TWO DISCUSSIONS ON NEW SEASONAL ADJUSTMENT METHODS

Agustin Maravall

Banco de España - Servicio de Estudios
Documento de Trabajo nº 9704

TWO DISCUSSIONS ON NEW SEASONAL ADJUSTMENT METHODS

Agustín Maravall

Banco de España - Servicio de Estudios
Documento de Trabajo nº 9704

In publishing this series the Banco de España seeks to disseminate studies of interest that will help acquaint readers better with the Spanish economy.

The analyses, opinions and findings of these papers represent the views of their authors; they are not necessarily those of the Banco de España.

The Banco de España is disseminating some of its principal reports via INTERNET and INFOVÍA.

The respective WWW server addresses are: http://www.bde.es and http://www.bde.inf.

ISSN: 0213-2710
ISBN: 84-7793-533-5
Depósito legal: M. 6662-1997
Imprenta del Banco de España

ABSTRACT

This document contains two discussions having to do with two new seasonal adjustment methods that are competing at present for possible replacement of the method X11 / X11ARIMA in European data producing agencies (mostly national statistical offices and central banks). The first discussion offers a critical review of X12ARIMA, the new U.S. Bureau of the Census method, still heavily based on the X11 approach. The second discussion is a reply to a critique made to SEATS, the alternative method, which represents a fairly drastic methodological change. The discussion clarifies some methodological points and deals with issues related to practical application of the program.

CONTENTS

COMMENTS ON X12ARIMA

1. GENERAL REMARK
2. FLEXIBILITY, IDEMPOTENCY, AND SPURIOUS RESULTS
3. TESTING AND INTERPRETATION OF DIAGNOSTICS
4. CONCLUDING REMARK

PRACTICAL EXPERIENCES WITH SEATS

1. SUMMARY OF METHODOLOGY
2. PRACTICAL EXPERIENCES WITH SEATS
3. FINAL REMARKS

COMMENTS ON X12ARIMA

1. GENERAL REMARK

The program X12ARIMA contains significant improvements over X11/X11ARIMA. Broadly, they can be classified into 3 groups:

(1) the development of REGARIMA;
(2) some new options for the filters;
(3) more diagnostics.

Clearly, (1) implies the belief that REGARIMA models are appropriate for time series. Group (2) implies an increased filter flexibility, where the filter selection depends more and more on the data. These improvements obviously represent a move towards a "model-based" (MB) approach, which is also implicit in Fidley et al. (1997): X12A "decomposes a time series into a product of (estimates) of a trend component, ...". The distinction between a "theoretical" component and its estimator forms the basis of a MB approach. (Still, how is the theoretical component? What is the estimation criterion?) Perhaps the reasons for not moving all the way to a MB method are the old fears having to do with the need for experts and with computing time, and, of course, the power of inertia. I believe that the fears are not anymore appropriate since, for example, programs TRAMO and SEATS are fully model-based, fully automatic, and faster than X12A. In the comment, I will not deal with REGARIMA. It is a good program, although a bit too slow and in need of a proper automatic model identification (AMI) and automatic outlier correction (AOC) procedure. My comments will center on flexibility and diagnosis; I will use as examples basic macroeconomic Spanish series.

2. FLEXIBILITY: IDEMPOTENCY AND SPURIOUS RESULTS

In the difficult field of finding "objective" criteria for comparison of seasonal adjustment methods, there are two that seem unquestionably desirable. One is idempotency, i.e., a seasonal adjustment method applied to the seasonally adjusted (SA) series it has produced should leave the SA series unchanged. The second requirement in that, when applied to white noise, the method should produce no spurious seasonality. Both properties show, of course, how flexible a filter is to adapt to the particular structure of a series.

With a reasonable AMI procedure, an MB approach would identify nonseasonal models in both cases. But let us move one step backwards and compare X12A to the simplest MB procedure, whereby only a default model is considered, namely the so-called Airline Model,

$$\nabla \nabla_{12}\, x(t) = (1 + \theta B)\,(1 + \Theta B^{12})a(t) + \mu \quad ; \qquad [1]$$

where $\nabla = 1-B$, $\nabla_{12} = 1-B^{12}$, B is the lag operator, and a(t) is a white-noise variable.

Thus X12A run by default is compared with a procedure that consists of fitting the default model, and using this model to obtain the filters for the component estimator (equivalent to running SEATS by default).

For the Consumer Price Index (CPI), Figure 1 compares the seasonal component of the original series with those obtained for the SA series, using X12A both times. The seasonality in the SA series, although not large, is nevertheless disturbing. Figure 2 performs the same comparison for the MB procedure we consider (the default model yields a good fit). To all purposes, the estimated seasonal component in the SA series is now zero. (Notice also how the seasonality estimated with the MB filter is noticeably more stable.) Figure 3 compares the seasonality estimated in the SA series by both methods.

As for spurious seasonality, I generated 50 white noise (0,1) independent series and applied the two methods. Figure 5 compares the 50 variances of the estimated $\hat{s}(t)$ series (the straight lines are the mean values: .211 for X12A and .101 for the MB method). For all 50 cases the variance of the spurious seasonal component for X12A is considerably larger than that for the MB case.

The flexibility of the MB approach is explained by a simple feature. When model [1] is fit to the SA series, Θ converges to -1 fast. Stopping its value at, say, -.98 (or -.99) the seasonal structure of the model in practice cancels out. Notice that, if deterministic seasonality were to be present, it would have been well captured, as shown in Figure 4, which magnifies the MB estimator of $\hat{s}(t)$. In a similar way, when model [1] is fit to white noise, both Θ and θ tend to -1, so that the regular unit root in practice also disappears, and the model reproduces well white noise. As a consequence, besides the good idempotency properties, in the MB case there is no need to worry about whether seasonality or trend are present; moreover, there is no need to worry about whether there may be deterministic or stochastic seasonality present: the model will handle it.

Therefore, the default model seems flexible enough to encompass a wide variety of simpler cases. But the robustness of the results also extends to larger models. Figure 6 exhibits the seasonal factors obtained from application of the default model and of the model

$$\nabla^2 \nabla_{12} x(t) = (1+ \theta_1 B+\theta_2 B^2)(1+\Theta B^{12}) a(t) \qquad [2]$$

to the ALP monetary aggregate. Since the default model provides an acceptable fit, model [2] may be seen as the result of overdifferencing. The two sets of seasonal factors are indistinguishable and, again, this is due to the fact that, in estimation, one of the roots of the regular MA polynomical goes to -1, and the effect of overdifferencing is cancelled out. Moderate overdifferencing causes, in fact, little damage.

3. TESTING AND INTERPRETATION OF DIAGNOSTICS

The mixing in X12A of MB and non-MB feature leads to some confusion. I shall illustrate this with some examples.

a) Testing for Trading Day (TD) and Easter Effects (EE)

Findley **et al.** (1997) state that the estimator of the irregular, "being an almost uncorrelated series", can be used for TD and EE estimation using simply OLS. Uncorrelatedness characterizes the component, but certainly not the estimator, which can be strongly correlated (see Maravall, 1995). Moreover, the spectrum of the irregular and of the SA series are used for detection of TD effects. Bur the estimator of the irregular (and of the SA series) is a noninvertible series and hence the use of finite AR models to estimate its spectrum makes little sense. In this respect, thus, I agree with the authors: REGARIMA seems the proper framework for TD and EE estimation. In other words, the MB way seems preferable to the old X11 spirit.

b) Sliding Spans and Revision Histories Diagnostics

These are new diagnostics. The first one basically consists of running succesively the program on overlapping subspans of the series, say

$$[(x(t-k), ,..., x(t)], [x(t-k+1), ..., x(t+1)] ,..., [x(t-k+h), ,..., x(t+h)] ,$$

and look at the variation of s(t|t+j) (the estimator of s(t) obtained with the subspan finishing at t+j), for j=0,1,...h. If that variation exceeds a limit k (recommended value of .03), month t is "unreliable"; if the percent of unreliable months is larger than 25%, then the series should not be adjusted (the variability is "much too high"). This type of diagnostics, in my view, tend to mix what should be the characteristics of a good extraction method with the analyst wishful thinking concerning the properties of a series.

Consider the decomposition x(t) = s(t) + n(t), where s(t) and n(t) denote the (orthogonals) seasonal component and SA series, respectively. Assume s(t) is generated by a perfectly acceptable seasonal component model of the type

$$(1+B+...+B^{11})\ s(t) = \theta_s(B)\ as(t), \qquad [3]$$

where as(t) is white noise with variance V(s), and the r.h.s. is a finite variance moving average (MA). Further, assume n(t) follows some ARIMA model. Let us fix the models for s(t) and n(t), except for V(s), which is systematically increased.

The new seasonal components obtained are all in exactly the same way reasonable seasonal components. When the same n(t) is added, a new series is obtained. For V(s) beyond some **a priori** fixed limit, why shouldn't the series be adjusted if, of course, an appropriate method is available? (The diagnostic reminds me of the one in X11A, whereby if the average forecasting error was greater than 12% the series should not be forecasted. If stock prices could be forecasted with 12% error, wouldn't it be insane for investors not to forecast?) The variability that should be acceptable is the variability the series displays. This dependence is even implicit in the paper when we are told that the recommended value of k will be too large for series with small seasonality, and too small for series with large seasonal movements.

The authors recommend that the span length be at least as large as the length of the filter, and that one should look at the variability starting with concurrent estimation. Since the filter can then be completed in one direction, the variability in the successive estimators of s(t) is, in essence, the total revision in the concurrent estimator. The MB approach permits us to address the variability issue in a more elegant and efficient manner (it is worth pointing out that the MB method yields a more direct measure of the uncertainty in the measurement of s(t), namely the standard error of the estimator.)

Let sf(t) and sc(t) denote the final estimator of s(t), obtained with the complete filter, and the concurrent estimator, $s(t|t)$, respectively. The MB equivalent of the condition for not adjusting the series becomes.

$$\text{Prob. } [\,|sf(t)-sc(t)|>k\,]>.25 \qquad [4]$$

or, letting d(t)=sf(t)-sc(t) denote the revision in the concurrent estimator and assuming k=.03, P(|d(t)|>.03)>.025. It is straightforward to express d(t) as a linear combination of a(t+j), with j>0 (see Maravall, 1995), from which the distribution of d(t) can be easily obtained.

Using as examples the total exports and imports series, for which the default model [1] provides also good fits, the probability in the l.h.s. of [4] was computed. The following table is obtained:

	Exports	Imports
Prob. that month t is unreliable	41%	33%
Frequency of unreliable months (sliding spans diagnostic)	38%	40%

Thus X12A would conclude that none of the 2 series should be adjusted. Figures 7 and 8 present the MB estimates of the two series of seasonal components. They display a "moving" structure, but can be estimated nicely; moreover, as shown in Figure 9, they can also be forecasted well.

In conclusion, the sliding span exercise seems a useful tool, not as a diagnostic on whether or not to adjust, but, in so far as stability is desirable, as an indication of whether direct or indirect adjustment of an aggregate may be preferable, or in helping in the selection of the filter length. Incidentally, in the MB approach, the sliding spans could serve as a diagnostic: if the frequency of unreliable months is markedly different from the probability that a month is unreliable, then something wrong could be suspected. (For the exports and imports series, both computations seem quite in agreement).

A similar reflection applies to the revision histories diagnostic. Given the stochastic structure of the series, there is an optimal revision process, and departures from it may be costly in terms of MSE. I agree with the authors in that if may be preferable to use trends if revisions are smaller, yet the statement needs qualification. Two features of revisions are important:

· the size of the revision;
· the length of the revision period;

and in general a smoother component does not decrease both. For the exports and imports series, the following table gives the variance of the revision in the concurrent estimator of the trend and of the SA series:

	Exports	Imports
Trend	.088	.093
SA series	.077	.061

(the variance are standarized by setting V(a)=1). It is seen that the trend revision is larger in both cases. The next table shows the percent reduction in the variance of the revision after 1 year of additional data has become available:

	Exports	Imports
Trend	92	91
SA series	31	22

It is seen that the trend estimator converges much faster to the final estimator. This trade-off between size and duration is often found in practice, and it is difficult to say what is preferable: a moderate revision that takes many years to be completed, or a large one that is removed in a few periods. Be that as it may, the imprecise criterion of having somewhat small revisions should be replaced by that of having optimal revisions (in the sense that MMSE estimators are obtained.)

4. CONCLUDING REMARK

I congratule Professor Findley and his team (and the U.S. Bureau of the Census) for the good and important work put into their new X12ARIMA program. Further, I hope they continue in the same direction (i.e. the MB direction). Then I think it is likely that X13 will be an ARIMA MB method.

Fig.1 IDEMPOTENCY:X12A
Seasonal Component

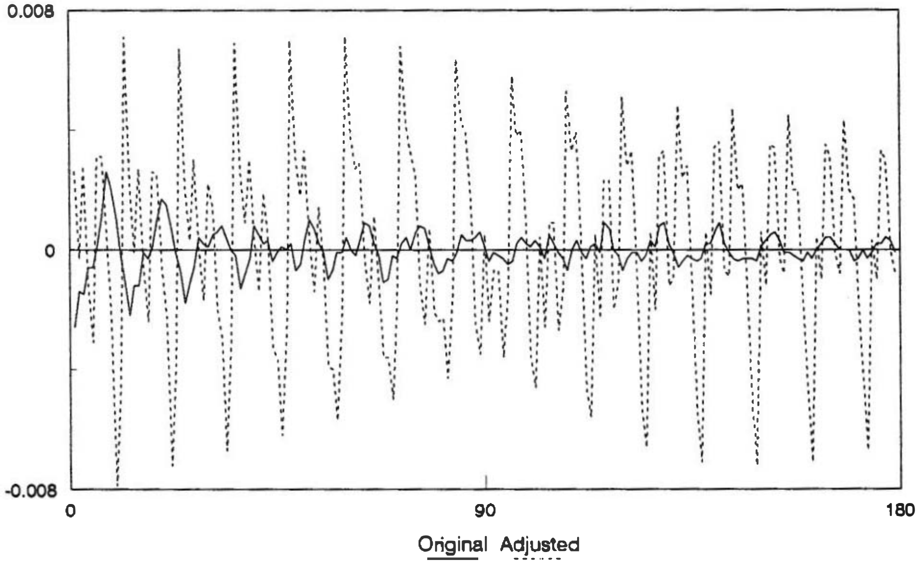


Fig.2 IDEMPOTENCY:SEATS
Seasonal Component

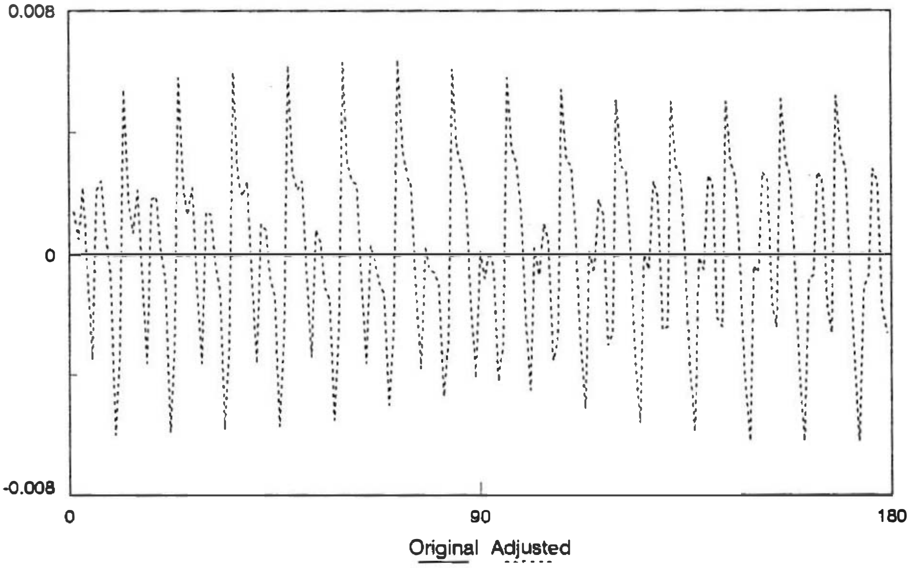


Fig.3 SEASONALITY IN SA SERIES
Seasonal Component

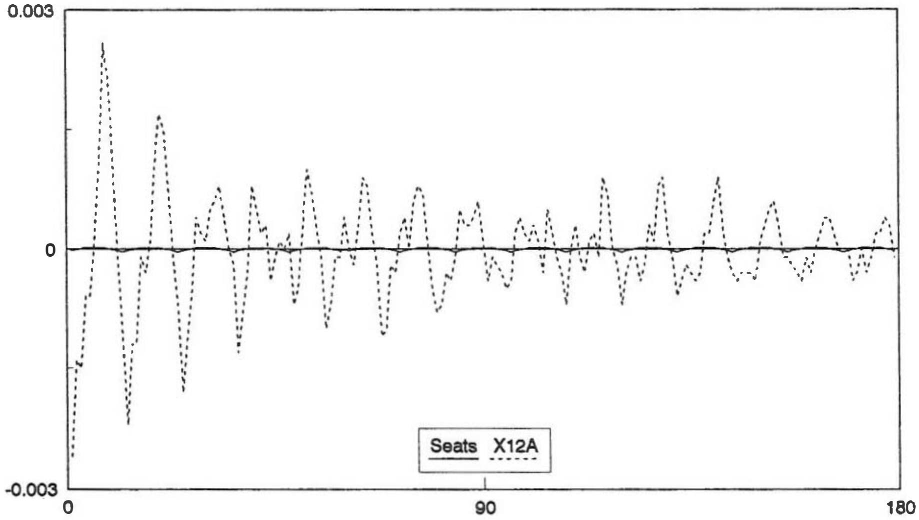


Fig.4 STABLE SEASONALITY:SEATS
Seasonal Component

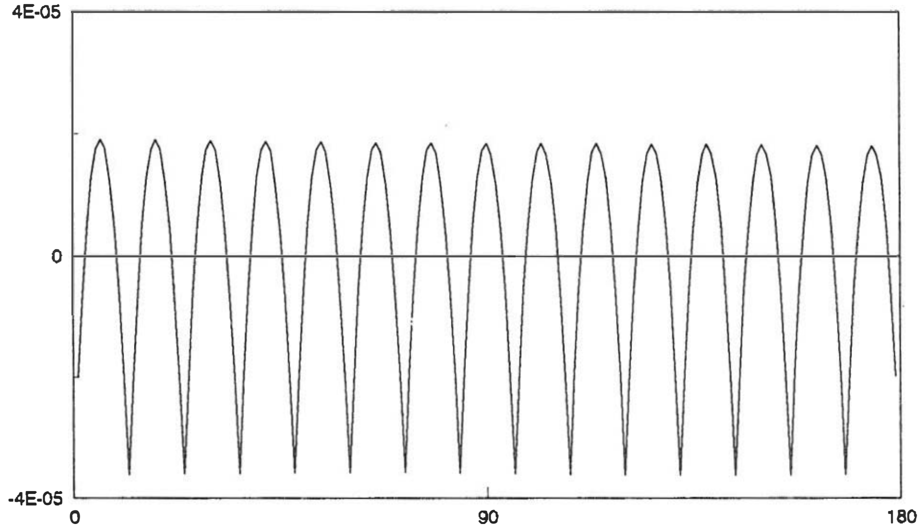


FIG.5 SPURIOUS SEASONALITY
Variance [s(t)]

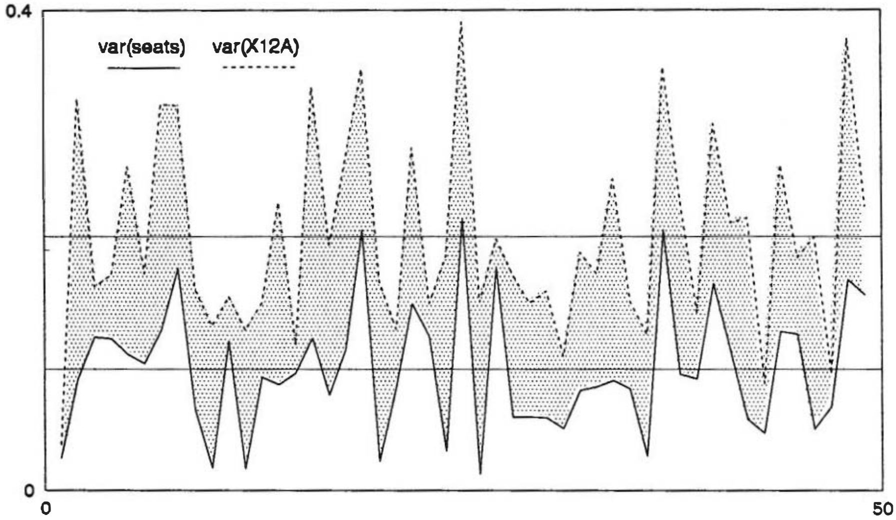


Fig.6 SEASONALITY AND OVERDIFFERENCING
Seasonal Factors

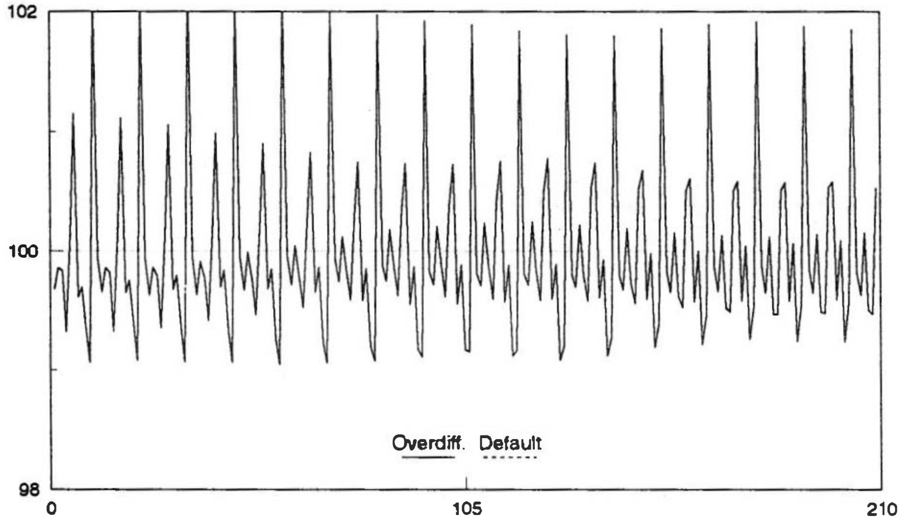


Fig.7 SEASONAL COMPONENT:EXPORTS

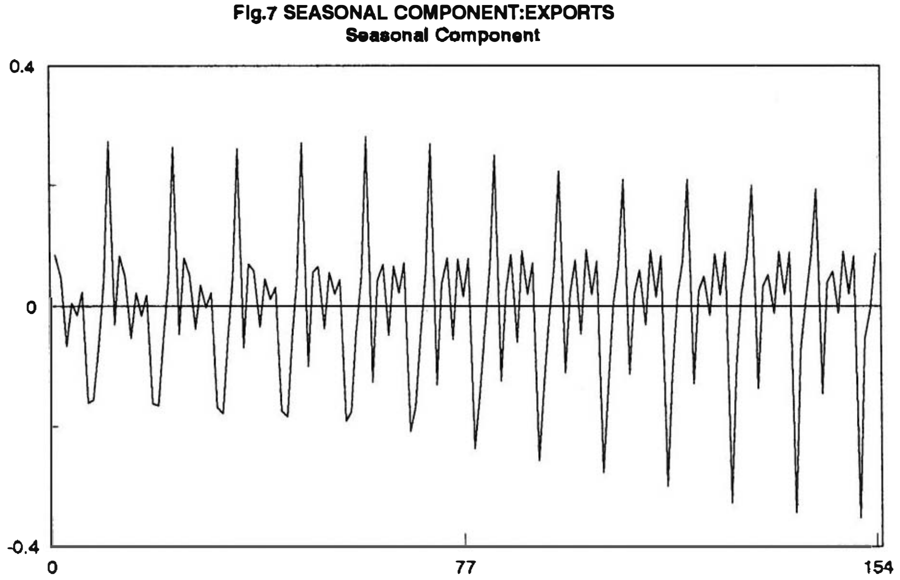
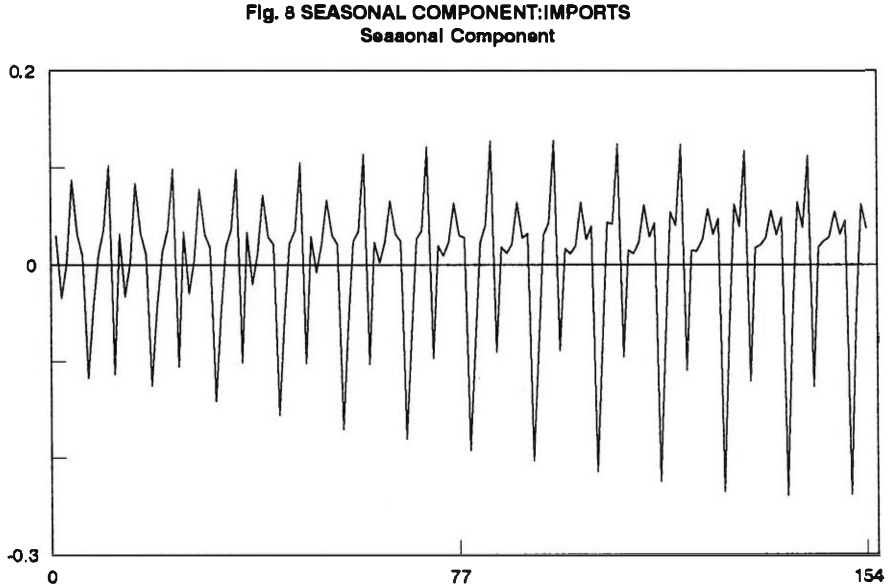


Fig. 8 SEASONAL COMPONENT:IMPORTS



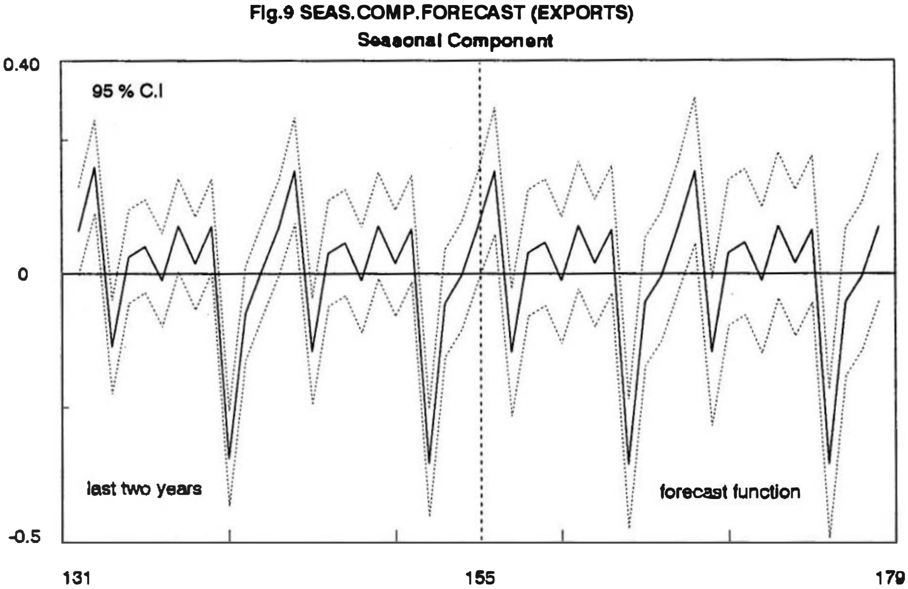


Fig.9 SEAS.COMP.FORECAST (EXPORTS)
Seasonal Component
0.40
0
-0.5
95 % C.I
last two years
forecast function
131
155
179

REFERENCES

FINDLEY, D.F., MONSELL, B., BELL, W.R., OTTO, M.C., CHEN, S. (1997), "New Capabilities and Methods of the X12ARIMA Seasonal Adjustment Program", **Journal of Business and Economics Statistics**, (forthcoming).

MARAVALL, A. (1995), "Unobserved Components in Economic Time Series", in Pesaran, H. and Wickens, M. (eds.), **The Handbook of Applied Econometrics**, vol. 1, Oxford: Basil Blackwell.

PRACTICAL EXPERIENCES WITH SEATS

In Stier (1996) the so-called ARIMA model-based (AMB) seasonal adjustment method as enforced in the program SEATS (Gómez and Maravall, 1996) is summarized and some "practical experiences" are reported. As Stier points out, the fact that its use has been recommended by Eurostat justifies a critical review. (The recommendation is the result of a detailed study and of three years of use, with careful monitoring, on many thousand series every month; see Eurostat, 1996a.) Unfortunately, Stier's review is plagued with methodological errors and with serious mistakes concerning correct use of the programs.

1. SUMMARY OF METHODOLOGY

Most of the paper is simply a summary of Maravall and Pierce (1987) and Maravall (1988). The summary is a direct transcription, with a few added comments which happen to be for the most part incorrect. For example:

a) Concerning the properties of the theoretical components, on page 315, we are told that the components provided by the AMB method have to be invertible. This is not true and, in fact, the canonical components are noninvertible.

b) In page 316 we are told that the covariance equations yield the restriction $\beta+\gamma=0$ and hence $\beta\leq 0$. Both restrictions are wrong. For example, the model with $\beta=.1$, $\gamma=.7$, $V_b=.039$, $V_c=.086$ and $V_I=.119$, aggregates into the "prototypical" model discussed by Stier (with $V_a=1$). It satisfies the covariance equations and represents a perfectly admissible decomposition. Yet $\beta+\gamma>0$ and $\beta>0$! This error reflects a misunderstanding of the identification problem and yields a totally distorted parameter space.

c) The next error (page 319) is a consequence of Stier's confusion between the theoretical component and its estimator. Stier presents as a weakness of the method the fact that the estimated irregular component has a spectral peak between the zero and the seasonal frequency, so that "the estimated irregular is no realization of white noise". But it should not be! This is not a property of SEATS; it is a property of almost any filter that removes nonstationary trend and seasonal components, and a well known feature (see, for example, Nerlove 1964). For the CENSUS method it is discussed in Cleveland and Tiao (1976), and for the AMB case in Maravall (1987). A similar effect appears when, for example, the Hodrick-Prescott filter is used to detrend a series seasonally adjusted with CENSUS (Maravall, 1995). An heuristic explanation is the following. The fact that, for the zero frequency, the irregular-to-trend ratio is zero implies that this frequency is not used for estimation of the irregular, which implies a spectral zero for the zero frequency. For the seasonal frequency, the irregular-to-seasonal ratio is again zero and, in a similar way, a spectral zero will show up. Since the spectrum is strictly positive between these two zeros, a peak should appear in between. A very simple example with deterministic seasonality may be illustrative. Consider the twice-a-year observed series

$$x(t) = s(t) + u(t) , \quad [1]$$

where s(t) is a deterministic seasonal component satisfying

$$s(t) + s(t-1) = 0 , \quad [2]$$

and u(t) is a white noise irregular. The seasonally adjusted series is equal to the irregular, and a standard filter to compute it would be, for example, (Kendall, 1976)

$$\hat{u}(t) = [x(t-1) + 2x(t) + x(t+1)] \,/\, 4$$

which, using [1] and [2] becomes $\hat{u}(t) = (1/4)\ [(1+B)^2\ u(t+1)]$, that is, a moving average with a spectral zero at the seasonal frequency (root B = -1). Proceeding in a similar way for the case of a deterministic trend, the spectral zero at the zero frequency is easily derived. Putting the two together, the peak of Stier is obtained.

Perhaps what may cause Stier's confusion is the apparent paradox that, since the estimator of the irregular displays autocorrelation, this may seem to imply forecastability, which would be incompatible with a theoretically white-noise irregular. A closer look at the estimator of the irregular shows that it is a linear filter of present and future innovations; in his example, $\hat{u}(t) = V_u\,[a(t) - a(t+2)]$, and hence any forecast $E_t\,\hat{u}(t+K)$ is always zero; thus there is no paradox.

d) In the example used in page 320, the canonical decomposition maximizes the variance of the trend revision, and it is said immediately after that the analysis can be generalized for any ARIMA process. This is somewhat misleading and, as shown in Maravall and Planas (1996), the revision of the trend may also be minimized for a canonical trend; it depends on the structure of the series.

e) Then, there are several minor mistakes. For example, in page 317, the discussion of the identification criteria is confusing. The structural component model of Harvey (1989) -solution (b) of Stier- reaches identification precisely through solution (a), i.,e., by restricting the orders of the MA polynomials in the models for the components. The distinction between (a) and (b) is thus misleading. In p. 322 Stier states "Identification is again reached via minimization of V_u"; on the contrary, it is reached via maximation of V_u. On p. 323, the expression for the Wiener-Kolmogrov filter is wrong: $\alpha(B)\alpha(F)$ should be replaced by $\beta(B)\beta(F)$,... and so on.

But Stier does not critizice the methodology as such; he repeatedly mentions the "theoretical elegance" of the approach. The problem according to him lies in its application. As we shall see, it is in this context that the most important mistake is made.

2. PRACTICAL EXPERIENCES WITH SEATS

The main mistake of Stier's discussion is not due to the fact that the programs may be complex and require careful use. Rather the contrary. In what follows we shall stick to the use of TRAMO and SEATS in their strictly and fully automatic manner, without any need for specialized skills. Stier has mixed up the "default model" with "automatic use". Nowhere in the user instructions (Gómez and Maravall, 1996) we "recommend the default model (i.e., the Airline Model of Box and Jenkins, 1970) for Routine Seasonal Adjustment" or for large-scale use in general. These cases are explained in a special section (Section 3: Routine Use on Many Series), and only one parameter is involved (the parameter RSA). In all cases a procedure for an automatic model identification (AMI) and automatic outlier detection (AOD) is involved. We fully agree with Stier in that the Airline Model is absurd for many series. A recent study by Eurostat (1996b) on 13,227 series (indicators for all activities, and for the 15 EU member states, USA and Japan) showed that approximately for 50% of them the default model was adequate. This proportion justifies using the Airline Model as the default one, but never as the only model!

We proceed now to review the 4 examples presented.

a) Stier claims that for one simulation of an ARIMA (1,1,1) model, the default model of SEATS estimates some seasonality. With just one simulated example, the claim is preposterous. Among say, 100 random samples one is bound to find some displaying seasonal autocorrelation (r_{12}) that appears to be significant (from the tails of the distribution of $\hat{r}_{12}$ under H_o:$r_{12}=0$; see Box and Pierce, 1970). In these cases, SEATS would capture some seasonality since, to start with, it is in the data. For Stier's example, $\hat{r}_{12}=.15$, not far thus from a borderline case. In any event, the (seasonal) Airline Model will naturally always produce some number for the estimator $\hat{s}(t)$ (the same would be true for OLS with seasonal dummies or for X11). The point is: how large this spurious seasonality is likely to be? In a more rigurous context and using many more series, Fischer (1995) looked at the problem and found that SEATS clearly outperformed other methods (a similar result, when X12 ARIMA is the alternative, is reported in Maravall, 1997). When there is no significant seasonality, even if one estimates the Airline model, a seasonal IMA (1,1) structure with θ_{12} close to -1 is typically obtained, and the two seasonal factors ∇_{12} and $(1+\theta_{12}B^{12})$ in practice cancel out and the seasonality obtained is, to all effects, negligible.

Be that as it may, as mentioned earlier, the correct way to proceed is to use SEATS after TRAMO. Simply setting RSA = 3 or 4 (perhaps the most frequently used fully automatic procedures) the program performs AMI and AOD, as well as several pretests. Proceeding in this way with the example of Stier, TRAMO informs us that the series should be modelled in the levels, that there are no outliers (nor trading day or easter effects), and that the orders of the seasonal polynomials in the model are (0,0,0); as a consequence, the seasonality estimated by SEATS is zero for all periods. NO spurious seasonality is estimated.

As for the regular part, TRAMO identifies a (0,1,1)-model instead of the (1,1,1) used to generate the series. Direct estimation of this latter model yields $\hat{\phi}=-.70$, $\hat{\theta}=-.50$, but the correlation between the two parameter estimates is .95! Thus it makes sense that a parsimonious AMI procedure reduces the parameters to only one. Incidentally, a look at the diagnostics of the model identified by TRAMO shows only one problem: lack of randomness of the residual autocorrelation, due to a sequence of (small) positive autocorrelations for low lags. To the careful analyst this may suggest the (1,1,1) specification.

Back to the main point of spurious seasonality, to better assess the performance of TRAMO-SEATS we simulated 300 N(0,1) random samples and, from these, series x(t), given by

$$(1-8B)\nabla x(t)=(1-.4B)a(t)$$

were generated. In 86% of the cases, SEATS detected that the model had no seasonality whatsoever. In 9% of the cases it estimated a negligible (not significant) one, with an Airline model for which θ_{12} approached -1. (As an example, Figure 1 displays the original series and the adjusted series for a case with $\theta_{12}=-.84$). Only in 5% of the cases there was some significant seasonality (in perfect agreement with a standard test size). Figure 2 exhibits the case with maximum variance of the spurious seasonal component; seasonal adjustment amounts to little more than a mild smoothing of the series. It has little influence on inference, as evidenced in Figure 3, where it is seen that it does not affect the forecasts. In what concerns spurious seasonality, TRAMO-SEATS perform quite reliably. We encourage Stier to check this on more series and in comparison with other procedures, such as X12ARIMA or Berlin BV4.

It should be mentioned that, being conceived mostly as a seasonal adjustment program, SEATS is biased towards finding models that provide stable estimators of the seasonal component. Thus, for example, if the nonseasonally differenced series displays $r_{12} = .4$, and perhaps an insignificant $r_{24} = .2$, instead of the seasonal MA(1) structure, which decomposes poorly, a seasonal IMA(1,1) structure, which provides better estimators of the seasonal, is used. As mentioned before, as the seasonality goes to zero, θ_{12} goes to -1 and no damage is done. Besides, in exactly the same way, purely deterministic seasonality would also be well captured when present.

b) Next Stier considers the series "New orders, manufacturing industry" and criticizes that the filters for the trend in SEATS are different when 48, 72 and 100 observations are used. Given that the trend filter depends on the model, and that the model parameter estimates depend on the sample size (T), the filter depends on T. The use of any model would produce that result, and that would affect not only trend estimation (even when, for example, OLS are used), but also forecasting, simulation or whatever use is made ot the model; the distribution of the estimators are functions of T. From a general point of view, it is sensible that the amount of information affects our knowledge (otherwise, why do we use data?) Of course, revising estimators causes inconveniences but, ultimately, it is simply the inconvenience of learning. Stier states that "the definition of the components of the series varies with the lenght of the series". But the model does not really change; what changes are the estimators.

The same series is used to illustrate how the use of TRAMO is of little help, but again, the argument is flawed from the beginning. No matter which automatic option is chosen, the Airline Model is always rejected, and TRAMO selects a (2,1,0) (0,1,1) model. Stier reports a different model, which apparently yields an abnormally erratic trend. We ignore where his results come from, since we have not been able to duplicate them. Given that he used a very old version of the programs, perhaps he ran into an old bug. This is regrettable, since updates of the programs, were periodically placed in the Internet site given in the user instructions. Criticism of the beta version of a program based on (at best) an old bug that has been already removed is of no help.

Using, as before, the automatic procedure of TRAMO-SEATS, one obtains the trend of Figure 4 (very different from Figure 10 of Stier), a rather reasonable trend for a series with a relatively fast-changing slope. For this series, however, TRAMO detects, at the pretesting stage, significant Trading Day and Easter effects, ignored by Stier. Incorporating them (still within a purely automatic use of the programs) yields a similar, perhaps slightly smoother, trend.

A remark on trends seems worth making, since it is with trends that Stier's paper is mostly concerned, although nowhere in the paper is there a definition of a trend. What Stier seems to have in mind is a very smooth trend, obtained with a filter with a very narrow peak around the zero frequency. These are the trends typically used in "economic business-cycle analysis", which is quite a different issue than seasonal adjustment. In the latter, the trend is a "short-term" trend. In X11 it has some power over cyclical frequencies, and in STAMP or in SEATS the same may happen. This is due to the fact that, in both approaches, the width of the peak around zero of the trend filter adapts to the width of the peak around zero in the spectrum of the actual series, and, on occasion, the peak is wide. This type of trend is often denoted trend-cycle, to emphasize precisely that short-term character. (Section 2.4.6 of the SEATS Manual clarifies this point.)

Incidentally, the use of fixed trend filters with a narrow pass band is not free from many of Stier's criticims. For example, applying his low pass filter to his deterministic example (Figure 12 in his paper), the nonsense trend component of Figure 5 is obtained. Also, if properly centered, those types of filters induce revisions that converge very slowly. Furthermore, they induce a strong phase effect in the trend concurrent estimator.

c) We are next told that "a fundamental condition with SEATS is that the series are integrated". This, again, is false, and here I cannot imagine the source of the error. Stier can try to decompose the model $(1-.7B)x(t) = a(t)$, to mention the simplest example. A different issue is the fact that highly stationary trends or seasonals make little sense, since they tend to converge fast to a constant or to zero.

Stier generates a purely deterministic series and correctly points out that the default model of SEATS does not exactly reproduce his deterministic components. Given that SEATS assumes the series to be stochastic, this is not really surprising. But, as seen in Figure 6, it is nevertheless remarkable how stable the trend estimated by the simple default model turns out to be (in fact, both parameters θ_1 and θ_{12} become close to -1). The figure does not quite reproduce Stier's Figure 13 and the same holds true for the seasonal component, considerably more stable than the one in Figure 15 of Stier. Ultimately, since for a long time it has been known that quadratic deterministic trends and deterministic sine functions are not adequate for time series found in reality, the criticism seems of little relevance. In any case, the relatively small seasonal component and the dominance of the irregular in the series created by Stier causes problems for X11, but not for SEATS

d) The last example considered by Stier is the Airline Passenger series of Box and Jenkins (1970). He questions the appropriateness of the default model of TRAMO-SEATS, namely the $(0,1,1)-(0,1,1)_{12}$ model, points out that the program AUTOBOX choses "the simpler" $(1,0,1)-(0,1,0)_{12}$ -plus- mean model, and that this last model produces a smoother trend. We fail to see why the model is simpler, since it implies estimation of one more parameter (3 versus 2). More relevantly, estimation of the second model yields unacceptable residuals: the Ljung-Box Q statistics with 24 autocorrelations is 23.98 for the default model and 47.34 for the second model (In this example we use the default model and no automatic procedure.) Besides, the residual variance for the default model is .037 and increases to .042 for the second model. Further, contrary to the case of the default model, residuals display excess kurtosis and fail several Normality tests. Moreover, an out-of-sample forecast test for the default model yields the value $F(24,107) = .82$. In summary, while the default model fits well the data, the model suggested by Stier is clearly rejected. Of course, an I(2) trend is more flexible than an I(1) trend, still, Figure 7 compares the trends estimated with the two models: Does the difference justifies the use of a bad model? Why not just use a straight line? (By the way, as TRAMO detects, the series contains a significant trading day effect and some possible outlier(s). Incorporating these effects also leads to an Airline-type model.)

Using the same example, Stier performs a final comparison. Letting $p(t|T)$ denote the estimator of the trend for period t when the last observation is x(T), he compares the series of 25 concurrent estimators p(t|t) with the heterogenous series of preliminary estimators p(t|144) for t = 120, 121, ..., 144. What that comparison means is not clear to me. SEATS produces a large revision in the concurrent estimator of the trend around period t=135, which, as TRAMO finds a few periods later, is associated with an outlier. TRAMO automatically corrects the series and the drop in the trend disappears. What the example shows is that outliers do create problems, and this is an important reason for using TRAMO as a preadjustment program. In general, the behavior of SEATS with respect to revisions in comparison with other methods has been analyzed with some care in Fischer (1995) and Balchin (1995), and SEATS was found to perform well. For example, Balchin finds "small to medium" revisions for SEATS (and nonconvergent revisions for STAMP.) Still, the criteria of minimum revisions is ambigous and should be interpreted with care. Whether it causes inconveniences or not, revisions should depend on the correlations between the future and the present or the past. In some cases the future will provide a much better understanding of the present; in other cases it will have little to say. For the Airline series, SEATS tells us that the variance of the total revision the concurrent estimator will undergo is slightly smaller than 1/6 of the 1-period-ahead-forecast error variance (or, equivalently, that the standard error of that revision is equal to 1.5% of the level of the series.) Whether this is too much or too little is difficult to answer; what we can say, however, is that the revision is optimal in a well-defined sense (of course, one may prefer another definition, but then it should be made explicit.) SEATS also tells us that 94% of the revision variance will disappear in one year and that, in any case, the forecast of the trend is considerably more precise than that of the observed series; these are facts of applied interest.

3. FINAL REMARKS

In his final remarks Stier mentions that different ARIMA models will produce different results, thereby introducing a degree of subjectivity. Used in an automatic way, the results will always be identical. Of course, people with different criteria may choose different models, but the AMB approach provides the framework to compare them as well as the associated results in a precise way (and hence provides the basis for improvement). In general, some element of subjectivity is unavoidable, be that in the selection of a model for forecasting, or in the choice of a method for seasonal adjustment (to the point that, for example, seasonal adjustment of the monetary aggregates at the Federal Reserve Board used to be X11 plus "judgmental" corrections; see Maravall and Pierce, 1983).

As for the summary of the evidence, contrary to Stier's assertion, if properly used with TRAMO, the behavior of SEATS with respect to the problem of spurious seasonality is excellent. Furthermore, the decomposition of the New-Orders series turns out to be quite reasonable, and has nothing to do with the results reported by Stier. As for the statement that "the trend component of SEATS shows oscillative properties" (which can be explained by the gain of the trend filter), as a general statement it is false. First, Stier considers only a few cases of the default model. There are many models with very stable trends. This is true even for the default model, and, as an example, the continuous line in Figure 8 displays the gain of the trend component filter for the default model when $\theta_1 = -.9$ and $\theta_{12} = .1$; obviously, no "oscillative behavior" will be induced. As mentioned before, the trend filter in SEATS adapts to the width of the spectral peak around zero of the observed series. This is clearly seen in Figure 8, where the dashed line corresponds to the example with a deterministic trend and the dotted line corresponds to the New-Orders example, a series with a relatively moving trend; both examples have been already discussed (notice that the dotted line in Figure 8 has little to do with Figure 11 of Stier).

The last paragraph contains the only constructive conclusion of the paper, and we fully support it: it would be interesting to compare TRAMO-SEATS with the Beveridge-Nelson type of approach in Breitung (1994) and with STAMP. The first approach mentioned, offers, in our opinion, a serious conceptual problem. Seasonality is typically removed because it is thought to be mostly caused by events (weather, calendar, habits,...) different than those driving the business cycle or the underlying longer-term evolution of the economy (for example, technological progress). Breitung's approach assumes perfectly correlated components, an unappealing assumption. In any case, we believe that relevant comparisons should be made on many more series and using the programs properly (possibly, in an automatic way).

Of course, one can always find or build examples that produce awkward results. Perfection is an asymptotic property, and we are forced to choose among imperfect methods (or models). Clearly, the choice should be based on evidence. In this respect, the four examples of Stier help little, since they are the result of combining a poor understanding of the method with a wrong use (that starts by confusing automatic use with default model) of what must have been rather obsolete versions of the programs. Given that the programs still are at the beta (experimental) stage, perhaps it might have helped if Stier could have asked us to check some of the strange results he was obtaining; this may have avoided some basic mistakes. Be that as it may, we appreciate Stier's interest and hope he continues his criticism (although with the correct versions!). We are sure that he will discover errors or weaknesses that can be of help to us. Ultimately, we insist, the real test should involve a more systematic and complete comparison with well-defined alternative methods, based on a minimally meaningful set of real world series.

FIG.1: AUTOMATIC USE OF T/S: SPURIOUS SEASONALITY
Ex. of negligible seasonality

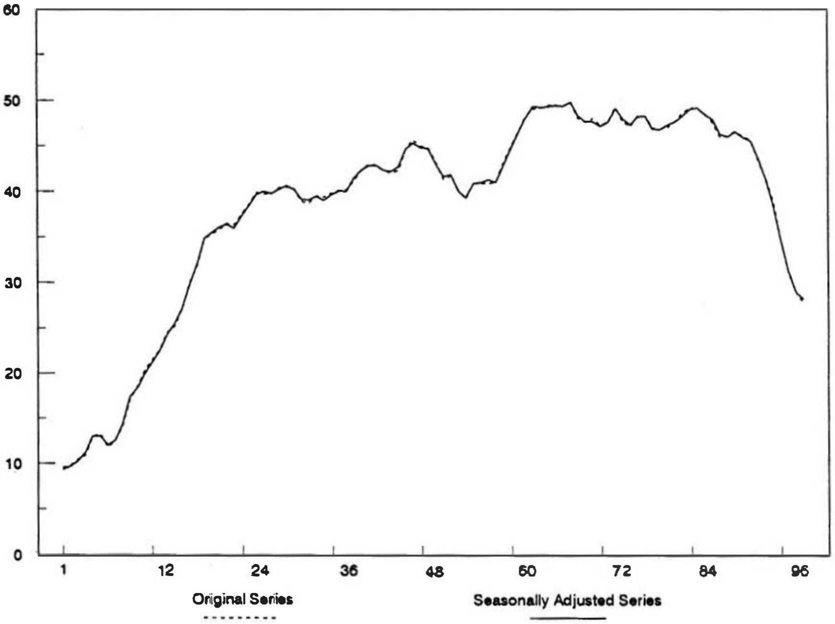


FIG.2: AUTOMATIC USE OF T/S: SPURIOUS SEASONALITY
Worst Example

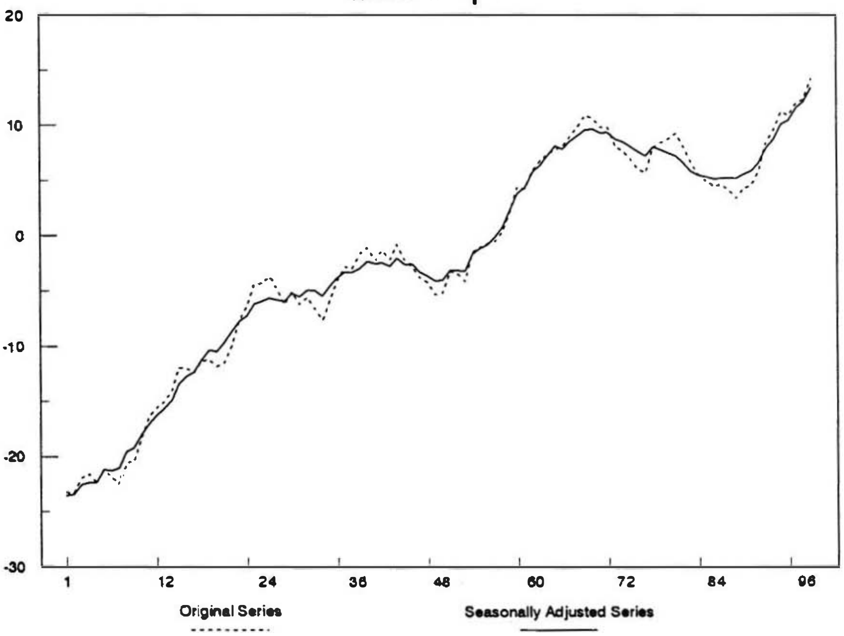


FIG.3: AUTOMATIC USE OF T/S: SPURIOUS SEASONALITY
Worst Example: Forecast

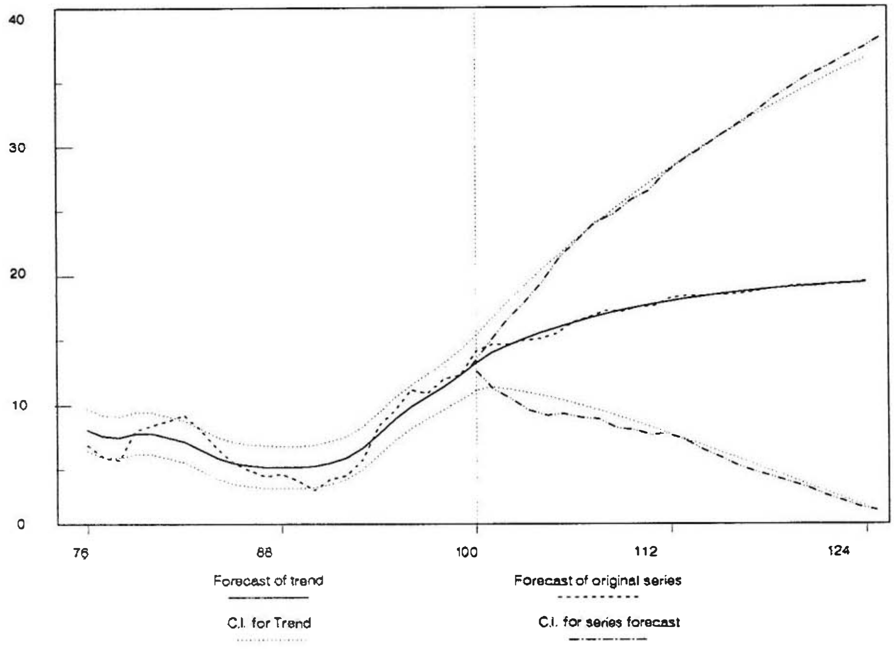


FIG.4:AUTOMATIC USE OF T/S: NEW-ORDERS EXAMPLES

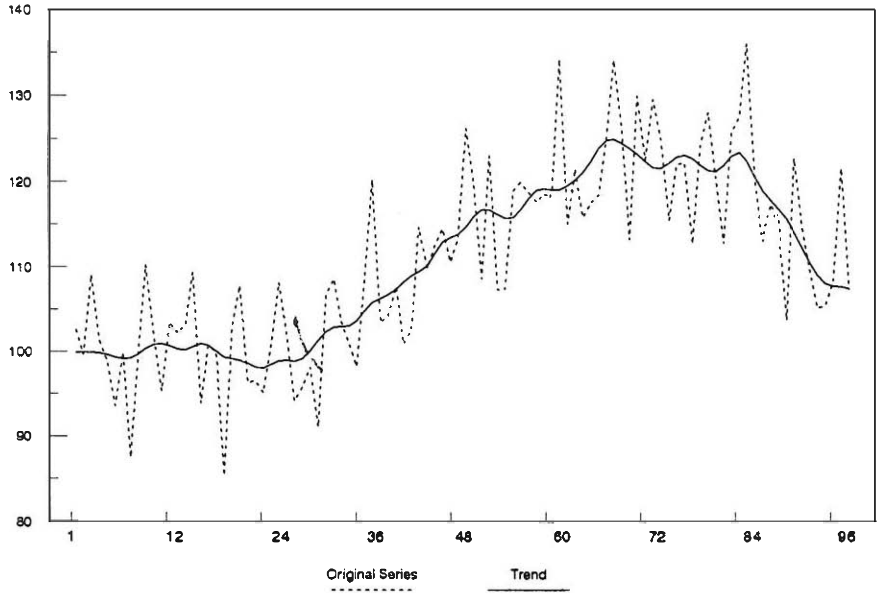


FIG.5: SIMULATED SERIES
TREND FROM LOW PASS FILTERS AND TREND FROM SEATS

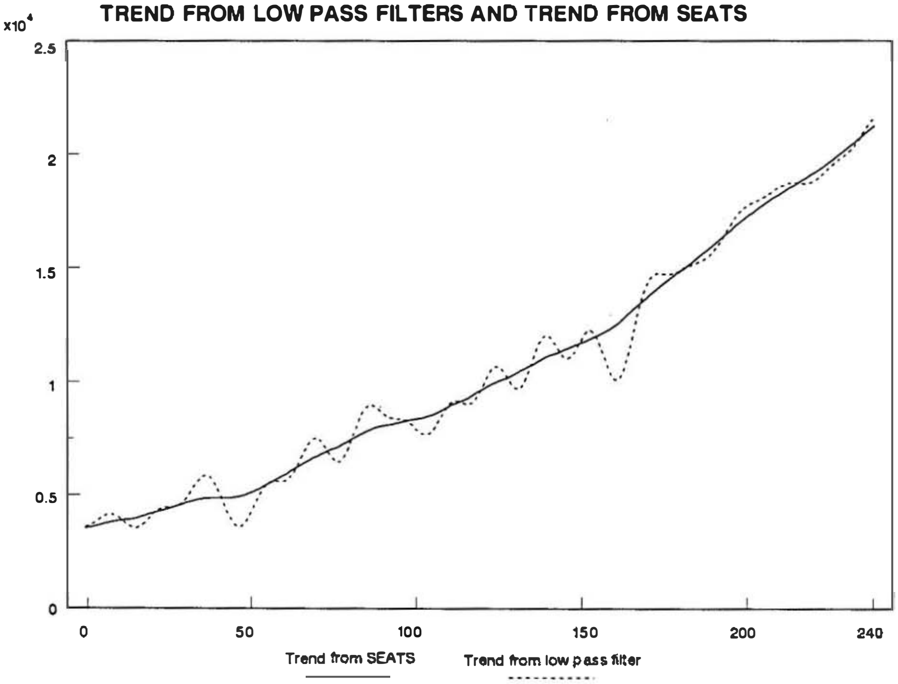


FIG.6: AUTOMATIC USE OF S/T DETERMINISTIC EXAMPLES

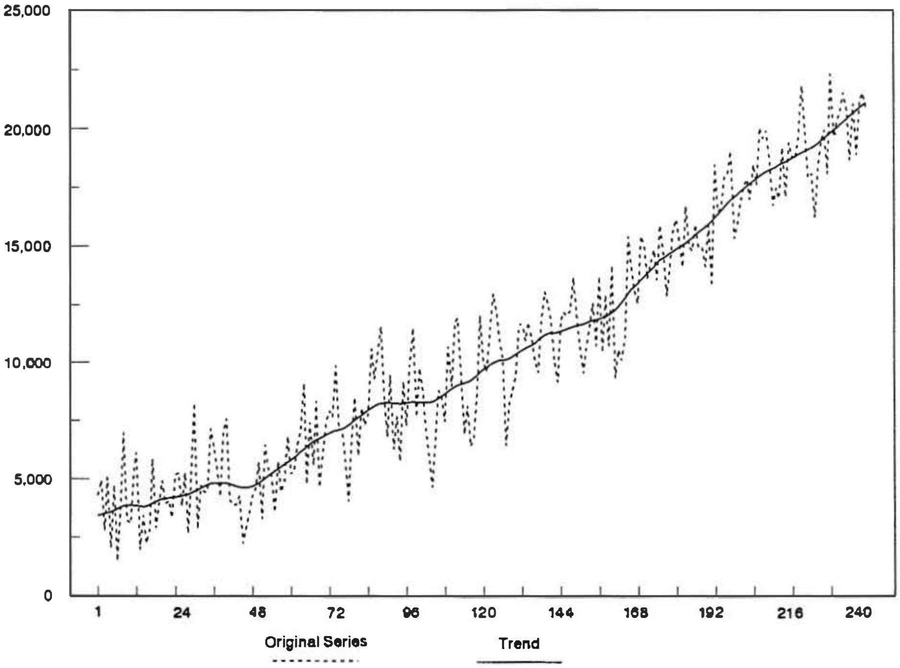


FIG.7: AIRLINE EXAMPLES: TRENDS OF THE TWO MODELS

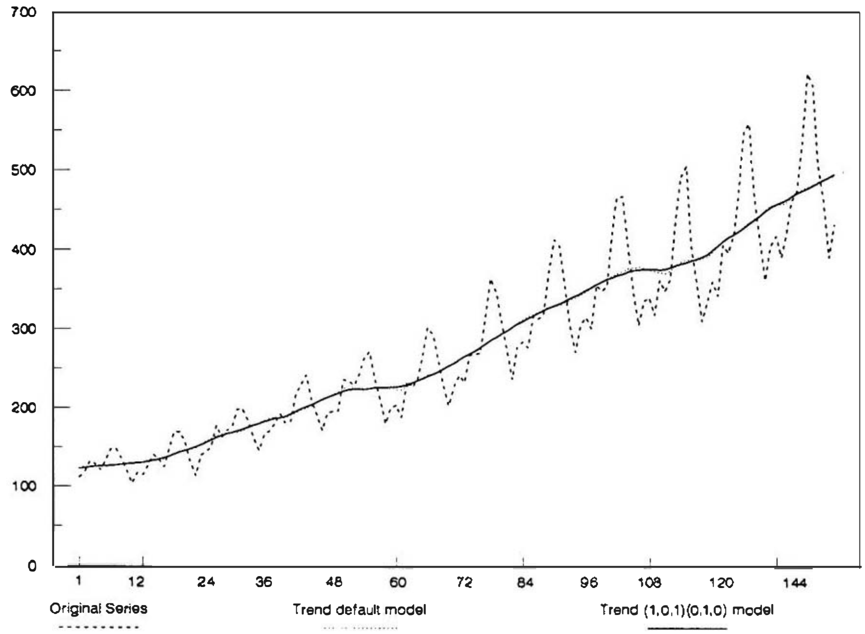
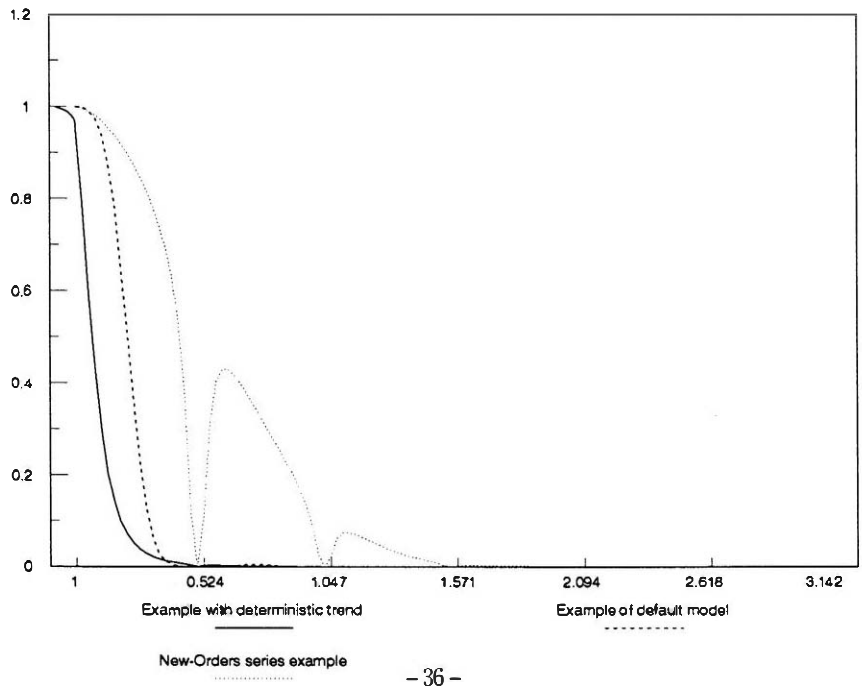


FIG.8: GAIN OF TREND FILTERS IN SEATS



REFERENCES

BALCHIN, S. (1995), "A description of the Seasonal Adjustment Methods X11, X11ARIMA, X12ARIMA, GLAS, STL, SEATS, STAMP and microCAPTAIN". (For the GSS Seasonal Adjustment Taskforce) Central Statistical Office.

BOX, G.E.P. and JENKINS, G.M. (1970), **Time Series Analysis: Forecasting and Control,** San Francio: Holden-Day.

BOX, G.E.P. and PIERCE, D. (1970), "Distribution of Residual Autocorrelations in Autoregressive Integrated Moving Average Time Series Models" **Journal of the American Statistical Association** 65, 1509-1526.

BREITUNG, J. (1994), "A Model Based Seasonal Adjustment Method Using the Beveridge-Nelson Decomposition", **Allg. Statistisches Archiv** 78, 365-385.

CLEVELAND, W.P. and TIAO, G.C. (1976), "Decomposition of Seasonal Time Series: A Model for the X-11 Program", **Journal of the American Statistical Association** 71, 581-587.

EUROSTAT (1996a), **Methodology of Industrial Short-Term Indicators; Rules and Recommendations**, Luxembourg: Official Publications of the European Communities.

EUROSTAT (1996b), "Is the ARIMA Model Adapted to Economic Time Series?" (Presented at the Informal Group on Seasonal Adjustment) Eurostat/D3/SA/04.

FISHER, B. (1995), **Decomposition of Time Series: Comparing Different Methods in Theory and Practice**, version 2.1, Luxembourg: Eurostat, April 1995.

GÓMEZ, V. and MARAVALL, A. (1996), "Programs TRAMO and SEATS; Instructions for the User", Working Paper 9628, Servicio de Estudios, Banco de España.

HARVEY, A.C. (1989), **Forecasting, Structural Time Series Models and the Kalman Filter**, Cambridge: Cambridge University Press.

KENDALL, M. (1976), **Time Series**, London: Griffin and Co.

MARAVALL, A. (1997), "Comments on the Seasonal Adjustment Program X12ARIMA", **Journal of Business and Economic Statistics** (forthcoming).

MARAVALL, A. (1995), "Unobserved Components in Economic Time Series, in Pesaran, H. and Wickens, M. (eds.), **The Handbook of Applied Econometrics**, Oxford: Basil Blackwell.

MARAVALL, A. (1987), "On Minimum Mean Squared Error Estimation of the Noise in Unobserved Components Models", **Journal of Business and Economic Statistics** 5, 115-120.

MARAVALL, A. (1988), "The Use of Arima Models in Unobserved Components Estimation", in Barnet, W., Berndt, E. and White, H. (eds.), **Dynamic Econometric Modeling**, Cambridge: Cambridge University Press.

MARAVALL, A. and PIERCE, D.A. (1987), "A Prototypical Seasonal Adjustment Model", **Journal of Time Series Analysis** 8, 177-193. Reprinted in S. Hylleberg (ed.), **Modelling Seasonality**, Oxford University Press, 1992.

MARAVALL, A. and PIERCE, D.A. (1983), "Preliminary-Data Error and Monetary Aggregate Targeting", **Journal of Business and Economics Statistics** 1, 179-186.

MARAVALL, A. and PLANAS, C. (1996), "Estimation Error and the Specification of Unobserved Components Models", Working Paper 9608, Servicio de Estudios, Banco de España (forthcoming in **Proceedings of the American Statistical Association**, Business and Economics Section.)

NERLOVE, M. (1964), "Spectral Analysis of Seasonal Adjustment Procedures", **Econometrica** 32, 241-286.

STIER, W. (1996), "Zur Kanonischen Zerlegung von Zeitreihen-Bemerkungen zum Verfahren SEATS", **Allg. Statistisches Archiv**, 80, 313-331.

WORKING PAPERS (1)

9525 **Aurora Alejano y Juan M.ª Peñalosa:** La integración financiera de la economía española: efectos sobre los mercados financieros y la política monetaria.

9526 **Ramón Gómez Salvador y Juan J. Dolado:** Creación y destrucción de empleo en España: un análisis descriptivo con datos de la CBBE.

9527 **Santiago Fernández de Lis y Javier Santillán:** Regímenes cambiarios e integración monetaria en Europa.

9528 **Gabriel Quirós:** Mercados financieros alemanes.

9529 **Juan Ayuso Huertas:** Is there a trade-off between exchange rate risk and interest rate risk? (The Spanish original of this publication has the same number.)

9530 **Fernando Restoy:** Determinantes de la curva de rendimientos: hipótesis expectacional y primas de riesgo.

9531 **Juan Ayuso and María Pérez Jurado:** Devaluations and depreciation expectations in the EMS.

9532 **Paul Schulstad and Ángel Serrat:** An Empirical Examination of a Multilateral Target Zone Model.

9601 **Juan Ayuso, Soledad Núñez and María Pérez-Jurado:** Volatility in Spanish financial markets: The recent experience.

9602 **Javier Andrés e Ignacio Hernando:** ¿Cómo afecta la inflación al crecimiento económico? Evidencia para los países de la OCDE.

9603 **Barbara Dluhosch:** On the fate of newcomers in the European Union: Lessons from the Spanish experience.

9604 **Santiago Fernández de Lis:** Classifications of Central Banks by Autonomy: A comparative analysis.

9605 **M.ª Cruz Manzano Frías y Sofía Galmés Belmonte:** Credit Institutions' Price Policies and Type of Customer: Impact on the Monetary Transmission Mechanism. (The Spanish original of this publication has the same number.)

9606 **Malte Krüger:** Speculation, Hedging and Intermediation in the Foreign Exchange Market.

9607 **Agustín Maravall:** Short-Term Analysis of Macroeconomic Time Series.

9608 **Agustín Maravall and Christophe Planas:** Estimation Error and the Specification of Unobserved Component Models.

9609 **Agustín Maravall:** Unobserved Components in Economic Time Series.

9610 **Matthew B. Canzoneri, Behzad Diba and Gwen Eudey:** Trends in European Productivity and Real Exchange Rates.

9611 **Francisco Alonso, Jorge Martínez Pagés y María Pérez Jurado:** Weighted Monetary Aggregates: an Empirical Approach. (The Spanish original of this publication has the same number.)

9612 **Agustín Maravall and Daniel Peña:** Missing Observations and Additive Outliers in Time Series Models.

9613 **Juan Ayuso and Juan L. Vega:** An empirical analysis of the peseta's exchange rate dynamics.

9614 **Juan Ayuso Huertas:** Un análisis empírico de los tipos de interés reales *ex-ante* en España.

9615 **Enrique Alberola Ila:** Optimal exchange rate targets and macroeconomic stabilization.

9616 **A. Jorge Padilla, Samuel Bentolila and Juan J. Dolado:** Wage bargaining in industries with market power.

9617 **Juan J. Dolado and Francesc Marmol:** Efficient estimation of cointegrating relationships among higher order and fractionally integrated processes.

9618 **Juan J. Dolado y Ramón Gómez:** La relación entre vacantes y desempleo en España: perturbaciones agregadas y de reasignación.

9619 **Alberto Cabrero and Juan Carlos Delrieu:** Construction of a composite indicator for predicting inflation in Spain. (The Spanish original of this publication has the same number.)

9620 **Una-Louise Bell:** Adjustment costs, uncertainty and employment inertia.

9621 **M.ª de los Llanos Matea y Ana Valentina Regil:** Indicadores de inflación a corto plazo.

9622 **James Conklin:** Computing value correspondences for repeated games with state variables.

9623 **James Conklin:** The theory of sovereign debt and Spain under Philip II.

9624 **José Viñals and Juan F. Jimeno:** Monetary Union and European unemployment.

9625 **María Jesús Nieto Carol:** Central and Eastern European Financial Systems: Towards integration in the European Union.

9626 **Matthew B. Canzoneri, Javier Vallés and José Viñals:** Do exchange rates move to address international macroeconomic imbalances?

9627 **Enrique Alberola Ila:** Integración económica y unión monetaria: el contraste entre Norteamérica y Europa.

9628 **Víctor Gómez and Agustín Maravall:** Programs TRAMO and SEATS.

9629 **Javier Andrés, Ricardo Mestre y Javier Vallés:** Un modelo estructural para el análisis del mecanismo de transmisión monetaria: el caso español.

9630 **Francisco Alonso y Juan Ayuso:** Una estimación de las primas de riesgo por inflación en el caso español.

9631 **Javier Santillán:** Política cambiaria y autonomía del Banco Central.

9632 **Marcial Suárez:** Vocábula (Notas sobre usos lingüísticos).

9633 **Juan Ayuso and J. David López-Salido:** What does consumption tell us about inflation expectations and real interest rates?

9701 **Víctor Gómez, Agustín Maravall and Daniel Peña:** Missing observations in ARIMA models: Skipping strategy versus outlier approach.

9702 **José Ranón Martínez Resano:** Los contratos DIFF y el tipo de cambio.

9703 **Gabriel Quirós Romero:** Una valoración comparativa del mercado español de deuda pública.

9704 **Agustín Maravall:** Two discussions on new seasonal adjustment methods.

---

(1) Previously published Working Papers are listed in the Banco de España publications catalogue.

Queries should be addressed to: Banco de España
Sección de Publicaciones. Negociado de Distribución y Gestión
Telephone: 338 51 80
Alcalá, 50. 28014 Madrid