APPLYING AND INTERPRETING MODEL-BASED SEASONAL ADJUSTMENT. THE EURO-AREA INDUSTRIAL PRODUCTION SERIES.

APPLYING AND INTERPRETING MODEL-BASED SEASONAL ADJUSTMENT. THE EURO-AREA INDUSTRIAL PRODUCTION SERIES.

Agustín Maravall Herrero

Domingo Pérez Cañete

BANCO DE ESPAÑA

(*) [Hacer click para insertar el texto de la primera nota]

(**) [Hacer click para insertar el texto de la segunda nota]

Documentos de Trabajo. N.º 11XX
2011

Abstract

The recent economic crisis has altered the dynamics of economic series and, as a consequence, introduced uncertainty in seasonal adjustment of recent years. This problem was discussed in recent workshops at the European Central Bank and at Eurostat in the context of adjustment of the Euro Area Industrial Production (EPI) series.

Because a seasonal component is unobserved and undefined, it is difficult to compare results from different adjustment methods. Within the regARIMA model-based approach, however, a framework for systematic analysis is indeed present. The EPI series is analyzed under the TRAMO-SEATS framework. The purpose of the analysis is not to compare alternative methods, but to show how the results of the model-based analysis can be exploited at the identification, diagnostics, and inference stages of modeling.

Despite the uncertainty induced by the crisis (and the revisions to the unadjusted data), the automatic procedure, with ramps to capture the spectacular 2008 drop in the series, provides excellent and stable results.

INDEX

1 Introduction

The economic crisis of the last three years has had an impact on the behavior of economic series, and models and methods used by statisticians and economists have had a difficult time dealing with it. The European Statistical System (ESS) Guidelines for Seasonal Adjustment (Eurostat, 2009), aimed at harmonizing and providing guidelines for seasonal adjustment within the ESS, was completed as the crisis unfolded. The Guidelines did not specifically deal with crisis like the present one, and seasonal adjusters have felt disoriented. In July 2010, a workshop was held at the European Central Bank (ECB) on implementation of the Guidelines and on ways to treat the crisis; Fifteen central banks and twenty statistical offices were present. One of the sessions centered on seasonal adjustment of the Euro Area Industrial Production (EIP) series and this paper expands on the comment we contributed.

One point under discussion was the behavior of the monthly seasonally adjusted (SA) EIP series in recent periods. The official adjustment carried out by Eurostat for the period January 1990-March 2010 (243 observations) is displayed in Figure 1. Because the economic recession has introduced significant uncertainties, some ECB statisticians conducted alternative test calculations which led them to suspect that official SA monthly rates of growth for the past few months were too high (Eiglsperger and Haine, 2010). Figure 2 compares the SA monthly rates of Eurostat with those of the Eiglsperger and Haine (EH) procedure for the last years of the period considered. The slope of the recovery in the last months of the sample is clearly higher for the Eurostat SA series.

While Eurostat uses the TRAMO-SEATS (TS) method in the DEMETRA interface, EH tests were performed with X12ARIMA (X12A). TS and X12A are widely used procedures and are the two methods considered in the ESS Guidelines. The programs and documentation can be downloaded from the US Bureau of the Census and Bank of Spain web sites (www.census.gov/srd/www/x12a, and www.bde.es/webbde/en/secciones/servicio/software/econom.html, respectively). An overview of the two methodologies can be found in Findley et al. (1998), Gómez and Maravall (2001a, b) and Findley (2005).

Both methods consist of two steps. First, the series is pre-adjusted; second, it is seasonally adjusted. The preadjustment step determines a possible transformation of the data, and identifies (often automatically) a regression(reg)-ARIMA model. This model is used to extend the series with forecasts, to identify and estimate outliers, calendar and other effects captured through regression variables. In the second step, a filter is applied to the extended series -net of outliers and regression effects- to obtain estimators of the (stochastic) SA series, as well as of the seasonal, trend-cycle, and irregular components. Combining them with the outliers and regression effects, the final estimators of the components are obtained.

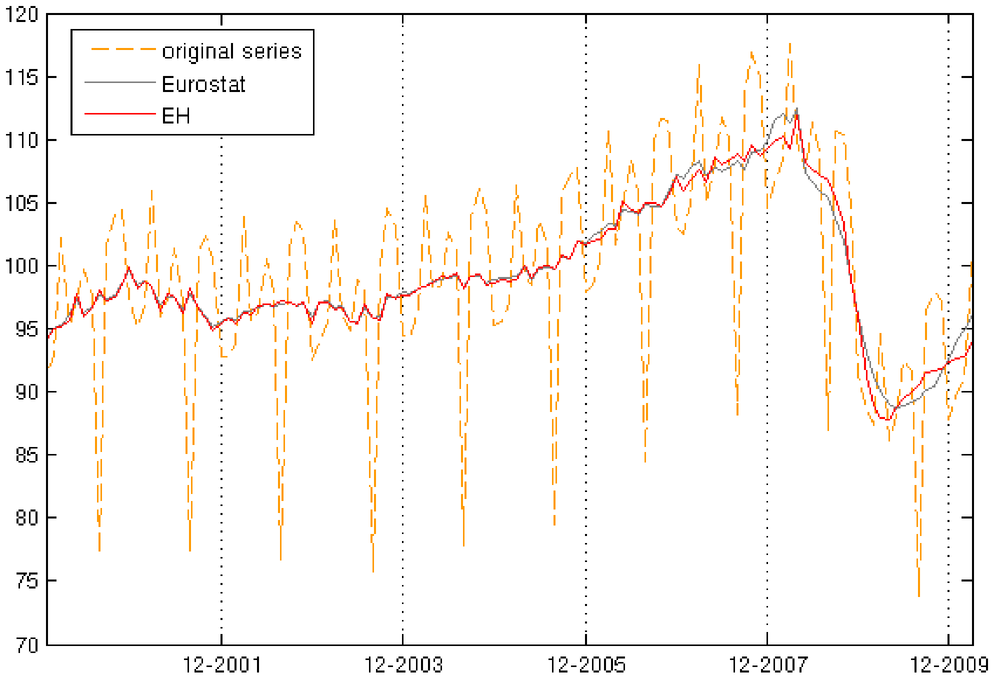


Figure 1: The official SA series and the alternative proposal.

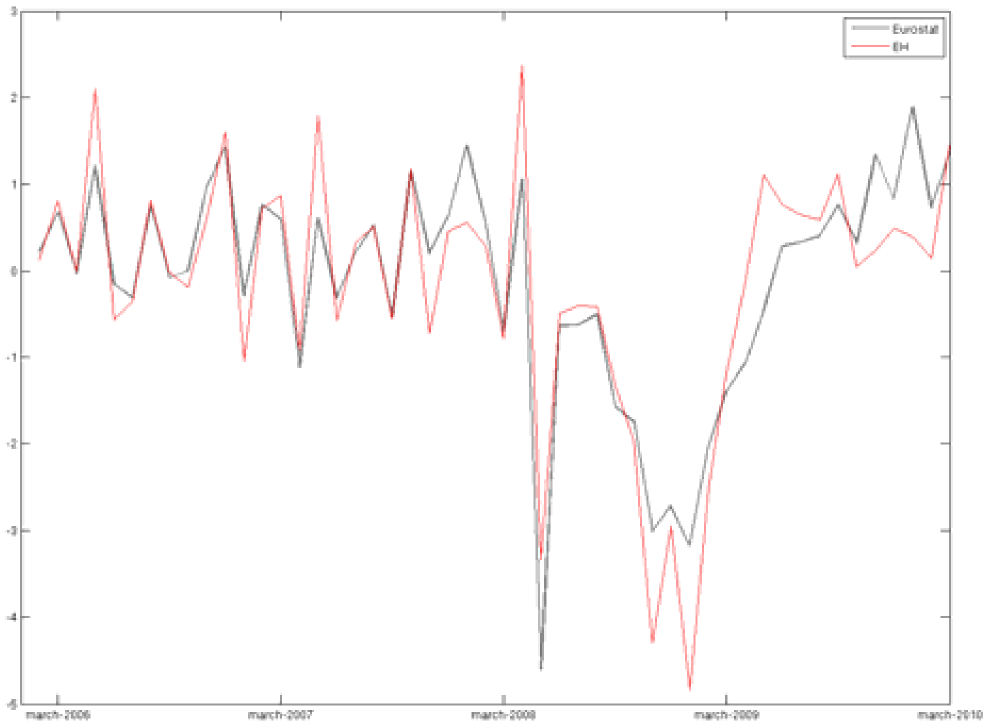


Figure 2: Rates of growth of the SA series.

The previous general structure is common to both methods. The similarities in the pre-adjustment step are many (e.g., the "automdl" option of X12A is based on the automatic model identification procedure in TRAMO). The adjustment step, on the contrary, follows different methodologies. The filter of X12A is selected from a set of a priori designed filters (X11 in essence). SEATS follows the so-called ARIMA-model-based (AMB) approach (see Burman, 1980, Hillmer and Tiao, 1982, and Bell and Hillmer, 1984). The filter in SEATS yields the minimum-mean-squared-error (MMSE) estimator of a theoretically specified component. This filter is often referred to as the Wiener-Kolmogorov (WK) filter. The component model is derived from the ARIMA model identified and fitted to the observed series. The two approaches represent different philosophies and this has repercussions in terms of analysis, diagnostics, and interpretation of results. Due to the continuity in the X11-X12A family, 50 years of use have made users familiar with its output. This is not quite the case for TS and for how its output can be used to evaluate seasonal adjustment.

Because a seasonal component is never observed and no universal definition is available, it is difficult to compare results from different adjustment methods in a systematic manner. Within the model-based approach, however, a framework for systematic analysis is indeed present. In this paper, the EPI series is analyzed under the TS model-based framework. No comparison of different methods is made; our aim is to show how the results can be exploited at the identification, diagnostics, and inference stages of modeling.

For the rest of the paper, unless otherwise specified, model-based diagnostics and inference are derived and computed under the assumption that the model is correct. The results should be interpreted thus as a “best case” approximation in which model parameters are “known”. (Some improvements that take into account parameter estimation error are presented in Blakely and McElroy, 2011 and Bell, 2005.)

2 The series and the "alternative" procedure

The seasonal adjustment method suggested by EH is X12A with the following specifications: log transformation, two level shift outliers (Nov 2008 and Jan 2009), no correction for calendar effects, and (0,1,1) $(0,1,1)_{12}$ ARIMA model orders (i.e., those of an Airline model). As shown in Figure 3a, the seasonal factors obtained are reasonably stable and, as seen in the first row of Table 1, the residuals pass the Behra-Jacque test for Normality.

Table 1: Model diagnostics.

	Residual SE in 10^{-2}	Residual diagnostics							OS Forecast RMSE (in 10^{-2})
		Q	N	Sk	K	Q2	Runs	QS	
EH	1.31	54.3	1.5	-0.8	-0.9	46.6	0.9	0.0	1.8
TS-0	1.28	20.0	2.3	-0.9	1.2	42.7	2.4	0.2	1.7
TS-2LS	1.21	29.1	2.1	-1.0	1.0	36.6	2.2	0.0	1.6
TS-2R	1.11	27.6	1.0	0.6	-0.8	15.2	1.6	0.3	0.9
Approx. 95% Critical Value	---	32	6	2	2	37	2	6	---

SE: Standard Error; Q: Box-Ljung test for 24 autocorrelation; N: Behra-Jarque Normality test; Sk: t-test for skewness; K: t-test for Kurtosis; Q2: McLeod and Li squared residuals test for nonlinearity; Runs: t-test for randomness in signs; QS: Pierce test for Seasonality; RMSE (OS Forecast): Out-of-Sample 1-period-ahead forecast root mean square error (last 12 observations).

But the output of X12A flags some problems: (a) the residuals are autocorrelated; (b) the in-sample forecast errors are large; (c) a peak at the TD frequency is detected in the spectra of the residuals, differenced log SA series, and irregular component; (d) the autocorrelation function (ACF) of the squared residuals indicates some nonlinearity. Direct inspection of Figure 4a shows, for the last two years, an accumulation of large negative residuals followed by a sequence of mostly positive ones. This abnormal behavior coincides with the crisis period.

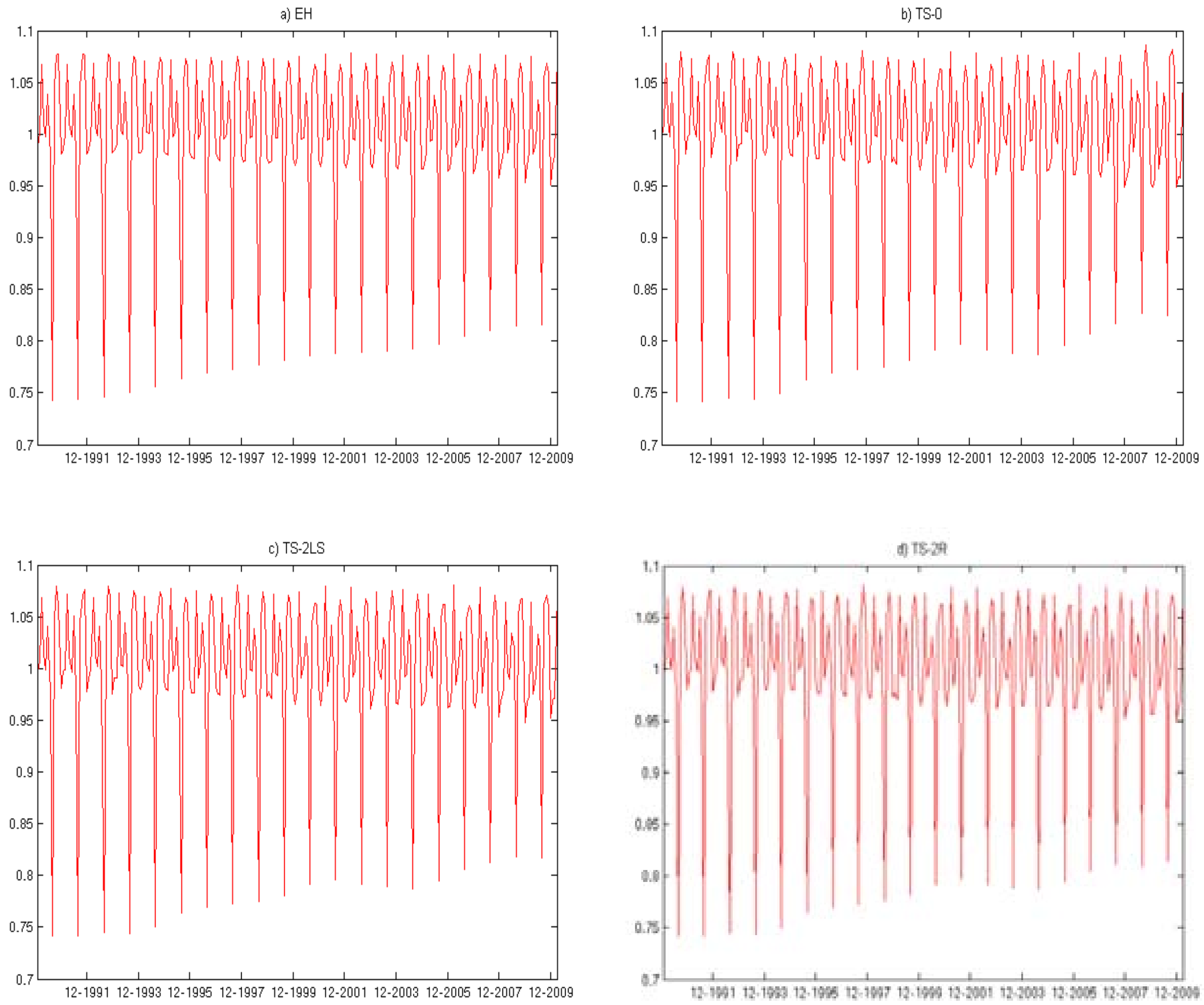


Figure 3: Seasonal Factors.

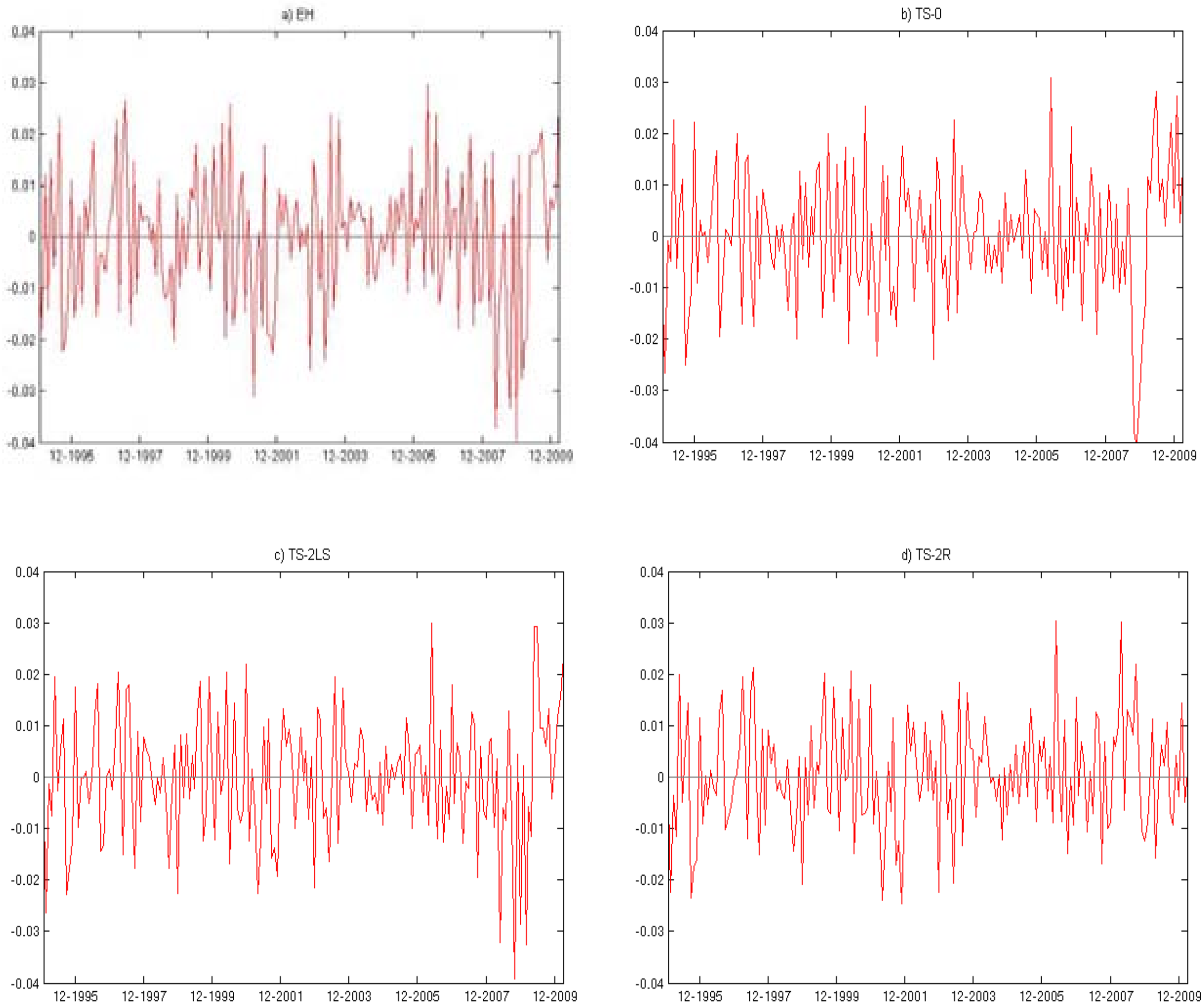


Figure 4: RegARIMA model residuals.

To understand the model choice, X12A was run automatically with the automdl option and all pretests included. The series is modeled in levels, five outliers are detected (one is the LS Nov 2008 outlier), significant calendar effects are found, and the identified ARIMA is a (1,1,2) $(0,1,1)_{12}$ model. The EH regARIMA model and the one automatically detected by X12A are considerably different, and the automatic result provides a better fit. The residuals show no evidence of autocorrelation and all Normality tests, as well as the Q2 test on the squared residuals, are comfortably passed. However, a few problems remain: the in-sample forecast errors are still large, a peak in the residual spectrum is found at the TD frequency, and the M4 X12A statistics signals too much autocorrelation in the irregular component. The problems, however, do not seem major. The forecast problem is mostly associated with the recent crisis; the TD peak is moderate; and the M4 statistics is of little interest (MMSE estimators of a white noise irregular in an ARIMA series ought to be autocorrelated; see Maravall, 1987).

The EH choice of logs versus levels seems acceptable: the ratio of the two likelihoods (appropriately adjusted) is extremely close to 1, thus the pre-test cannot discriminate between the two transformations; besides, logs improve normality. As for TD and Easter effects, the EIP series is obtained from aggregation of country series supposed to have already been adjusted for calendar effects. Indirect TD adjustment is thus taken for granted, although it does not correct the series properly. Still, compared to the seasonal component, calendar effects are of secondary importance. Concerning the ARIMA model, possibly the Airline model was motivated because it is simple and robust. Given that in X12A the model is only used for preadjustment and 1-year forecast extension of the series, it is thought that the effect of model misspecification on the seasonal factors is likely to be small and, ultimately, seasonal factors are all that matter. Yet, in the absence of a definition (or model), how do we know how the seasonal component should be?

In justifying their model, EH stressed that seasonality should be stable and uncontaminated by the economic crisis. Considering that more stable seasonal components will likely imply more unstable SA series, in so far as one of the main uses of SA series is to help cyclical assessment, forcing seasonal stability would not help. Further, to assume that a profound crisis does not affect seasonality of industrial production seems unrealistic. Thus the priority given to seasonal stability has fuzzy grounds. Some series will have more moving seasonality than others and the seasonal component should adapt accordingly. In TS this adaptation is achieved through the width of the spectral peaks at seasonal frequencies in the model for the seasonal component and by the size of its innovation. (The dilemma “stability versus optimality” is discussed in Maravall, 1998.)

In summary, the EH proposal presents problems. (Both the EH and Eurostat procedures have been improved after the Workshop and are now closer.) The automatic run suggests that the problems could be properly dealt with working with X12A. We will not pursue this issue and will look instead at the approach of TS. The aim will not be to find the best possible model, but to show how one can proceed within the model-based framework, and how quality of the seasonal adjustment can be addressed. (The analysis will also use results of a new version of TS not yet released. Some of them, such as the phase delay function and the autoregressive spectrum used in Section 5, have benefited from the X12A developers help.)

Although the X12A and TS filters represent different methodologies, it has often been pointed out that, for many series, they yield similar results. Running TS with the EH specifications, as seen in Figure 5, the difference between the X12A and TS SA series is negligible even for the crisis years. Thus the EH SA series can be approximately seen as the MMSE estimator of the SA series implied by the fitted Airline model, and the output of TS could provide inference such as, for example, the SE of the X12A SA series (interpreted as a model-based estimator) and of its rates of growth or forecasts. In our case, however, the presence of autocorrelation in the ARIMA residuals would invalidate the model and the inference. (Be that as it may, given that the X11 filter is not model based, model-based inferences need not be relevant.)

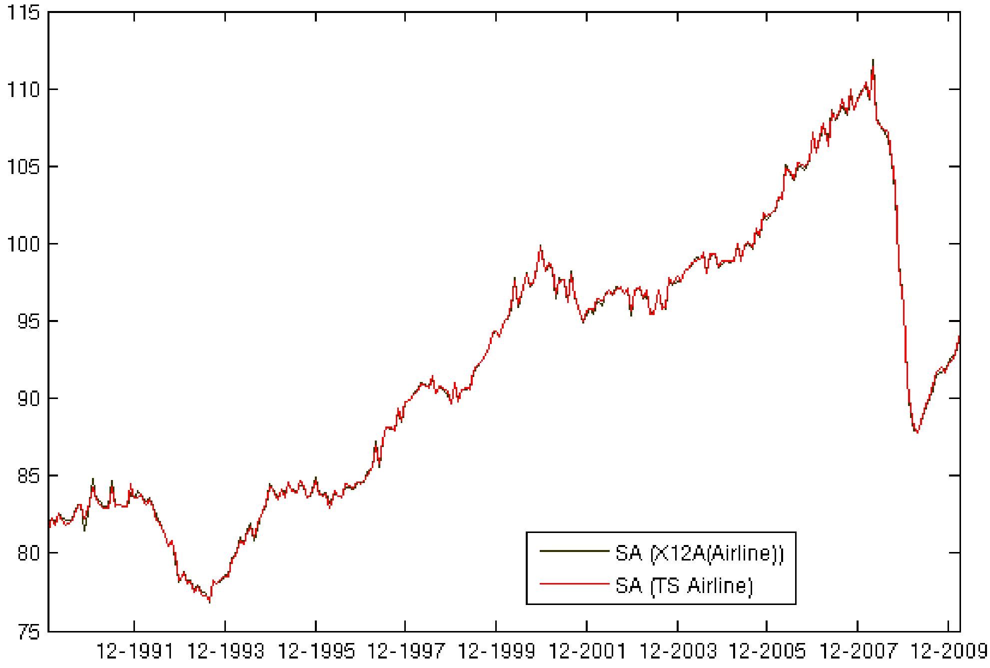


Figure 5: X12A and TS-0 SA series: Airline (EH) model.

3 TRAMO-SEATS automatic (model-based) procedure

3.1 Summary of the procedure

Instead of DEMETRA, we use program TSW, (a Windows version of TS, also available at the Bank of Spain website) and start with the automatic option that pretests for a 6 parameter TD variable. This option estimates a general model of the type

$$y(t) = \sum_{i=1}^{n_0} \alpha_i \, D_i(B) \, d_i(t) + \sum_{i=1}^{n_c} \beta_i \, cal_i(t) + x(t), \qquad (1)$$

where B denotes the lag operator, $d_i(t)$ is a dummy variable that indicates the position of the i-th outlier, and $D_i(B)$ is a polynomial in B reflecting the type of outlier. For an additive outlier (AO), $D_i(B) = 1$; for a transitory change (TC), $D_i(B) = 1/(1 - 0.7\,B)$, and for a levels shift (LS), $D_i(B) = 1/(1 - B)$. "Cal" denotes a calendar effect variable, such as TD or Easter effect, β_i is its associated coefficient, and $x(t)$ is an invertible ARIMA model, say

$$\varphi(B)\, x(t) = \theta(B)\, a(t)\,, \qquad \varphi(B) = \phi(B)\, \delta(B), \qquad (2)$$

where $a(t)$ is a zero-mean white-noise innovation with variance $Var(a)$, and $\phi(B), \theta(B)$ and $\delta(B)$ are finite polynomials in B, the last one containing the unit AR roots. Thus,

$$\delta(B) = \nabla^d \, \nabla_s^{d_s}\,, \qquad \nabla = 1 - \mathrm{B}, \qquad \nabla_{\mathrm{s}} = 1 - \mathrm{B}^{\mathrm{s}} = \nabla\, \mathrm{S}, \qquad (3)$$

where $\mathrm{S} = 1 + B + \cdots + B^{s-1}$, s denotes the number of observations per year, and $d = 0, 1, 2, 3$, and $d_s = 0, 1$ are the possible choices for the difference orders. (The procedure allows for missing observations in $y(t)$.) In TS, after testing for a log transformation, possible outliers and calendar effects, as well as the ARIMA model for $x(t)$, are automatically identified. The resulting model is estimated by exact maximum likelihood.

When used for seasonal adjustment, first, removal of the regression effects from $y(t)$, yields an estimate of the stochastic component $x(t)$. In what follows it is assumed that the stochastic component is equal to this estimate, which is then split into an uncorrelated stochastic seasonal component - $s(t)$- and a stochastic SA series - $n(t)$- as in

$$x(t) = s(t) + n(t)\,. \qquad (4)$$

First, the roots of $\phi(B)$ and $\delta(B)$ are assigned to the seasonal component or to the SA series. This yields the factorizations $\phi(B) = \phi_s(B)\ \phi_n(B)$; $\delta(B) = \delta_s(B)\ \delta_n(B)$, where the subindex indicates the allocation of roots to the components. For example, when $d_s = 1$, (3) implies $\delta_s(B) = S$, and $\delta_n(B) = \nabla^{d+1}$.

Letting $\varphi_s(B) = \phi_s(B)\, \delta_s(B)$, and $\varphi_n(B) = \phi_n(B)\, \delta_n(B)$, a partial fraction expansion eventually yields the decomposition

$$\frac{\theta(B)}{\varphi(B)}\, a(t) = \frac{\theta_s(B)}{\varphi_s(B)}\, a_s(t) + \frac{\theta_n(B)}{\varphi_n(B)}\, a_n(t)\,, \qquad (5)$$

where a_s and a_n are the (uncorrelated) innovations in the seasonal component and SA series. The partial fractions decomposition is done in the frequency domain: the pseudo-spectrum of $x(t)$ is partitioned into additive spectra associated with the terms in the r.h.s. of (5). Identification of a unique decomposition is achieved by imposing the canonical condition that the minimum of the seasonal component pseudo-spectrum be zero (see Burman, 1980).

The WK filter for estimating $s(t)$ is given by the ratio of the $s(t)$ and $x(t)$ pseudo-spectra, and yields the MMSE estimator –also the conditional mean- of $s(t)$ given $x(t)$. The filter is centered, symmetric, and convergent; its derivation requires an infinite realization of $x(t)$ in the direction of the past and of the future. To apply the filter to a finite realization, model (2) is used to extend $x(t)$ with forecasts and backcasts and, the full effect of the infinite filter can be captured with a moderate number of them. As new observations become available, forecasts will be updated and eventually replaced by observations. As a consequence, the estimator of $s(t)$ near the end of the series is preliminary and will be revised. For long enough series, the filter for periods in the central years will have converged and the estimator will be final (or "historical"). This estimator can be assumed to have been obtained with the complete WK filter applied to the observed series.

The duration of the revision process of a preliminary estimator depends on the ARIMA model identified for the series. (For EIP case the revision of a concurrent estimator is completed in 2 years, and hence historical estimators can be assumed for the central 16 years.)

Spectral factorization provides the time-domain expression of the components models, say

$$\varphi_s(B)\, s(t) = \theta_s(B)\ a_s(t)\,, \qquad (6a)$$

$$\varphi_n(B)\, n(t) = \theta_n(B)\ a_n(t)\,. \qquad (6b)$$

Letting F denote the forward shift operator (i.e., $F = B^{-1}$) and replacing the ratio of pseudo-spectra by the ratio of autocovariance generating functions (ACGF), the time domain expression of the WK filter becomes (after simplification)

$$\nu_s(B, F)\ = k_s\ \frac{\theta_s(B)\ \theta_s(F)\ \varphi_n(B)\ \varphi_n(F)}{\theta(B)\ \theta(F)}\ , \qquad (7)$$

where $k_s = Var\ (a_s)/Var\ (a)$, so that the final estimator is given by

$$\hat{s}(t) = \nu_s(B, F)\ x(t)\,, \qquad (8)$$

where $\nu_s(B, F)$ is the ACGF of the stationary ARMA model

$$\theta(B)\, z_t = \theta_s(B)\, \varphi_n(B)\, b(t)\,, \qquad Var\ (b) = k_s\,. \qquad (9)$$

From (2), (7), and (8), $\hat{s}(t)$ can be expressed in terms of the innovations in $x(t)$,

$$\varphi_s(B)\ \theta(F)\ \hat{s}(t) = k_s\ \theta_s(B)\ \theta_s(F)\, \varphi_n(F)\, a(t)\,, \qquad (10)$$

so that the ACGFs and spectra of $s(t)$ and $\hat{s}(t)$ will be different. In particular, if $n(t)$ is non-stationary, the model for $\hat{s}(t)$ will be noninvertible, displaying spectral zeros associated with the unit roots in $\varphi_n(B)$. (Changing the subscript s for n in (7) and (8), the previous derivation yields $\nu_n(B, F)$ and $\hat{n}(t)$.)

When the series $x(t)$ is split into more components (as in: trend-cycle + seasonal + stochastic TD + irregular component) the procedure is a straightforward extension, with all components made canonical.

3.2 The application

The automatic TS procedure yields the SA series of Figure 6. The regARIMA model obtained selects the logs, finds significant calendar effects, and identifies the $(3,1,0)\ (0,1,1)_{12}$ ARIMA model

$$(1 + 0.04\ B - 0.20\ B^2 - 0.35\ B^3)\ \nabla\ \nabla_{12}\ \mathrm{x(t)} = (1 - 0.41\ B^{12})\ \mathrm{a(t)}$$

with $Var(a) = 0.013^2$. We shall refer to this model as the "TS-0 model". (In previous versions of TS that had lower critical values for outlier detection, a moderate LS was detected for May 2008.)

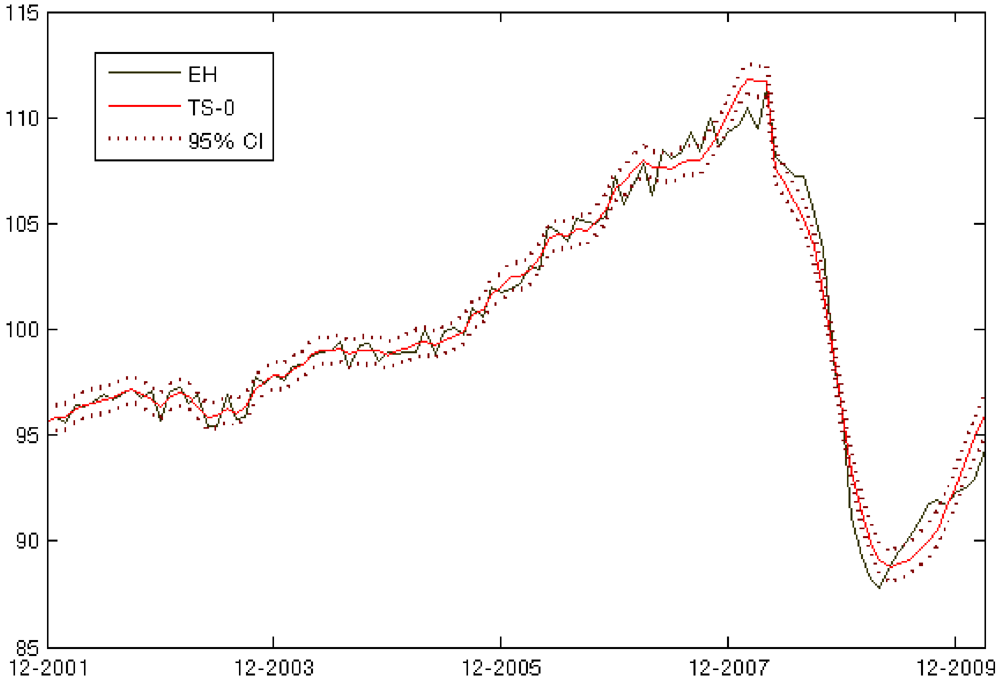


Figure 6: EH and TS-0 SA series (with confidence interval for the latter).

The difference between the result of the automatic modeling of X12A and TS is perhaps puzzling because the *automdl* procedure of X12A is based on TRAMO. When adapting the code to X12A, modifications were made in the default sequence of estimation algorithms, in the critical value for outlier detection, in the main information criterion for model comparison, in some model restrictions, etc. (see Monsell, 2009). These modifications may induce differences in results. In our example, these differences are: (a) TS selects logs, X12A levels. As mentioned before, the likelihood ratio is very close to 1, so that trivial differences might tip the coin; (b) TS detects no outlier, X12A detects 5, mostly due to the fact that the use of levels tends to increase the number of outliers detected (when logs are imposed, the automatic X12A procedure detects no outlier); (c) X12A selects a $(1,1,2)\ (0,1,1)_{12}$ model, while TS chooses a $(3,1,0)\ (0,1,1)_{12}$ one. Thus the stationary regular part of the model for X12A has an ARMA(1,2) structure, with a real AR root (0.75) and a pair of complex conjugates MA roots; the AR(3) model of TRAMO factors into a real root (0.80) and a pair of complex conjugate ones. Figure 7 compares the pseudo-spectra for the two models. The most relevant difference is a small amount of variance centered between the fourth and fifth harmonic, near the TD frequency, associated with the complex roots of the AR(3) polynomial in TS-0.

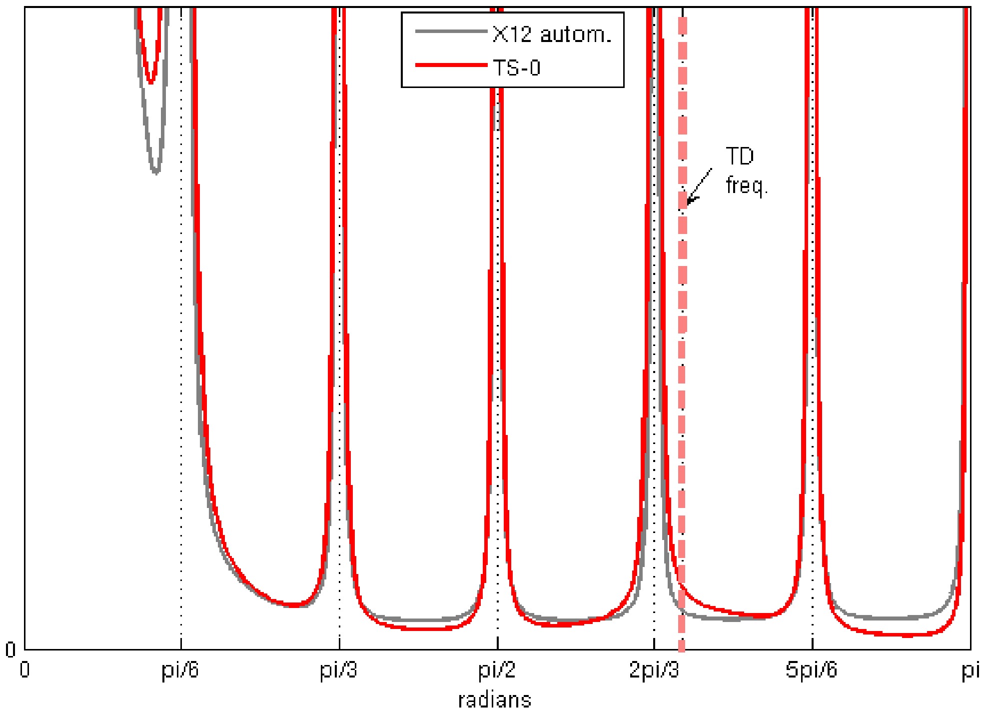


Figure 7: Pseudo-spectra of automatically identified models.

4 Calendar effects

4.1 Stochastic trading day

There is evidence that calendar effects are still present in the aggregate of the working day adjusted series provided by the individual countries. Given that the quality of calendar adjustment varies considerably across countries, the result may be understandable. More intriguing is that, after the automatic run of X12A has removed the calendar effects remaining in the indirectly adjusted series, the spectrum of the residuals still presents a peak for the TD frequency. This is not the case with TS. The pair of complex conjugate roots in the AR(3) polynomial have frequency = 2.22 radians, not far from 2.19 (or 4.2 cycles per year), which is the main frequency used in the TS and X12A spectral TD checks (frequency will always be expressed in radians.) These roots may suggest a stochastic TD effect not removed by the deterministic specification.

TS models this effect as a stationary ARMA(2,2), where the AR contains the roots close to the TD frequency, and the MA(2) is obtained from the model decomposition. Because of the canonical condition, the MA(2) will have two real roots (at the 0 and π frequencies) and the first one will be a unit root (i.e., a spectral zero). For the EIP series, the stochastic TD effect, c(t), follows the stationary model

$$(1 + 0.84B + 0.47B^2)\, c(t) = (1 - B)\,(1 + 0.58B)\, a_c(t)\ , \qquad (11)$$

with $a_c(t)$ white noise with $Var(a_c) = 0.0018^2$. Its spectrum is shown in Figure 8. Figures 9 and 10 compare the deterministic and stochastic TD component estimators: broadly, both look noisy, specially the stochastic trading day component.

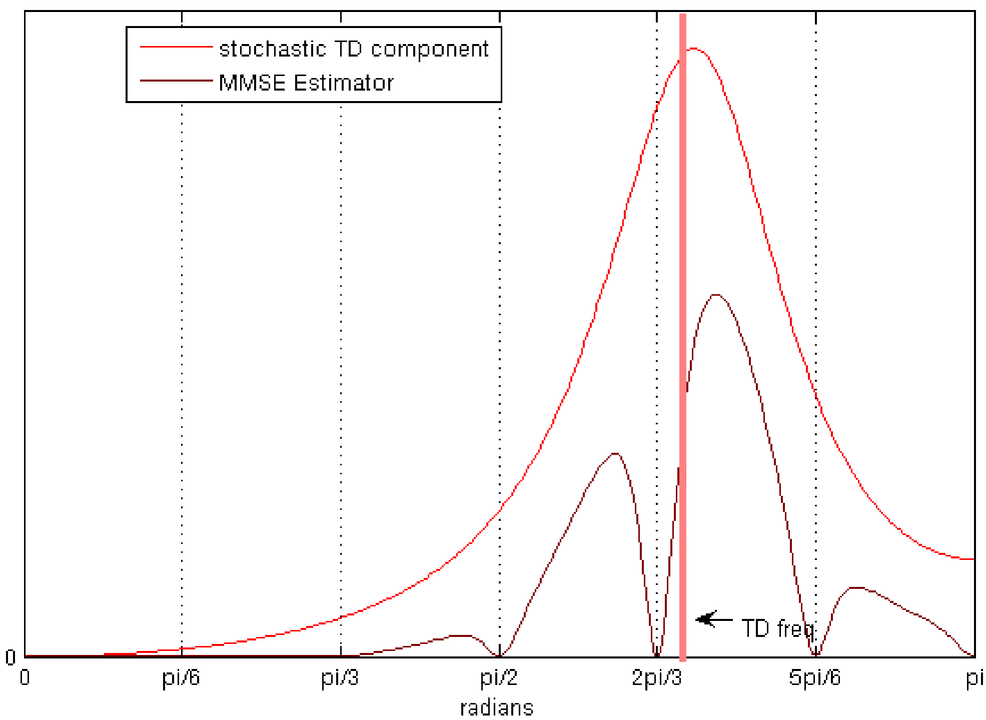


Figure 8: Spectra of the Stochastic Trading Day component model and of its theoretical estimator.

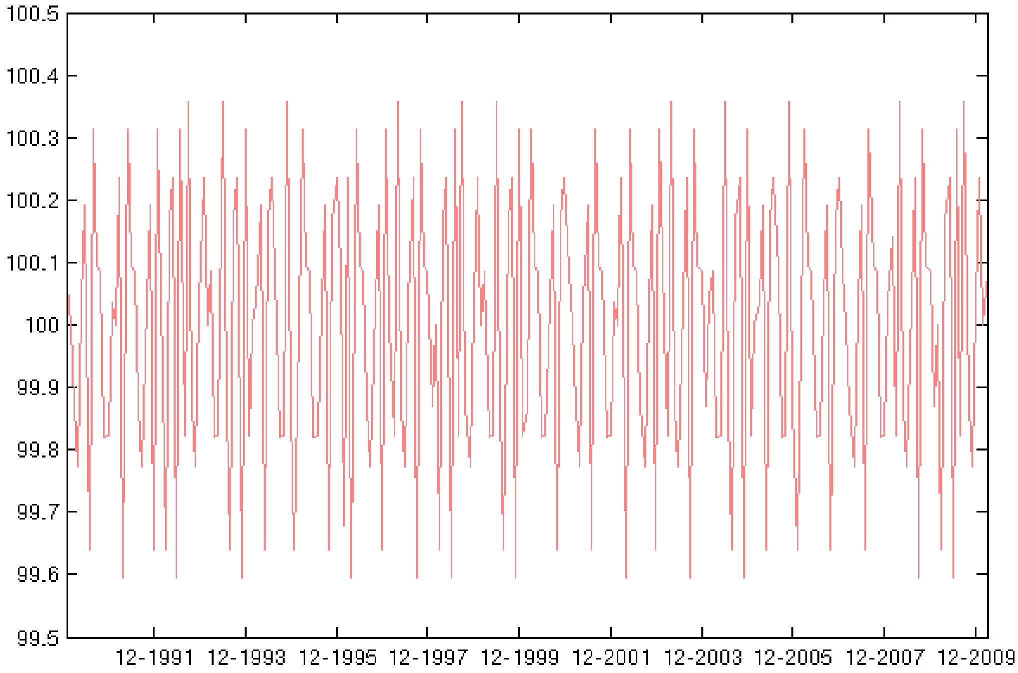


Figure 9: Deterministic trading day effect.

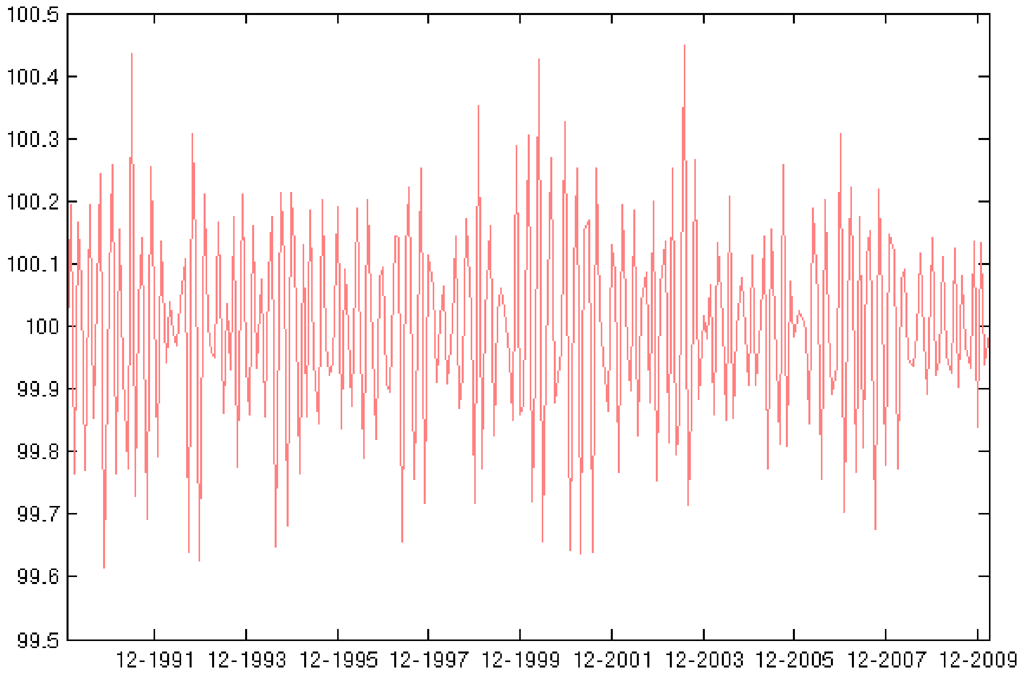


Figure 10: Stochastic trading day effect.

Our TD adjustment consists, first, of the removal of a deterministic TD through regression and, second, when the resulting series still exhibits a spectral peak at the TD frequency, removal of this peak through MMSE estimation of an ARMA(2,2) component. Purely stochastic TD components, whereby random walk models on the TD regression coefficients are specified have been used by Harvey (1989) and Bell (2004). In our approach, the stochastic TD is captured with a single stochastic component plus the deterministic TD effects, while in the random walk approach the mean of each TD coefficient evolves over time subject to the constraint that, at any point in time, the seven stochastic TD coefficients sum to zero.

4.2 A comment on trading-day spectral diagnostics

Spectral peaks in the residuals or in the SA series are useful diagnostic tools for TD effects. However, TD detection by looking at the spectrum of the unadjusted series is trickier: because of the proximity between the main TD frequency and the 4th seasonal harmonic, the TD effect may well pass unnoticed. In TS, a stochastic TD can also be detected through the presence of a pair of AR roots associated with the main TD frequency. This complex root will typically be stationary and an ambiguity emerges.

First, the frequency of the AR roots is not the same as the frequency of the spectral peak they induce (Jenkins and Watts, 1968). Consider, for example, an AR(2) model with polynomial $(1 + \phi_1 B + \phi_2 B^2)$ that contains a pair of complex conjugate roots.

The frequency of the root is

$$\lambda_r = \arccos |\phi_1|/2\sqrt{\phi_2} \ ,$$

while the maximum of the model spectrum is achieved at the frequency

$$\lambda_s = \arccos \left[-\phi_1(1 + \phi_2)/4\phi_2\right] .$$

The two frequencies are different and the difference $(\lambda_r - \lambda_s)$ depends only on the modulus of the root, or equivalently, on ϕ_2. As the root approaches non-stationarity, the difference goes to zero. This "frequency displacement" effect is aggravated because, when $c(t)$ follows the stationary model (11) and its estimator $\hat{c}(t)$ follows a model similar to (10) (with s replaced by c), it is easily checked that the frequencies of the spectral peaks in $c(t)$ and $\hat{c}(t)$ are not the same.

As a consequence, for a stationary stochastic TD, there are three relevant frequencies. One is the frequency of the AR root, λ_r, that will also be driving the “eventual” ACF and forecast function (i.e., the ACF and forecast function for lags and horizon beyond the model MA order). The others are the frequencies associated with the spectral peaks of the model for the component, λ_s, and of the model for its MMSE estimator, $\hat{\lambda}_s$.

In the model-based approach, the three frequencies can be derived from the ARIMA model for the series. For the EIP series and the stochastic TD component given by (11), it is found that $\lambda_r = 2.22$, $\lambda_s = 2.23$, and $\hat{\lambda}_s = 2.29$. As seen in Figure 8, these frequencies are slightly above the TD frequency used in the spectral diagnostics. Thus, if a fixed frequency is used in the checks, the associated interval (in SEATS at present, 2.15 – 2.30) should allow for the previous distortions.

The discussion applies in an identical way to stationary seasonal components, but given that highly stationary seasonal components are unlikely, it will seldom be relevant.

5 Some frequency domain results

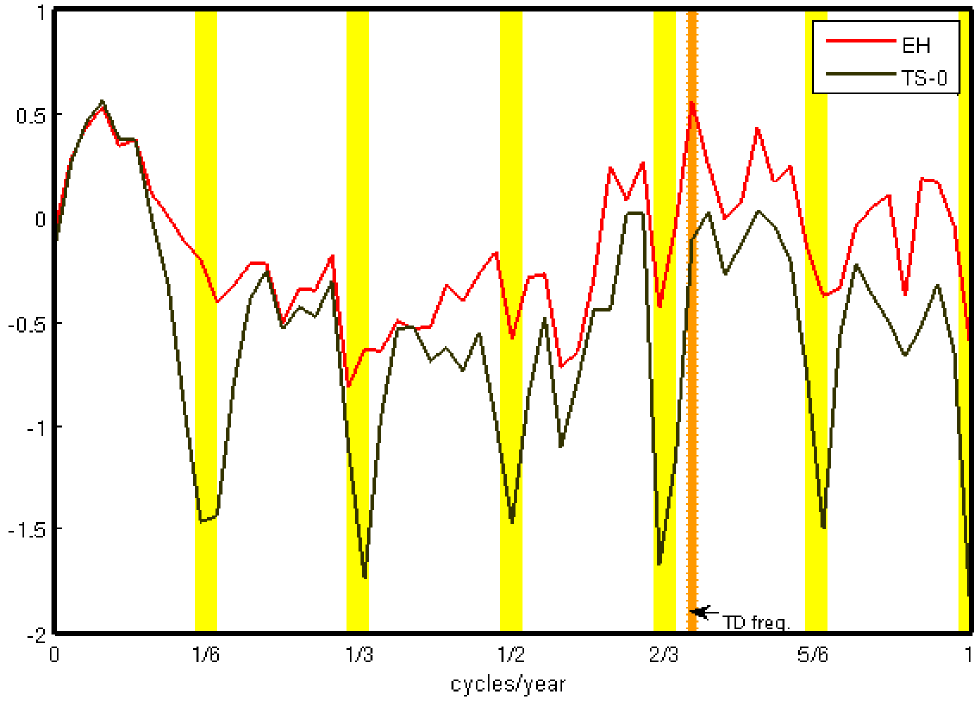
Model TS-0 contains 11 parameters; their estimators however have small correlations, never exceeding 0.17 (in absolute value). As seen in Table 1, the TS-0 model improves upon the EH model in that no residual autocorrelation is detected. Still, the Run test for randomness in the residual's signs flags a problem, evidenced by direct inspection of Figure 4b. The last 12 residuals are positive, which seems unacceptable and implies underestimation of the series for the last months.

From the point of view of seasonal adjustment, a consequence of not having introduced LS outliers to capture the spectacular plunge at the turn of the year 2008 is a more moving estimate of seasonality. As shown in Figures 3a and 3b, the increased variability is nevertheless moderate.

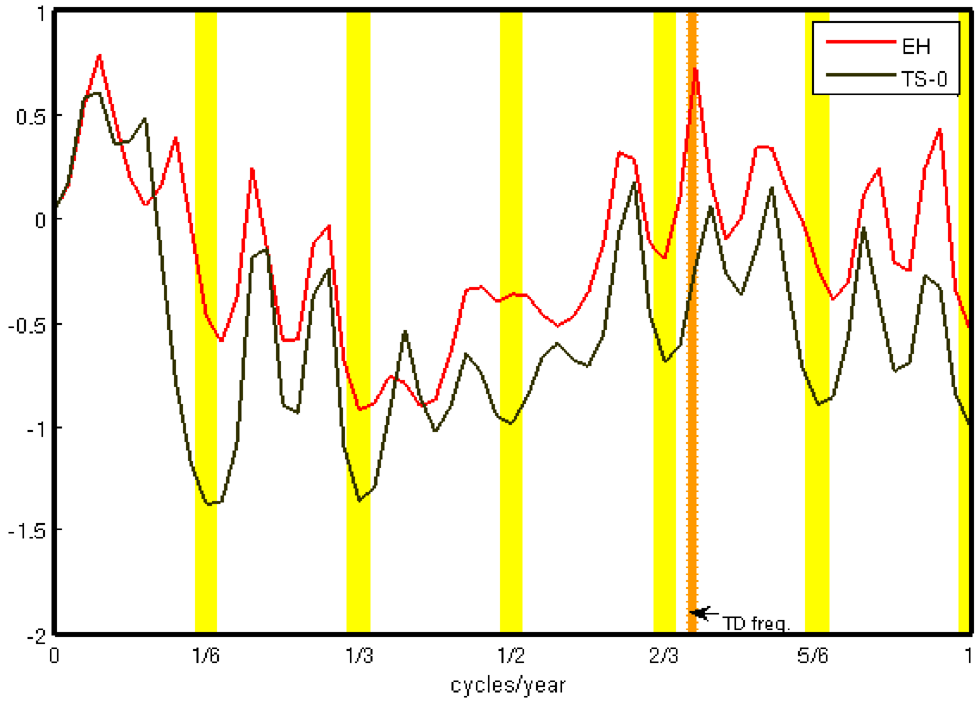
Spectra of the differenced SA series (in logs) are displayed in Figures 11a and 11b. The first one is a parametric AR(30) spectra, similar to the one in X12A. The second one is a non-parametric Tukey spectrum. The AR spectrum is appropriate for detecting peaks at the seasonal and TD frequencies (see Childers, 1976). The Tukey spectrum, on the other hand, is more informative in what concerns spectral minima. In the AMB approach, the r.h.s. of the equation for the historical estimator of the SA series (i.e., equation (10) exchanging s and n) contains the polynomial $\varphi_s(F)$. Hence non-stationary seasonal components induce zeros for seasonal frequencies in the spectrum of the SA historical estimator. Given that seasonal ARIMA models typically contain a seasonal differencing, all seasonal frequencies will present that feature.

For preliminary estimators, the spectral minima (at the seasonal frequencies) are close to (but not quite) zero. Thus the spectrum of an estimated SA series that includes preliminary estimators at both ends will be close to (though not strictly) non-invertible. Given that AR models are always invertible, the spectral minima will be poorly captured (and pushed away from zero, Maravall, 1995). The spectral dips (or near-zeros) at the seasonal frequencies are likely to be better captured with the Tukey spectrum. The EH SA series shows a slight distortion in the frequency of the local minimum near some seasonal frequencies (e.g., 3rd and 5th harmonics in the AR(30) spectrum) and the two spectra show that more variance has been removed at the seasonal frequencies with the TS-0 model, in accordance with its more moving seasonal. Accordingly, as Figure 6 showed, the EH SA series is more erratic.

Using the AMB SEATS decomposition of the EH model (as Figure 5 shows, very close to the X12A one), Figure 12 plots the finite-filter phase delay functions (adapted from the one provided to us by the X12A developers) of the EH and TS-0 concurrent estimators of the SA series. Although, as shown in Findley and Martin (2006), smoother SA series tend to imply larger phase delays, in our case the opposite result is obtained. Within the range of frequencies of cyclical interest, the TS-0 estimator induces a smaller phase delay, that extends from a few days (long-term cycles) to half-a-month (2-year cycle).



a) Non-parametric Tukey (log) spectrum.



b) AR (log) spectrum.

Figure 11: Spectrum of SA series (differenced and mean corrected).

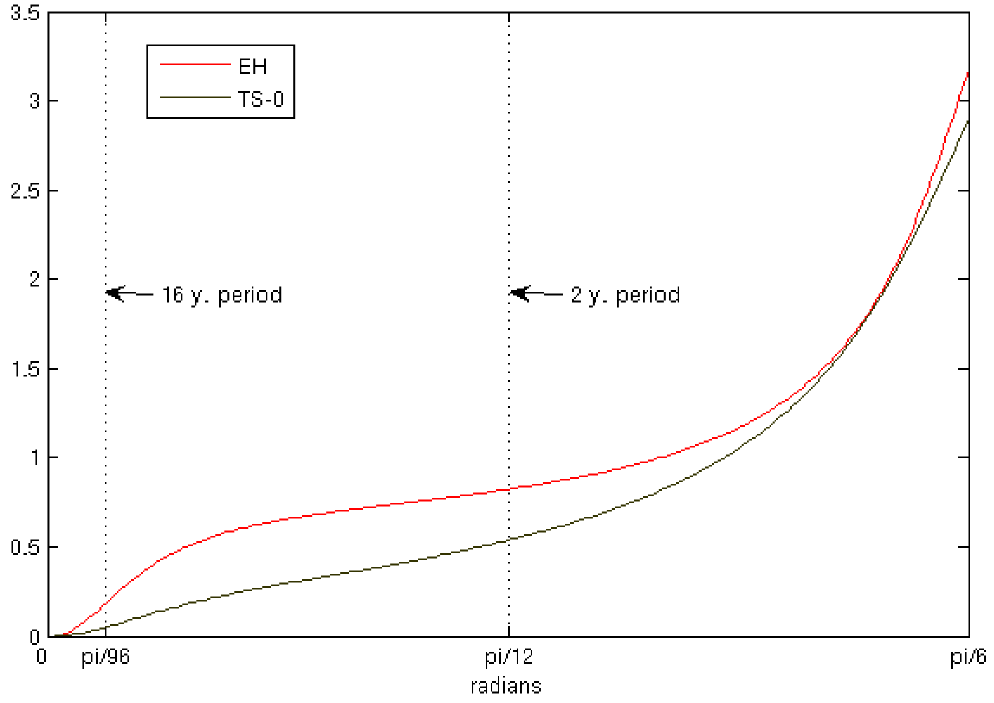


Figure 12: Phase delay of finite SA concurrent filter (in months).

6 Seasonal adjustment errors and revisions

It is difficult to assess the significance of the difference between two SA figures without having information on their estimation errors. TS provides SE of the SA series for the semi-infinite filter realization. (McElroy, 2008a, presents a matrix extension for the finite filter case; Bell, 2005, considers the effect of adding sampling and parameter estimation errors.) The SA series estimation error is the sum of two types of orthogonal errors. One is the error in the historical estimator, that typically characterizes the central years of the period considered. For months at both ends of the series, preliminary estimators will suffer revisions as they converge to the historical estimator. The difference between the preliminary and historical estimator is the revision error.

Let $x(t)$ follow model (2), and consider the decomposition of $\mathrm{x(t)}$ as in (4). Assume $\mathrm{s(t)}$ and $\mathrm{n(t)}$ follow the models (6a) and (6b). The error in the historical estimator, $\mathrm{s(t)} - \hat{\mathrm{s}}\mathrm{(t)}$, has the ACGF of the model

$$\theta(\mathrm{B})\mathrm{z(t)} = \theta_{\mathrm{n}}(\mathrm{B})\theta_{\mathrm{s}}(\mathrm{B})\mathrm{b(t)}, \qquad (12)$$

with $Var(b) = Var(a_n)\, Var(a_s)/Var(a)$; see Pierce (1979). As for the revision error, let $\mathrm{s(t|t+k)}$ denote the estimator of $\mathrm{s(t)}$ when $\mathrm{x(t+k)}$ is the last available observation. (When $\mathrm{k} < 0$, $\mathrm{s(t|t+k)}$ yields a forecast.) Model (10) can be expressed as the partial fractions decomposition

$$\hat{\mathrm{s}}\mathrm{(t)} = k_s \, \frac{\theta_{\mathrm{s}}(\mathrm{B})\theta_{\mathrm{s}}(\mathrm{F})\varphi_{\mathrm{n}}(\mathrm{F})}{\varphi_{\mathrm{s}}(\mathrm{B})\theta(\mathrm{F})}\,\mathrm{a(t)} = \left[\frac{\mathrm{M(B)}}{\varphi_{\mathrm{s}}(\mathrm{B})} + \mathrm{F}\frac{\mathrm{N(F)}}{\theta(\mathrm{F})}\right]\mathrm{a(t)}$$

(see Maravall, 1994, and Bell and Martin, 2004). The series expansion of the fraction in F is convergent; the one in B is not. Thus $\hat{\mathrm{s}}\mathrm{(t)}$ can be expressed as

$$\hat{\mathrm{s}}\mathrm{(t)} = \psi_{\mathrm{B}}(\mathrm{B})\mathrm{a(t)} + \psi_{\mathrm{F}}(\mathrm{F})\mathrm{a(t+1)}\,, \qquad (13)$$

where $\psi_{\mathrm{B}}(\mathrm{B})$ contains the effect of the starting conditions and of the innovations up to period t, and $\psi_{\mathrm{F}}(\mathrm{B}) = N(F)/\theta(F)$ is a convergent filter of innovations posterior to t. Because the expectation at time t of $\mathrm{a(t+k)}$ when $k > 0$ is zero, the concurrent estimator is $\hat{\mathrm{s}}\mathrm{(t|t)} = \psi_{\mathrm{B}}(\mathrm{B})\mathrm{a(t)}$, and its associated revision is $\mathrm{r(t|t)} = \hat{\mathrm{s}}\mathrm{(t)}\ \hat{\mathrm{s}}\mathrm{(t|t)} = \psi_{\mathrm{F}}(\mathrm{F})\ \mathrm{a(t+1)}$. Therefore, r(t|t) follows the model

$$\theta(\mathrm{F})\mathrm{r(t|t)} = \mathrm{N(F)a(t+1)}, \qquad (14)$$

from which its variance and autocorrelations can be computed.

The derivation extends easily to any preliminary estimator, including forecasts. In particular, and taking conditional expectations at times t_1 and t_2 $(t_2 > t_1 \geq t)$ in (13), the revision in the preliminary estimator $\hat{\mathrm{s}}(\mathrm{t}|\mathrm{t}_1)$ caused by the new observations $\mathrm{x}(\mathrm{t}_1 + 1), \ldots, \mathrm{x}(\mathrm{t}_2)$, is an MA($\mathrm{t}_2 - \mathrm{t}_1$) process (Pierce, 1980). (Finite sample extensions are considered in McElroy and Gagnon, 2008, and a model-based diagnostic based on them is developed in McElroy and Wildi, 2010.)

The ACGF of the total estimation error is the sum of the ACGFs of the historical and revision errors. All three ACGFs correspond to stationary ARMA models with the AR polynomial equal to $\theta(\mathrm{B})$. From the models, inferences can be drawn having to do with the precision of the estimators and their rates of growth. Further, some general properties can be derived: For example, under the standardization $Var(a) = 1$, when the model for $\mathrm{x(t)}$ contains the seasonal difference $(1 - \mathrm{B}^{12})$, if $\theta(\mathrm{B})$ contains the factor $(1 + \theta_{12}\mathrm{B}^{12})$ with θ_{12} close to -1, the model will present a stable seasonal component that will require small revision. But due to the close-to-unit root in $\theta(\mathrm{B})$, these revisions will be slow to converge to zero. On the contrary, when $\theta_{12} \approx 0$, the revision will tend to be large but converge fast.

For the TS-0 model, it is found that the concurrent estimator of the SA series follows the model

$$(1 - 0.78B)\, \nabla^2\, \hat{\mathrm{n}}(\mathrm{t|t}) = 0.72\, (1 - 0.53\mathrm{B} + 0.58\mathrm{B}^2 - 0.02\mathrm{B}^3)\, \mathrm{a(t)}\,, \qquad (15)$$

where the MA(3) polynomial has three real roots (0.92, 0.57, and 0.03). The model is close to (though not quite) non-invertible. The revision this estimator will suffer follows the model

$$(1 - 0.41F^{12})\, \mathrm{r(t|t)} = 0.37\, \theta_{\mathrm{r}}(\mathrm{F})\, \mathrm{a(t+1)}\,,$$

with $\theta_{\mathrm{r}}(\mathrm{F})$ a polynomial of degree 16, with three real roots (0.92, 0.55, and 0.14) and five pairs of complex conjugate roots. The SD of the total estimation error of the concurrent SA series estimator is equal to 0.69 percent points (p.p.) and a 95% confidence interval for its month-to-month rate of growth is given by (± 1.16 p.p.). Given that the historical estimation error is never observed, of more practical interest is the error with respect to the best measure that can eventually be achieved, i.e., the revision error. For the concurrent estimator, the SD of the revision is 0.60 p.p.; after one year 78% of its variance has been removed, and the concurrent estimator has practically converged to the historical estimator in two years. The revision error is relatively large, but converges to zero reasonably fast.

Figure 6 plots the 95% confidence interval of the TS-0 estimator of the SA series. The months with a significant disagreement between the EH and TS-0 estimators are numerous, more noticeably since the year 2007. The two adjustments seem “significantly different”.

7 Distribution of MMSE estimators

The M4 diagnostic, failed by the automatic option of X12A in Section 2, serves to illustrate model-based translation of an X12A quality assessment statistics. The M4 statistics indicates too much autocorrelation in the estimator of the irregular component. Not knowing what the proper autocorrelation should be, the diagnostics indicates that an estimator with autocorrelation beyond a certain amount would indicate poor quality. A similar diagnostic can be made in the model-based framework, but a more complete analysis of the irregular is now possible.

In general, TS decomposes $x(t)$ into the sum of orthogonal components all of which follow ARIMA-type models. Denoting by $u(t)$ the irregular component and by $s(t)$ the sum of all other components, the decomposition can be expressed as

$$x(t) = s(t) + u(t),$$

and a derivation similar to the one in Section 3.1 shows that the final estimator of $u(t)$ can be expressed as

$$\theta(F)\hat{u}(t) = k_u \varphi(F) a(t), \qquad (16)$$

with $k_u = \mathrm{Var}(u)/\mathrm{Var}(a)$ and the ACF of $\hat{u}(t)$ is that of the "inverse" model of (2) (Bell and Hillmer, 1984). For close to non-invertible models the autocorrelations can be remarkably high, although they will not be associated with trend or seasonal frequencies due to the spectral zeros induced by the unit roots in $\varphi(F)$.

SEATS estimates $u(t)$ as the residual after all other components have been removed. Let this estimator be $\tilde{u}(t)$. If the model is correct, the sample variance and ACF of $\tilde{u}(t)$ should be in agreement with the theoretical ones implied by (16), and this agreement can be assessed by means of the SE of the estimated autocorrelations. In this way simple tests are obtained, that are, in essence, misspecification tests (see Maravall, 1987, 2003).

The comparison between theoretical and empirical second-order moments of the estimator extends to other components. Thus, for example, using TS to decompose the EH model it is obtained that the variance of the stationary transformation of the estimated trend-cycle is 0.016, with $SE = 0.003$, while according to (10) –with s standing now for the trend- it should be 0.026. This discrepancy may indicate underestimation of the trend-cycle component. Findley, McElroy, and Wills, 2004, and McElroy, 2008b, show that the SEATS test –based on final estimators- are biased towards underestimation, and provide a finite sample improvement (not incorporated to SEATS yet).

For the TS-0 case, Table 2 compares the sample variance and autocorrelations of the irregular component estimator with those implied by model (16), together with the associated SE. No discrepancy is found.

Table 2: Irregular component: Autocorrelations and Variance.

Autocorrelation	Model for Estimator	Sample Estimate	SE
Lag-1	-0.35	-0.36	0.07
Lag-2	-0.08	-0.06	0.09
Lag-3	-0.23	-0.24	0.07
Lag-4	0.17	0.10	0.07
Variance	0.008	0.007	0.001

As already mentioned, the frequency domain expression of the WK filter that provides the final estimator of $s(t)$ is the ratio of the pseudo-spectrum of $s(t)$ to that of $x(t)$. Letting $f_*(\lambda)$ denote a pseudo-spectrum, it follows from (8) that

$$\mathrm{f}_{\hat{\mathrm{n}}}(\lambda) = \left[\frac{\mathrm{f_n}(\lambda)}{\mathrm{f_x}(\lambda)}\right]^2 \mathrm{f_x}(\lambda) = \left(\frac{\mathrm{f_n}(\lambda)}{\mathrm{f_x}(\lambda)}\right) \mathrm{f_n}(\lambda) \leq \mathrm{f_n}(\lambda) \ ,$$

$$\mathrm{f}_{\hat{\mathrm{s}}}(\lambda) = \left(\frac{\mathrm{f_s}(\lambda)}{\mathrm{f_x}(\lambda)}\right) \mathrm{f_s}(\lambda) \leq \mathrm{f_s}(\lambda) \ .$$

The difference $\mathrm{f_x}(\lambda) - [\mathrm{f}_{\hat{\mathrm{n}}}(\lambda) + \mathrm{f}_{\hat{\mathrm{s}}}(\lambda)]$ (> 0) is the cross-spectrum or, in the time domain, the cross-covariance function. Thus, although the components are assumed uncorrelated, MMSE yields correlated estimators. For the two component decomposition of (4), the crosscovariance generating function between $\hat{\mathrm{s}}(\mathrm{t})$ and $\hat{\mathrm{n}}(\mathrm{t})$ is the ACGF of the stationary ARMA model given by (12).

An implication is that the error in the historical estimator of $\mathrm{n}(\mathrm{t})$ (and of $s(\mathrm{t})$) is equal to the lag-0 crosscovariance between $\hat{\mathrm{n}}(\mathrm{t})$ and $\hat{\mathrm{s}}(\mathrm{t})$. (Thus, in the absence of explicit models for the series and components, the sample lag-0 crosscovariance could provide a rough estimate of the SA series historical estimation error variance. For the EIP case, the theoretical value of the lag-0 crosscovariance is 0.0034 while the sample estimate is 0.0040.)

The appearance of crosscovariances between the estimators of components assumed uncorrelated has often been signaled as an inconvenience of the standard unobserved components MMSE model-based approach. (Although the result would apply in general to any components estimated as linear functions of the same observations). In the TS approach, given that the crosscovariances between component´s estimators are always finite, when at least one of the components is non-stationary the corresponding crosscorrelations would tend to zero and hence their sample estimates will be small. If $d+d_s > 1$ – a condition satisfied by the vast majority of series that are seasonally adjusted – the convergence to zero will also be true for the crosscorrelations between the components rates of growth. Thus, in practice, the inconvenience will likely be of little relevance.

Given that typically the SA series and seasonal component are non-stationary, cross-correlations have to be computed for their stationary transformations. Proceeding as with the autocorrelations, the theoretical crosscorrelations between the component estimators should be close to their sample estimates. Table 3 exhibits the lag-0 theoretical and sample crosscorrelation for the TS-0 procedure. Except for the stochastic TD-irregular case, correlations are small and the comparison shows no anomalies. The relatively high correlation between the stochastic TD and irregular component may cause instability in the detection of the former.

Table 3: Crosscorrelation between stationary transformation of estimators.

	Model for Estimator	Sample Estimate	SE
Trend-cycle/Seasonal	-0.09	-0.15	0.05
Seasonal/Irregular	0.06	0.06	0.02
Trend-cycle/Irregular	-0.17	-0.16	0.03
Seasonal/Stoch. TD	0.07	0.08	0.01
Trend-cycle/ Stoch. TD	0.05	0.05	0.01
Stoch. TD/Irregular	0.79	0.78	0.02

Remark: The role of the moving average polynomial in the observed series model

In the model-based approach, while the models for the series and the components are models in B (i.e., the past explains the present), those of the revision error and irregular estimator are models in F (the future explains the present). As expressions (12), (14), and (16) indicate, the historical estimation error, the revision, and the irregular component estimator (the residual of the decomposition) are stationary and autocorrelated (although their forecast will always be 0). In the three cases, the ACF is that of an ARMA model with AR polynomial equal to $\theta(B)$, the MA polynomial of the original series model.

The autocorrelation functions of the “errors” will have different starting conditions, as a result of the different MA parts of models (12), (14), and (16), after which they will follow the same difference equation. Therefore, a slow converging revision error, for example, will be associated with a highly autocorrelated irregular component estimator. The $\theta(B)$ polynomial is also the AR polynomial in (9), the model whose ACF is the WK filter; its roots will determine the speed of convergence of the filter, i.e., how many terms are needed for a finite truncation.

8 Diagnostics and quality assessment

The diagnostics of the previous section are based on statistics that, if the regARIMA model is correctly specified, should be close to some theoretical values. Ultimately, these comparisons are misspecification tests. Obviously, they have implications for seasonal adjustment. For example, a high value of the seasonal autocorrelation QS statistics in the model residuals indicates that the model has not captured seasonality properly. Or, if the sample and theoretical variances of the SA estimator are in clear disagreement, one may suspect some adjustment failure.

It may happen, however, that with no evidence of misspecification, the decomposition may be questionable. As a simple example, the model

$$(1 + \phi_{12}\, B^{12})\, x_t = a_t$$

with $\phi_{12} = -0.3$ has peaks in the spectrum at all seasonal frequencies, but the seasonal correlation is small and, in practice, only lasts one year. The AMB estimator would imply (as it should) a very erratic seasonality, unrecognizable as such through direct inspection. Would it make sense to seasonally adjust the series? Possibly not, because the concept of seasonality would seem to be associated with longer lasting inertia. In other words, there are criteria that should be satisfied on a priori grounds by a seasonal component or a SA series for the adjustment to be justified. Some important criteria are the following.

(1) The seasonal correlation should last for several years. In TSW, this criterion would be satisfied if $d_s = 1$ or $\phi_{12} < k$ $(k = -0.5$ by default).
(2) The seasonal component should not move too fast. Given that the moving features are caused by the seasonal innovation, this could translate into $\mathrm{Var}(a_{st}) < k\ \mathrm{Var}(a)$.
(3) Historical estimation of seasonality should be reliable and revisions in preliminary estimators should not be excessive. In TSW these requirements could set also a limit to the variance of the final and preliminary estimators.
(4) The delay induced by the phase shift in the concurrent SA series estimators in the frequency range of cyclical interest should not be too large.

One could easily introduce additional criteria (for example, on the autocorrelation of the irregular estimator, on the crosscorrelation between trend-cycle and seasonal component, on the speed of convergence of the revision, etc.) These quality measures, however, are judgmental. In the seasonal AR(1) example, most people would agree with not adjusting when $\phi_{12} = -0.3$, but what if $\phi_{12} = -0.8$? The cutting point is unclear.

In the model-based approach, the variances of the component innovation, estimation error, and revisions, as well as the revision speed of convergence, depend on the proper model for the series. Thus, for example, some series will require large revisions, other will require small ones, and the vague criterion of minimizing revisions is replaced by the criterion of “optimal revisions,” tailored to the stochastic properties of the series.

The characteristics of the adjustment, however, can be of help at the model selection stage. If seasonal adjustment is the application of interest and several models seem to provide acceptable fits, looking for smaller estimation error and smaller revision of the SA series would be sensible. An example is provided in the next section.

9 Outliers and intervention variables

9.1 TRAMO-SEATS automatic with level shifts

The EH and TS-0 models do not provide a satisfactory adjustment. The TS-0 model overcomes some of the problems but does a poor job with the last 12 months of observations. The large negative residuals at the beginning of the recent recession and the large positive ones during the recovery period suggests the convenience of introducing outliers to capture the big drop at the turn of the year 2008. Thus the two LS outliers of the EH model are specified and automatic model identification is performed. The resulting model will be referred to as the “TS-2LS model”.

The ARIMA model obtained is very similar to the TS-0 case. The model orders are $(3,1,0)\ (0,1,1)_{12}$ and calendar effects are significant. The regular AR(3) polynomial presents the same pair of complex conjugate roots, with similar moduli, and frequency close to the TD one. Table 1 indicates that the residuals are not autocorrelated and can be accepted as Normally distributed with zero skewness and no excess kurtosis. However, the randomness in signs of residuals test is again failed, and Figure 4c shows that the residuals for the last two years exhibit the unpleasant pattern of the EH and TS-0 residuals. Still, a slight improvement is detected in the out-of-sample performance for the last year of observations in Table 1.

The EIP series provides a good example of the limitations of the automatic outlier detection procedure (based on AO, TC, and LS outliers) when the series experiences explosive behavior and increased uncertainty in the recent year(s). The TS automatic procedure was applied to the first $(243 - \mathrm{j},\ \mathrm{j} = 0,1,\dots,100)$ observations of the series. For series that end before Nov 2008, only two marginally significant outliers are detected. Both last for a few periods, and eventually disappear. Then, the outlier for Nov 2008 is concurrently detected as an AO and becomes an LS one month later. But, after Nov 2008 outlier detection becomes unstable, with LS outliers for Oct, Nov, and Dec 2008, and for Jan 2009, appearing and disappearing in several combinations. These outliers disappear throughout 2009, possibly as a result of the accumulation of abnormally large values at the end of the series. The three types of outliers considered have a difficult time capturing the crisis in a parsimonious way. In the example, to model the recent crisis as sudden drops in two particular months seems somewhat inadequate.

9.2 TRAMO-SEATS automatic with ramps

Direct inspection of the trend-cycle component obtained with the TS-0 model shows a spectacular drop in the series, that starts at about March 2008 and lasts for about one year. Moreover, the rate of decrease is different for the first and last months within that period. These considerations suggest that ramps may be more appropriate than LS to capture the recession. In TS, inclusion of outliers in the model is seen as a price paid for maintaining the Normality assumption that underlies the likelihood function (also needed for testing and inference). Intervention variables are intended to capture some known special effect. Outliers should be kept to a minimum and do not require an ex-post explanation. Often, however, an explanation is found, in which case it makes sense to replace the outlier with an intervention variable. Thus two ramps are entered as intervention variables, to capture the drop at the beginning of the crisis.

From a first guess provided by the trend-cycle inspection, a comparison of BICs (or likelihoods) for various ramps specifications leads to the choice of March and September as the starting dates, and of 5 and 7 months for the lengths, respectively. TS was run in automatic mode ($RSA = 5$) with the two ramps. As before, the ARIMA model obtained is a $(3,1,0)\ (0,1,1)_{12}$ for the logs, with a 6-variable TD, Easter Effect, and the two specified ramps. The AR(3) polynomial contains a pair of complex conjugate roots with modulus 0.70 and frequency 2.22 that produce again a stochastic TD with an ARMA(2,2) model. The complete model will be referred to as the "TS-2R model".

Table 1 shows that, of the 4 models considered, the TS-2R case produces the best results, and the RMSE of the out-of-sample forecast for the last 12 months shows a remarkable improvement. Figure 4d plots the TS-2R model residuals: the abnormal behavior at the end of the series is no longer present.

The model for the SA series $\mathrm{n(t)}$ is found to be

$$(1 - 0.53\mathrm{B})\nabla^2\mathrm{n(t)} = (1 - 1.49\mathrm{B} + 0.62\mathrm{B}^2 - 0.09\mathrm{B}^3)\mathrm{a_n(t)},$$

with innovation standard deviation $\sigma_{\mathrm{n}} = 0.0057$. Noticing that one of the roots of the MA polynomial is (1-0.92B), the model for the SA series is close to a $(1,1,2)$ ARIMA model plus mean. (Interestingly enough, these are the same orders as those of the regular part of the model identified by X12A in Section 2.)

Some relevant features of the seasonal adjustment for the three TS cases considered are contained in Table 4. First, given that the standard deviation of the innovation in a component measures stability (close to zero innovation variances imply close to deterministic components), the TS-2R model provides the most stable SA series and seasonal component. The innovation standard deviation also provides a measure of the size of the one-period-ahead forecast error of the component.

Table 4: Features of the decomposition.

Model	SE of innovation (in 10^{-2})			Concurrent estimator (in p.p.)		% Convergence of revision in one year	SE of m-to-m SA rog (in 10^{-2})
	Seasonal	Stoch. TD	SA series	SE of Estimator	SE of Revision		
TS-0	0.44	0.18	0.68	0.69	0.60	76	0.58
TS-2LS	0.43	0.18	0.63	0.61	0.52	78	0.56
TS-2R	0.41	0.17	0.57	0.54	0.44	78	0.55

Second, estimation precision of the SA series can be measured by the SE of its concurrent estimator. Table 4 shows that the TS-2R model yields the most precise estimator. For the month-to-month rate of growth, the width of a 95% confidence interval is about $\pm$ 1p.p.

Third, an additional desirable feature of the SA series is that they require small revisions. TS measures the SE of the revisions in preliminary estimators of the SA series. Table 4 shows that the TS-2R model has the smallest revisions in the concurrent estimator; its variance nevertheless is three times the variance of the historical estimation error. Convergence is practically unaffected: for the three TS models, close to 80% of the revision is completed in the first year; 100% during the second.

Concerning seasonal adjustment of the EIP series, what the previous analysis has shown is that improvements in the regARIMA model improve the adjustment, that some intervention is needed to deal with the recent crisis, and that ramps do a better job than LS outliers. Judging from the results obtained with the TS-2R model for the original period of concern (Oct 2009-Mar 2010), Table 5 indicates that Eurostat overestimated the SA rates of growth, while the EH model underestimated them.

Table 5: Month-to-month SA series rates of growth (in p.p.) for Oct 2009-Feb 2010.

(March estimate)	Oct	Nov	Dec	Jan	Feb
EH	0.0	0.2	0.5	0.4	0.2
Eurostat	0.3	1.3	0.8	1.9	0.7
TS-2R	0.3	0.7	0.6	0.9	0.4

Besides seasonal adjustment, other components can be of help in forecasting and cyclical assessment. Information similar to that having to do with the SA series is provided in TSW for the trend-cycle component, thus the relative merits of both components can be compared. In our example, the trend-cycle component is the SA series without a relatively small white-noise irregular component. Except for a mild improvement in the SE of the monthly rate of growth, and a mild increase in the phase delay (see Figure 13), the trend-cycle yields very similar results.

Further, the trend-cycle component (or the SA series) can be decomposed in turn into long-term trend and cycle by means of a model-based adaptation of the Hodrick-Prescott filter (Kaiser and Maravall, 2005). This decomposition is shown in Figure 14.

Models for these new components –compatible with the overall model for the series- are also derived and, for example, the model for the cycle is

$$(1 - 1.92\mathrm{B} + 0.93\mathrm{B}^2)\,(1 - 0.53\mathrm{B})\,\mathrm{z}(\mathrm{t}) = (1 - 0.26\mathrm{B} - 0.95\mathrm{B}^2 + 0.32\mathrm{B}^3)\mathrm{a_z}(\mathrm{t})\,,$$

with $\mathrm{Var}(\mathrm{z}) = 0.04\mathrm{Var}(\mathrm{a})$. These models permit the analyst to extend the analysis of the previous sections (e.g., distribution of estimators, of estimation and revision errors, and forecasting) to the long-term trend and cycle components.

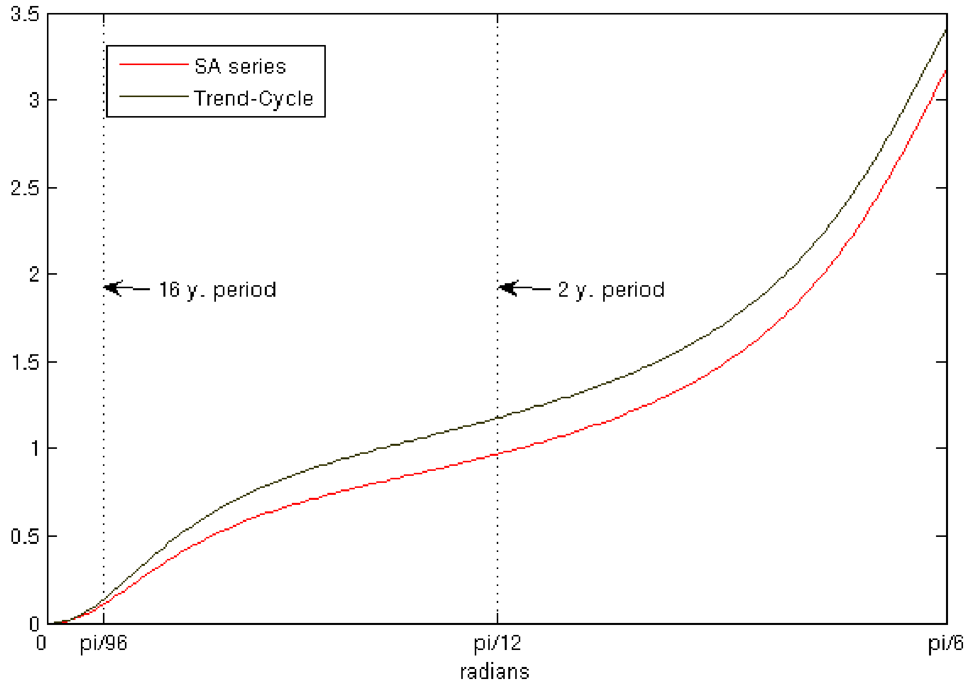


Figure 13: Phase delay of finite SA series and trend-cycle concurrent filters: TS-2R model (in months).

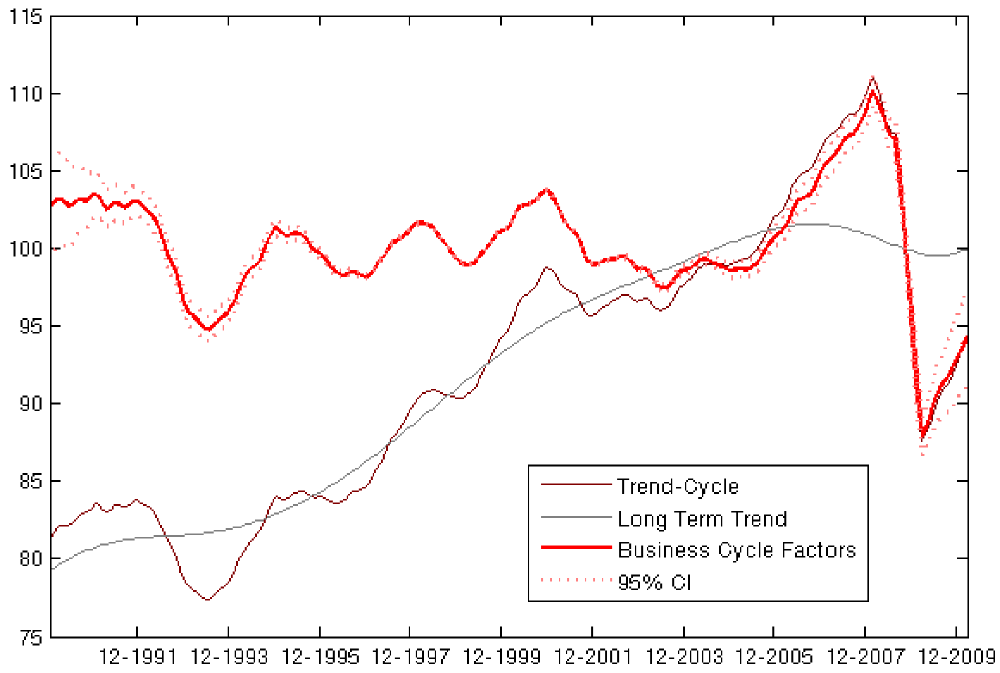


Figure 14: Long-term trend and business-cycle factors.

10 Stability of the model-based method

An objection often made to model-based adjustment is that changes in the model as new observations are added may produce instability in the results. To maintain a balance between instability and misspecification, it is often recommended that, as a rule, the model be re-identified perhaps once a year and, for the periods before the next re-identification, parameters be re-estimated every month, keeping fixed the general model specifications.

Since work in this paper was started, 9 new observations have become available (period April-December 2010) and we use them to analyze out-of-sample stability of the results. Unfortunately, each new release includes revisions of the unadjusted series for previous periods. Typically, the revision affects the most recent months, although in May 2010 it affected the entire series and exhibited a strongly seasonal pattern for the first half (see Figure 15).

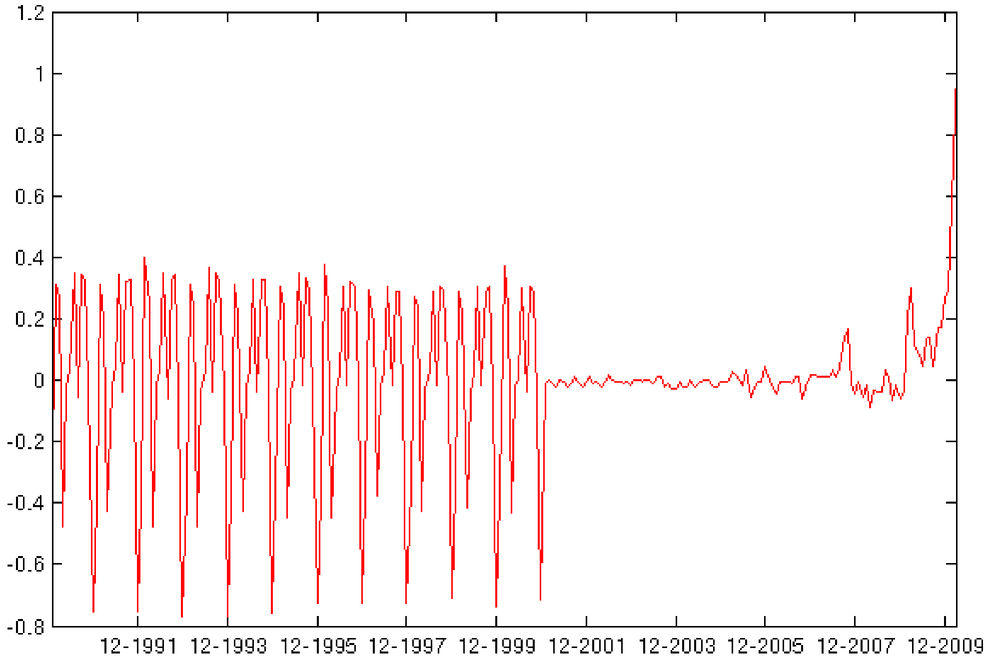


Figure 15: Revisions in unadjusted series (p.p.): period Apr-Dec 2010.

Adding the two ramps of the previous section, the automatic procedure was applied to the series finishing in each of the ten months for March-December 2010. (The sequence of concurrent releases of the series were used, thus revisions in unadjusted data were ignored.) The ten series produce the same model (for the logs), namely, the ARIMA $(3,1,0)\ (0,1,1)_{12}$ model with 6 TD variables, Easter effect, and no outlier. A summary of results is presented in Table 6.

The model fit, parameter estimates, residual diagnostics, and basic features of the decomposition barely change. Figure 16 plots the one-period-ahead forecast errors of the (log) series for the nine new months; they are small and random, with $\mathrm{SE} = 0.0130$, in line with the in-sample residual SE. Despite the uncertainty created by the crisis and despite the revisions in the unadjusted data, automatic re-estimation every month provides stable and acceptable results. The results are identical to those that would have been obtained with the rule mentioned above to balance stability versus misspecification.

Table 6: Results of monthly automatic model re-identification. Interval of variation for period March-November 2010

(a) Model fit

	SE(res)	BIC	Real AR root	Complex AR root		Seasonal MA parameter	Ramp coeff (t-val)	
				mod	period		1st ramp	2nd ramp
min	0.0110	-8.75	0.52	0.70	2.83	-0.37	-3.8	-11.1
max	0.0111	-8.77	0.54	0.73	2.85	-0.39	-4.0	-11.4

(b) Residual diagnostics

	Q	N	Sk	Kur	Q2	Runs	QS	RMSE (in 10^{-2}) (OS forecast)
min	26.0	1.0	0.6	0.8	15	1.2	0.2	0.7
max	27.6	1.8	0.7	1.2	16.7	1.8	0.3	1.0
CV	31.4	6	2	2	34	2	6	1.8

(c) Features of the adjustment

	SE of innov (in 10^{-2})			SE of concurrent estimator (pp)	SE of revision (pp)	% convergence of revision in one year	SE of m-to-m SA rog (in 10^{-2})
	Seasonal	Stoch TD	SA series				
min	0.39	0.17	0.56	0.53	0.43	75	0.54
max	0.41	0.18	0.58	0.54	0.44	78	0.56

11 Conclusion

In the context of a recent debate among some European statisticians having to do with seasonal adjustment of the Euro Area Industrial Production Index, it is shown how the regARIMA model-based method of programs TRAMO-SEATS can be efficiently used, the results can be analyzed, and the quality of the adjustment –as well as the information the adjustment provides- can be improved.

In particular, it is seen how, adding two ramps to deal with the abnormal drop at the beginning of the recent crisis, automatic use of the programs (that includes automatic model identification) provides excellent and stable results (despite non-trivial revisions in the unadjusted data).

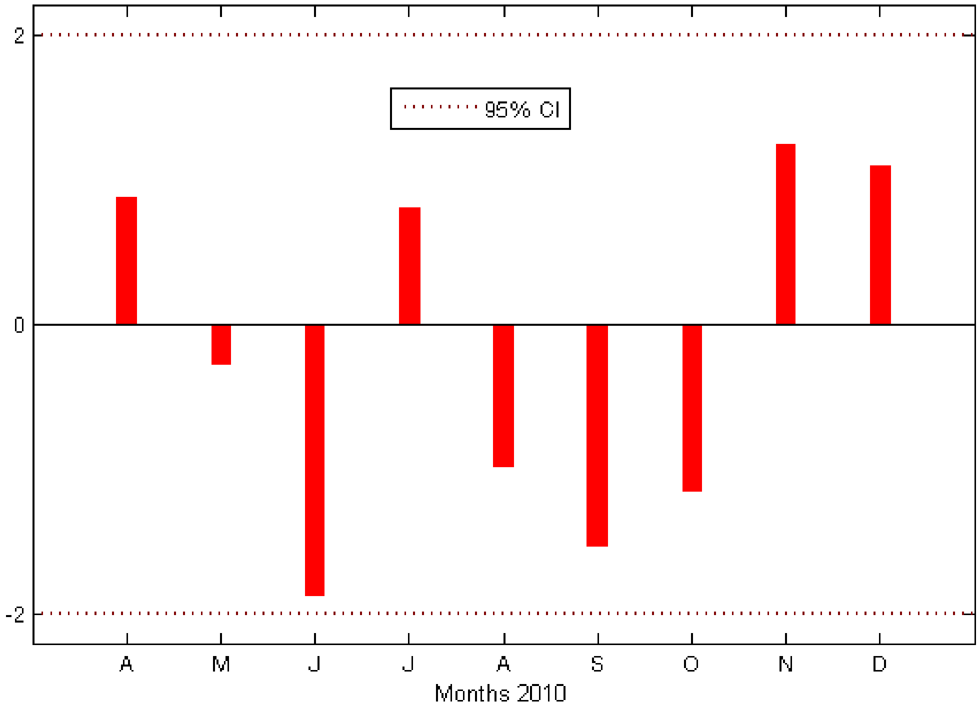


Figure 16: Out-of-sample forecast errors: period Apr-Dec 2010.

BIBLIOGRAPHY

BELL, W.R. (2005). Some consideration of seasonal adjustment variances. *2005 Proceedings of the American Statistical Association*: Alexandria, VA.

– (2004). On RegComponent time series models and their applications. In Harvey A.C., Koopman S.J. and Shepard, N. (eds.): *State space and unobserved component models: Theory and applications*, 248-283. Cambridge, UK: Cambridge University Press.

BELL, W.R. and HILLMER, S.C. (1984). Issues involved with the seasonal adjustment of economic time series. *Journal of Business and Economic Statistics*, 2:291-320.

BELL, W.R. and MARTIN, D.E.K. (2004). Computation of asymmetric signal extraction filters and mean squared error for ARIMA component models. *Journal of Time Series Analysis*, 25:603-623.

BLAKELY, C. and MCELROY, T. (2010). *An empirical evaluation of signal extraction goodness –of- fit diagnostic tests*. Mimeo. Statistical Research Division. U.S. Bureau of the Census.

BURMAN, J.P. (1980). Seasonal adjustment by signal extraction. *Journal of the Royal Statistical Society*, A, 143:321-337.

CHILDERS, D.G. (ed.) (1976). *Modern spectrum analysis*. New York: IEEE Press.

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.

EIGLSPERGER, M. and HAINE, W. (2010). *Seasonal adjustment of Euro Area GDP and industrial production and the statistical treatment of the crisis. Workshop on the Implementation of ESS Guidelines on Seasonal Adjustment*. Mimeo. European Central Bank. May 2010.

EUROSTAT (2009). *ESS guidelines on Seasonal Adjustment*. Luxembourg: Office for Official Publications of the European Communities.

FINDLEY, D.F. (2005). Some recent developments and directions in seasonal adjustment. *Journal of Official Statistics*, 21(2):343-365.

FINDLEY, D.F. and MARTIN, D.E.K. (2006). Frequency domain analyses of SEATS and X-11/12-ARIMA seasonal adjustment filters for short and moderate-length time series. *Journal of Official Statistics*, 22(1).

FINDLEY, D.F., MCELROY, T.S. and WILLS, K.C. (2004). Modifications of SEAT's diagnostic for detecting over –or underestimation- of seasonal adjustment decomposition components. *2004 Proceedings of the American Statistical Association*. Alexandria, VA.

FINDLEY, D.F., MONSELL, B.C., BELL, W.R., OTTO, M.C. and CHEN, B.C. (1998). New capabilities and methods of the X12A ARIMA seasonal adjustment program (with discussion). *Journal of Business and Economic Statistics*, 12:127-177.

GÓMEZ, V. and MARAVALL, A. (2001b). Seasonal adjustment and signal extraction in economic time series. In Peña D., Tiao G.C. and Tsay, R.S. (eds.): *A course in time series analysis*, chapter 8. New York: J. Wiley and Sons.

– (2001a). Automatic modeling methods for univariate series. In Peña D., Tiao G.C. and Tsay, R.S. (eds.): *A course in time series analysis*, chapter 7. New York: J. Wiley and Sons.

HARVEY, A.C. (1989). *Forecasting, structural time series models and the Kalman filter*. Cambridge, UK: Cambridge University Press.

HILLMER, S.C. and TIAO, G.C. (1982). An ARIMA model based approach to seasonal adjustment. *Journal of the American Statistical Association*, 77:63-70.

JENKINS, G.M. and WATTS, D.G. (1968). *Spectral analysis and its applications*. San Francisco: Holden Day.

KAISER, R. and MARAVALL, A. (2005). Combining filter design with model-based filtering: An application to business-cycle estimation. *International Journal of Forecasting*, 21:691-710.

MARAVALL, A. (2003). A class of diagnostics in the ARIMA-model-based decomposition of a time series. *Seasonal Adjustment*. European Central Bank. November 2003.

– (1998). Comment on the X12ARIMA seasonal adjustment method. *Journal of Business and Economic Statistics*, 16:155-160.

– (1995). Unobserved components in economic time series. In Pesaran, H. and Wickens, M. (eds.): *The Handbook of Applied Econometrics*, chapter 1, pages 12-72. Oxford: Basil Blackwell.

– (1987). On minimum mean squared error estimation of the noise in unobserved component models. *Journal of Business and Economic Statistics*, 5:115-120.

McELROY, T. S. (2008a). Matrix formulas for non-stationary signal extraction. *Econometric Theory*, 24:1-22.

– (2008b). Statistical properties of model-based signal extraction diagnostic tests. *Communications in Statistics, Theory and Methods*, 37:591-616.

McELROY, T. S. and WILDI, M. (2010). Signal extraction revision variances as a goodness-of-fit measure. Statistical Research Division Research Report. March 2010.

McELROY, T. S. and GAGNON, R. (2008). Finite sample revision variances for ARIMA model-based signal extraction. *Journal of Official Statistics*, 24:451-467.

MONSELL, B.C. (2009). Update on the development of X-13ARIMA-SEATS. *Joint Statistical Meetings Proceedings*. Alexandria: VA.

PIERCE, D.A. (1980). Data revisions in moving average seasonal adjustment procedures. *Journal of Econometrics*, 14:95-114.

– (1979). Signal extraction error in non-stationary time series. *Annals of Statistics*, 7:1303-1320.