Bernhard Hammer
0103294

Klaus Prettner
0201418

Seasonal Adjustment Programs

Chapter 4:
"The Econometric Analysis of Seasonal Time Series"
(Ghysels, Osborn, 2001)

1. Introduction

A procedure that filters the seasonal fluctuations from a time series is called a seasonal adjustment program. There exist a lot of different procedures but most of the statistical agencies use quite standardized techniques where the most important ones are the Census X-11 and its successor the Census X-12 program and in addition the TRAMO/SEATS procedure. We will discuss the main features of those techniques in detail.
First of all it should be mentioned that all seasonal adjustment procedures rely on a decomposition of the data series into orthogonal unobserved components which can be done in a lot of different ways and in more or less detailed form. We will present the two most common methods of decompositions used by seasonal adjustment programs which are also quite intuitively clear: In both cases the original series is decomposed into a trend and cycle component, y_t^{tc}, into a seasonal component, y_t^s, and into an irregular noise component, y_t^i, but in the first case the different components are multiplied with each other whereas in the second case they are summed:

$y_t = y_t^{tc} * y_t^s * y_t^i$ multiplicative version, (m)

$y_t = y_t^{tc} + y_t^s + y_t^i$ (log) additive version, (a)

If in the second case the data is in logarithmic form one refers to this case as the log additive version instead of the additive version. There also exists a pseudo additive decomposition, which is better than a multiplicative one, if in some months parts of the series have zero values. However lots of extensions can be made to those simple decompositions, for example to account for deterministic effects like holidays, trading days, length of months etc. but this is not necessary to understand the principles of seasonal adjustment programs.

After a series has been decomposed into estimates of the different parts, seasonal adjustment is done by dividing the series by an estimate of the seasonal component in the multiplicative case and by subtracting this estimate in the (log) additive case.

2. Census X-11

The Census X-11 procedure is the most widely used seasonal adjustment program. Its roots go back to the early 1930's and it was developed during the next decades until in the 1960's it was really an operable algorithm. Basically Census X-11 consits of different MA- filters that are sequentially applied to the data. All in all the X 11 Program can be approximated by a linear filter if no extreme value corrections have to be performed and if forecasting and backcasting is unnecessary since enough data points are available. This is the case for the default options and in Census X-12 there are new procedures implemented that deal with these two possible sources for nonlinearities.
The basic algorithm can be described by the following three steps in case of a monthly cycle, the seasonal cycle is treated in a similar way:

1. Initial estimates:

First of all an initial estimate for the trend (business)cycle component is obtained by the following centered MA filter, where SM refers to the fact that it is a monthly series:

$$y_t^{tc}(1) = SM(L)y_t = \left(\frac{1}{24}\right)(1+L)(1+L+L^2+...+L^{11})L^{-6}y_t \qquad (1)$$

which seemed to be more accessible to us in following form:

$$y_t^{tc}(1) = SM(L)y_t = \left(\frac{1}{24}\right)\left[(L^{-6}+L^{-5}+...+L^5)y_t + (L^{-5}+L^{-4}+...+L^6)y_t\right] \qquad (2)$$

Next the initial seasonal + noise part can be obtained for the two different representations as:

$$\begin{aligned} y_t^{si}(1) &= \frac{y_t}{y_t^{tc}(1)} \qquad (m) \\ y_t^{si}(1) &= y_t - y_t^{tc}(1) \qquad (a) \end{aligned} \qquad (3)$$

Now it is possible to calculate a first estimate of the seasonal part by applying the following filter for S = 12:

$$\tilde{y}_t^s(1) = M_1(L)y_t^{si}(1) = \left(\frac{1}{9}\right)\left(L^S + 1 + L^{-S}\right)^2 y_t^{si}(1) \tag{4}$$

However in this expression the seasonal components do not sum up to unity therefore the filter is applied once again to get the initial seasonal factor as:

$$\begin{aligned} y_t^s(1) &= \frac{\tilde{y}_t^s(1)}{SM(L)\tilde{y}_t^s(1)} && (m) \\ y_t^s(1) &= \tilde{y}_t^s(1) - SM(L)\tilde{y}_t^s(1) && (a) \end{aligned} \tag{5}$$

After this initial seasonal factor is obtained one can calculate the initial seasonal adjustment as:

$$\begin{aligned} y_t^{sa}(1) &= \frac{y_t}{\tilde{y}_t^s(1)} && (m) \\ y_t^{sa}(1) &= y_t - \tilde{y}_t^s(1) && (a) \end{aligned} \tag{6}$$

2. Intermediate estimates:

The first thing that has to be done is to detrend the initially obtained series by a Henderson filter of the order H. The choice of H will be described below but the default value is H = 6 leading to a thirteen-term (2H+1) Henderson MA-filter of the following form:

$$\begin{aligned} & y_t^{tc}(2) = HM(L)y_t^{sa}(1) = \\ & [-0.019L^6 - 0.028L^5 + 0.066L^3 + 0.147L^2 + 0.214L^1 + 0.240L^0 + \\ & + 0.214L^{-1} + 0.147L^{-2} + 0.066L^{-3} - 0.028L^{-5} - 0.019L^{-6}]y_t^{sa}(1) \end{aligned} \tag{7}$$

where the lags and leads of order 4 have a weight of 0. Afterwards the seasonal adjustment can be refined by the following steps.
The second seasonal + noise part is obtained as:

$$\begin{aligned} y_t^{si}(2) &= \frac{y_t}{y_t^{tc}(2)} && (m) \\ y_t^{si}(2) &= y_t - y_t^{tc}(2) && (a) \end{aligned} \tag{8}$$

Analogous to the first estimate of the seasonal part the second preliminary estimate is calculated by applying the seasonal MA- filter:

$$\tilde{y}_t^s(2) = M_2(L) y_t^{si}(2) = \left(\frac{1}{15}\right)\left(\sum_{j=-1}^{1} L^{jS}\right)\left(\sum_{j=-2}^{2} L^{jS}\right) y_t^{si}(2) \qquad (9)$$

where again the seasonal components do not sum up to unity and therefore the procedure is applied again:

$$\begin{aligned} y_t^s(2) &= \frac{\tilde{y}_t^s(2)}{SM(L)\tilde{y}_t^s(2)} && (m) \\ y_t^s(2) &= \tilde{y}_t^s(2) - SM(L)\tilde{y}_t^s(2) && (a) \end{aligned} \qquad (10)$$

yielding the second seasonal factor. Therefore the second seasonal adjustment can be done by:

$$\begin{aligned} y_t^{sa}(2) &= \frac{y_t}{y_t^s(2)} && (m) \\ y_t^{sa}(2) &= y_t - y_t^s(2) && (a) \end{aligned} \qquad (11)$$

In principle the algorithm could stop here but fort he sake of completeness also the last step is described which enables to calculate the final trend-(business)cycle and the irregular component.

3. Final Henderson Trend and Final Irregular

Once more the order of the Henderson filter has to be chosen, but this doesn't need to coincide with the previously chosen value of H. Then the final trend-(business)cycle component is:

$$y_t^{tc}(3) = HM(L) y_t^{sa}(2) \qquad (12)$$

and the final irregular part becomes:

$$\begin{aligned} y_t^i(3) &= \frac{y_t^{sa}(2)}{y_t^{tc}(3)} && (m) \\ y_t^i(3) &= y_t^{sa}(2) - y_t^{tc}(3) && (a) \end{aligned} \qquad (13)$$

altogether yielding the final estimated decomposition:

$$y_t = y_t^{tc}(3) * y_t^s(2) * y_t^i(3) \qquad (m)$$
$$y_t = y_t^{tc}(3) + y_t^s(2) + y_t^i(3) \qquad (a) \qquad (14)$$

To summarize, all three steps together lead to the linear approximation of the monthly X-11 filter with the following weights for a total of 68 lags and leads:

Table 4.1. *Filter weights of the linear monthly X-11 filter*

Lags and leads					
0	0.819	23	0.011	46	−0.003
1	0.019	24	−0.121	47	−0.005
2	0.018	25	0.013	48	−0.005
3	0.017	26	0.013	49	−0.003
4	0.016	27	0.013	50	−0.001
5	0.015	28	0.013	51	0.002
6	0.014	29	0.012	52	0.003
7	0.013	30	0.008	53	0.003
8	0.014	31	0.005	54	0.002
9	0.015	32	0.004	55	0.001
10	0.018	33	0.003	56	0.000
11	0.020	34	0.003	57	−0.001
12	−0.179	35	0.004	58	−0.001
13	0.021	36	−0.063	59	−0.001
14	0.020	37	0.005	60	−0.001
15	0.018	38	0.007	61	−0.001
16	0.016	39	0.008	62	−0.001
17	0.015	40	0.008	63	0.001
18	0.012	41	0.008	64	0.001
19	0.009	42	0.005	65	0.001
20	0.009	43	0.002	66	0.000
21	0.009	44	0.001	67	0.000
22	0.010	45	−0.001	68	0.000

Source: Ghysels (1984, Table A.3.3.).

Figure 2.1: weights of the monthly linear X-11 filter (Ghysels, Osborn, 2001; p. 99)

One can easily see that every lag and lead that is a multiple of 12 has a negative sign. A similar table for the quarterly X-11 filter is presented in Ghysels and Osborn (Ghysels, Osborn, 2001; p. 102) as well.

2.1 Possible Sources of Nonlinearity

All in all in earlier stages of development seasonal adjustment procedures were seen as linear data transformations in earlier stages of the development of seasonal filters. However recently this question is addressed again and the answer is not so clear anymore. There are some possible sources of nonlinearity in the X-11 procedure:

- If the multiplicative decomposition is used, as it is often the case, X-11 uses arithmetic means rather than geometric means

- There is an outlier detection algorithm implemented that basically consists of rules of thumb for the confidence bands whether observations should be included or not
- There is an automatic algorithm that chooses the order of the Henderson filter as new raw data is added (this algorithm applies in both cases when the Henderson filter is used) Also this algorithm is based on rules of thumb
- If disaggregated data is available and one is interested in the aggregate a possible source of nonlinearity is, that the disaggregated data is already seasonally adjusted, and afterwards the aggregate of those series is obtained. If the seasonal adjustment procedure is linear and uniform the two operations are interchangeable, otherwise additional nonlinearities are introduced.

2. Census X-12

The core procedure of the Census X-12 procedure is the same as those of Census X-11 but some improvements were implemented:

- A better treatment of back- and forecasted time series was made possible by implementations of Statistics Canada which also allowed for smaller revisions of seasonally adjusted series when new data got available
- New diagnostic tools were introduced (for example to investigate the stability of the seasonal factors if new data is available)
- The most important improvement is the application of the so called regARIMA program for a preadjustment of the data series. This preadjustment contains for example trading day effects, length of month variables, outlier detection/removal and so on

3. Practical application of Census X-12

To show an example for a seasonal adjustment procedure we applied the Census X-12 program that is implemented in Eviews 5.0 to the index of the Volume of Austrian GDP from 1970 (1) to 1997 (4) were the base year lies outside the considered range and is the year 2000. As options we chose the additive version of the procedure and automatic Henderson filter selection. The original and the transformed series are shown in the next two diagrams:

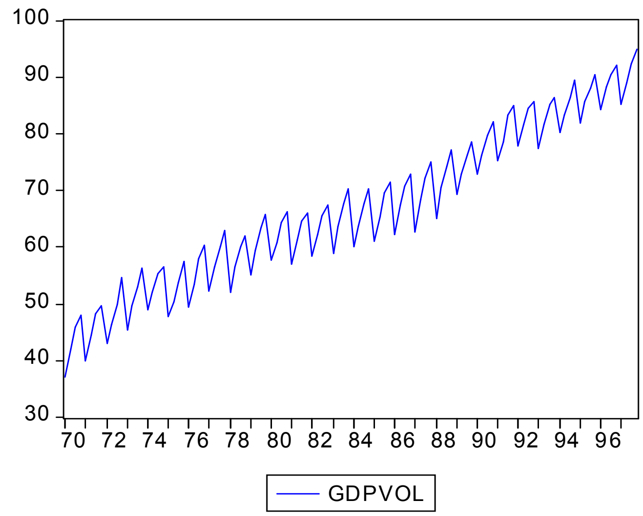

Figure 3.1: Unadjusted index of Austrian GDP from 1970 (1) to 1997 (4)

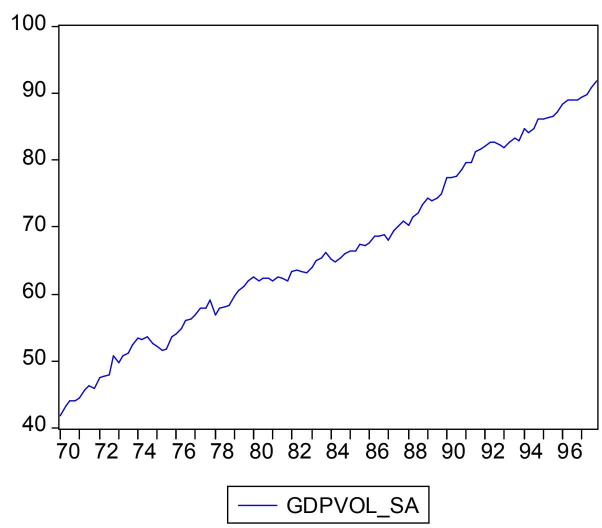

Figure 3.2: Seasonally adjusted index of Austrian GDP from 1970 (1) to 1997 (4)

So it is easily seen that in Figure 3 the seasonal movements are almost completely removed.

4. TRAMO/SEATS

4.1 Motivation:

The described family of filters is called „ad-hoc“ filters in the literature. The choice of weights is done "a priory", in the sense that it is independent of the actual series it is applied to.[1] To point out possible shortcomings of these methods and to motivate the use of a model-based approach, such as it is used in the programs TRAMO/SEATS, it is useful to have a look at the spectrum of a seasonal time series. Recall that every time series with T observations can be perfectly "explained" and reproduced by the sum of $T/2$ cosine functions of the type

$$r_t = A^t \cos(\omega t + B) \qquad (4.1)$$

where A denotes the amplitude, B the phase, and ω the frequency. The set of function is constructed by starting with the fundamental frequency $\omega = \frac{2\pi}{T}$, which completes one cycle in T periods, and its harmonics $\omega = \frac{2\pi}{T} j$, where $j = 1, \ldots, T/2$. For simplicity we assume that T is even. The function (4.1) can then be expressed as:

$$r_{jt} = a_j \cos \omega_j t + b_j \sin \omega_j t \qquad (4.2)$$

The coefficients a and b are related to the amplitude by $a_j^2 + b_j^2 = A_j^2$. The observed values z_t can then be written as:

$$z_t = \sum_{j=1}^{T/2} r_{jt} \qquad (4.3)$$

We now have a set of periodic functions with different frequencies and amplitudes. These functions can be grouped in intervals of frequencies by summing up the squared amplitudes of the functions which fall in the same interval. Doing so, a histogram of the contributions of each frequency to the overall variation of the series is obtained. Letting the intervals go to zero the histogram becomes a continuous function which is denoted as sample spectrum. It shows for each frequency the contribution to the variation of the series.

The smoothed spectrum of the quarterly measured Austrian GDP is shown in Figure 4.1:

[1] Though, this is not true in this strict form. Especially the X12 filters can be seen as a move from "ad hoc" filtering to a model based approach.

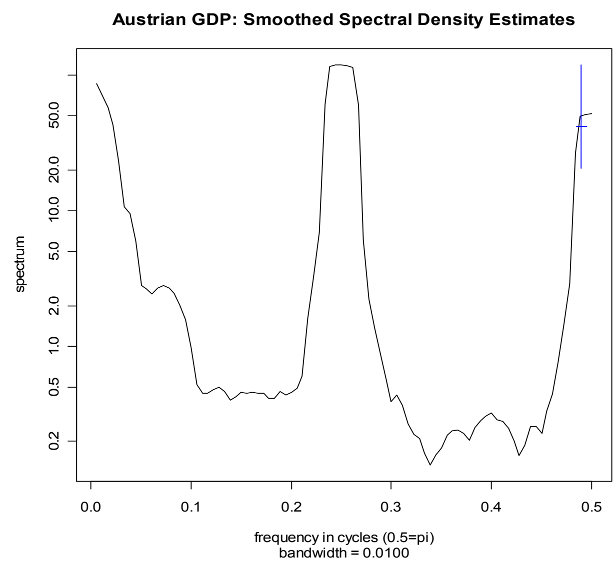

Figure 4.1: smoothed spectrum of the quarterly measured Austrian GDP

We have a peak at the zero frequency, and at the seasonal frequencies π and $\pi/2$. The width of the peaks depends on "how stochastic" the process is. A deterministic seasonal process will simply have a spike at the seasonal frequencies.

Ad hoc filters clean or adjust the series from the variance which falls in a certain band around the frequencies which are regarded as noise. Seasonal adjustment cleans the series from the variation which can be attributed to the seasonal frequencies. Since the width of this band is chosen a priory, this can lead to under- or to over-fitting, depending on the time series. In the first case not all of the variation which is due to seasonality is cleaned from the series, in the second case variation which is not part of seasonality is cleaned from the series. It seems obvious that the filter should depend to a higher degree on the structure of the time series it is applied to.

4.2 Introduction to TRAMO/SEATS

One package of programs which use such a model based adjustment method is TRAMO/SEATS. TRAMO does the pre-adjustment of the series and stands for "**T**ime Series **R**egression with **A**RIMA Noise, **M**issing Observations and **O**utliers". It does a similar job as regARIMA in the X12 program. SEATS stand for "**S**ignal **E**xtraction in **A**RIMA **T**ime **S**eries" ,and decomposes the series in its unobserved components following an ARIMA based method. It extracts the signal of interest from the series, thus those components which are of interest, and cleans the series from the noise, the seasonal components.

The TRAMO and SEATS packages were developed at the Bank of Spain by Victor Gomez and Agustin Maravall, with the programming help of Gianluca Caporello. A Windows version of the programs is downloadable from the homepage of the Bank of

Spain. There exists also a package for implementation in MATLAB and they are included in newer versions of EViews as well. Eurostat released the program package DEMETRA, which includes X12 and the TRAMO/SEATS programs.

4.3 TRAMO

The pre-adjustment of the series is an important part of seasonal adjustment or signal extraction. TRAMO adjusts the series for outliers which can not be explained by the underlying normality of the ARIMA model. It further interpolates missing values and provides forecasts. Effects which are adjusted by TRAMO can be forms of:

Outliers:

- additive outliers, thus isolated deviations
- level shifts, which imply step changes in the mean level
- transitory changes

Calendar Effects:

This term refer to effects of calendar dates such as working days in a period, holidays or the location of Easter.

Intervention Variables:

This are special unusual events, like natural disasters, change of the base index, new regulations,…

If z is the vector of observations, TRAMO fits the regression model $z_t = y'_t \beta + x_t$. The regression variables $y = (y_{1t}, ..., y_{nt})$ should capture the above mentioned effects. They can be entered by the user or generated by the program. The x_t follow a general ARIMA process:

$$\phi(B)\delta(B)x_t = \theta(B)a_t \tag{4.4}$$

The polynomial $\delta(B)$ contains the unit root associated with differencing, $\phi(B)$ the stationary autoregressive roots. $\theta(B)$ is the invertible moving average polynomial.

4.4 SEATS

Usually SEATS receives from TRAMO the original series, the “deterministic” effects (outlier, calendar effects,…) and the “linearized” (in the sense that it can be assumed to be generated by a linear process) and interpolated series. TRAMO also already identifies an ARIMA model to the stationary (and pre-adjusted) series.

Decomposition:

The decomposition is done multiplicative or additive. The latter can be achieved by taking logs, in the following discussion we will use the additive decomposition:

$$x_t = \sum_i x_{it} \tag{4.5}$$

where x_{it} represents a component.

SEATS considers

- the trend-cycle component
- the seasonal component
- the transitory component
- the irregular component

The trend-cycle component captures the low frequency variation. It displays a peak at the spectral frequency 0. The seasonal component captures the spectral peaks at the seasonal frequencies, the irregular component is erratic white noise and therefore has a flat spectrum. The transitory component is a zero mean, stationary component that picks up those fluctuations which should not contaminate the trend or seasonal component and are not white noise.
SEATS also contains an ARIMA estimation routine, since there is the possibility that the model passed over by TRAMO does not accept an admissible decomposition. SEATS performs a control on AR and MA roots. The AR roots are fixed when they are in preset interval around 1. SEATS then uses the ARIMA model to filter the linearized series to obtain the residuals. It also estimates the residuals which are lost through differencing. These residuals are then subject to diagnostics such as autocorrelation, presence of seasonality, randomness of signs, skewness, kurtosis, normality and nonlinearity.
Afterwards the program proceeds to decompose the ARIMA model. The decomposition assumes orthogonal components of which every one will have an ARIMA expression of the form

$$\phi_i(B)x_{it} = \theta_i(B)a_{it} \tag{4.6}$$

for each component $i = 1,..,k$. $\phi_i(B)$ and $\theta_i(B)$ are finite polynomials in B of order p_i and q_i, with no roots in common and all roots on or outside the unit circle. The following assumptions are made:

- The variables a_{it} are independent white noise processes
- The ϕ_i processes are prime

- The θ_i polynomials do not share unit roots in common

The first assumption is based on the belief that the forces which drive or cause the single components are different. For example, that the causes seasonality (weather), has nothing to do with the trend. The second assumption is sensible given that different components are associated with different spectral peaks. The last assumption guarantees the invertibility of the model.

The spectrum is decomposed into additive spectra associated with different components. These are determined by the AR roots of the model. The factorization can be written as

$$\Phi(B) = \phi_p(B)\phi_s(B)\phi_c(B) \tag{4.7}$$

where $\phi_p(B)$, $\phi_s(B)$ and $\phi_c(B)$ are the AR polynomials associated with the trend, seasonal and transitory roots, respectively. Since aggregation of ARIMA models yield an ARIMA model, it should be clear that aggregation of the models unobserved components should lead to the estimated model for the adjusted series.

The MA polynomial can be obtained through the relationship

$$\theta(B)a_t = \sum_{i=1}^{k} \phi_{ni}(B)\theta_i(B)a_{it} \tag{4.8}$$

where the $\phi_{ni}(B)$ are the product of all ϕ_j, $j = 1,...,k$, not including ϕ_i. Equations (4.5) and (4.6), together with the assumptions above are referred to as Unobserved Component ARIMA model.

In general, there is no unique UCARIMA representation that can generate it. The AR polynomials are obtained through factorization, but the MA polynomials and innovation variances are not identified. This under-identification problem is solved by assuming that $q_i \le p_i$.[2] The compositions then differ in the way white noise is allocated among the components. In order to uniquely identify the components, all white noise is added to the irregular component. This is called the canonical decomposition.

4.5 Constructing the Filter

After identifying the UC-model an optimal filter can be constructed. Optimal in the sense, that it minimizes the mean squared error $E[(s-\hat{s})^2 \mid X_t]$ where s denotes the real signal and $\hat{s}$ its estimate. The minimum mean squared error is obtained by the

[2] It is possible to estimate also models with $Q > P$, but it is recommended to favor balanced models because of their better decomposition properties.

conditional expectation of the unobserved s. Since the joint distribution is multivariate normal, the conditional expectation is a linear combination of the elements x_t .

The conditional expectation can be computed by the Kalman Filter or the Wiener-Kolmogorov (WK) Filter. In the stationary case these filters are equivalent. If one writes the model for x_t, and the model for the signal in their MA expression as $\Psi(B)a_t$ and $\Psi_s(B)a_{st}$, where $\Psi(B)=\frac{\theta(B)}{\phi(B)}$ and $\Psi_s(B)=\frac{\theta_s(B)}{\phi_s(B)}$. and denotes the forward operator with F, then $\hat{s}$ is obtained by the symmetric filter

$$\hat{s}=[\frac{\sigma_s^2}{\sigma_a^2}\frac{\Psi_s(B)\Psi_s(F)}{\Psi(B)\Psi(F)}]x_t=\vartheta(B,F)x_t \tag{4.9}$$

The filter can be expressed in the frequency domain as

$$\widetilde{\vartheta}(\omega)=\frac{g_s(\omega)}{g_x(\omega)} \tag{4.10}$$

This function is also referred to as gain of the filter. The spectrum of the estimator of the signal is given by

$$g_{\hat{s}}(\omega)=[\frac{g_s(\omega)}{g_x(\omega)}]^2 g_x(\omega) \tag{4.11}$$

The squared gain of the filter determines how the variance of the series contributes to the variance of the signal. If the ratio is 1, the variance corresponding to these frequencies is completely passed through to the signal. The frequencies where the ratio is 0 are cut out and do not contribute to the estimated spectrum of the signal.

An example how the squared gain of the filter for the quarterly GDP series could look like is given in Figure 4.2:

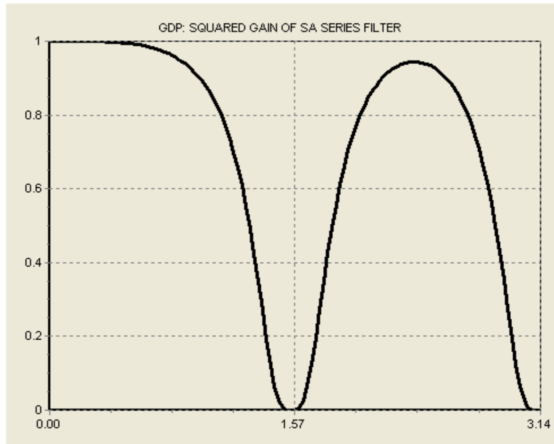

Figure 4.2: example of how the squared gain of the quarterly series could look like

The part of the variance corresponding to the seasonal frequencies is filtered out, while the other part of the variance is passed through and contributes to the formation of the signal.

In the non-stationary case the spectra have to be replaced by pseudo-spectra. Note that the frequency-domain representation remains valid, despite ∞ peaks in the spectrum, $\tilde{\vartheta}(\omega)$ is everywhere well defined.

It is clear that in a symmetric (or two-sided) filter the signal for the most recent and therefore most interesting observations would not be available. SEATS solves this by producing several years of forecasts. SEATS also produces standard errors of the estimates and forecasts as well as standard errors for the revision the estimator and forecasts will undergo. Revisions are the corrections when new observations become available. They are necessary until the estimate for the signal is based solely on observed values.

4.6 Conclusion

TRAMO/SEATS tailors the filter to the observed series according to the underlying ARIMA model. At Eurostat a lot of research is done which compares the performance of the X12 family of filters and model based approaches. There is a tendency which favors model based methods, however for most tasks all these filters do their job well. In addition the “real” data generating process is unknown and therefore it is impossible to justify which procedure leads to a better fit of the adjusted series to the generated data without a seasonal cycle. Figure 4.3 shows the seasonal adjusted series for the Austrian GDP using the X12 filter and TRAMO/SEATS. For most parts of the series the differences between the two methods are negligible small. However, adjusting quarterly GDP for seasonality is quite a standard application and huge differences would be very surprising.

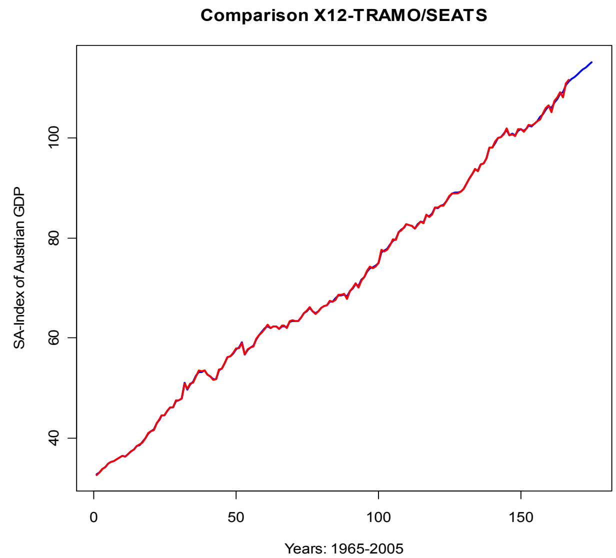

Figure 4.3: Comparison of Census X12 and TRAMO/SEATS for the quarterly Austrian GDP series

Literature:

Caporello, G. Maravall, A.: Software package TSW available on: http://www.bde.es/servicio/software/econome.htm

EViews 5.0: Software package available on: http://www.eviews.com/download/download.html

Ghysels, E. Osborn, D. R. (2001): "*The Econometric Analysis of Seasonal Time Series"* Cambridge University Press

Gomez, V., Maravall, A.: Seasonal Adjustment and Signal Extraction in Economic Time Series. On: http://www.bde.es/servicio/software/papers.htm

Maravall, A. & Kaiser, M.: Notes on Time Series Analysis ARIMA Models and Signal Extraction. On: http://www.bde.es/servicio/software/papers.htm

R Development Core Team (2006): "R: A language and environment for statistical computing", R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0 package available on: http://www.R-project.org