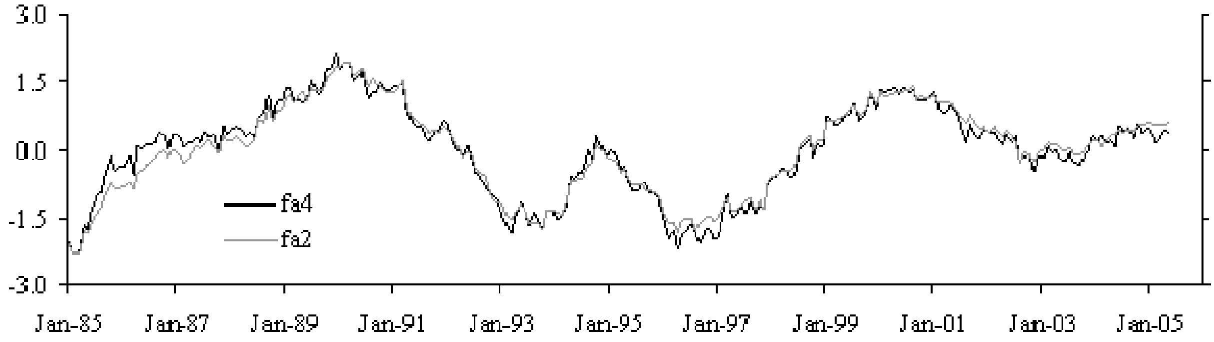


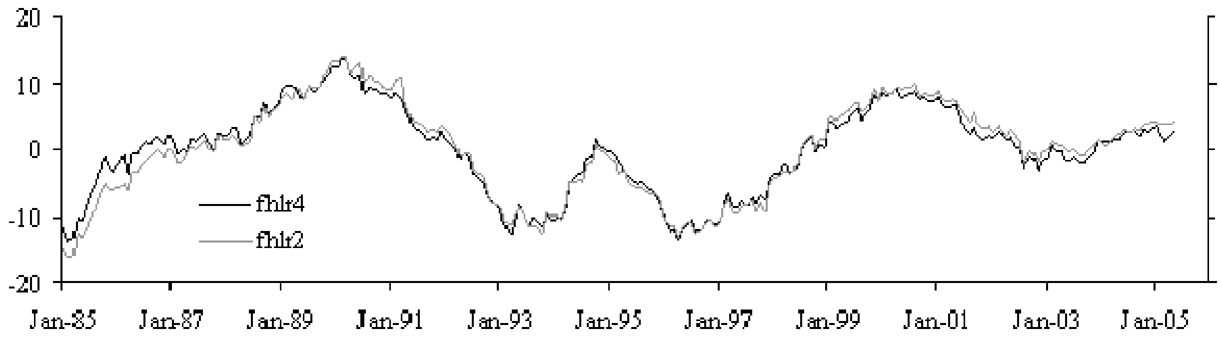
GRAPH C3: Services

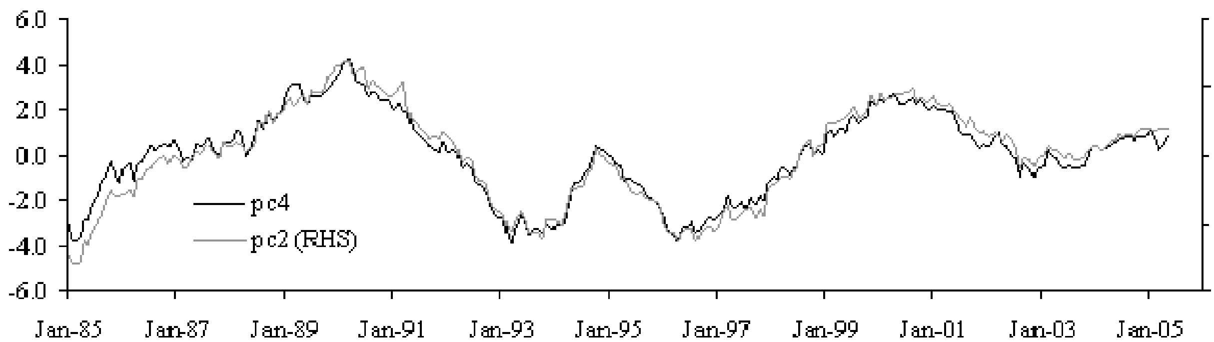


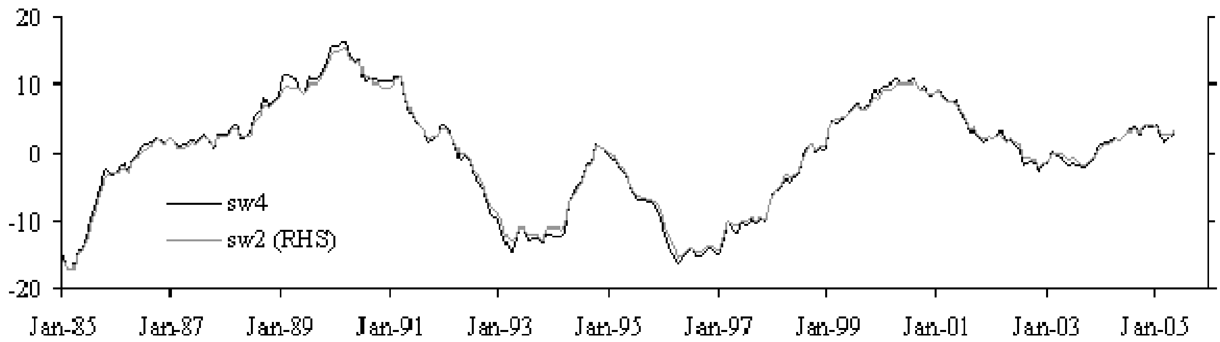
GRAPH C4: Building











GRAPH C5: Retail