Forecasting Business and Consumer Surveys Indicators. A Time Series Models Competition

Moreno, Rosina; Clar, Miquel; Duque, Juan Carlos

Forecasting Business and Consumer Surveys Indicators. A Time Series Models Competition

Journal:	*Applied Economics*
Manuscript ID:	APE-05-0424.R1
Journal Selection:	Applied Economics
JEL Code:	C53 - Forecasting and Other Model Applications <
Keywords:	forecast competition, indicators, business surveys, monitoring forecasts

Forecasting Business and Consumer Surveys Indicators. A Time Series Models Competition

Miquel Clar
Juan-Carlos Duque
Rosina Moreno
Grup d'Anàlisi Quantitativa Regional (AQR)
Espai de Recerca en Economia
Facultat de Ciències Econòmiques i Empresarials, Universitat de Barcelona

Abstract
The objective of this paper is to compare different time series methods for the short-run forecasting of Business and Consumer Survey Indicators. We consider all available data taken from the Business and Consumer Survey Indicators for the Euro area between 1985 and 2002. The main results of the forecast competition are offered not only for raw data but we also consider the effects of seasonality and removing outliers on forecast accuracy. In most cases the univariate autoregressions were not outperformed by the other methods. As for the effect of seasonal adjustment methods and the use of data from which outliers have been removed, we obtain that the use of raw data has little effect on forecast accuracy. The forecasting performance of qualitative indicators is important since enlarging the observed time series of these indicators with forecast intervals may help in interpreting and assessing the implications of the current situation and can be used as an input in quantitative forecast models.

Keywords: Comparative methods, Evaluating forecast, Forecasting competition, Indicators, Monitoring forecasts, Business Surveys.

Address for correspondence
Grup d'Anàlisi Quantitativa Regional (AQR)
Espai de Recerca en Economia
Facultat de Ciències Econòmiques i Empresarials, Universitat de Barcelona
Avda. Diagonal 690 - 08034 Barcelona, Espanya
Tel: +34 93 4021012
Fax: +34 93 4021821
E-mail: rmoreno@ub.edu

1. Introduction

The amount of information provided by Business and Consumer Surveys concerning agents' perceptions and expectations of their environment means that they are now recognised as a crucial instrument for gathering economic information in today's ever-changing environment. Data obtained from Business and Consumer Surveys are often used in forecasting models and in testing different expectation schemes – Kauppi et al (1996), Batchelor and Dua (1998), Mourougane and Roma (2002) and Nardo (2003)-. The speed with which the results of these surveys are made available and the wide range of variables included make them extremely useful for decision-making -see Stuart (1985) for a deep discussion on the value of Business and Consumer Surveys-. The remarkable growth in business surveys in Europe since the early 1960s, and the need for them to be carried out and presented in a comparable way, led to the implementation of the Joint Harmonised EU Programme by the Commission in 1961.

The present paper tries to compare different time series methods for the short-run forecasting of Business and Consumer Survey Indicators. Certainly, forecast competitions have been considered in the economic literature although focused in quantitative variables such as industrial production (Byers and Peel, 1995; Simpson et al, 2001), output growth and employment (Clements and Smith, 2000) and exchange rates (Clements and Smith, 2001; Boreo and Marrocu, 2002), as well as for general macroeconomic time series (Stock and Watson, 1999). On the other hand, as far as we know, forecast competitions have not been conducted for the case of qualitative variables. However, we consider that this kind of exercise can be useful to analyse which forecasting method presents the best behaviour. The usefulness of this comparison is twofold. First, it will allow having the best qualitative forecast to evaluate whether a series will move up or down in all kind of processes which can be modified according to changing economic conditions (McIntosh and Dorfman, 1992) or to predict business cycle turning points (Diebold and Rudebusch, 1989). And second, it will guarantee that the best forecast could be used as an explanatory variable in quantitative forecast models (Biart and Praet, 1987; Parigi and Schlitzer, 1995).

The objective of this paper is therefore to compare different time series methods for the short-run forecasting of Business and Consumer Survey Indicators. This objective can be summarised in the following two questions: Is it possible to forecast qualitative indicators? And, if so, which is the best procedure for conducting such forecasts? In order to answer these questions, we considered all available data taken from the Business and Consumer Survey Indicators for the Euro area between, in the main, January 1985 and December 2002. The dataset analysed includes 38 indicators (33 of which are monthly and 5 quarterly) and 6 composite indicators.[1]

First, we considered raw data in order to test the forecast accuracy of five different sets of models: autoregressions, ARIMA, Self-exciting threshold autoregressions (SETAR), Markov switching regime models and vector autoregressions (traditional VAR and also VAR models considering the joint evolution of different indicators). Some of the conclusions obtained are related to the high volatility presented by indicators of this type, so that their statistical properties need to be taken into account when concluding.

An additional aspect considered in this comparison of forecasts is related to the balance forecasts. As survey data are derived from qualitative questions and based on subjective evaluations, the results are usually presented in terms of balances, which show the difference between the positive and negative percentages of answers. Since it is this balance that analysts take into account and the information that is usually forecast, we examine whether it is better to forecast the balance directly or rather to forecast the negative and positive answers first and, then, calculate the balance. Additionally, we examine the composite indicators, which are calculated from business surveys in which answers are weighted to various questions in the survey. When looking at these aggregate indicators, we analyse whether it is preferable to forecast them directly or to obtain the forecast by weighting the forecasts for the different components.

Finally, we consider the effects of seasonal adjustment procedures and the removal of outliers on forecast accuracy. In the case of seasonal adjustment procedures, we apply different methods to the indicators so as to obtain seasonally adjusted data (using Tramo/Seats - TS, X12 and Dainties - DA) and trend cycle estimation (using TS, X12 and Wavelets - WAV) in order to evaluate differences. In the case of outliers, we consider how their presence might affect the results by using TS to remove them (additive, transitory changes and level shifts) from the original series, and to what extent the effects of outliers should be removed each time a new observation is available.

The outline of the paper is as follows. The next section presents the models that are considered in the forecast competition. The database and the design of the forecast competition experiment is given in section 3. Section 4 offers the main results of the forecast competition not only for raw data but we also consider the effects of seasonality and removing outliers on forecast accuracy. Finally, section 5 concludes.

2. Models for forecasting Business and Consumer Survey Indicators

In order to assess alternative methods and models for forecasting Business and Consumer Surveys Indicators, we chose to focus on five different sets of model: autoregressions (AR), ARIMA, Self-exciting threshold autoregressions (SETAR), Markov switching regime models (MK) and vector autoregressions (traditional VAR and also VAR models considering the joint evolution of different indicators).

Autoregressions

The widely known autoregressive model (also known as the distributed-lags model) explains the behaviour of the endogenous variable as a linear combination of its own past values:

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + ... + \phi_p Y_{t-p} + \varepsilon_t .$$

The key question is how to determine the number of lags that should be included in the model. For monthly -quarterly- data we considered different models with a minimum number of 1 lag up to a maximum of 24 -8- (including all the intermediate lags), selecting the model with the lowest Akaike Information Criteria (AIC) value.

ARIMA models

Since the study conducted by Box and Jenkins (1970), ARIMA models have been widely used and their forecast performance has also been confirmed. The general expression of an ARIMA model is the following:

$$x_t^{\lambda} = \frac{\Theta_s(L^s)\theta(L)}{\Phi_s(L^s)\phi(L)\Delta_s^D\Delta^d}\varepsilon_t,$$

$\Theta_s(L^s) = (1 - \Theta_s L^s - \Theta_{2s}L^{2s} - \ldots - \Theta_{Qs}L^{Qs})$ is a seasonal moving average polynomial, $\Phi_s(L^s) = (1 - \Phi_s L^s - \Phi_{2s}L^{2s} - \ldots - \Phi_{Ps}L^{Ps})$ is a seasonal autoregressive polynomial, $\theta(L) = (1 - \theta_1 L^1 - \theta_2 L^2 - \ldots - \theta_q L^q)$ is a regular moving average polynomial, $\phi(L) = (1 - \phi_1 L^1 - \phi_2 L^2 - \ldots - \phi_p L^p)$ is a regular autoregressive polynomial, λ is the value of the Box-Cox (1964) transformation, Δ_s^D is the seasonal difference operator, Δ^d is the regular difference operator, *S* is the periodicity of the time series under consideration, and ε_t is the innovation assumed to behave as a white noise.

In order to use models of this kind for forecasting, the proper model has to be identified (i.e., giving values to the order of the different polynomials, to the difference operator, etc.). For monthly data, we considered models with up to 12 AR and MA terms (4 in the case of quarterly data) selecting the model with the lowest AIC value. The statistical goodness of the selected model was also checked.

TAR models

In the case of the ARIMA model the relationship between the current value of a variable and its lags is supposed to be linear and constant over time. However, when looking at real data it can be seen that expansions tend to be more prolonged over time than recessions (Hansen, 1997). In fact, in the behaviour of most economic variables there seems to be a cyclical asymmetry that linear models are unable to capture (Clements and Smith, 1999). A Self-Excited Threshold Autoregressive model (SETAR) for the time series X_t can be summarised as follows:

$$\begin{array}{ll} B(L)\cdot X_t + u_t & \text{if } X_{t-k} \leq X\text{, and} \\ \zeta(L)\cdot X_t + v_t & \text{if } X_{t-k} > X\text{,} \end{array}$$

where u_t and v_t are white noises, $B(L)$ and $\zeta(L)$ are autoregressive polynomials, the value k is known as delay and the value X is known as threshold. This two-regime self-exciting threshold autoregressive process is estimated using monthly and quarterly data for each indicator and the Monte Carlo procedure is used to generate multi-step forecasts. The delay values selected are those that minimise the sum of squared errors among values between 1 and 12 for monthly data and 1 and 4 for quarterly data. The values of the threshold are given by the variation of the variable being analysed.

Markov switching regime models

Threshold autoregressive models are perhaps the simplest generalisations of linear autoregressions. In fact, these models were built on developments of traditional ARMA time series models. As an alternative to these, time series regime-switching models assume that the distribution of the variable is known conditional on a particular regime or state occurring. When the economy changes from one regime to another, a substantial change occurs in the series. Hamilton (1989) presented the Markov regime-switching model in which the unobserved regime evolves over time as a 1st-order Markov process. The regime completely governs the dynamic behaviour of the series. This implies that once we assume conditions on a particular regime occurring, and assume a particular parameterisation of the model, we can write down the density of the variable of interest. However, as the regime is strictly unobservable, statistical inferences have to be drawn concerning the likelihood of each regime occurring at any point in time. So, we need to obtain the transition probabilities from one regime to the other.

Three approaches have been adopted in estimating these models (Potter, 1999). First, Hamilton (1989) developed a non-linear filter to evaluate the likelihood function of the model and, then, he directly maximised the likelihood function. Second, in a later article, Hamilton (1990) constructed an EM algorithm which proves particularly useful should all the parameters switch. Finally, Albert and Chib (1993) developed a Bayesian approach to estimation.

Here, we employ a Markov-switching threshold autoregressive model (MK-TAR) in which we are able to allow for different regime-dependent intercepts, autoregressive parameters, and variances. The estimation of the models is carried out by maximum likelihood using the Hamilton (1989) filter[2] together with Kim's smoothing filter (1994).

Having estimated the probabilities of expansion and recession, we could then construct the following model for the time series X_t:

$$
\begin{array}{ll}
B(L)\cdot X_t + u_t & \text{if } P[Expansion/X_{t-k}] \leq P\text{, and} \\
\zeta(L)\cdot X_t + v_t & \text{if } P[Recession/X_{t-k}] > P\text{,}
\end{array}
$$

where, as in the SETAR models, u_t and v_t are white noises, $B(L)$ and $\zeta(L)$ are autoregressive polynomials, the value k is the estimated delay and the value P is the estimated threshold[3]. The selected delay values are those that minimise the sum of squared errors for values between 1 and 12 for monthly data and 1 and 4 for quarterly data. The values of the threshold are given by the variation of the probability.

VAR models

In these models, each variable depends on a certain number of lags of the other variables under analysis (Sims, 1982). The idea is that the positive, neutral and negative answers to each question can be considered jointly. Moreover, as the sum of the percentages of positive (P), neutral (E) and negative (M) answers would, by definition, total one hundred, this restriction could also be introduced in the model improving its forecasting accuracy:

$$\begin{bmatrix} P_t \\ E_t \\ M_t \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix} \cdot \begin{bmatrix} P_{t-1} \\ E_{t-1} \\ M_{t-1} \end{bmatrix} + \ldots + \begin{bmatrix} Z_{11} & Z_{12} & Z_{13} \\ Z_{21} & Z_{22} & Z_{23} \\ Z_{31} & Z_{32} & Z_{33} \end{bmatrix} \cdot \begin{bmatrix} P_{t-p} \\ E_{t-p} \\ M_{t-p} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \end{bmatrix}.$$

In order to use models of this kind for forecasting, the proper model must first be identified (i.e., give values to the number of lags p). For monthly data, we considered models with up to 24 lags (8 in the case of quarterly data), selecting the model with the lowest AIC value. The statistical goodness of the selected model was also checked.

3. Database and design of the forecast competition experiment

The immediate relevance of the results of harmonised business and consumer surveys is one of their main properties, given that they are published shortly after the termination of the month to which they refer. The survey results are presented either in the form of balances for particular questions or as synthetic indicators. They are able to describe the panorama of the current economic situation much quicker than the quantitative indicators and the macroeconomic magnitudes from national accounts, though the latter tend to be more precise in their descriptions.

The European Commission draws up and publishes a wide range of indicators calculated on the basis of the results sent in by more than 40 institutes in 25 countries in the framework of the Joint Harmonised EU Programme of Business and Consumer Surveys. Particular attention is paid to indicators for the Euro area. The EU Programme currently includes surveys for industry, construction, the retail trade, services, investment and consumers.

For our analysis, we considered all the information available for the Business and Consumer Survey Indicators in the Euro area. The dataset analysed includes 38 indicators (33 of which are monthly and 5 quarterly) and 6 composite indicators. The starting date of these indicators differs but most of them begin in January 1985 (or in the first quarter of 1985). The latest period to be included in the analysis is December 2002 (or the final quarter of 2002). More details on the dataset can be found in Table 1.

Since the objective of the paper was to assess alternative methods and models for forecasting Business and Consumer Survey Indicators, we initially considered raw data (in all cases, non seasonally adjusted levels of each category of the variables were used) in order to test the forecast accuracy of the five different sets of models presented in the previous section.

In order to evaluate the relative forecasting accuracy of the models, each model was estimated for all the indicators included up to 2000.12 (or 2000.IV for quarterly indicators) and forecasts for 1, 2, 3, 6 and 12 months (or 1, 2, 4 quarters) in the future were computed. The model specifications are based on information up to 2000.12 or 2000.IV and, thereafter, the models were re-estimated in each month or quarter and the forecasts were computed with these estimation results. Given the availability of actual values up to 2002.12 or 2002.III or 2002.IV, we were able to compute the forecast errors for each indicator and method in a recursive way (i.e., for the 1 month forecast horizon, 24 forecast errors were computed for each indicator). In order to summarise this information, the Root Mean Square Error (RMSE) and Mean

Absolute Percentage Error (MAPE) were computed. These values provide useful information for analysing the forecast accuracy of each method, and enabled us to rank the methods according to their values.

Before showing the results of this comparative exercise, it should be stressed that the statistical properties of the Business and Consumer Survey Indicators differ substantially from those of the main macroeconomic variables (GDP, CPI, Industrial Production, Industrial Producer Prices, etc.). In Table 2, the variation coefficient is shown. It is worth noting that the variation coefficient values are extremely high for some indicators. This can be interpreted as evidence of the high volatility of the indicators and, with this in mind, the forecast accuracy of the methods considered can be expected to be lower than that for other macroeconomic variables.

A further result of interest is that if we examine the average values of the variation coefficient for categories (i.e., positive answers, negative answers, balance, etc.), the highest value corresponds to the balance. A possible explanation for the higher variance of the balance is purely statistical. The variance of the balance can be decomposed as follows:

$$Var(b) = Var(p-m) = Var(p) + Var(m) - 2 \cdot Cov(p,m) .$$

So, if the covariance between positive and negative answers is negative (i.e. they move in opposite directions), the variance of the balance would be higher than the variance of the other two components. In 87% of cases, the sign of this covariance is negative, so the variance of the balance is higher (in most cases) than the other two components (Table 3a). To test whether it was significantly higher, we applied a test of equality of variances between the balance and the other two components. In 76% of the cases, the null hypothesis of equality of variances was rejected (Table 3b).

Additionally, we computed some of the most commonly used methods to test the unit root hypothesis: the augmented Dickey and Fuller test (1979), the Phillips and Perron test (1988) and the Kwiatkowski, Phillips, Schmidt and Yongcheol test (1992) (Table 4a). The Perron test (1989) was also adapted so as to allow for different types of structural change: in the level (crash model) or in the slope (breaking trend) (Table 4b). Since the variables under consideration can only take values between 0 and 100, a priori we would expect most of them to be I(0). The most striking result is that in many cases the variables are eventually considered to be I(1) - 85% if the structural break is not considered and 52% when it is [4].

4. Results: forecast accuracy comparison

4.1. Main results for raw data

In Table 5 we present a summary of the recursive forecasts of the main indicators (Economic Sentiment -v1-, Industrial Confidence -v2-, Consumer Confidence -v12-, Construction Confidence -v28-, Retail Trade Confidence -v33- and Services Confidence -v39- Indicators) for the different models. The table refers to raw data. For each indicator, the best model is highlighted.

Before analysing the results, it should be pointed out that composite indicators are usually calculated from business surveys. For example, the Economic Sentiment Indicator is obtained by weighting the answers to different questions on the survey. When dealing with these aggregates, there are two possible courses of action: forecasting them directly or obtaining the forecast by weighting the forecasts for the different components. To the best of our knowledge, no research has been undertaken in order to determine which method provides the best results. Therefore, we chose to forecast composite indicators in two ways: what we call a *direct* (computation) method and by using the forecasts of the components, an *indirect* (computation) method.

Among the conclusions that can be extracted from Table 5, we would like to point out the following. First, indirect, as opposed to direct, methods seem to perform best. Thus, it would appear to be better to forecast these indicators from the forecasts of the components than directly. Second, among the direct methods, the AR model outperformed the rest of the models in almost all cases. By contrast, the ARIMA and TAR models were never found to be the best forecasters. Third, among the indirect methods, the AR and VAR models were the ones that provided the lowest RMSE (with the exception of the Economic Sentiment Indicator, where the ARIMA model outperformed the rest of the models). Once again, the TAR model presented the highest values for the RMSE.

As for the results of the forecast comparison for the remaining indicators, details figures are shown in Table 6, in which the average RMSE is shown for each type of answer and where the best model is highlighted in each case. The main conclusions of this analysis can be summarised as follows. As far as the forecast accuracy of the different methods is concerned, in most cases the univariate autoregressions are not outperformed by the other methods. In fact, only the forecast errors from the VAR and Markov models were lower than those obtained from AR models in some cases. It should also be stressed that unrestricted VAR models usually work better than restricted VAR models and that the errors displayed by the ARIMA, TAR and VAR models for different indicators were generally higher.

Furthermore, taking into account that the variables considered (positive, neutral and negative answers but not the balance) can only take values between 0 and 100, the forecast errors are quite high even in the case of the best model. As expected, the forecasts errors increased for longer horizons in most cases and, in general, variables corresponding to questions with a higher number of possible answers were better forecast than the rest. In most cases, the size of the errors was higher for the balance than for the components (this is related to the higher volatility of these variables), which is a common result for most composite indicators (one notable exception is the Economic Sentiment Indicator).

A further aspect to be considered is related to that of balance forecasts. As survey data are derived from qualitative questions and based on subjective evaluation, the results are usually presented in terms of balances, which show the difference between positive and negative percentages of answers. The balance is the information that analysts take into account and the information that is usually forecast. But, is it better to forecast the balance directly or to forecast negative and positive answers and then calculate the balance? In order to answer this question we replicated the forecasting comparison described above, but on this occasion we computed the balance from the forecasts of negative and positive answers using the AR, ARIMA, TAR and VAR models. The results (in the last block of Table 6, referred to as 'b') show that it is usually better to forecast the balance from the forecasts of positive and negative answers rather than by doing it directly, and that the AR model outperforms the rest of the models in almost all cases.

4.2. The effects of seasonality on data revision and on forecast accuracy

Seasonal adjustment methods are usually applied to these indicators. However, an interesting point that has not been analysed to date is the extent to which the method chosen for seasonal adjustment - TS, X12, DA, WAV, among others - might affect the values of the series under consideration. In order to remedy this situation, we applied all these methods of seasonal adjustment so as to obtain seasonally adjusted data (using TS, X12 and DA) and trend cycle estimations (using TS, X12 and WAV) for a number of qualitative indicators [5]. From the results shown in Table 7, several conclusions can be drawn:

First, TS and X12 results were very similar but DA results (seasonal adjustment) and WAV results (trend cycle-estimation) differed greatly. Specifically, differences were found to be greater in trend-cycle estimations than in seasonal adjustments (Table 7a and 7b).

Second, in order to analyse how the seasonal adjustment method chosen affected the revision of the series, seasonally adjusted data and trend-cycle estimations were computed in a recursive way adding one more observation from 2001.12 to 2002.12. The results (see Table 7c and 7d) showed that:

- There was no relationship between the size of the revision in a given observation and the number of time periods between this observation and the last observation available.
- There were great differences in the size of revision for the different variables in each seasonal adjustment method considered.
- There were no revisions of seasonally adjusted data using the DA method.
- TS revisions of seasonally adjusted data were greater than X12 revisions.
- Trend-cycle estimations using WAV method showed dramatic revisions.
- TS revisions of trend-cycle data were greater than X12 revisions.

Third, the choice of seasonal adjustment method was found to affect forecasts. Thus, we were interested in determining whether it was better to forecast raw data and then apply a seasonal adjustment method or not. As pointed out by Commission staff, the information from Business and Consumer Surveys has little seasonality, and, therefore the results should not vary significantly. To confirm this, we computed the Kruskal-Wallis test (also in Table 7e) for all the qualitative variables in order to verify the importance of seasonality. In almost 87% of cases, the null hypothesis of non-seasonality was not rejected, that is, most series did not present a seasonal component. This, it would seem, despite the large number of studies discussing this matter, that too much attention has been given to this issue in the context of Business and Consumer Surveys.

4.3. The effects on forecast accuracy of removing outliers using TS

The presence of outliers may well affect the results discussed in the previous sections. Therefore, TS can be used in order to remove the outliers from the original series. TS uses an automatic procedure to detect and eliminate outliers from a series. Three different types of outliers are considered here: Additive outliers (AO), Transitory changes (TC) and Level shifts (LS).

In Table 8a the variation coefficient for the qualitative variables from the industrial survey and the balances for the other surveys are shown. From these statistics, it can be seen that for some series, TS did not detect any outliers (i.e., the Economic Sentiment Indicator). However, in most cases, the volatility was substantially lower, although in some the value of the variation coefficient increased (although the standard deviation was lower, the value of the mean –which is clearly affected by the presence of extreme values- was even lower). As before, we now computed the value of the covariance between positive and negative answers and we tested to see if the variance of the balance differed from that of positive or negative answers. As observed in Table 8b, in 95% of the cases (as opposed to 87% when seasonality was not taken into account), the covariance was negative, so that the null hypothesis of equality of variances between the balance and the other two components (positive and negative answers) was rejected in 82% of the cases (compared with 76%) -see Table 8c-. Finally, the results on the unit root tests did not change at all when the outliers were removed from the variables, so that the main results remained unchanged.

Using these data, we repeated the comparison of forecasts as in the previous sections. The main results are shown in Table 8d. The main conclusions to be drawn from this analysis can be summarised as follows: First, RMSE values were the same or lower in nearly all variables when using data from which the outliers had been removed using TS. However, there were differences between models: while in the AR, ARIMA, TAR and Markov models the values of RMSE were markedly lower than those for raw data, the results for VAR (unrestricted and restricted) were, some times, worse. Finally, the VAR models consisting of different indicators performed much better when the outliers were removed.

Second, when comparing the RMSE for the balance computed from forecasts with positive and negative answers, we observed that the results for the models analysed (AR, ARIMA, TAR and VAR) - once the outliers had been removed - were, in most cases, higher than those for raw data. Similarly, when the composite indicators were computed from forecasts based on their components, here again the models performed worse once the outliers had been removed.

A practical issue that needs to be borne in mind when forecasting these series in real time concerns the effects of incorporating new observations on outlier detection using TS. In other words, is it necessary to remove the effects of outliers each time a new observation is available? To analyse this, we adopted a similar approach to that regarding the effects of seasonal adjustment procedures. In order to evaluate the effects of incorporating new observations, we identified the time periods during which an outlier is found in a recursive way by adding one more observation from 2001.1 to 2002.12. The results of this analysis are shown in Table 9. From these results, we can conclude that the number and type of outliers clearly depend on the available sample. Although there was some 'persistence' in the moment and type of detected outliers, some significant changes in the results might also be recorded using TS once an additional observation is included (sometimes due to changes in the underlying model, but also to the dynamics of the series itself). So, if we take into account the results regarding the forecast accuracy of the various models, and although this has the effect of increasing the computational cost, it would appear to be necessary to apply the TS to remove the effects of outliers each time an additional observation becomes available.

In addition, we should stress that when comparing the various seasonal adjustment methods, the TS and X12 results were very similar, but the DA (seasonal adjustment) and WAV results (trend cycle-estimation) presented marked differences. In fact, these differences were greater for the trend-cycle estimations than the seasonal adjustments, whereas there were considerable differences in the size of revision for the different variables for each seasonal adjustment method considered.

Additionally, after removing the outliers using TS the forecast accuracy of the various methods was similar to that observed with raw data. The number and type of outliers identified with TS clearly depends on the available sample. Although there is some 'persistence' in the moment and type of detected outliers, significant changes can be reflected in the results.

5. Conclusions

The objective of this paper is to compare time series methods for the short-run forecasting of Business and Consumer Survey Indicators. We analysed a dataset for the Euro area, which included 38 indicators (33 of which were monthly and 5 quarterly) and 6 composite indicators, mostly between January 1985 and December 2002. In order to test the forecast accuracy, we used five different sets of models: autoregressions, ARIMA, Self-exciting threshold autoregressions, Markov switching regime models and vector autoregressions (traditional VAR and also VAR models considering the joint evolution of different indicators).

As far as the forecast accuracy of the different methods is concerned, in most cases the univariate autoregressions were not outperformed by the other methods. In fact, only the forecast errors from the VAR and Markov models were, in some cases, lower than those from the AR models. It should also be stressed that unrestricted VAR models usually worked better than restricted VAR models and that the errors displayed by the ARIMA, TAR and VAR models comprising a range of indicators were generally higher. However, the size of the forecast errors was high even in the case of the best model.

Interestingly, both in the case of composite indicators and indicators which belong to a balance category, our results show that it is preferable to forecast them indirectly. As for the effect of the seasonal adjustment methods that are typically applied to these indicators, we used a range of methods (TS, X12, Dainties and Wavelets) that suggest that the use of raw data has little effect on forecast accuracy. A similar conclusion is obtained when using data from which outliers have been removed using TS, as here again forecast errors were found to be similar for nearly all variables.

References

Albert, J. & Chib, S. (1993). Bayesian inference via Gibbs sampling of autoregressive time series subject to Markov mean and variance shifts, *Journal of Business and Economic Statistics*, 11, 1-15.

Batchelor, R.A. & Dua, P. (1998). Improving macro-economic forecasts. The role of consumer confidence, *International Journal of Forecasting*, 14, 71-81.

Biart, M. & Praet, P. (1987). The contribution of opinion surveys in forecasting aggregate demand in the four main EC countries, *Journal of Economic Psychology*, 8, 409-428.

Boero, G. & Marrocu, E. (2002). The performance of non-linear exchange rate models: A forecasting comparison, *Journal of Forecasting*, 21, 513-542.

Box, G. & Cox, D. (1964). An analysis of transformation, *Journal of the Royal Statistical Society Series B*, 26, 2.

Box, G.E.P. & Jenkins, G.M. (1970). *Time series analysis: Forecasting and Control*, Holden-Day, San Francisco.

Byers, J.D. & Peel, D.A. (1995). Forecasting industrial production using non-linear methods, *Journal of Forecasting*, 14, 325-336.

Clements, M.P. & Smith, J. (1999). A Monte Carlo study of the forecasting performance of empirical SETAR models, *Journal of Applied Econometrics*, 14, 123-141.

Clements, M.P. & Smith, J. (2000). Evaluating the forecast densities of linear and non-linear models: Application to output growth and unemployment, *Journal Forecasting*, 19, 255-276.

Clements, M.P. & Smith, J. (2001). Evaluating forecasts from SETAR models of exchange rates, *Journal of International Money and Finance*, 20, 133-148.

Dickey, D.A. & Fuller, W.A. (1979). Distribution of the estimators for autoregressive time series with a unit root, *Journal of American Statistical Association*, 74, 427-431.

Diebold, F.X. & Rudebusch, G.D. (1989). Scoring the leading indicators, *Journal of Business*, 62, 369-391.

Hamilton, J. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle, *Econometrica*, 57, 357-384.

Hamilton, J. (1990). Analysis of time series subject to changes in regime, *Journal of Econometrics*, 45, 39-70.

Hansen, B. (1997). *Inference in threshold models*, mimeo.

Kauppi, E., Lassila, J. & Teräsvirta, T. (1996). Short-term forecasting of industrial production with business survey data: experience from Finland's great depression 1990-1993, *International Journal of Forecasting*, 12, 373-381.

Kim, C.J. (1994). Dynamic linear models with Markov-Switching, *Journal of Econometrics,* 60, 1-22.

Kwiatkowski, D., Phillips, P.C.B., Schmidt, P. & Yongcheol, S. (1992). Testing the null hypothesis of stationary against the alternative of a unit root: How sure are we that economic time series have a unit root?, *Journal of Econometrics*, 54, 159-178.

McIntosh, C.S. & Dorfman, J.H. (1992). Qualitative forecast evaluation: A comparison of two performance measures, *American Journal of Agricultural Economics*, 74, 209-214.

Mourougane, A. & Roma, M. (2002). Can confidence indicators be useful to predict short term real GDP growth?, Working Paper No. 133, *European Central Bank*.

Nardo, M. (2003). The quantification of qualitative survey data: A critical assessment, *Journal of Economic Surveys*, 17, 645-668.

Parigi, G. & Schlitzer, G. (1995). Quarterly forecasts of the Italian business-cycle by means of monthly economic indicators, *Journal of Forecasting*, 14, 117-141.

Perron, P. (1989). The great crash, the oil price shock and the unit root hypothesis, *Econometrica*, 57, 1361-1401.

Phillips, P.C.B. & Perron, P. (1988). Testing for a unit root in time series regression, *Biometrika*, 75, 335-346.

Potter, S. (1999). *Nonlinear time series modelling: An introduction*, Federal Reserve Bank of New York, mimeo.

Simpson, P.W., Osborn, D.R. & Sensier, M. (2001). Forecasting UK industrial production over the business cycle, *Journal of Forecasting*, 20, 405-424.

Sims, C. (1982). Policy analysis with econometric models, *Brookings Papers on Economic Activity*, 1, 107-164.

Stock, J.H. & Watson, M.W. (1999). A comparison of linear and nonlinear univariate models for forecasting macroeconomic time series. In: R.F. Engle & H. White (Eds.), *Cointegration, Causality, and Forecasting. A Festschrift in Honour of Clive W.J. Granger* (pp. 1-44). Oxford: Oxford University Press.

Stuart, O.D.J. (1985). Are Business and Consumer Surveys Still of Value?, *Journal for Studies in Economics and Econometrics*, 0, iss. 23, 29-50.

Notes:

1. More details on the dataset can be found in Table 1.
2. The Hamilton filter is an iterative procedure which provides estimates of the probability that a given state is prevailing at each point in time given its previous history. These estimates are dependent upon the parameter values given to the filter. Running the filter through the entire sample provides a log likelihood value for the particular set of estimates used. This filter is then repeated to optimise the log likelihood of obtaining the MLE estimates of the parameters. With the maximum likelihood parameters, the probability of state 0 at each point in time is calculated and these are the probabilities of recession and expansion.
3. An alternative approach would have consisted in imposing the value of P and k instead of estimating them. These models are known as Markov Switching Autoregressive Models (MS-AR) and, in general, the values of P are 0.7 or 0.8 and the values of k, 0 or 1.
4. It is also interesting to note that all the tests only coincide in 10 series.
5. The seasonality analysis was conducted for variables v1, v2, v3p, v3e, v3m, v3b, v4p, v4e, v4m and v4b.

We acknowledge financial support from ECFIN/2002/A3-01. We would like to thank the European Commission for their helpful comments and support. The usual disclaimer applies.

Table 1. List of Business and Consumer Surveys Indicators for the Euro area

	Description	Freq.		Last obs	Obs.	Categories
v1	**Economic Sentiment Indicator**	month	jan-85	dec-02	216	
v2	**Industrial Confidence Indicator (v7+v4-v6)/3**	month	jan-85	dec-02	216	
v3	Production trend observed in recent months	month	jan-85	dec-02	216	P E M B
v4	Assessment of order-book levels	month	jan-85	dec-02	216	P E M B
v5	Assessment of export order-book levels	month	jan-85	dec-02	216	P E M B
v6	Assessment of stocks of finished products	month	jan-85	dec-02	216	P E M B
v7	Production expectations for the months ahead	month	jan-85	dec-02	216	P E M B
v8	Selling price expectations for the months ahead	month	jan-85	dec-02	216	P E M B
v9	Employment expectations for the months ahead	month	jan-85	dec-02	216	P E M B
v10	*New orders in recent months*	*quarter*	*1985-I*	*2002-IV*	72	P E M B
v11	*Export expectations for the months ahead*	*quarter*	*1985-I*	*2002-IV*	72	P E M B
v12	**Consumer Confidence Indicator (v14+v16-v19+v23)/4**	month	jan-85	dec-02	216	
v13	Financial situation over last 12 months	month	jan-85	dec-02	216	PP P E M MM N B
v14	Financial situation over next 12 months	month	jan-85	dec-02	216	PP P E M MM N B
v15	General economic situation over last 12 months	month	jan-85	dec-02	216	PP P E M MM N B
v16	General economic situation over next 12 months	month	jan-85	dec-02	216	PP P E M MM N B
v17	Price trends over last 12 months	month	jan-85	dec-02	216	PP P E M MM N B
v18	Price trends over next 12 months	month	jan-85	dec-02	216	PP P E M MM N B
v19	Unemployment expectations over next 12 months	month	jan-85	dec-02	216	PP P E M MM N B
v20	Major purchases at present	month	jan-85	dec-02	216	PP E MM N B
v21	Major purchases over next 12 months	month	jan-85	dec-02	216	PP P E M MM N B
v22	Savings at present	month	jan-85	dec-02	216	PP P M MM N B
v23	Savings over next 12 months	month	jan-85	dec-02	216	PP P M MM N B
v24	Statement on financial situation of household	month	jan-85	dec-02	216	PP P E M MM N B
v25	*Intention to buy a car within the next 2 years*	*quarter*	*1990-I*	*2002-IV*	52	PP P M MM N B
v26	*Purchase or build a home within the next 2 years*	*quarter*	*1990-I*	*2002-IV*	52	PP P M MM N B
v27	*Home improvements over the next 12 months*	*quarter*	*1990-I*	*2002-IV*	52	PP P M MM N B

(Continues next page)

	Description	Freq.	First obs	Last obs	Obs	Categories			
v28	**Construction Confidence Indicator (v30+v31)/2**	month	jan-85	dec-02	216				
v29	Trend of activitiy compared with preceding months	month	jan-85	dec-02	216	P	E	M	B
v30	Assessment of order books	month	jan-85	dec-02	216	P	E	M	B
v31	Employment expectations for the months ahead	month	jan-85	dec-02	216	P	E	M	B
v32	Price expectations for the months ahead	month	jan-85	dec-02	216	P	E	M	B
v33	**Retail Trade Confidence Indicator (v34-v35+v37)/3**	month	jan-86	dec-02	204				
v34	Present business situation	month	jan-85	dec-02	216	P	E	M	B
v35	Assessment of stocks	month	jan-85	dec-02	216	P	E	M	B
v36	Orders placed with suppliers	month	feb-85	dec-02	215	P	E	M	B
v37	Expected business situation	month	jan-86	dec-02	204	P	E	M	B
v38	Employment	month	abr-85	dec-02	213	P	E	M	B
v39	**Services Confidence Indicator (v40+v41+v42)/3**	month	abr-95	dec-02	93				
v40	Assessment of business climate	month	abr-95	dec-02	93	P	E	M	B
v41	Evolution of demand in recent months	month	abr-95	dec-02	93	P	E	M	B
v42	Evolution of demand expected in the months ahead	month	abr-95	dec-02	93	P	E	M	B
v43	Evolution of employment in recent months	month	abr-95	dec-02	93	P	E	M	B
v44	Evolution of employment expected in the months ahead	month	jan-97	dec-02	72	P	E	M	B

The letters refer to positive answers (pp and p), neutral answers (e), negative answers (mm and m), non answers (n), balance (b) and composite indicators (i).

Table 2. Variation coefficient for the Business and Consumer Surveys Indicators

	Average variation coefficient	
	Monthly indicators	Quarterly indicators
pp	27.22	13.90
p	31.33	15.10
e	16.03	4.38
m	33.75	17.02
mm	33.76	2.55
n	19.52	36.44
b	348.47	118.68
i	67.26	

Table 3a. Analysis of the sign of the covariance between positive and negative answers

Covariance	Negative sign	Positive sign	TOTAL
Month	28	5	33
Quarter	5	0	5
TOTAL	33	5	38

Covariance	Negative sign	Positive sign	TOTAL
Month	74%	13%	87%
Quarter	13%	0%	13%
TOTAL	87%	13%	100%

Table 3b. Tests of equality of variance

H_0: Equality of variance
H_A: Non-equality of variance

Statistic	Rejection of the Null	Non-rejection of the Null	TOTAL
Month	27	6	33
Quarter	2	3	5
TOTAL	29	9	38

Statistic	Rejection of the Null	Non-rejection of the Null	TOTAL
Month	71%	16%	87%
Quarter	5%	8%	13%
TOTAL	76%	24%	100%

Table 4a. Results of the tests for unit root hypothesis. Unit root tests without structural break. (I(1) in 85% of the cases)

Order	Variable	ADF	PP(k=4)	PP(k=8)	KPSSa	KPSSb
I(0)	pp	3	1	4	9	6
	p	4	4	13	18	10
	e	1	6	4	16	5
	m	4	3	12	21	9
	mm			1	8	3
	n	1	5	1	1	4
	b	4	9	11	24	10
	v1				1	1
	v2	1		1	1	1
	v12				1	
	v28				1	
	v33					1
Total I(0)		18	28	47	101	50
I(1)	pp	9	11	8	3	6
	p	23	23	14	9	17
	e	25	20	22	10	21
	m	23	24	15	6	18
	mm	12	12	11	4	9
	n	11	7	11	11	8
	b	24	19	17	4	18
	v1	1	1	1		
	v2		1			
	v12	1	1	1		1
	v28	1	1	1		1
	v33	1	1	1	1	
Total I(1)		131	121	102	48	99
Total general		149	149	149	149	149

ADF: augmented Dickey and Fuller test (1979); PP: Phillips and Perron test (1988); KPSS: Kwiatkowski, Phillips, Schmidt and Yongcheol test (1992).

Table 4b. Results of the tests for unit root hypothesis. Perron's (1989) test with structural break. (I(1) in 52 % of the cases)

Order	Variable	Crash model (a)	Crash model (b)	Breaking trend (a)	Breaking trend (b)
I(0)	pp	4	4	5	5
	p	13	12	13	12
	e	17	17	17	17
	m	14	13	12	13
	mm	5	4	4	4
	n	9	8	9	9
	b	11	11	10	10
	v1				
	v2	1	1	1	1
	v12				
	v28				
	v33				
Total I(0)		74	70	71	71
I(1)	pp	8	8	7	7
	p	14	15	14	15
	e	9	9	9	9
	m	13	14	15	14
	mm	7	8	8	8
	n	3	4	3	3
	b	17	17	18	18
	v1	1	1	1	1
	v2				
	v12	1	1	1	1
	v28	1	1	1	1
	v33	1	1	1	1
Total I(1)		75	79	78	78
Total general		149	149	149	149

Table 5. Average RMSE - Recursive forecasts from January 2001 to December 2002. Raw data for main composite indicators

Economic Sentiment Indicator (v1)

		1 month	2 months	3 months	6 months	12 months
direct methods	AR	***0.45****	0.77	1.02	1.51	1.52
	ARIMA	1.92	3.21	4.36	7.27	6.25
	TAR	4.65	7.14	9.33	15.07	20.51
	MK-TAR	0.61	0.75	1.19	2.20	6.12
indirect methods	AR	2.73	2.83	2.92	3.08	1.70
	ARIMA	0.46	***0.46****	***0.42****	***0.41****	***0.46****
	TAR	0.71	0.84	1.06	1.80	3.03
	VAR	4.38	4.55	4.70	4.98	3.49

**** Best model***

Industrial Confidence Indicator (v2)

		1 month	2 months	3 months	6 months	12 months
direct methods	AR	2.03	3.97	5.43	7.99	4.68
	ARIMA	10.54	21.70	29.85	40.20	39.19
	TAR	21.88	31.29	39.02	61.76	95.72
	MK-TAR	3.68	4.31	7.33	10.93	87.78
	VAR	3.39	5.63	6.90	12.25	17.82
indirect methods	AR	2.04	3.54	4.68	6.22	3.69
	ARIMA	2.21	4.14	5.52	7.60	7.97
	TAR	4.77	6.78	8.45	14.65	22.69
	VAR	***0.08****	***2.15****	***3.29****	***3.28****	***2.28****

**** Best model***

Consumer Confidence Indicator (v12)

		1 month	2 months	3 months	6 months	12 months
direct methods	AR	***1.68****	***2.71****	***3.65****	5.49	3.77
	ARIMA	9.13	13.57	18.02	30.84	25.56
	TAR	14.31	18.75	23.70	39.32	47.02
	MK-TAR	2.68	4.42	6.77	10.38	90.78
	VAR	3.02	5.24	7.67	13.62	24.60
indirect methods	AR	1.78	2.79	3.69	5.64	***3.27****
	ARIMA	5.42	4.54	3.93	***3.00****	3.67
	TAR	6.09	6.54	7.14	9.22	11.99
	VAR	9.47	11.66	13.42	15.31	12.92

**** Best model***

Construction Confidence Indicator (v28)

		1 month	2 months	3 months	6 months	12 months
direct methods	AR	2.01	2.36	2.38	2.97	***2.59****
	ARIMA	15.82	26.90	36.08	50.42	44.52
	TAR	26.32	38.53	50.42	73.45	100.06
	VAR	5.00	8.01	11.38	22.41	38.31
indirect methods	AR	1.97	***2.09****	***2.15****	***2.82****	2.88
	ARIMA	12.62	14.07	15.31	18.56	22.76
	TAR	13.06	14.62	16.96	25.21	32.69
	VAR	***0.07****	4.93	9.38	15.92	11.37

**** Best model***

Retail Trade Confidence Indicator (v33)

		1 month	2 months	3 months	6 months	12 months
direct methods	AR	2.59	3.08	3.32	4.68	8.04
	ARIMA	12.05	14.40	15.64	21.17	31.50
	TAR	17.73	24.04	26.92	39.67	59.06
	MK-TAR	2.56	3.63	***2.41****	4.15	***4.00****
	VAR	3.70	4.70	4.83	***4.00****	11.80
indirect methods	AR	***2.47****	3.17	3.50	5.13	8.82
	ARIMA	23.79	23.40	23.01	20.79	10.25
	TAR	29.27	32.48	34.80	40.68	40.94
	VAR	2.96	***2.72****	4.05	5.01	7.46

***** Best model***

Services Confidence Indicator (v39)

		1 month	2 months	3 months	6 months	12 months
direct methods	AR	5.44	10.31	15.51	26.48	43.33
	ARIMA	22.17	38.48	53.52	76.93	74.57
	TAR	56.92	76.78	86.97	126.94	182.96
	VAR	6.20	9.33	12.04	17.95	32.69
indirect methods	AR	5.61	9.94	13.97	20.49	35.87
	ARIMA	4.43	7.53	10.46	16.32	20.39
	TAR	12.00	16.22	18.57	25.46	41.09
	VAR	***0.69****	***2.34****	***7.59****	***14.37****	***18.90****

***** Best model***

Table 6. Average RMSE - Recursive forecasts from January 2001 to December 2002. Raw data for qualitative indicators according to the type of answer

pp

	1 month	2 months	3 months	6 months	12 months	1 quarter	2 quarters	4 quarters
AR	**1.00***	**1.48***	**1.91***	**2.74***	**3.23***	**0.61***	**0.57***	**0.61***
ARIMA	4.15	5.60	6.71	8.79	9.53	1.79	1.85	2.17
TAR	10.23	24.04	56.91	836.09	193363.17	2.55	3.36	4.35
MK-TAR	1.53	2.01	3.06	5.21	10.94	na	na	na

**** Best model***

p

	1 month	2 months	3 months	6 months	12 months	1 quarter	2 quarters	4 quarters
AR	2.07	2.77	**3.40***	**4.41***	**5.48***	**1.87***	**2.01***	**2.29***
ARIMA	9.94	12.68	15.07	18.91	19.76	6.30	7.46	7.74
TAR	17.83	24.60	31.22	48.50	72.76	10.95	11.73	17.79
MK-TAR	**1.92***	3.38	4.27	10.00	42.15	3.69	3.76	5.08
VAR unr	2.00	**2.77***	3.48	5.01	6.25	2.48	2.72	2.91
VAR rest	2.41	3.41	4.31	6.06	6.91	2.46	2.79	2.90

**** Best model***

e

	1 month	2 months	3 months	6 months	12 months	1 quarter	2 quarters	4 quarters
AR	2.22	2.87	3.37	4.28	5.21	3.12	3.52	3.47
ARIMA	10.42	13.05	15.00	17.95	18.30	9.40	10.13	7.64
TAR	36.59	39.57	42.46	92.09	167.79	12.38	14.58	19.95
MK-TAR	2.08	**2.46***	4.01	5.78	9.27	na	na	na
VAR unr	**1.88***	2.53	**3.03***	**3.82***	**4.42***	**1.50***	**1.83***	**2.19***
VAR rest	2.30	3.00	3.54	4.64	5.47	1.68	1.91	2.26

**** Best model***

m

	1 month	2 months	3 months	6 months	12 months	1 quarter	2 quarters	4 quarters
AR	**1.92***	2.77	**3.46***	**4.75***	**5.41***	**2.15***	**3.13***	**3.43***
ARIMA	9.67	13.72	16.81	22.32	23.98	7.02	10.47	11.00
TAR	18.11	39.66	51.40	112.63	781.91	10.53	14.04	15.11
MK-TAR	2.21	**2.59***	3.65	5.15	12.09	3.74	4.34	6.92
VAR unr	2.03	3.00	3.74	5.21	6.19	2.79	3.58	4.24
VAR rest	2.23	3.15	3.91	5.46	6.39	2.53	3.22	3.61

**** Best model***

mm

	1 month	2 months	3 months	6 months	12 months	1 quarter	2 quarters	4 quarters
AR	**0.72***	**1.00***	**1.19***	**1.69***	**2.34***	**1.24***	**1.22***	**1.40***
ARIMA	3.66	4.83	5.64	7.89	10.01	3.60	3.50	3.70
TAR	5.17	6.37	7.71	13.06	23.38	4.53	3.78	4.82
MK-TAR	0.96	1.12	2.08	4.40	14.34	1.28	1.34	1.72

**** Best model***

n

	1 month	2 months	3 months	6 months	12 months	1 quarter	2 quarters	4 quarters
AR	0.45	0.58	0.67	0.90	1.19	**0.58***	**0.62***	**1.10***
ARIMA	2.12	2.41	2.58	3.10	3.47	1.45	1.06	1.72
TAR	4.91	6.92	9.70	25.89	526.08	2.21	3.32	5.99
MK-TAR	**0.28***	**0.29***	**0.42***	**0.46***	**0.73***	na	na	na

**** Best model***

b

		1 month	2 months	3 months	6 months	12 months	1 quarter	2 quarters	4 quarters
direct methods	AR	3.40	4.90	6.33	8.94	12.00	4.05	4.90	5.38
	ARIMA	16.36	23.15	28.27	37.98	41.40	13.84	18.37	19.15
	TAR	29.00	39.76	49.61	70.67	89.30	20.48	23.94	24.43
	MK-TAR	3.63	6.46	8.76	17.37	67.04	6.74	7.18	11.37
	VAR	3.89	5.85	7.94	12.74	16.51			
indirect methods	AR	**3.29***	4.80	6.14	**8.41***	**9.64***	**3.93***	**4.79***	**5.15***
	ARIMA	3.31	**4.74***	**5.97***	8.50	11.19	4.68	6.73	8.77
	TAR	5.94	11.02	15.78	96.28	20317.70	6.07	6.63	6.87
	VAR unr	4.98	6.09	7.07	9.22	10.83	5.12	5.70	6.34
	VAR rest	5.47	6.84	8.14	10.59	11.61	4.73	5.31	5.63

**** Best model***

Table 7. Evaluation of the different seasonal adjustment methods

a) Mean squared error (MSE) between results from seasonally adjusted data

Average MSE between methods. (v1 … v4b)

		TS	X12	DA
TS	Average	0.000	0.008	0.173
	S.D.	0.000	0.004	0.074
X12	Average		0.000	0.193
	S.D.		0.000	0.089
DA	Average			0.000
	S.D.			0.000

b) Mean squared error (MSE) between results from trend-cycle estimation

Average MSE between methods. (v1 … v4b)

		TS	X12	WAV
TS	Average	0.000	0.010	0.369
	S.D.	0.000	0.000	0.016
X12	Average		0.000	0.354
	S.D.		0.000	0.011
WAV	Average			0.000
	S.D.			0.000

c) Revisions of trend-cycle series

S.D of mean absolute deviation

	TS	X12	WAV
V1	0.000	0.000	0.105
V2	0.041	0.193	4.131
V3P	0.003	0.002	0.965
V3E	0.012	0.002	0.209
V3M	0.023	0.004	1.911
V3B	0.030	0.010	5.495
V4P	0.003	0.035	0.922
V4E	0.016	0.001	1.065
V4M	0.022	0.001	3.685
V4B	0.035	0.015	7.998
V3B*	0.033	0.010	5.536
V4B*	0.035	0.004	5.427

* Indirect method.

d) Revisions of seasonally-adjusted series

S.D. of mean absolute deviation

	TS	X12	DA
V1	0.00	0.00	0.00
V2	0.25	0.02	0.00
V3P	0.01	0.01	0.00
V3E	0.00	0.01	0.00
V3M	0.00	0.01	0.00
V3B	0.03	0.03	0.00
V4P	0.04	0.00	0.00
V4E	0.00	0.01	0.00
V4M	0.00	0.01	0.00
V4B	0.02	0.02	0.00
V3B*	0.02	0.02	0.00
V4B*	0.00	0.02	0.00

* Indirect method.

e) Forecasting raw-data or seasonally adjusted data? The Kruskal-Wallis test for detecting seasonality (summary of the results)

H_0: non-seasonality
H_A: seasonality

	Rejection of the Null	Non-	TOTAL
Month	22	148	170
Quarter	4	22	26
TOTAL	26	170	196

	Rejection of the Null	Non-	TOTAL
Month	11.22%	75.51%	86.73%
Quarter	2.04%	11.22%	13.27%
TOTAL	13.27%	86.73%	100.00%

Table 8. The effects of removing outliers using TS

a) Average variation coeffcient

	Monthly indicators
P	28.92
E	5.86
M	32.64
B	979.29

b) Analysis of the sign of the covariance between positive and negative answers

Covariance	Negative sign	Positive sign	TOTAL
Month	32	1	33
Quarter	4	1	5
TOTAL	36	2	38

Covariance	Negative sign	Positive sign	TOTAL
Month	84%	3%	87%
Quarter	11%	3%	13%
TOTAL	95%	5%	100%

c) Tests of equality of variance

H_0: Equality of variance
H_A: Non-equality of variance

Statistic	Rejection of the Null	Non-rejection of the Null	TOTAL
Month	29	4	33
Quarter	2	3	5
TOTAL	31	7	38

Statistic	Rejection of the Null	Non-rejection of the Null	TOTAL
Month	76%	11%	87%
Quarter	5%	8%	13%
TOTAL	82%	18%	100%

d) The effects on forecasting accuracy

		Economic Sentiment Indicator (v1)				
		1 month	2 months	3 months	6 months	12 months
direct methods	AR	**0.45***	0.77	1.02	1.51	1.52
	ARIMA	1.92	3.21	4.36	7.27	6.25
	TAR	4.51	6.72	8.73	14.25	20.04
	MK-TAR	0.61	0.75	1.20	2.20	6.15
indirect methods	AR	2.83	2.94	3.04	3.22	1.84
	ARIMA	0.47	**0.44***	**0.41***	**0.40***	**0.48***
	TAR	0.77	0.85	1.04	1.61	2.34
	VAR	4.56	4.73	4.89	5.19	3.67

**** Best model***

p

	1 month	2 months	3 months	6 months	12 months	1 quarter	2 quarters	4 quarters
AR	**1.50***	**1.95***	**2.50***	**3.19***	3.38	**3.25***	**3.88***	**4.68***
ARIMA	8.52	11.42	13.65	18.01	15.99	11.22	13.92	15.08
TAR	15.53	20.80	26.23	45.54	76.03	20.94	21.29	34.48
MK-TAR	2.02	2.09	2.56	3.34	4.66	6.10	5.85	10.38
VAR unr	1.78	2.49	3.28	4.46	4.65	5.19	6.50	5.87
VAR rest	1.76	2.36	3.06	4.20	**3.14***	4.06	5.01	4.98

**** Best model***

e

	1 month	2 months	3 months	6 months	12 months	1 quarter	2 quarters	4 quarters
AR	**1.70***	2.30	**2.75***	3.33	**3.74***	**2.51***	**2.95***	**2.76***
ARIMA	8.53	11.41	13.81	16.64	16.20	7.17	7.96	5.72
TAR	13.31	16.98	19.71	27.24	34.88	10.17	12.37	18.09
MK-TAR	1.70	**2.12***	2.96	**3.31***	5.03	na	na	na
VAR unr	1.89	2.54	3.02	3.67	4.23	3.42	3.73	3.08
VAR rest	4.76	4.83	4.92	5.00	4.40	2.88	3.38	3.40

**** Best model***

m

	1 month	2 months	3 months	6 months	12 months	1 quarter	2 quarters	4 quarters
AR	1.54	2.32	3.07	4.53	3.92	**4.30***	6.15	6.27
ARIMA	8.68	13.16	16.76	22.93	15.55	14.55	21.98	22.36
TAR	14.48	21.38	28.02	49.10	83.71	22.46	30.87	35.20
MK-TAR	**0.68***	**0.63***	**0.81***	**0.94***	**0.66***	6.73	7.65	12.68
VAR unr	1.70	2.71	3.61	5.40	6.08	6.80	9.48	8.49
VAR rest	1.56	2.27	2.95	4.45	3.60	4.34	**6.00***	**6.03***

**** Best model***

b

		1 month	2 months	3 months	6 months	12 months	1 quarter	2 quarters	4 quarters
direct methods	AR	**2.57***	**3.64***	**4.78***	**6.89***	**5.39***	7.17	9.18	**9.13***
	ARIMA	14.55	21.81	27.82	40.06	42.39	26.40	37.28	37.60
	TAR	26.94	39.04	49.91	88.28	138.05	43.16	53.28	53.19
	MK-TAR	6.54	7.46	8.83	10.71	12.90	11.16	11.49	19.76
	VAR	3.10	4.34	5.77	9.48	12.17			
indirect methods	AR	4.09	4.97	6.00	8.06	7.36	**6.69***	**8.98***	9.24
	ARIMA	4.08	5.28	6.33	9.14	9.39	8.30	13.12	16.61
	TAR	5.73	6.87	8.14	12.69	20.76	12.20	13.44	13.61
	VAR unr	4.18	5.43	6.91	9.75	11.14	11.56	15.68	14.16
	VAR rest	4.37	5.45	6.68	9.36	7.27	7.77	10.48	10.47

**** Best model***

Table 9.

The next tables should be interpreted as follows: the number and type of detected outliers when adding a new observation are shown in columns, while in rows, the moment of time for each outlier can be found.

v2	jan-01	feb-01	mar-01	apr-01	may-01	jun-01	jul-01	aug-01	set-01	oct-01	nov-01	des-01	jan-02	feb-02	mar-02	apr-02	may-02	jun-02	jul-02	aug-02	set-02	oct-02	nov-02	des-02
01-																			AO	AO				
10-																			AO	AO				
07-	AO	AO	AO	AO	AO	AO	AO		AO			AO												AO
06-	AO	AO	AO	AO	AO	AO	AO		AO			AO												AO
10-													LS	LS	LS	LS	LS	LS				LS	LS	
11-																		TC						
12-												AO												
01-																			LS	LS				

AO: Additive Outlier, TC: Transitory Change, LS: Level shift.

v3b[6]	jan-01	feb-01	mar-01	apr-01	may-01	jun-01	jul-01	aug-01	set-01	Oct-01	nov-01	des-01	jan-02	feb-02	mar-02	apr-02	may-02	jun-02	jul-02	aug-02	set-02	oct-02	nov-02	des-02
03-91							TC	TC	TC					TC	TC	TC	TC	TC	TC	TC	TC	TC	TC	TC
05-97														LS										

AO: Additive Outlier, TC: Transitory Change, LS: Level shift.

v3m	jan-01	feb-01	mar-01	apr-01	may-01	jun-01	jul-01	aug-01	set-01	oct-01	nov-01	des-01	jan-02	feb-02	mar-02	apr-02	may-02	jun-02	jul-02	aug-02	set-02	oct-02	nov-02	des-02
03-91																			TC		TC	TC	TC	TC
07-96					AO	AO	AO	AO	AO	AO	AO	AO	AO	AO	AO	AO	AO	AO	AO	AO	AO	AO	AO	AO
05-00																			AO		AO	AO	AO	AO
05-01																			LS		LS	LS	LS	LS

AO: Additive Outlier, TC: Transitory Change, LS: Level shift.

[6] No outlier is detected for v3p.

v3e	jan-01	feb-01	mar-01	apr-01	may-01	jun-01	jul-01	aug-01	set-01	oct-01	nov-01	des-01	jan-02	feb-02	mar-02	apr-02	may-02	jun-02	jul-02	aug-02	set-02	oct-02	nov-02	des-02
02-91															TC									
11-92	LS		LS					LS							LS									

AO: Additive Outlier, TC: Transitory Change, LS: Level shift.

v8b	jan-01	feb-01	mar-01	apr-01	may-01	jun-01	jul-01	aug-01	set-01	oct-01	nov-01	des-01	jan-02	feb-02	mar-02	apr-02	may-02	jun-02	jul-02	aug-02	set-02	oct-02	nov-02	des-02
07-85	LS																							
01-88	TC	TC	TC	TC	TC	TC	TC	TC	TC	TC	TC	TC	TC	TC	TC	TC	TC	TC	TC	TC	TC	TC	TC	TC
02-93								TC										TC						TC
01-97	AO	AO	AO	AO	AO	AO	AO	AO	AO	AO	AO	AO	AO	AO	AO	AO	AO	AO	AO	AO	AO	AO		AO
06-01																		AO						
07-01																								AO

AO: Additive Outlier, TC: Transitory Change, LS: Level shift.

v8m[7]	jan-01	feb-01	mar-01	apr-01	may-01	jun-01	jul-01	aug-01	set-01	oct-01	nov-01	des-01	jan-02	feb-02	mar-02	apr-02	may-02	jun-02	jul-02	aug-02	set-02	oct-02	nov-02	des-02
07-85		LS	LS	LS		LS			LS	LS	LS		LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS
03-86												TC	TC	TC	TC	TC	TC	TC	TC	TC	TC	TC	TC	TC
05-86																			AO					
09-87																			AO					
01-88	AO	AO	AO	AO	TC	AO		AO	AO	AO	AO	TC	TC	TC	TC	TC	TC	TC	TC	TC	TC	TC	TC	TC
11-88							TC						TC	TC	TC	TC	TC	TC	TC	TC	LS	TC	TC	TC
03-89																AO			TC	AO				
05-90																			TC					
01-93	AO											AO												
10-98	AO																							
03-99	AO	AO	AO	AO	AO	AO	AO	AO	AO	AO	AO	AO												
01-00	AO	AO			AO				AO			AO												
02-00			AO	AO		AO				AO	AO								AO					
07-01							AO	AO	AO	AO	AO	AO				AO			AO	AO	AO	AO	AO	AO

AO: Additive Outlier, TC: Transitory Change, LS: Level shift.

v8e	jan-01	feb-01	mar-01	apr-01	may-01	jun-01	jul-01	aug-01	set-01	oct-01	nov-01	des-01	jan-02	feb-02	mar-02	apr-02	may-02	jun-02	jul-02	aug-02	set-02	oct-02	nov-02	des-02
03-99	AO	AO	AO	AO	AO	AO	AO	AO	AO	AO	AO												AO	AO
12-01																							AO	AO

AO: Additive Outlier, TC: Transitory Change, LS: Level shift.

[7] No outlier is detected for v8p.

Forecasting Business and Consumer Surveys Indicators. A Time Series Models Competition

Miquel Clar
Juan-Carlos Duque
Rosina Moreno
Grup d'Anàlisi Quantitativa Regional (AQR)
Espai de Recerca en Economia
Facultat de Ciències Econòmiques i Empresarials, Universitat de Barcelona

Abstract

The objective of this paper is to compare different time series methods for the short-run forecasting of Business and Consumer Survey Indicators. We consider all available data taken from the Business and Consumer Survey Indicators for the Euro area between 1985 and 2002. The main results of the forecast competition are offered not only for raw data but we also consider the effects of seasonality and removing outliers on forecast accuracy. In most cases the univariate autoregressions were not outperformed by the other methods. As for the effect of seasonal adjustment methods and the use of data from which outliers have been removed, we obtain that the use of raw data has little effect on forecast accuracy. The forecasting performance of qualitative indicators is important since enlarging the observed time series of these indicators with forecast intervals may help in interpreting and assessing the implications of the current situation and can be used as an input in quantitative forecast models.

Keywords: Comparative methods, Evaluating forecast, Forecasting competition, Indicators, Monitoring forecasts, Business Surveys.

Address for correspondence

Grup d'Anàlisi Quantitativa Regional (AQR)
Espai de Recerca en Economia
Facultat de Ciències Econòmiques i Empresarials, Universitat de Barcelona
Avda. Diagonal 690 - 08034 Barcelona, Espanya
Tel: +34 93 4021012
Fax: +34 93 4021821
E-mail: rmoreno@ub.edu

1. Introduction

The amount of information provided by Business and Consumer Surveys concerning agents' perceptions and expectations of their environment means that they are now recognised as a crucial instrument for gathering economic information in today's ever-changing environment. Data obtained from Business and Consumer Surveys are often used in forecasting models and in testing different expectation schemes – Kauppi et al (1996), Batchelor and Dua (1998), Mourougane and Roma (2002) and Nardo (2003)-. The speed with which the results of these surveys are made available and the wide range of variables included make them extremely useful for decision-making -see Stuart (1985) for a deep discussion on the value of Business and Consumer Surveys-. The remarkable growth in business surveys in Europe since the early 1960s, and the need for them to be carried out and presented in a comparable way, led to the implementation of the Joint Harmonised EU Programme by the Commission in 1961.

The present paper tries to compare different time series methods for the short-run forecasting of Business and Consumer Survey Indicators. Certainly, forecast competitions have been considered in the economic literature although focused in quantitative variables such as industrial production (Byers and Peel, 1995; Simpson et al, 2001), output growth and employment (Clements and Smith, 2000) and exchange rates (Clements and Smith, 2001; Boero and Marrocu, 2002), as well as for general macroeconomic time series (Stock and Watson, 1999). On the other hand, as far as we know, forecast competitions have not been conducted for the case of qualitative variables. However, we consider that this kind of exercise can be useful to analyse which forecasting method presents the best behaviour. The usefulness of this comparison is twofold. First, it will allow having the best qualitative forecast to evaluate whether a series will move up or down in all kind of processes which can be modified according to changing economic conditions (McIntosh and Dorfman, 1992) or to predict business cycle turning points (Diebold and Rudebusch, 1989). And second, it will guarantee that the best forecast could be used as an explanatory variable in quantitative forecast models (Biart and Praet, 1987; Parigi and Schlitzer, 1995).

The objective of this paper is therefore to compare different time series methods for the short-run forecasting of Business and Consumer Survey Indicators. This objective can be summarised in the following two questions: Is it possible to forecast qualitative indicators? And, if so, which is the best procedure for conducting such forecasts? In order to answer these questions, we considered all available data taken from the Business and Consumer Survey Indicators for the Euro area between, in the main, January 1985 and December 2002. The dataset analysed includes 38 indicators (33 of which are monthly and 5 quarterly) and 6 composite indicators. This gives a total of 216 observations for the monthly series and 73 for the quarterly series.[1]

First, we considered raw data in order to test the forecast accuracy of five different sets of models: autoregressions, ARIMA, Self-exciting threshold autoregressions (SETAR), Markov switching regime models and vector autoregressions (traditional VAR and also VAR models considering the joint evolution of different indicators). Some of the conclusions obtained are related to the high volatility presented by indicators of this type, so that their statistical properties need to be taken into account when concluding.

An additional aspect considered in this comparison of forecasts is related to the balance forecasts. As survey data are derived from qualitative questions and based on subjective evaluations, the results are usually presented in terms of balances, which show the difference between the positive and negative percentages of answers. Since it is this balance that analysts take into account and the information that is usually forecast, we examine whether it is better to forecast the balance directly or rather to forecast the negative and positive answers first and, then, calculate the balance. Additionally, we examine the composite indicators, which are calculated from business surveys in which answers are weighted to various questions in the survey. When looking at these aggregate indicators, we analyse whether it is preferable to forecast them directly or to obtain the forecast by weighting the forecasts for the different components.

Finally, we consider the effects of seasonal adjustment procedures and the removal of outliers on forecast accuracy. In the case of seasonal adjustment procedures, we apply different methods to the indicators so as to obtain seasonally adjusted data (using Tramo/Seats - TS, X12 and Dainties - DA) and trend cycle estimation (using TS, X12 and Wavelets - WAV) in order to evaluate differences. In the case of outliers, we consider how their presence might affect the results by using TS to remove them (additive, transitory changes and level shifts) from the original series, and to what extent the effects of outliers should be removed each time a new observation is available.

The outline of the paper is as follows. The next section presents the models that are considered in the forecast competition. The database and the design of the forecast competition experiment is given in section 3. Section 4 offers the main results of the forecast competition not only for raw data but we also consider the effects of seasonality and removing outliers on forecast accuracy. Finally, section 5 concludes.

2. Models for forecasting Business and Consumer Survey Indicators

In order to assess alternative methods and models for forecasting Business and Consumer Surveys Indicators, we chose to focus on five different sets of model: autoregressions (AR), ARIMA, Self-exciting threshold autoregressions (SETAR), Markov switching regime models (MK) and vector autoregressions (traditional VAR and also VAR models considering the joint evolution of different indicators).

Autoregressions

The widely known autoregressive model (also known as the distributed-lags model) explains the behaviour of the endogenous variable as a linear combination of its own past values:

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + ... + \phi_p Y_{t-p} + \varepsilon_t \, .$$

The key question is how to determine the number of lags that should be included in the model. For monthly -quarterly- data we considered different models with a minimum number of 1 lag up to a maximum of 24 -8- (including all the intermediate lags), selecting the model with the lowest Akaike Information Criteria (AIC) value.[2]

ARIMA models

Since the study conducted by Box and Jenkins (1970), ARIMA models have been widely used and their forecast performance has also been confirmed.[3] The general expression of an ARIMA model is the following:

$$\Phi_s\left(L^s\right)\phi(L)(1-L^s)^D(1-L)^d x_t^\lambda = \Theta_s\left(L^s\right)\theta(L)\varepsilon_t ,$$

where L is the lag operator, $\Theta_s\left(L^s\right)=\left(1-\Theta_s L^s-\Theta_{2s}L^{2s}-\ldots-\Theta_{Qs}L^{Qs}\right)$ is a Q-order seasonal moving average polynomial, $\Phi_s\left(L^s\right)=\left(1-\Phi_s L^s-\Phi_{2s}L^{2s}-\ldots-\Phi_{Ps}L^{Ps}\right)$ is a P-order seasonal autoregressive polynomial, $\theta(L)=\left(1-\theta_1 L^1-\theta_2 L^2-\ldots-\theta_q L^q\right)$ is a q-order regular moving average polynomial, $\phi(L)=\left(1-\phi_1 L^1-\phi_2 L^2-\ldots-\phi_p L^p\right)$ is a p-order regular autoregressive polynomial, D is the seasonal difference order, d is the regular difference order, λ is the value of the Box-Cox (1964) transformation, *S* is the periodicity of the time series under consideration (S=4 for quarterly data, and S=12 for monthly data), and ε_t is the disturbance assumed to behave as a white noise.

In order to use models of this kind for forecasting, the proper model has to be identified (i.e., giving values to the order of the different polynomials, to the difference operator, etc.). For monthly data, we considered models with up to 12 AR and MA terms (4 in the case of quarterly data) selecting the model with the lowest AIC value. The statistical goodness of the selected model was also checked.

TAR models

In the case of the ARIMA model the relationship between the current value of a variable and its lags is supposed to be linear and constant over time. However, when looking at real data it can be seen that expansions tend to be more prolonged over time than recessions (Hansen, 1997). In fact, in the behaviour of most economic variables there seems to be a cyclical asymmetry that linear models are unable to capture (Clements and Smith, 1999). A Self-Excited Threshold Autoregressive model (SETAR) for the time series X_t can be summarised as follows:

$$\begin{array}{ll} B(L)\cdot X_t + u_t & \text{if } X_{t-k} \leq X\text{, and} \\ \zeta(L)\cdot X_t + v_t & \text{if } X_{t-k} > X\text{,} \end{array}$$

where u_t and v_t are white noises, $B(L)$ and $\zeta(L)$ are autoregressive polynomials, the value k is known as delay and the value X is known as threshold. This two-regime self-exciting threshold autoregressive process is estimated using monthly and quarterly data for each indicator and the Monte Carlo procedure is used to generate multi-step forecasts. The delay values selected are those that minimise the sum of squared errors among values between 1 and 12 for monthly data and 1 and 4 for quarterly data.[4] The values of the threshold are given by the variation of the variable being analysed.

Markov switching regime models

Threshold autoregressive models are perhaps the simplest generalisations of linear autoregressions. In fact, these models were built on developments of traditional ARMA time series models. As an alternative to these, time series regime-switching models assume that the distribution of the variable is known conditional on a particular regime or state occurring. When the economy changes from one regime to another, a substantial change occurs in the series. Hamilton (1989) presented the Markov regime-switching model in which the unobserved regime evolves over time as a 1st-order Markov process. The regime completely governs the dynamic behaviour of the series. This implies that once we assume conditions on a particular regime occurring, and assume a particular parameterisation of the model, we can write down the density of the variable of interest. However, as the regime is strictly unobservable, statistical inferences have to be drawn concerning the likelihood of each regime occurring at any point in time. So, we need to obtain the transition probabilities from one regime to the other.

Three approaches have been adopted in estimating these models (Potter, 1999). First, Hamilton (1989) developed a non-linear filter to evaluate the likelihood function of the model and, then, he directly maximised the likelihood function. Second, in a later article, Hamilton (1990) constructed an EM algorithm which proves particularly useful should all the parameters switch. Finally, Albert and Chib (1993) developed a Bayesian approach to estimation.

Here, we employ a Markov-switching threshold autoregressive model (MK-TAR) in which we are able to allow for different regime-dependent intercepts, autoregressive parameters, and variances. The estimation of the models is carried out by maximum likelihood using the Hamilton (1989) filter[5] together with Kim's smoothing filter (1994).

Having estimated the probabilities of expansion and recession, we could then construct the following model for the time series X_t:

$$\begin{array}{ll} B(L)\cdot X_t + u_t & \text{if } P[Expansion/X_{t-k}] \leq P\text{, and} \\ \zeta(L)\cdot X_t + v_t & \text{if } P[Recession/X_{t-k}] > P\text{,} \end{array}$$

where, as in the SETAR models, u_t and v_t are white noises, $B(L)$ and $\zeta(L)$ are autoregressive polynomials, the value k is the estimated delay and the value P is the estimated threshold[6]. The selected delay values are those that minimise the sum of squared errors for values between 1 and 12 for monthly data and 1 and 4 for quarterly data. The values of the threshold are given by the variation of the probability.

VAR models

In these models, each variable depends on a certain number of lags of the other variables under analysis (Sims, 1982). The idea is that the positive, neutral and negative answers to each question can be considered jointly. Moreover, as the sum of the percentages of positive (P), neutral (E) and negative (M) answers would, by definition, total one hundred, this restriction could also be introduced in the model improving its forecasting accuracy:

$$\begin{bmatrix} P_t \\ E_t \\ M_t \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix} \cdot \begin{bmatrix} P_{t-1} \\ E_{t-1} \\ M_{t-1} \end{bmatrix} + \ldots + \begin{bmatrix} Z_{11} & Z_{12} & Z_{13} \\ Z_{21} & Z_{22} & Z_{23} \\ Z_{31} & Z_{32} & Z_{33} \end{bmatrix} \cdot \begin{bmatrix} P_{t-p} \\ E_{t-p} \\ M_{t-p} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \end{bmatrix}.$$

In order to use models of this kind for forecasting, the proper model must first be identified (i.e., give values to the number of lags p). For monthly data, we considered models with up to 24 lags (8 in the case of quarterly data), selecting the model with the lowest AIC value. The statistical goodness of the selected model was also checked.

3. Database and design of the forecast competition experiment

The immediate relevance of the results of harmonised business and consumer surveys is one of their main properties, given that they are published shortly after the termination of the month to which they refer. The survey results are presented either in the form of balances for particular questions or as synthetic indicators. They are able to describe the panorama of the current economic situation much quicker than the quantitative indicators and the macroeconomic magnitudes from national accounts, though the latter tend to be more precise in their descriptions.

The European Commission draws up and publishes a wide range of indicators calculated on the basis of the results sent in by more than 40 institutes in 25 countries in the framework of the Joint Harmonised EU Programme of Business and Consumer Surveys. Particular attention is paid to indicators for the Euro area. The EU Programme currently includes surveys for industry, construction, the retail trade, services, investment and consumers.

For our analysis, we considered all the information available for the Business and Consumer Survey Indicators in the Euro area. The dataset analysed includes 38 indicators (33 of which are monthly and 5 quarterly) and 6 composite indicators. The starting date of these indicators differs but most of them begin in January 1985 (or in the first quarter of 1985). The latest period to be included in the analysis is December 2002 (or the final quarter of 2002). More details on the dataset can be found in Table 1.

Since the objective of the paper was to assess alternative methods and models for forecasting Business and Consumer Survey Indicators, we initially considered raw data (in all cases, non seasonally adjusted levels of each category of the variables were used) in order to test the forecast accuracy of the five different sets of models presented in the previous section.

In order to evaluate the relative forecasting accuracy of the models, each model was estimated for all the indicators included up to 2000.12 (or 2000.IV for quarterly indicators) and forecasts for 1, 2, 3, 6 and 12 months (or 1, 2, 4 quarters) in the future were computed. The model specifications are based on information up to 2000.12 or 2000.IV and, thereafter, the models were re-estimated in each month or quarter and the forecasts were computed with these estimation results. Given the availability of actual values up to 2002.12 or 2002.III or 2002.IV, we were able to compute the forecast errors for each indicator and method in a recursive way (i.e., for the 1 month forecast horizon, 24 forecast errors were computed for each indicator). In order to summarise this information, the Root Mean Square Error (RMSE) and Mean

Absolute Percentage Error (MAPE) were computed. These values provide useful information for analysing the forecast accuracy of each method, and enabled us to rank the methods according to their values.

Before showing the results of this comparative exercise, it should be stressed that the statistical properties of the Business and Consumer Survey Indicators differ substantially from those of the main macroeconomic variables (GDP, CPI, Industrial Production, Industrial Producer Prices, etc.). In Table 2, the variation coefficient is shown. It is worth noting that the variation coefficient values are extremely high for some indicators. This can be interpreted as evidence of the high volatility of the indicators and, with this in mind, the forecast accuracy of the methods considered can be expected to be lower than that for other macroeconomic variables.

A further result of interest is that if we examine the average values of the variation coefficient for categories (i.e., positive answers, negative answers, balance, etc.), the highest value corresponds to the balance. A possible explanation for the higher variance of the balance is purely statistical. The variance of the balance can be decomposed as follows:

$$Var(b) = Var(p-m) = Var(p) + Var(m) - 2 \cdot Cov(p,m) .$$

So, if the covariance between positive and negative answers is negative (i.e. they move in opposite directions), the variance of the balance would be higher than the sum of the variances of the other two components. In 87% of cases, the sign of this covariance is negative, so the variance of the balance is higher (in most cases) than the variance of the sum of variances of the two components (Table 3a). To test whether it was significantly higher, we applied a test of equality of variances between the balance and the other two components. In 76% of the cases, the null hypothesis of equality of variances was rejected (Table 3b).

Additionally, we computed some of the most commonly used methods to test the unit root hypothesis: the augmented Dickey and Fuller test (1979), the Phillips and Perron test (1988) and the Kwiatkowski, Phillips, Schmidt and Yongcheol test (1992) (Table 4a). The Perron test (1989) was also adapted so as to allow for different types of structural change: in the level (crash model) or in the slope (breaking trend) (Table 4b). Since the variables under consideration can only take values between 0 and 100, a priori we would expect most of them to be I(0). The most striking result is that in many cases the variables are eventually considered to be I(1) - 85% if the structural break is not considered and 52% when it is [7].

4. Results: forecast accuracy comparison

4.1. Main results for raw data

In Table 5 we present a summary of the recursive forecasts of the main indicators (Economic Sentiment -v1-, Industrial Confidence -v2-, Consumer Confidence -v12-, Construction Confidence -v28-, Retail Trade Confidence -v33- and Services Confidence -v39- Indicators) for the different models. The table refers to raw data. For each indicator, the best model is highlighted.

Before analysing the results, it should be pointed out that composite indicators are usually calculated from business surveys. For example, the Economic Sentiment Indicator is obtained by weighting the answers to different questions on the survey. When dealing with these aggregates, there are two possible courses of action: forecasting them directly or obtaining the forecast by weighting the forecasts for the different components. To the best of our knowledge, no research has been undertaken in order to determine which method provides the best results. Therefore, we chose to forecast composite indicators in two ways: what we call a *direct* (computation) method and by using the forecasts of the components, an *indirect* (computation) method.

Among the conclusions that can be extracted from Table 5, we would like to point out the following. First, indirect, as opposed to direct, methods seem to perform best. Thus, it would appear to be better to forecast these indicators from the forecasts of the components than directly. Second, among the direct methods, the AR model outperformed the rest of the models in almost all cases. By contrast, the ARIMA and TAR models were never found to be the best forecasters. Third, among the indirect methods, the AR and VAR models were the ones that provided the lowest RMSE (with the exception of the Economic Sentiment Indicator, where the ARIMA model outperformed the rest of the models). Once again, the TAR model presented the highest values for the RMSE.

As for the results of the forecast comparison for the remaining indicators, details figures are shown in Table 6, in which the average RMSE is shown for each type of answer and where the best model is highlighted in each case. The main conclusions of this analysis can be summarised as follows. As far as the forecast accuracy of the different methods is concerned, in most cases the univariate autoregressions are not outperformed by the other methods. In fact, only the forecast errors from the VAR and Markov models were lower than those obtained from AR models in some cases. It should also be stressed that unrestricted VAR models usually work better than restricted VAR models and that the errors displayed by the ARIMA, TAR and VAR models for different indicators were generally higher.

Furthermore, taking into account that the variables considered (positive, neutral and negative answers but not the balance) can only take values between 0 and 100, the forecast errors are quite high even in the case of the best model. As expected, the forecasts errors increased for longer horizons in most cases and, in general, variables corresponding to questions with a higher number of possible answers were better forecast than the rest. In most cases, the size of the errors was higher for the balance than for the components (this is related to the higher volatility of these variables), which is a common result for most composite indicators (one notable exception is the Economic Sentiment Indicator).

A further aspect to be considered is related to that of balance forecasts. As survey data are derived from qualitative questions and based on subjective evaluation, the results are usually presented in terms of balances, which show the difference between positive and negative percentages of answers. The balance is the information that analysts take into account and the information that is usually forecast. But, is it better to forecast the balance directly or to forecast negative and positive answers and then calculate the balance? In order to answer this question we replicated the forecasting comparison described above, but on this occasion we computed the balance from the forecasts of negative and positive answers using the AR, ARIMA, TAR and VAR models. The results (in the last block of Table 6, referred to as 'b') show that it is usually better to forecast the balance from the forecasts of positive and negative answers rather than by doing it directly, and that the AR model outperforms the rest of the models in almost all cases.

4.2. The effects of seasonality on data revision and on forecast accuracy

Seasonal adjustment methods are usually applied to these indicators. However, an interesting point that has not been analysed to date is the extent to which the method chosen for seasonal adjustment - TS, X12, DA, WAV, among others - might affect the values of the series under consideration. In order to remedy this situation, we applied all these methods of seasonal adjustment so as to obtain seasonally adjusted data (using TS, X12 and DA) and trend cycle estimations (using TS, X12 and WAV) for a number of qualitative indicators [8]. From the results shown in Table 7, several conclusions can be drawn:

First, TS and X12 results were very similar but DA results (seasonal adjustment) and WAV results (trend cycle-estimation) differed greatly. Specifically, differences were found to be greater in trend-cycle estimations than in seasonal adjustments (Table 7a and 7b).

Second, in order to analyse how the seasonal adjustment method chosen affected the revision of the series, seasonally adjusted data and trend-cycle estimations were computed in a recursive way adding one more observation from 2001.12 to 2002.12. The results (see Table 7c and 7d) showed that:

- There was no relationship between the size of the revision in a given observation and the number of time periods between this observation and the last observation available.
- There were great differences in the size of revision for the different variables in each seasonal adjustment method considered.
- There were no revisions of seasonally adjusted data using the DA method.
- TS revisions of seasonally adjusted data were greater than X12 revisions.
- Trend-cycle estimations using WAV method showed dramatic revisions.
- TS revisions of trend-cycle data were greater than X12 revisions.

Third, the choice of seasonal adjustment method was found to affect forecasts. Thus, we were interested in determining whether it was better to forecast raw data and then apply a seasonal adjustment method or not. As pointed out by Commission staff, the information from Business and Consumer Surveys has little seasonality, and, therefore the results should not vary significantly. To confirm this, we computed the Kruskal-Wallis test (also in Table 7e) for all the qualitative variables in order to verify the importance of seasonality. In almost 87% of cases, the null hypothesis of non-seasonality was not rejected, that is, most series did not present a seasonal component. This, it would seem, despite the large number of studies discussing this matter, that too much attention has been given to this issue in the context of Business and Consumer Surveys.

4.3. The effects on forecast accuracy of removing outliers using TS

The presence of outliers may well affect the results discussed in the previous sections. Therefore, TS can be used in order to remove the outliers from the original series. TS uses an automatic procedure to detect and eliminate outliers from a series (Gómez and Maravall, 1997). Three different types of outliers are considered here: Additive outliers (AO), Transitory changes (TC) and Level shifts (LS).

In Table 8a the variation coefficient for the qualitative variables from the industrial survey and the balances for the other surveys are shown. From these statistics, it can be seen that for some series, TS did not detect any outliers (i.e., the Economic Sentiment Indicator). However, in most cases, the volatility was substantially lower, although in some the value of the variation coefficient increased (although the standard deviation was lower, the value of the mean –which is clearly affected by the presence of extreme values- was even lower). As before, we now computed the value of the covariance between positive and negative answers and we tested to see if the variance of the balance differed from that of positive or negative answers. As observed in Table 8b, in 95% of the cases (as opposed to 87% when seasonality was not taken into account), the covariance was negative, so that the null hypothesis of equality of variances between the balance and the other two components (positive and negative answers) was rejected in 82% of the cases (compared with 76%) -see Table 8c-. Finally, the results on the unit root tests did not change at all when the outliers were removed from the variables, so that the main results remained unchanged.

Using these data, we repeated the comparison of forecasts as in the previous sections. The main results are shown in Table 8d. The main conclusions to be drawn from this analysis can be summarised as follows: First, RMSE values were the same or lower in nearly all variables when using data from which the outliers had been removed using TS. However, there were differences between models: while in the AR, ARIMA, TAR and Markov models the values of RMSE were markedly lower than those for raw data, the results for VAR (unrestricted and restricted) were, some times, worse. Finally, the VAR models consisting of different indicators performed much better when the outliers were removed.

Second, when comparing the RMSE for the balance computed from forecasts with positive and negative answers, we observed that the results for the models analysed (AR, ARIMA, TAR and VAR) - once the outliers had been removed - were, in most cases, higher than those for raw data. Similarly, when the composite indicators were computed from forecasts based on their components, here again the models performed worse once the outliers had been removed.

A practical issue that needs to be borne in mind when forecasting these series in real time concerns the effects of incorporating new observations on outlier detection using TS. In other words, is it necessary to remove the effects of outliers each time a new observation is available? To analyse this, we adopted a similar approach to that regarding the effects of seasonal adjustment procedures. In order to evaluate the effects of incorporating new observations, we identified the time periods during which an outlier is found in a recursive way by adding one more observation from 2001.1 to 2002.12. The results of this analysis are shown in Table 9. From these results, we can conclude that the number and type of outliers clearly depend on the available sample. Although there was some 'persistence' in the moment and type of detected outliers, some significant changes in the results might also be recorded using TS once an additional observation is included (sometimes due to changes in the underlying model, but also to the dynamics of the series itself). So, if we take into account the results regarding the forecast accuracy of the various models, and although this has the effect of increasing the computational cost, it would appear to be necessary to apply the TS to remove the effects of outliers each time an additional observation becomes available.

In addition, we should stress that when comparing the various seasonal adjustment methods, the TS and X12 results were very similar, but the DA (seasonal adjustment) and WAV results (trend cycle-estimation) presented marked differences. In fact, these differences were greater for the trend-cycle estimations than the seasonal adjustments, whereas there were considerable differences in the size of revision for the different variables for each seasonal adjustment method considered.

Additionally, after removing the outliers using TS the forecast accuracy of the various methods was similar to that observed with raw data. The number and type of outliers identified with TS clearly depends on the available sample. Although there is some 'persistence' in the moment and type of detected outliers, significant changes can be reflected in the results.

5. Conclusions

The objective of this paper is to compare time series methods for the short-run forecasting of Business and Consumer Survey Indicators. We analysed a dataset for the Euro area, which included 38 indicators (33 of which were monthly and 5 quarterly) and 6 composite indicators, mostly between January 1985 and December 2002. In order to test the forecast accuracy, we used five different sets of models: autoregressions, ARIMA, Self-exciting threshold autoregressions, Markov switching regime models and vector autoregressions (traditional VAR and also VAR models considering the joint evolution of different indicators).

As far as the forecast accuracy of the different methods is concerned, in most cases the univariate autoregressions were not outperformed by the other methods. In fact, only the forecast errors from the VAR and Markov models were, in some cases, lower than those from the AR models. It should also be stressed that unrestricted VAR models usually worked better than restricted VAR models and that the errors displayed by the ARIMA, TAR and VAR models comprising a range of indicators were generally higher. However, the size of the forecast errors was high even in the case of the best model.

Interestingly, both in the case of composite indicators and indicators which belong to a balance category, our results show that it is preferable to forecast them indirectly. As for the effect of the seasonal adjustment methods that are typically applied to these indicators, we used a range of methods (TS, X12, Dainties and Wavelets) that suggest that the use of raw data has little effect on forecast accuracy. A similar conclusion is obtained when using data from which outliers have been removed using TS, as here again forecast errors were found to be similar for nearly all variables.

References

Albert, J. & Chib, S. (1993). Bayesian inference via Gibbs sampling of autoregressive time series subject to Markov mean and variance shifts, *Journal of Business and Economic Statistics*, 11, 1-15.

Batchelor, R.A. & Dua, P. (1998). Improving macro-economic forecasts. The role of consumer confidence, *International Journal of Forecasting*, 14, 71-81.

Biart, M. & Praet, P. (1987). The contribution of opinion surveys in forecasting aggregate demand in the four main EC countries, *Journal of Economic Psychology*, 8, 409-428.

Boero, G. & Marrocu, E. (2002). The performance of non-linear exchange rate models: A forecasting comparison, *Journal of Forecasting*, 21, 513-542.

Box, G. & Cox, D. (1964). An analysis of transformation, *Journal of the Royal Statistical Society Series B*, 26, 2.

Box, G.E.P. & Jenkins, G.M. (1970). *Time series analysis: Forecasting and Control*, Holden-Day, San Francisco.

Byers, J.D. & Peel, D.A. (1995). Forecasting industrial production using non-linear methods, *Journal of Forecasting*, 14, 325-336.

Clements, M.P. & Smith, J. (1999). A Monte Carlo study of the forecasting performance of empirical SETAR models, *Journal of Applied Econometrics*, 14, 123-141.

Clements, M.P. & Smith, J. (2000). Evaluating the forecast densities of linear and non-linear models: Application to output growth and unemployment, *Journal Forecasting*, 19, 255-276.

Clements, M.P. & Smith, J. (2001). Evaluating forecasts from SETAR models of exchange rates, *Journal of International Money and Finance*, 20, 133-148.

Debenedictis, L.F. (1997). A vector autoregressive model of the British Columbia regional economy, *Applied Economics*, 29(7), 877-888.

Dickey, D.A. & Fuller, W.A. (1979). Distribution of the estimators for autoregressive time series with a unit root, *Journal of American Statistical Association*, 74, 427-431.

Diebold, F.X. & Rudebusch, G.D. (1989). Scoring the leading indicators, *Journal of Business*, 62, 369-391.

Feng, H. & Liu, J. (2003). A SETAR model for Canadian GDP: non-linearities and forecast comparisons, *Applied Economics*, 35(18), 1957-1964.

Gómez, V. & Maravall, A. (1997). Programs TRAMO (Time Series Regression with ARIMA Noise, Missing Observations, and Outliers) and SEATS (Signal Extraction in ARIMA Time Series) Instructions for the User, *Banco de España*.

Hamilton, J. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle, *Econometrica*, 57, 357-384.

Hamilton, J. (1990). Analysis of time series subject to changes in regime, *Journal of Econometrics*, 45, 39-70.

Hansen, B.E. (1997). *Inference in threshold models*, mimeo.

Kauppi, E., Lassila, J. & Teräsvirta, T. (1996). Short-term forecasting of industrial production with business survey data: experience from Finland's great depression 1990-1993, *International Journal of Forecasting*, 12, 373-381.

Kim, C.J. (1994). Dynamic linear models with Markov-Switching, *Journal of Econometrics,* 60, 1-22.

Kwiatkowski, D., Phillips, P.C.B., Schmidt, P. & Yongcheol, S. (1992). Testing the null hypothesis of stationary against the alternative of a unit root: How sure are we that economic time series have a unit root?, *Journal of Econometrics*, 54, 159-178.

McIntosh, C.S. & Dorfman, J.H. (1992). Qualitative forecast evaluation: A comparison of two performance measures, *American Journal of Agricultural Economics*, 74, 209-214.

Mourougane, A. & Roma, M. (2002). Can confidence indicators be useful to predict short term real GDP growth?, Working Paper No. 133, *European Central Bank*.

Nardo, M. (2003). The quantification of qualitative survey data: A critical assessment, *Journal of Economic Surveys*, 17, 645-668.

Parigi, G. & Schlitzer, G. (1995). Quarterly forecasts of the Italian business-cycle by means of monthly economic indicators, *Journal of Forecasting*, 14, 117-141.

Perron, P. (1989). The great crash, the oil price shock and the unit root hypothesis, *Econometrica*, 57, 1361-1401.

Phillips, P.C.B. & Perron, P. (1988). Testing for a unit root in time series regression, *Biometrika*, 75, 335-346.

Potter, S. (1999). Nonlinear time series modelling: An introduction, *Journal of Economic Surveys*, 13, 505-528.

Seitz, H. (1988). The estimation of inflation forecasts from business survey data, *Applied Economics*, 20, 427-438.

Simpson, P.W., Osborn, D.R. & Sensier, M. (2001). Forecasting UK industrial production over the business cycle, *Journal of Forecasting*, 20, 405-424.

Sims, C. (1982). Policy analysis with econometric models, *Brookings Papers on Economic Activity*, 1, 107-164.

Stock, J.H. & Watson, M.W. (1999). A comparison of linear and nonlinear univariate models for forecasting macroeconomic time series. In: R.F. Engle & H. White (Eds.), *Cointegration, Causality, and Forecasting. A Festschrift in Honour of Clive W.J. Granger* (pp. 1-44). Oxford: Oxford University Press.

Stuart, O.D.J. (1985). Are Business and Consumer Surveys Still of Value?, *Journal for Studies in Economics and Econometrics*, 0, iss. 23, 29-50.

Notes:

1. More details on the dataset can be found in Table 1. Although the number of observations in any empirical analysis should be ideally as large as possible, the type of analysis performed in this paper is considered to be valid over 50 observations.
2. We provide the AIC since it is one of the criteria most widely used when comparing time-series models. However, the sensitivity of the results has been analysed so as to check whether the results differ significantly if other criteria were used. We have also obtained the Hannan-Quinn criterion as well as the Schwartz criterion (or Bayesian information criterion) for all the models and a wide range of variables. The selected model is the same in all the cases, confirming the robustness of our results.
3. The univariate ARIMA models are usually used in the forecasting literature for comparative purposes when other forecasting methods are analysed (e.g. Debenedictis, 1997; Feng and Liu, 2003).
4. We have assumed one year in order to minimise the sum of squared errors since variables in Business and Consumer Surveys only ask for agents' perceptions and expectations of their environment in some next months or as a maximum of one year ahead. Thus, since the short term is the one being analysed in these surveys, this is the one that is taken in this paper when comparing the different time series methods.
5. The Hamilton filter is an iterative procedure which provides estimates of the probability that a given state is prevailing at each point in time given its previous history. These estimates are dependent upon the parameter values given to the

filter. Running the filter through the entire sample provides a log likelihood value for the particular set of estimates used. This filter is then repeated to optimise the log likelihood of obtaining the MLE estimates of the parameters. With the maximum likelihood parameters, the probability of state 0 at each point in time is calculated and these are the probabilities of recession and expansion.

6. An alternative approach would have consisted in imposing the value of P and k instead of estimating them. These models are known as Markov Switching Autoregressive Models (MS-AR) and, in general, the values of P are 0.7 or 0.8 and the values of k, 0 or 1.
7. It is also interesting to note that all the tests only coincide in 10 series.
8. The seasonality analysis was conducted for variables v1, v2, v3p, v3e, v3m, v3b, v4p, v4e, v4m and v4b.

We acknowledge financial support from ECFIN/2002/A3-01. We would like to thank the European Commission for their helpful comments and support. The usual disclaimer applies.

Table 1. List of Business and Consumer Surveys Indicators for the Euro area

	Description	Freq.		Last obs	Obs.	Categories
v1	**Economic Sentiment Indicator**	month	jan-85	dec-02	216	
v2	**Industrial Confidence Indicator (v7+v4-v6)/3**	month	jan-85	dec-02	216	
v3	Production trend observed in recent months	month	jan-85	dec-02	216	P E M B
v4	Assessment of order-book levels	month	jan-85	dec-02	216	P E M B
v5	Assessment of export order-book levels	month	jan-85	dec-02	216	P E M B
v6	Assessment of stocks of finished products	month	jan-85	dec-02	216	P E M B
v7	Production expectations for the months ahead	month	jan-85	dec-02	216	P E M B
v8	Selling price expectations for the months ahead	month	jan-85	dec-02	216	P E M B
v9	Employment expectations for the months ahead	month	jan-85	dec-02	216	P E M B
v10	*New orders in recent months*	*quarter*	*1985-I*	*2002-IV*	72	P E M B
v11	*Export expectations for the months ahead*	*quarter*	*1985-I*	*2002-IV*	72	P E M B
v12	**Consumer Confidence Indicator (v14+v16-v19+v23)/4**	month	jan-85	dec-02	216	
v13	Financial situation over last 12 months	month	jan-85	dec-02	216	PP P E M MM N B
v14	Financial situation over next 12 months	month	jan-85	dec-02	216	PP P E M MM N B
v15	General economic situation over last 12 months	month	jan-85	dec-02	216	PP P E M MM N B
v16	General economic situation over next 12 months	month	jan-85	dec-02	216	PP P E M MM N B
v17	Price trends over last 12 months	month	jan-85	dec-02	216	PP P E M MM N B
v18	Price trends over next 12 months	month	jan-85	dec-02	216	PP P E M MM N B
v19	Unemployment expectations over next 12 months	month	jan-85	dec-02	216	PP P E M MM N B
v20	Major purchases at present	month	jan-85	dec-02	216	PP E MM N B
v21	Major purchases over next 12 months	month	jan-85	dec-02	216	PP P E M MM N B
v22	Savings at present	month	jan-85	dec-02	216	PP P M MM N B
v23	Savings over next 12 months	month	jan-85	dec-02	216	PP P M MM N B
v24	Statement on financial situation of household	month	jan-85	dec-02	216	PP P E M MM N B
v25	*Intention to buy a car within the next 2 years*	*quarter*	*1990-I*	*2002-IV*	52	PP P M MM N B
v26	*Purchase or build a home within the next 2 years*	*quarter*	*1990-I*	*2002-IV*	52	PP P M MM N B
v27	*Home improvements over the next 12 months*	*quarter*	*1990-I*	*2002-IV*	52	PP P M MM N B

(Continues next page)

	Description	Freq.	First obs	Last obs	Obs	Categories			
v28	**Construction Confidence Indicator (v30+v31)/2**	month	jan-85	dec-02	216				
v29	Trend of activitiy compared with preceding months	month	jan-85	dec-02	216	P	E	M	B
v30	Assessment of order books	month	jan-85	dec-02	216	P	E	M	B
v31	Employment expectations for the months ahead	month	jan-85	dec-02	216	P	E	M	B
v32	Price expectations for the months ahead	month	jan-85	dec-02	216	P	E	M	B
v33	**Retail Trade Confidence Indicator (v34-v35+v37)/3**	month	jan-86	dec-02	204				
v34	Present business situation	month	jan-85	dec-02	216	P	E	M	B
v35	Assessment of stocks	month	jan-85	dec-02	216	P	E	M	B
v36	Orders placed with suppliers	month	feb-85	dec-02	215	P	E	M	B
v37	Expected business situation	month	jan-86	dec-02	204	P	E	M	B
v38	Employment	month	abr-85	dec-02	213	P	E	M	B
v39	**Services Confidence Indicator (v40+v41+v42)/3**	month	abr-95	dec-02	93				
v40	Assessment of business climate	month	abr-95	dec-02	93	P	E	M	B
v41	Evolution of demand in recent months	month	abr-95	dec-02	93	P	E	M	B
v42	Evolution of demand expected in the months ahead	month	abr-95	dec-02	93	P	E	M	B
v43	Evolution of employment in recent months	month	abr-95	dec-02	93	P	E	M	B
v44	Evolution of employment expected in the months ahead	month	jan-97	dec-02	72	P	E	M	B

The letters refer to positive answers (pp and p), neutral answers (e), negative answers (mm and m), non answers (n), balance (b) and composite indicators (i).

Table 2. Variation coefficient for the Business and Consumer Surveys Indicators

	Average variation coefficient	
	Monthly indicators	Quarterly indicators
pp	27.22	13.90
p	31.33	15.10
e	16.03	4.38
m	33.75	17.02
mm	33.76	2.55
n	19.52	36.44
b	348.47	118.68
i	67.26	

Table 3a. Analysis of the sign of the covariance between positive and negative answers

Covariance	Negative sign	Positive sign	TOTAL
Month	28	5	33
Quarter	5	0	5
TOTAL	33	5	38

Covariance	Negative sign	Positive sign	TOTAL
Month	74%	13%	87%
Quarter	13%	0%	13%
TOTAL	87%	13%	100%

Table 3b. Tests of equality of variance $(F = \frac{S_X^2}{S_Y^2} \sim F_{N_X-1,N_Y-1})$

H_0: Equality of variance ($\sigma_X^2 = \sigma_Y^2$)
H_A: Non-equality of variance

Statistic	Rejection of the Null	Non-rejection of the Null	TOTAL
Month	27	6	33
Quarter	2	3	5
TOTAL	29	9	38

Statistic	Rejection of the Null	Non-rejection of the Null	TOTAL
Month	71%	16%	87%
Quarter	5%	8%	13%
TOTAL	76%	24%	100%

Table 4a. Results of the tests for unit root hypothesis. Unit root tests without structural break. (I(1) in 85% of the cases)

Order	Variable	ADF	PP(k=4)	PP(k=8)	KPSSa	KPSSb
I(0)	pp	3	1	4	9	6
	p	4	4	13	18	10
	e	1	6	4	16	5
	m	4	3	12	21	9
	mm			1	8	3
	n	1	5	1	1	4
	b	4	9	11	24	10
	v1				1	1
	v2	1		1	1	1
	v12				1	
	v28				1	
	v33					1
Total I(0)		18	28	47	101	50
I(1)	pp	9	11	8	3	6
	p	23	23	14	9	17
	e	25	20	22	10	21
	m	23	24	15	6	18
	mm	12	12	11	4	9
	n	11	7	11	11	8
	b	24	19	17	4	18
	v1	1	1	1		
	v2		1			
	v12	1	1	1		1
	v28	1	1	1		1
	v33	1	1	1	1	
Total I(1)		131	121	102	48	99
Total general		149	149	149	149	149

ADF: augmented Dickey and Fuller test (1979); PP: Phillips and Perron test (1988); KPSS: Kwiatkowski, Phillips, Schmidt and Yongcheol test (1992).

Table 4b. Results of the tests for unit root hypothesis. Perron's (1989) test with structural break. (I(1) in 52 % of the cases)

Order	Variable	Crash model (a)	Crash model (b)	Breaking trend (a)	Breaking trend (b)
I(0)	pp	4	4	5	5
	p	13	12	13	12
	e	17	17	17	17
	m	14	13	12	13
	mm	5	4	4	4
	n	9	8	9	9
	b	11	11	10	10
	v1				
	v2	1	1	1	1
	v12				
	v28				
	v33				
Total I(0)		74	70	71	71
I(1)	pp	8	8	7	7
	p	14	15	14	15
	e	9	9	9	9
	m	13	14	15	14
	mm	7	8	8	8
	n	3	4	3	3
	b	17	17	18	18
	v1	1	1	1	1
	v2				
	v12	1	1	1	1
	v28	1	1	1	1
	v33	1	1	1	1
Total I(1)		75	79	78	78
Total general		149	149	149	149

Table 5. Average RMSE - Recursive forecasts from January 2001 to December 2002. Raw data for main composite indicators

Economic Sentiment Indicator (v1)

		1 month	2 months	3 months	6 months	12 months
direct methods	AR	***0.45****	0.77	1.02	1.51	1.52
	ARIMA	1.92	3.21	4.36	7.27	6.25
	TAR	4.65	7.14	9.33	15.07	20.51
	MK-TAR	0.61	0.75	1.19	2.20	6.12
indirect methods	AR	2.73	2.83	2.92	3.08	1.70
	ARIMA	0.46	***0.46****	***0.42****	***0.41****	***0.46****
	TAR	0.71	0.84	1.06	1.80	3.03
	VAR	4.38	4.55	4.70	4.98	3.49

**** Best model***

Industrial Confidence Indicator (v2)

		1 month	2 months	3 months	6 months	12 months
direct methods	AR	2.03	3.97	5.43	7.99	4.68
	ARIMA	10.54	21.70	29.85	40.20	39.19
	TAR	21.88	31.29	39.02	61.76	95.72
	MK-TAR	3.68	4.31	7.33	10.93	87.78
	VAR	3.39	5.63	6.90	12.25	17.82
indirect methods	AR	2.04	3.54	4.68	6.22	3.69
	ARIMA	2.21	4.14	5.52	7.60	7.97
	TAR	4.77	6.78	8.45	14.65	22.69
	VAR	***0.08****	***2.15****	***3.29****	***3.28****	***2.28****

**** Best model***

Consumer Confidence Indicator (v12)

		1 month	2 months	3 months	6 months	12 months
direct methods	AR	***1.68****	***2.71****	***3.65****	5.49	3.77
	ARIMA	9.13	13.57	18.02	30.84	25.56
	TAR	14.31	18.75	23.70	39.32	47.02
	MK-TAR	2.68	4.42	6.77	10.38	90.78
	VAR	3.02	5.24	7.67	13.62	24.60
indirect methods	AR	1.78	2.79	3.69	5.64	***3.27****
	ARIMA	5.42	4.54	3.93	***3.00****	3.67
	TAR	6.09	6.54	7.14	9.22	11.99
	VAR	9.47	11.66	13.42	15.31	12.92

**** Best model***

Construction Confidence Indicator (v28)

		1 month	2 months	3 months	6 months	12 months
direct methods	AR	2.01	2.36	2.38	2.97	***2.59****
	ARIMA	15.82	26.90	36.08	50.42	44.52
	TAR	26.32	38.53	50.42	73.45	100.06
	VAR	5.00	8.01	11.38	22.41	38.31
indirect methods	AR	1.97	***2.09****	***2.15****	***2.82****	2.88
	ARIMA	12.62	14.07	15.31	18.56	22.76
	TAR	13.06	14.62	16.96	25.21	32.69
	VAR	***0.07****	4.93	9.38	15.92	11.37

**** Best model***

Retail Trade Confidence Indicator (v33)

		1 month	2 months	3 months	6 months	12 months
direct methods	AR	2.59	3.08	3.32	4.68	8.04
	ARIMA	12.05	14.40	15.64	21.17	31.50
	TAR	17.73	24.04	26.92	39.67	59.06
	MK-TAR	2.56	3.63	***2.41****	4.15	***4.00****
	VAR	3.70	4.70	4.83	***4.00****	11.80
indirect methods	AR	***2.47****	3.17	3.50	5.13	8.82
	ARIMA	23.79	23.40	23.01	20.79	10.25
	TAR	29.27	32.48	34.80	40.68	40.94
	VAR	2.96	***2.72****	4.05	5.01	7.46

***** Best model***

Services Confidence Indicator (v39)

		1 month	2 months	3 months	6 months	12 months
direct methods	AR	5.44	10.31	15.51	26.48	43.33
	ARIMA	22.17	38.48	53.52	76.93	74.57
	TAR	56.92	76.78	86.97	126.94	182.96
	VAR	6.20	9.33	12.04	17.95	32.69
indirect methods	AR	5.61	9.94	13.97	20.49	35.87
	ARIMA	4.43	7.53	10.46	16.32	20.39
	TAR	12.00	16.22	18.57	25.46	41.09
	VAR	***0.69****	***2.34****	***7.59****	***14.37****	***18.90****

***** Best model***

Table 6. Average RMSE - Recursive forecasts from January 2001 to December 2002. Raw data for qualitative indicators according to the type of answer

pp

	1 month	2 months	3 months	6 months	12 months	1 quarter	2 quarters	4 quarters
AR	**1.00***	**1.48***	**1.91***	**2.74***	**3.23***	**0.61***	**0.57***	**0.61***
ARIMA	4.15	5.60	6.71	8.79	9.53	1.79	1.85	2.17
TAR	10.23	24.04	56.91	836.09	193363.17	2.55	3.36	4.35
MK-TAR	1.53	2.01	3.06	5.21	10.94	na	na	na

**** Best model***

p

	1 month	2 months	3 months	6 months	12 months	1 quarter	2 quarters	4 quarters
AR	2.07	2.77	**3.40***	**4.41***	**5.48***	**1.87***	**2.01***	**2.29***
ARIMA	9.94	12.68	15.07	18.91	19.76	6.30	7.46	7.74
TAR	17.83	24.60	31.22	48.50	72.76	10.95	11.73	17.79
MK-TAR	**1.92***	3.38	4.27	10.00	42.15	3.69	3.76	5.08
VAR unr	2.00	**2.77***	3.48	5.01	6.25	2.48	2.72	2.91
VAR rest	2.41	3.41	4.31	6.06	6.91	2.46	2.79	2.90

**** Best model***

e

	1 month	2 months	3 months	6 months	12 months	1 quarter	2 quarters	4 quarters
AR	2.22	2.87	3.37	4.28	5.21	3.12	3.52	3.47
ARIMA	10.42	13.05	15.00	17.95	18.30	9.40	10.13	7.64
TAR	36.59	39.57	42.46	92.09	167.79	12.38	14.58	19.95
MK-TAR	2.08	**2.46***	4.01	5.78	9.27	na	na	na
VAR unr	**1.88***	2.53	**3.03***	**3.82***	**4.42***	**1.50***	**1.83***	**2.19***
VAR rest	2.30	3.00	3.54	4.64	5.47	1.68	1.91	2.26

**** Best model***

m

	1 month	2 months	3 months	6 months	12 months	1 quarter	2 quarters	4 quarters
AR	**1.92***	2.77	**3.46***	**4.75***	**5.41***	**2.15***	**3.13***	**3.43***
ARIMA	9.67	13.72	16.81	22.32	23.98	7.02	10.47	11.00
TAR	18.11	39.66	51.40	112.63	781.91	10.53	14.04	15.11
MK-TAR	2.21	**2.59***	3.65	5.15	12.09	3.74	4.34	6.92
VAR unr	2.03	3.00	3.74	5.21	6.19	2.79	3.58	4.24
VAR rest	2.23	3.15	3.91	5.46	6.39	2.53	3.22	3.61

**** Best model***

mm

	1 month	2 months	3 months	6 months	12 months	1 quarter	2 quarters	4 quarters
AR	**0.72***	**1.00***	**1.19***	**1.69***	**2.34***	**1.24***	**1.22***	**1.40***
ARIMA	3.66	4.83	5.64	7.89	10.01	3.60	3.50	3.70
TAR	5.17	6.37	7.71	13.06	23.38	4.53	3.78	4.82
MK-TAR	0.96	1.12	2.08	4.40	14.34	1.28	1.34	1.72

**** Best model***

n

	1 month	2 months	3 months	6 months	12 months	1 quarter	2 quarters	4 quarters
AR	0.45	0.58	0.67	0.90	1.19	**0.58***	**0.62***	**1.10***
ARIMA	2.12	2.41	2.58	3.10	3.47	1.45	1.06	1.72
TAR	4.91	6.92	9.70	25.89	526.08	2.21	3.32	5.99
MK-TAR	**0.28***	**0.29***	**0.42***	**0.46***	**0.73***	na	na	na

**** Best model***

b

		1 month	2 months	3 months	6 months	12 months	1 quarter	2 quarters	4 quarters
direct methods	AR	3.40	4.90	6.33	8.94	12.00	4.05	4.90	5.38
	ARIMA	16.36	23.15	28.27	37.98	41.40	13.84	18.37	19.15
	TAR	29.00	39.76	49.61	70.67	89.30	20.48	23.94	24.43
	MK-TAR	3.63	6.46	8.76	17.37	67.04	6.74	7.18	11.37
	VAR	3.89	5.85	7.94	12.74	16.51			
indirect methods	AR	**3.29***	4.80	6.14	**8.41***	**9.64***	**3.93***	**4.79***	**5.15***
	ARIMA	3.31	**4.74***	**5.97***	8.50	11.19	4.68	6.73	8.77
	TAR	5.94	11.02	15.78	96.28	20317.70	6.07	6.63	6.87
	VAR unr	4.98	6.09	7.07	9.22	10.83	5.12	5.70	6.34
	VAR rest	5.47	6.84	8.14	10.59	11.61	4.73	5.31	5.63

**** Best model***

Table 7. Evaluation of the different seasonal adjustment methods

a) Mean squared error (MSE) between results from seasonally adjusted data

Average MSE between methods. (v1 … v4b)

		TS	X12	DA
TS	Average	0.000	0.008	0.173
	S.D.	0.000	0.004	0.074
X12	Average		0.000	0.193
	S.D.		0.000	0.089
DA	Average			0.000
	S.D.			0.000

b) Mean squared error (MSE) between results from trend-cycle estimation

Average MSE between methods. (v1 … v4b)

		TS	X12	WAV
TS	Average	0.000	0.010	0.369
	S.D.	0.000	0.000	0.016
X12	Average		0.000	0.354
	S.D.		0.000	0.011
WAV	Average			0.000
	S.D.			0.000

c) Revisions of trend-cycle series

S.D of mean absolute deviation

	TS	X12	WAV
V1	0.000	0.000	0.105
V2	0.041	0.193	4.131
V3P	0.003	0.002	0.965
V3E	0.012	0.002	0.209
V3M	0.023	0.004	1.911
V3B	0.030	0.010	5.495
V4P	0.003	0.035	0.922
V4E	0.016	0.001	1.065
V4M	0.022	0.001	3.685
V4B	0.035	0.015	7.998
V3B*	0.033	0.010	5.536
V4B*	0.035	0.004	5.427

* Indirect method.

d) Revisions of seasonally-adjusted series

S.D. of mean absolute deviation

	TS	X12	DA
V1	0.00	0.00	0.00
V2	0.25	0.02	0.00
V3P	0.01	0.01	0.00
V3E	0.00	0.01	0.00
V3M	0.00	0.01	0.00
V3B	0.03	0.03	0.00
V4P	0.04	0.00	0.00
V4E	0.00	0.01	0.00
V4M	0.00	0.01	0.00
V4B	0.02	0.02	0.00
V3B*	0.02	0.02	0.00
V4B*	0.00	0.02	0.00

* Indirect method.

e) Forecasting raw-data or seasonally adjusted data? The Kruskal-Wallis test for detecting seasonality (summary of the results)

H_0: non-seasonality
H_A: seasonality

	Rejection of the Null	Non-	TOTAL
Month	22	148	170
Quarter	4	22	26
TOTAL	26	170	196

	Rejection of the Null	Non-	TOTAL
Month	11.22%	75.51%	86.73%
Quarter	2.04%	11.22%	13.27%
TOTAL	13.27%	86.73%	100.00%

Table 8. The effects of removing outliers using TS

a) Average variation coeffcient

	Monthly indicators
P	28.92
E	5.86
M	32.64
B	979.29

b) Analysis of the sign of the covariance between positive and negative answers

Covariance	Negative sign	Positive sign	TOTAL
Month	32	1	33
Quarter	4	1	5
TOTAL	36	2	38

Covariance	Negative sign	Positive sign	TOTAL
Month	84%	3%	87%
Quarter	11%	3%	13%
TOTAL	95%	5%	100%

c) Tests of equality of variance

H_0: Equality of variance
H_A: Non-equality of variance

Statistic	Rejection of the Null	Non-rejection of the Null	TOTAL
Month	29	4	33
Quarter	2	3	5
TOTAL	31	7	38

Statistic	Rejection of the Null	Non-rejection of the Null	TOTAL
Month	76%	11%	87%
Quarter	5%	8%	13%
TOTAL	82%	18%	100%

d) The effects on forecasting accuracy

Economic Sentiment Indicator (v1)

		1 month	2 months	3 months	6 months	12 months
direct methods	AR	**0.45***	0.77	1.02	1.51	1.52
	ARIMA	1.92	3.21	4.36	7.27	6.25
	TAR	4.51	6.72	8.73	14.25	20.04
	MK-TAR	0.61	0.75	1.20	2.20	6.15
indirect methods	AR	2.83	2.94	3.04	3.22	1.84
	ARIMA	0.47	**0.44***	**0.41***	**0.40***	**0.48***
	TAR	0.77	0.85	1.04	1.61	2.34
	VAR	4.56	4.73	4.89	5.19	3.67

**** Best model***

p

	1 month	2 months	3 months	6 months	12 months	1 quarter	2 quarters	4 quarters
AR	**1.50***	**1.95***	**2.50***	**3.19***	3.38	**3.25***	**3.88***	**4.68***
ARIMA	8.52	11.42	13.65	18.01	15.99	11.22	13.92	15.08
TAR	15.53	20.80	26.23	45.54	76.03	20.94	21.29	34.48
MK-TAR	2.02	2.09	2.56	3.34	4.66	6.10	5.85	10.38
VAR unr	1.78	2.49	3.28	4.46	4.65	5.19	6.50	5.87
VAR rest	1.76	2.36	3.06	4.20	**3.14***	4.06	5.01	4.98

**** Best model***

e

	1 month	2 months	3 months	6 months	12 months	1 quarter	2 quarters	4 quarters
AR	**1.70***	2.30	**2.75***	3.33	**3.74***	**2.51***	**2.95***	**2.76***
ARIMA	8.53	11.41	13.81	16.64	16.20	7.17	7.96	5.72
TAR	13.31	16.98	19.71	27.24	34.88	10.17	12.37	18.09
MK-TAR	1.70	**2.12***	2.96	**3.31***	5.03	na	na	na
VAR unr	1.89	2.54	3.02	3.67	4.23	3.42	3.73	3.08
VAR rest	4.76	4.83	4.92	5.00	4.40	2.88	3.38	3.40

**** Best model***

m

	1 month	2 months	3 months	6 months	12 months	1 quarter	2 quarters	4 quarters
AR	1.54	2.32	3.07	4.53	3.92	**4.30***	6.15	6.27
ARIMA	8.68	13.16	16.76	22.93	15.55	14.55	21.98	22.36
TAR	14.48	21.38	28.02	49.10	83.71	22.46	30.87	35.20
MK-TAR	**0.68***	**0.63***	**0.81***	**0.94***	**0.66***	6.73	7.65	12.68
VAR unr	1.70	2.71	3.61	5.40	6.08	6.80	9.48	8.49
VAR rest	1.56	2.27	2.95	4.45	3.60	4.34	**6.00***	**6.03***

**** Best model***

b

		1 month	2 months	3 months	6 months	12 months	1 quarter	2 quarters	4 quarters
direct methods	AR	**2.57***	**3.64***	**4.78***	**6.89***	**5.39***	7.17	9.18	**9.13***
	ARIMA	14.55	21.81	27.82	40.06	42.39	26.40	37.28	37.60
	TAR	26.94	39.04	49.91	88.28	138.05	43.16	53.28	53.19
	MK-TAR	6.54	7.46	8.83	10.71	12.90	11.16	11.49	19.76
	VAR	3.10	4.34	5.77	9.48	12.17			
indirect methods	AR	4.09	4.97	6.00	8.06	7.36	**6.69***	**8.98***	9.24
	ARIMA	4.08	5.28	6.33	9.14	9.39	8.30	13.12	16.61
	TAR	5.73	6.87	8.14	12.69	20.76	12.20	13.44	13.61
	VAR unr	4.18	5.43	6.91	9.75	11.14	11.56	15.68	14.16
	VAR rest	4.37	5.45	6.68	9.36	7.27	7.77	10.48	10.47

**** Best model***

Table 9.

The next tables should be interpreted as follows: the number and type of detected outliers when adding a new observation are shown in columns, while in rows, the moment of time for each outlier can be found.

v2	jan-01	feb-01	mar-01	apr-01	may-01	jun-01	jul-01	aug-01	set-01	oct-01	nov-01	des-01	jan-02	feb-02	mar-02	apr-02	may-02	jun-02	jul-02	aug-02	set-02	oct-02	nov-02	des-02
01-																			AO	AO				
10-																			AO	AO				
07-	AO	AO	AO	AO	AO	AO	AO		AO			AO												AO
06-	AO	AO	AO	AO	AO	AO	AO		AO			AO												AO
10-													LS	LS	LS	LS	LS	LS				LS	LS	
11-																		TC						
12-												AO												
01-																			LS	LS				

AO: Additive Outlier, TC: Transitory Change, LS: Level shift.

v3b[1]	jan-01	feb-01	mar-01	apr-01	may-01	jun-01	jul-01	aug-01	set-01	Oct-01	nov-01	des-01	jan-02	feb-02	mar-02	apr-02	may-02	jun-02	jul-02	aug-02	set-02	oct-02	nov-02	des-02
03-91							TC	TC	TC					TC	TC	TC	TC	TC	TC	TC	TC	TC	TC	TC
05-97														LS										

AO: Additive Outlier, TC: Transitory Change, LS: Level shift.

v3m	jan-01	feb-01	mar-01	apr-01	may-01	jun-01	jul-01	aug-01	set-01	oct-01	nov-01	des-01	jan-02	feb-02	mar-02	apr-02	may-02	jun-02	jul-02	aug-02	set-02	oct-02	nov-02	des-02
03-91																			TC		TC	TC	TC	TC
07-96					AO	AO	AO	AO	AO	AO	AO	AO	AO	AO	AO	AO	AO	AO	AO	AO	AO	AO	AO	AO
05-00																			AO		AO	AO	AO	AO
05-01																			LS		LS	LS	LS	LS

AO: Additive Outlier, TC: Transitory Change, LS: Level shift.

[1] No outlier is detected for v3p.

v3e	jan-01	feb-01	mar-01	apr-01	may-01	jun-01	jul-01	aug-01	set-01	oct-01	nov-01	des-01	jan-02	feb-02	mar-02	apr-02	may-02	jun-02	jul-02	aug-02	set-02	oct-02	nov-02	des-02
02-91															TC									
11-92	LS		LS					LS							LS									

AO: Additive Outlier, TC: Transitory Change, LS: Level shift.

v8b	jan-01	feb-01	mar-01	apr-01	may-01	jun-01	jul-01	aug-01	set-01	oct-01	nov-01	des-01	jan-02	feb-02	mar-02	apr-02	may-02	jun-02	jul-02	aug-02	set-02	oct-02	nov-02	des-02
07-85	LS																							
01-88	TC	TC	TC	TC	TC	TC	TC	TC	TC	TC	TC	TC	TC	TC	TC	TC	TC	TC	TC	TC	TC	TC	TC	TC
02-93								TC										TC						TC
01-97	AO	AO	AO	AO	AO	AO	AO	AO	AO	AO	AO	AO	AO	AO	AO	AO	AO	AO	AO	AO	AO	AO		AO
06-01																		AO						
07-01																								AO

AO: Additive Outlier, TC: Transitory Change, LS: Level shift.

v8m[2]	jan-01	feb-01	mar-01	apr-01	may-01	jun-01	jul-01	aug-01	set-01	oct-01	nov-01	des-01	jan-02	feb-02	mar-02	apr-02	may-02	jun-02	jul-02	aug-02	set-02	oct-02	nov-02	des-02
07-85		LS	LS	LS		LS			LS	LS	LS		LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS
03-86												TC	TC	TC	TC	TC	TC	TC	TC	TC	TC	TC	TC	TC
05-86																			AO					
09-87																			AO					
01-88	AO	AO	AO	AO	TC	AO		AO	AO	AO	AO	TC	TC	TC	TC	TC	TC	TC	TC	TC	TC	TC	TC	TC
11-88							TC						TC	TC	TC	TC	TC	TC	TC	TC	LS	TC	TC	TC
03-89																AO			TC	AO				
05-90																			TC					
01-93	AO											AO												
10-98	AO																							
03-99	AO	AO	AO	AO	AO	AO	AO	AO	AO	AO	AO	AO												
01-00	AO	AO			AO				AO			AO												
02-00			AO	AO		AO				AO	AO								AO					
07-01							AO	AO	AO	AO	AO	AO				AO			AO	AO	AO	AO	AO	AO

AO: Additive Outlier, TC: Transitory Change, LS: Level shift.

v8e	jan-01	feb-01	mar-01	apr-01	may-01	jun-01	jul-01	aug-01	set-01	oct-01	nov-01	des-01	jan-02	feb-02	mar-02	apr-02	may-02	jun-02	jul-02	aug-02	set-02	oct-02	nov-02	des-02
03-99	AO	AO	AO	AO	AO	AO	AO	AO	AO	AO	AO												AO	AO
12-01																							AO	AO

AO: Additive Outlier, TC: Transitory Change, LS: Level shift.

[2] No outlier is detected for v8p.

Sub-sample of variables taken in order to analyse differences between Information Criteric

DIRECT METHODS

AR **Worksheet: direct AR**	***Composite*** **raw data** v1 v2 v12 v28 v33 v39 **adjusted** v1

VAR **Worksheet: direct VAR**	All composite indicators raw and adjusted

INDIRECT METHODS

AR **Worksheet: indirect VAR**	***Composite*** **raw data** v1 v2 v12 v28 v33 v39 **adjusted** v1

VAR **Worksheet: indirect VAR**	***Composite*** **raw data** v1 v2 v12 v28 v33 v39 **adjusted** v1

VAR restricted

Indicators		
raw data	*monthly*	*quarterly*
Worksheet: VAR restricted	v3	v10
	v9	v11
	v16	
	v29	
	v36	

VAR unrestricted

Indicators	
raw data	*monthly*
Worksheet: VAR unrestricted	v4
	v8
	v16
	v32
	v34
	v38

on

Indicators		
raw data	*monthly*	*quarterly*
	v3p	v10b
	v4e	v25b
	v7p	v27b
	v13p	
	v15p	
	v19p	
	v23p	
	v31p	
	v37p	
	v13e	
	v18e	
	v20e	
	v29e	
	v34e	
	v37e	
	v43e	

		AIC	SIC	HQC	
COMPOSITE					AIC
	raw				SIC
	v1	-417.5935	-417.5755	-417.5862	HQC
	v2	-325.6424	-325.6254	-325.6355	
	v12	-306.8569	-306.8399	-306.85	
	v28	-452.8738	-452.8568	-452.8669	
	v33	-412.751	-412.7332	-412.7438	
	v39	-135.0856	-135.053	-135.0727	
	adjusted				
	v1	-417.5935	-417.5755	-417.5862	
INDICATORS					
	raw				
	monthly				
	v3p	-414.5567	-414.5397	-414.5498	
	v4e	-269.5258	-269.5088	-269.5189	
	v7p	-373.8947	-373.8777	-373.8878	
	v13e	-222.3528	-222.3344	-222.3453	
	v13p	-299.7041	-299.6869	-299.6972	
	v15p	-233.9184	-233.9014	-233.9115	
	v18e	-224.9946	-224.9765	-224.9873	
	v19p	-319.0595	-319.0425	-319.0526	
	v20e	-221.4247	-221.4077	-221.4178	
	v23p	-270.2258	-270.2083	-270.2187	
	v29e	-531.3928	-531.3758	-531.3859	
	v31p	-379.8651	-379.8481	-379.8582	
	v34e	-486.5867	-486.5696	-486.5798	
	v37e	-474.9216	-474.9038	-474.9143	
	v37p	-428.2354	-428.2176	-428.2282	
	v43e	-124.3674	-124.3347	-124.3544	
	quarterly				
	v10b	-209.9952	-209.9612	-209.9819	
	v25b	-71.1007	-71.0597	-71.0856	
	v27b	-64.7831	-64.7421	-64.768	

Akaike information criteria
Schwarz criterion or Bayesian information criterion
Hannan-Quinn criterion

COMPOSITE

	AIC	SIC	HQC
raw	-3938.2663	-3938.0887	-3938.1971
adjusted	-3863.5138	-3863.3362	-3863.4446

AIC	Akaike information criteria
SIC	Schwarz criterion or Bayesian information criterion
HQC	Hannan-Quinn criterion

COMPOSITE		AIC	SIC	HQC	
	raw				
	v1	-358.2062	-358.1892	-358.1993	v1
		-256.5865	-256.5694	-256.5796	
		-476.0159	-475.9988	-476.009	
		-225.2675	-225.2497	-225.2603	
		-396.5351	-396.5181	-396.5282	
		-419.6729	-419.6559	-419.666	
		-281.4471	-281.4301	-281.4402	
		-449.2374	-449.2203	-449.2305	
		-529.8743	-529.8573	-529.8674	
		-546.4626	-546.4455	-546.4557	
		-400.5203	-400.5033	-400.5134	
		-512.0821	-512.0651	-512.0752	
	v2	-476.0159	-475.9988	-476.009	v2
		-358.2062	-358.1892	-358.1993	
		-256.5865	-256.5694	-256.5796	
	v12	-225.2675	-225.2497	-225.2603	v12
		-396.5351	-396.5181	-396.5282	
		-419.6729	-419.6559	-419.666	
		-281.4471	-281.4301	-281.4402	
	v28	-449.2374	-449.2203	-449.2305	v28
		-529.8743	-529.8573	-529.8674	
	v33	-508.9507	-508.9329	-508.9435	v33
		-373.8325	-373.8146	-373.8252	
		-485.1502	-485.1324	-485.143	
	v39	-166.6995	-166.6668	-166.6865	v39
		-196.79	-196.7573	-196.777	
		-182.8637	-182.8311	-182.8508	
	adjusted				
	v1	-358.2062	-358.1892	-358.1993	
		-256.5865	-256.5694	-256.5796	
		-473.4402	-473.4231	-473.4333	
		-247.1958	-247.1775	-247.1883	
		-360.8181	-360.8011	-360.8112	
		-407.2603	-407.2432	-407.2534	
		-259.5352	-259.5181	-259.5283	
		-443.5743	-443.5573	-443.5674	
		-529.8743	-529.8573	-529.8674	

AIC	Akaike information criteria
SIC	Schwarz criterion or Bayesian information criterion
HQC	Hannan-Quinn criterion

Economic Sentiment Indicator

Industrial Confidence Indicator (v7+v4-v6)/3

Consumer Confidence Indicator (v14+v16-v19+v23)/4

Construction Confidence Indicator (v30+v31)/2

Retail Trade Confidence Indicator (v34-v35+v37)/3

Services Confidence Indicator (v40+v41+v42)/3

COMPOSITE		AIC	SIC	HQC
	raw			
	v1	-4235.1686	-4234.9643	-4235.0858
	v2	-1100.2962	-1100.2404	-1100.2736
		-1162.5494	-1162.4939	-1162.5269
		-1278.3017	-1278.2472	-1278.2796
	v12	-1229.053	-1228.9986	-1229.0309
		-887.3439	-887.2885	-887.3214
		-1028.621	-1028.5652	-1028.5983
		-1433.4271	-1433.3756	-1433.4062
	v28	-887.0438	-886.988	-887.0212
		-801.3756	-801.3209	-801.3534
	v33	-674.325	-674.2663	-674.3012
		-756.8375	-756.7789	-756.8137
		-639.7203	-639.6617	-639.6965
	v39	-407.1731	-407.0775	-407.1351
		-417.0883	-416.9927	-417.0503
		-437.4502	-437.3531	-437.4117
	adjusted			
	v1	-4154.8225	-4154.6181	-4154.7397

AIC	Akaike information criteria
SIC	Schwarz criterion or Bayesian information criterion
HQC	Hannan-Quinn criterion

INDICATORS		AIC	SIC	HQC
	raw			
	monthly			
	v3	-726.6391	-726.605	-726.6253
	v9	-479.1296	-479.0927	-479.1146
	v16	-487.9228	-487.8888	-487.909
	v29	-999.5321	-999.4981	-999.5183
	v36	-844.0026	-843.9684	-843.9887
	quarterly			
	v10	-285.8194	-285.7514	-285.7927
	v11	-238.1692	-238.1012	-238.1425

AIC Akaike information criteria
SIC Schwarz criterion or Bayesian information criterion
HQC Hannan-Quinn criterion

INDICATORS		AIC	SIC	HQC
	raw monthly			
	v4	-1162.5494	-1162.4939	-1162.5269
	v8	-1184.0465	-1183.9907	-1184.0238
	v16	-715.7464	-715.6953	-715.7257
	v32	-1152.9515	-1152.8957	-1152.9289
	v34	-685.9522	-685.8964	-685.9295
	v38	-755.4727	-755.4163	-755.4498

AIC Akaike information criteria
SIC Schwarz criterion or Bayesian information criterion
HQC Hannan-Quinn criterion