***Draft in progress ***

To do:
- submit SRs to change/improve SEASABS functionality: stricter balancing, include ref period (actual location) in knowledge.
- initiate research project on using X12/regARIMA for detecting/correcting SBs. need SR on new methodology.
- enable one to specify a SB location explicity from S*I chart and SEASABS takes care of the rest?
- refine/streamline manual methods below.

--

SEASONAL BREAK

What?

A Seasonal Break occurs when there is an abrupt and permanent change in the seasonal pattern of a time series. Before a seasonal break can be estimated and corrected, at least three years of data is needed to illustrate the permanent change in the seasonal pattern. The actual level of a series does not change when there is a seasonal break, only the yearly pattern.

Seasonal breaks can be caused changes in things such as:
- the coverage of a survey
- social traditions
- administrative practices
- technological innovations

Seasonal breaks can be either 'compensating' or 'non-compensating'. A compensating seasonal break occurs when the increase in the seasonal factor for one period(s) is offset by a decrease in the seasonal factor for another period(s). Thus a compensating seasonal break will have upward and downward movements that balance. A non-compensating seasonal break is where the change in seasonal pattern for one period is not offset by an opposite change in another period.

There are two types of non-compensating seasonal breaks, both of which are difficult to correct. The first is when all affected months have breaks moving in the same direction. In this case, all months not affected will need to have a balancing correction factor applied so that the level of the series remains unchanged. The other type of non-compensating break occurs when there are two months or quarters where there is both an increase and a decrease, but the changes are not compensating. A time series analyst would have to correct for this type of break manually as it cannot be automatically corrected in SEASABS.

Why do we correct it?

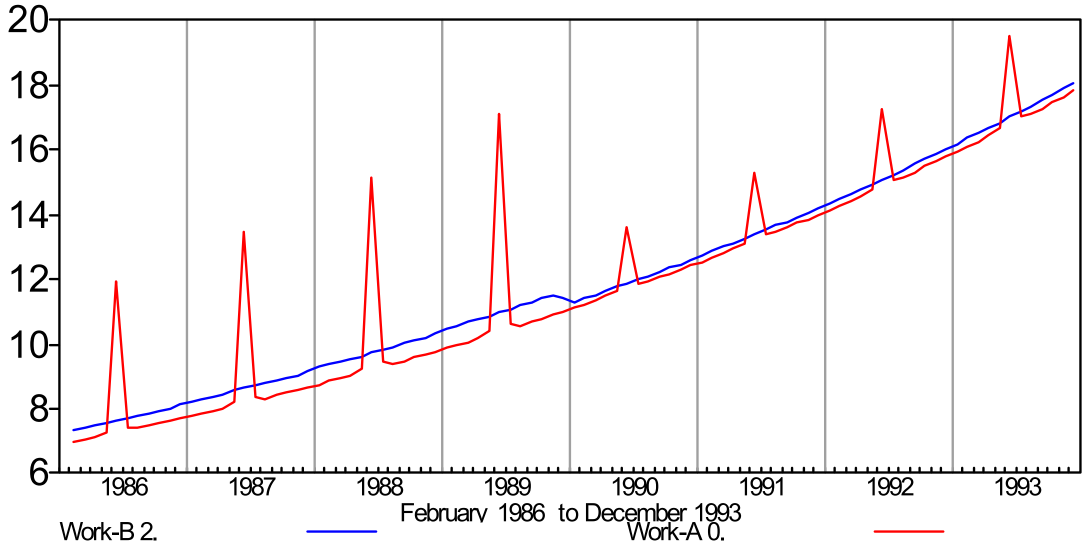
If there is an abrupt and permanent change in the seasonal pattern which is not corrected by a seasonal break, then the seasonal factors will be distorted. This can lead to a poor seasonal adjustment resulting in patterns in the irregular components, which is known as 'residual seasonality'.

When do we correct it?

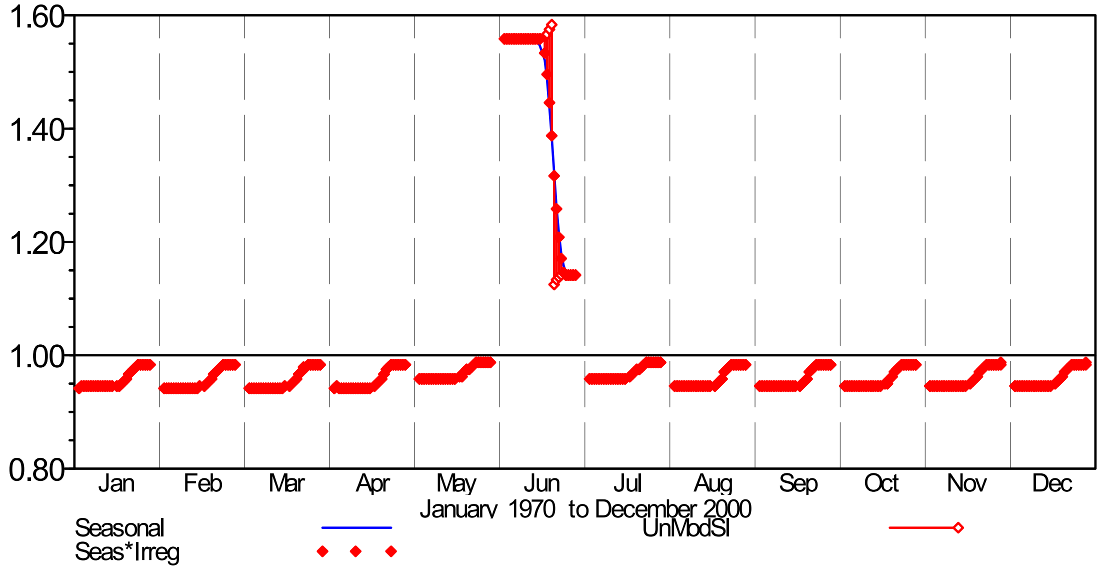
The following factors need to be considered when deciding whether to insert a seasonal break correction:

- Is there enough data? Usually at least three years is needed before a seasonal break correction is inserted.
- Is the change in seasonal pattern evident in the original data?
- Can appropriate correction factors be estimated? (A seasonal break correction must be 'balanced' - see the 'How do we correct it' section below.)
- Is a break evident from looking at the SI chart? A change in the seasonal pattern of a time series will cause a level shift in the seasonal factors, and hence a seasonal break will appear as a 'trend break' in the SI's.
- The subject matter area responsible for the data must provide a real-world reason for the change in the seasonal pattern and/or approve the correction.

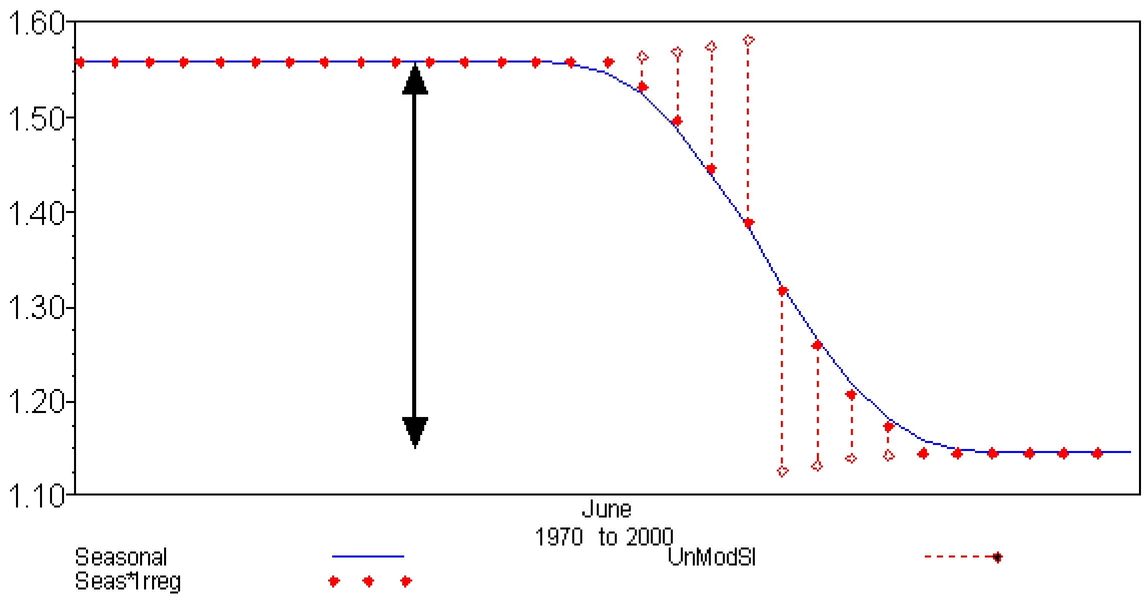
Example of a seasonal break



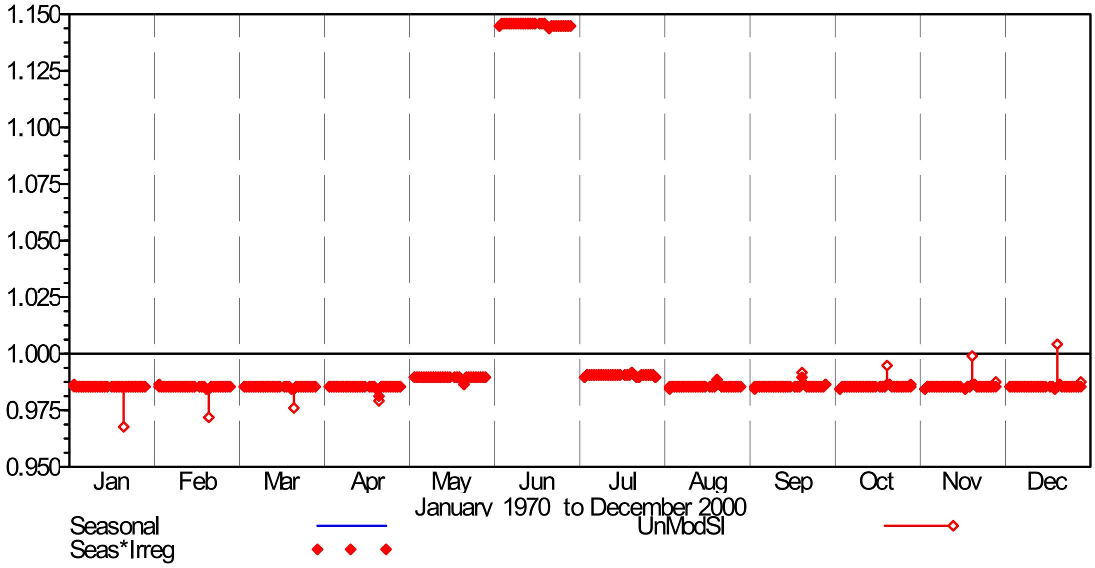
Graph 1: Example of an exagerated seasonal break in 1989/1990 for June. This is from a simulated (noiseless) series run through X11.



Graph 2: The corresponding S*I chart for the above. The tiny SBs in all other months in the above S*I chart are a consequence of the initial trend estimate from a centered moving average on the original used to estimate the S*I's, i.e. there is a level shift at 1989/1990 (at the location of our simulated SB; see graph 1). A level shift before and after the SB will give a change in mean S*I levels for all months at the same location since S*I = O/T.



Graph 3: This is a zoom-in of graph 2 for June. Note the internal corrections (dashed red lines) that are used by SEASABS to arrive at the blue seasonal moving average (SMA) line from which seasonal factors are taken. Our goal is to reduce the incidence of these internal corrections, i.e. make the blue line more representatative of the red (real) seasonal data points. Otherwise, seasonal factors will be distorted and residual seasonality will ensue.



<u>Graph 4:</u> After SB-corrected S*I chart using <u>manual method 1</u>. This attempts to <u>equalise</u> mean S*I levels before and after the break in all months.

If we don't correct it, what then?

If a series is displaying a change in the seasonal pattern, but no correction is to be inserted, then the application of a seasonal break correction should be revisited when another years worth of data becomes available. The presence of residual seasonality should also be monitored.

How do we correct it?

There are three ways:

- When a research analysis is run in SEASABS, a check for seasonal breaks is done.
    - o The seasonal break algorithm needs at least six years of data before it can detect seasonal breaks.
    - o If a seasonal break is detected in SEASABS, and the SMA is in agreement that there is a change in seasonal pattern and a correction needs to be entered, the analyst can accept the correction factors suggested as a starting point.
    - o They then need to check if the correction entered is balanced and won't change the level of the series, i.e. the B1 table totals with and without a break can be compared. If the table total has differed by more than 0.2% after entering the seasonal break correction, then the correction is unbalanced and the factors need to be improved (see below for how to calculate correction factors manually).
    - o ***Is there a document that outlines the different options and consequences when a research analysis detects possible prior corrections that need to be inserted?***

- If SEASABS does not detect a seasonal break, but a change in the seasonal pattern is evident and a correction is needed, the user can calculate the correction factors manually. Suggested methods are below.

- Using a model-based approach via regressors under X12/regARIMA. ***This is a future research project and not yet explored***

How do we correct for seasonal breaks manually?

There are two suggested methods. You should always start with 'method 1' below. 'Method 2' is optional and is recommended if after using method 1, the final trend cycle shows a systematic level shift (Trend Break) greater than the inherent volatility at the seasonal break (SB) location.

Manual Method 1: equalising S*I levels

As a first test, this method approximates SB factors from the sole criterion of equalising seasonal factor levels in all periods where a seasonal break can be justified. The main steps used to arrive at SB corrections (with reference to the simulated example in graph 1) is as follows:

1. Having identified the "break period of interest" (June), we now need to identify other periods in which to redistribute the "excess" seasonality to be 'removed' from June at years <=1989. This is so that we can balance the B1 totals later, i.e. so that they remain unchanged after application of all SB priors. There are two ways in which to pick these "balancing periods", depending on whether they are compensating or non-compensating:
(i) compensating periods are those where we can "gently" throw-in/remove a little more seasonality in order to compensate for the reduction/addition of seasonality in our main break period. In other words, these are periods where there is a hint of a break that goes more-or-less in the "opposite" direction to the break in our period of interest. We want to ensure that inclusion of seasonal break corrections are justified. Note that if one has a large seasonal break for one period (e.g. example above) then tiny breaks in other periods are inevitable (why? graph 1 may give you a clue). The example above obviously falls under this category. A rule of thumb when picking balancing periods is to pick those with the "noisiest" S*I values, or, deviate most from the blue SMA lines. I.e. basically those periods where inclusion of a SB might alleviate the discrepancy between S*I values and smoothed (seasonal factor) estimates from a SMA.
(ii) non-compensating periods are those which show no evidence of seasonal breaks going in the opposite sense that may be used for B1 balancing. Inclusion of a break may not be appropriately justified. However, we can still perform the B1 balancing by apportioning the excess seasonality removed from our main break equally into all other periods (i.e. spread it out thinly so that we avoid introducing any new undesired breaks). More weight should be given (that is, larger SB factors can be allowed) to periods with the "noisiest" S*I values, or, deviate most strongly from the blue SMA lines.

2. We then attempt to estimate initial SB factors for the periods identified in (1). A good approximation is the ratio of mean S*I level before/after the break years in each period. In the example above, this is before/after the

1989/1990 break years. The arrow in graph 3 indicates what we're after: the required factor for June from the uncorrected S*I chart is f ~ 1.58 / 1.16 ~ 1.36. One can arrive at more accurate measures by using the uncorrected D10 (seasonal factors) table. We do the same for every other period where a break can be justified. In general, when estimating SB factors f_j for some period j , there are three cases to consider:
(i) if there is no evidence for a break, then explicity f_j =1 (under a multiplicative decomposition model of course).
(ii) if seasonal factor levels are ~ constant before/after the break year for period j "SByr(j)" (=1989 in above example),

$$f_j \approx \frac{[\text{mean S*I at years } \leq \; SByr(j)]}{[\text{mean S*I at years } > \; SByr(j)]}.$$

The simulated case above is an example of this.
(iii) if there is <u>moving</u> seasonality, the f_j is really the size of the "jump", i.e. the size of the discontinuity at the break years: SByr(j) to Sbyr(j) + 1 for period j:

$$f_j \approx \frac{(\text{S*I})_{SByr(j)}}{(\text{S*I})_{SByr(j)+1}}.$$

3. Having estimated your initial SB factors, you now need to ensure they give "balanced" B1 totals. In other words, these 'approximate' factors need to be rescaled such that the <u>original</u> B1 total is still equal to the B1 total with these rescaled factors applied to each period. This can be performed using either of following methods:
(i) enter these initial SB factors together with B1 series from SEASABS into the *R* program described in Appendix I. This program performs exact balancing, i.e. such that the B1 total remains the same (within machine precision) after all SB factors are applied. Recall that SEASABS currently allows a "liberal" 2% variation in the B1 total. This does not seem methodologically correct. It may have to do with the inner workings of the automated SEASABS algorithm.
(ii) you can also perform "B1 balancing" by comparing the B1 table total in SEASABS before/after SB corrections are applied and ensuring the difference remains appropriately small. A good working measure is to ensure a variation of no more than ~0.2%. To meet this, you may need to iterate the process by toggling factor values and rechecking the B1 totals each time.

4. The rescaled (or "B1-balanced") SB factors from step 3 can now be entered into the SEASABS "Prior Corrections" GUI (e.g. Table 2 below). You are now ready to perform a "Research Analysis" run. Note that if at any time you tweak/readjust the SB factors in the SEASABS GUI, you must rebalance them (see step 3).

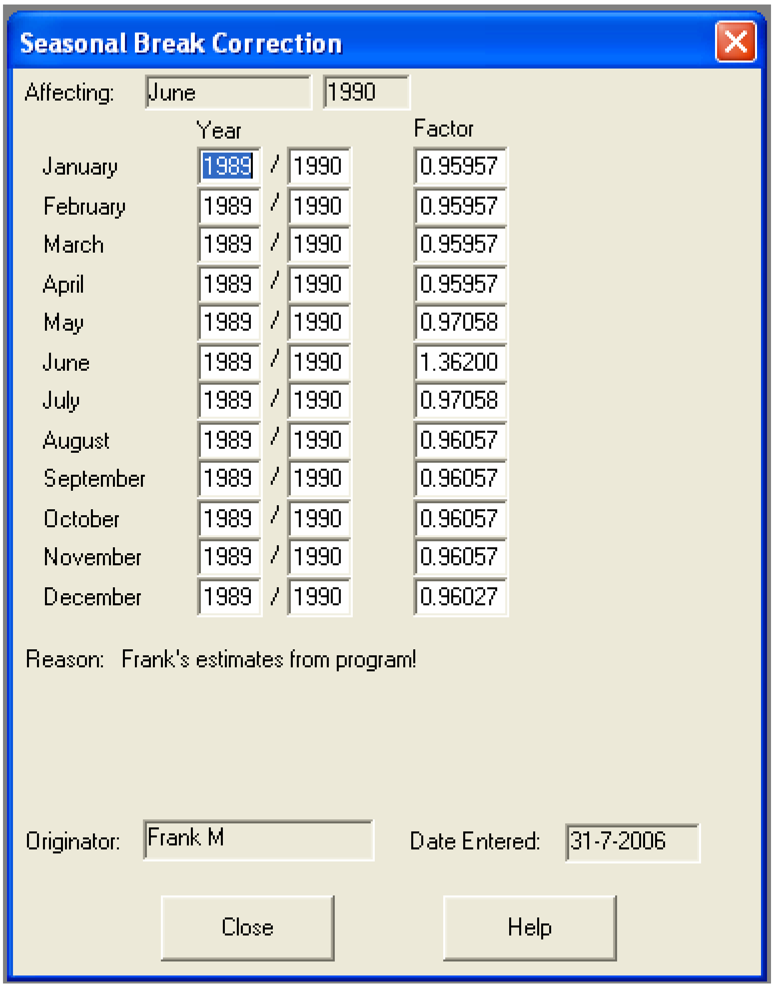
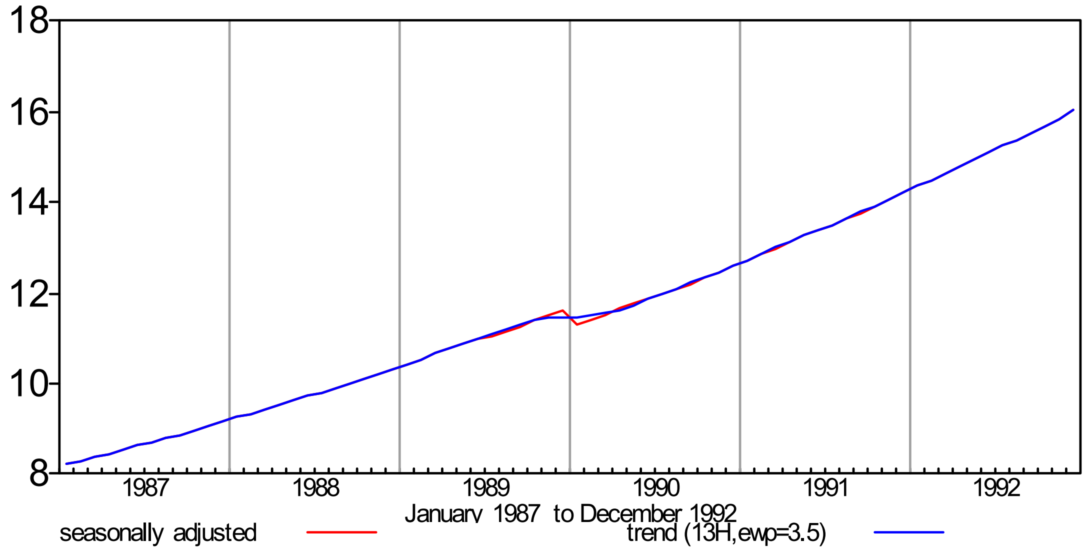


Table 2.

5. You can now perform a few sanity checks after correcting for SBs:
(i) peruse the S*I chart to ensure that all prior SBs were appropriately minimised and seasonal factor estimates are reliable for all periods (i.e. internal corrections to SMA derived seasonal factors are minimised); see graph 4 above.
(ii) ensure the B1 table total is ~ the same before and after SB correction.
(iii) peruse the D10 (seasonal factors) table and ensure they average to ~ 1 in every year.
(iv) examine graph of final seasonally adjusted overlayed with final trend cycle to ensure no systematic level shifts (Trend Breaks) greater than the inherent volatility were introduced (e.g. see graph 5 below). If so, then it is recommended you read through " *Manual Method 2: preventing/minimising potential level shifts* ".



Graph 5: This shows the SA and Trend estimates after correcting for the SB using manual method 1. Note the level shift around Dec/1989 - Jan/1990! This is a consequence of our criterion of equalising S*I levels for every month. In a nutshell, this is because in order to keep the original total the same before and after SB insertion, there must be a compensating increase in the trend before the seasonal break. I.e. given the multiplicative model, O = T.S.I, enforcing the same "S" before and after the break years 1989/1990 in all months means that "T" must be correspondingly higher before 1989 to account for the 'excess seasonality' (graph 1) removed from June and hence conserve "O". Such a level shift is rare and it is visible here because our simulation is highly unrealistic, i.e. it is perfectly noiseless - no irregular component was included.

Manual Method 2: preventing/minimising 'potential' level shifts

If a series is substantially volatile (which the above simulated series is not), then you may be able to get away in introducing a noticeable level shift from the criterion of strictly equalising S*I levels, as assumed in 'Manual method 1' above. I.e. it might get hidden amongst the volatility and be declared insignificant. A good working practice is to always check your SA and trend estimates for level shifts around the SB periods. If you see one, how do you minimise it or better still, remove it? One way is to relax the assumption of equalising S*I levels across all months. There are two key points:

- The goal is to ensure that the derived seasonal factors, S = SMA[S*I], are the best possible representation of the actual S*I's. That is, such that variation in I = S*I/SMA[S*I] = T*I/TMA[T*I] is minimised. We do not need to try too hard at "smoothing" seasonal breaks [i.e. by equalising all S*I levels], we just need to ensure that the SMA[S*I] curves "fit" well, while at all times preserving the B1 total. If we are too perfect in equalising S*I levels, then nusance level shifts can be introduced to compensate for (or that attempt to undo) our "removal" or "redistribution" of seasonality from/in the original.

- In essence, we want to redistribute total 'power' in the original such that:

(i) the resulting seasonal factor estimates from S=SMA(S*I) are reliable and representative enough in both the "period of interest" (i.e. that contains the main break to correct) and other periods used for "B1-balancing".
(ii) no systematic level shifts are introduced.

With these in mind, here's a general method, which is essentially an extension to 'Manual Method 1':

1. It is assumed you have first used 'Manual Method 1' above to estimate and apply SB corrections, and that you do indeed see a noticeable (significant) level shift around the SB year(s).

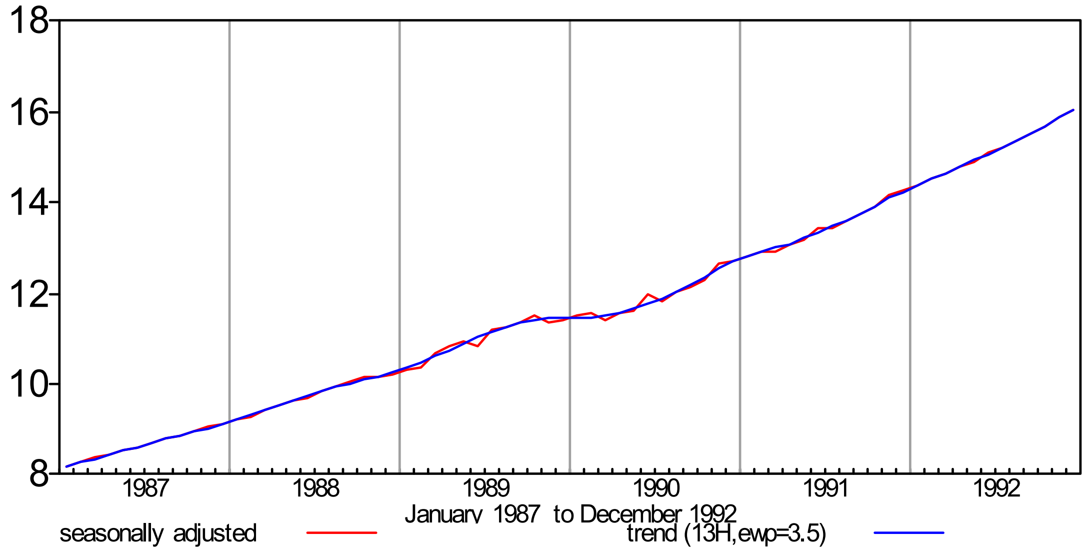
2. From your full set of SB factors (from 'Manual Method 1'), the procedure is to remove one or more factors (i.e. set them explicitly to 1) on either side of the "main break period" of interest (June in above example).

3. Rescale these factors to ensure B1 totals get balanced after application (using either the *R* program described in Appendix I, or, by comparing B1 table totals before/after the SB corrections are applied).

4. Enter these factors in the SEASABS "Prior Corrections" GUI and run a "Research".

5. Examine the quality of the seasonal factors "S" from the S=SMA(S*I) operation, i.e. that they are representative enough of the actual S*I values given the volatility. Also, examine the strength of any prior level shift: is this smaller/larger than the level shift that was present when you started with step (1)? If still noticeable (significant), go to step 6.

6. Continue iterating steps (2) - (5) by resetting/removing/tweaking different combinations of SB factors (in some "justifiable" manner of course) from your initial set. This is to be done until level shifts are appropriately minimised (or removed), but at the same time, reliability in seasonal factors is maintained. The trick is, how much "redistributing of seasonality" can we perform before we start to mess up seasonal factor estimates, i.e. introduce too many internal corrections on the S*I charts.



<u>Graph 6:</u> This shows the resulting SA and Trend estimates after correcting for SBs using manual method 2. Note the level shift around Dec/1989 - Jan/1990 is less apparent than that introduced by simply equalising all S*I levels (graph 5).

Appendix I: a program to ensure "balanced" B1 series

We have a 'quick and dirty' *R* function which reads as input a set of initial (approximate) SB factors estimated from a S*I chart and a B1 series from SEASABS. It then rescales the factors such that the B1 total remains unchanged after they are applied. Note that this program performs <u>exact</u> balancing, i.e. such that the B1 total remains the same (within machine precision) after the final SB factors are applied. Recall that SEASABS currently allows a 2% variation in the B1 total. This does not seem methodologically correct. It may have to do with the inner workings of the automated SEASABS algorithm.

Rescaling of input (approximate) SB factors to ensure "balanced" B1 totals is performed using the following procedure. Note that this only applies under a <u>multiplicative</u> decomposition model. The *R* program needs to be modified to handle series described by additive models.

$B1$ totals before SBs inserted $=$ $B1$ totals after SBs inserted:

$$\sum_{period\ j}^{N}\ \sum_{year\ i}^{SByr(j)} B1_{ij} = \left[f_1 \sum_{year\ i=1}^{SByr(1)} B1_{i1} + \sum_{i=SByr(1)+1}^{i\max(1)} B1_{i1} \right] + \left[f_2 \sum_{year\ i=1}^{SByr(2)} B1_{i2} + \sum_{i=SByr(2)+1}^{i\max(2)} B1_{i2} \right] + \cdots + \left[f_N \sum_{year\ i=1}^{SByr(N)} B1_{iN} + \sum_{i=SByr(N)+1}^{i\max(N)} B1_{iN} \right], \qquad (1)$$

where:

$N =$ number of periods (months or qtrs in a year).

$f_j =$ approximate prior SB factor for period j (see step 2 under "Manual Method 1.." for cases).

$i\max(j) =$ the maximum year for period j.

$SByr(j) =$ the 'seasonal break' year for period j [= last year in period before which S*I level changes abruptly, i.e. break location]

To ensure that the B1 total before application of our approximate seasonal break factors = B1 total after their application, we compute the residual (either excess or deficit) in this equality:

$$D = B1(total\ before) - B1(total\ after), \qquad (2)$$

and then equally apportion this difference amongst all those periods which required a seasonal break correction. If we denote the number of periods requiring a SB correction as $N_{SB} \le N,$ then we assume that the "amount of residual B1" that needs to be thrown back into any particular period j is $\approx\ D/N_{SB}$. If we take any period j from the right hand side of equation (1), our goal is to determine a factor g_j such that:

$$f_j \sum_{year\ i=1}^{SByr(j)} B1_{ij} + \sum_{i=SByr(j)+1}^{i\max(j)} B1_{ij} + \frac{D}{N_{SB}} = g_j f_j \sum_{year\ i=1}^{SByr(j)} B1_{ij} + \sum_{i=SByr(j)+1}^{i\max(j)} B1_{ij}. \qquad (3)$$

I.e. g_j is the factor by which an initial f_j needs to be multiplied in order to make the B1 totals balance when g_j is computed for all periods j that require a seasonal break. Simplifying and rearranging equation (3), we have:

$$g_j = \frac{D}{N_{SB}} \left[f_j \sum_{year\ i=1}^{SByr(j)} B1_{ij} \right]^{-1} + 1. \qquad (4)$$

Therefore, each period with a SB prior gets its own g_j "fudge" factor. The final "effective" SB factors that ensure "exact" balancing of the B1 totals are then given by:

$SB_j = g_j f_j.$

The R function that implements the above is located in the file **"S:\R\TSAlib\compute_SBfactors.r"**. Instructions and an example on how to call this function are at the top of this file.

Related BPG TOPICs

<u>SEASABS Help</u>
Correcting Seasonal Breaks
Seasonal Breaks
Seasonal Break Algorithm
Seasonal Break Dialog Box
Compensating Seasonal Breaks
Non-compensating Seasonal Breaks

- *insert in this section, under SEASABS Help header, titles that contain relevant Help for this TOPIC*
- *does the Help accurately describe our current practices?*
- *if NO suggest improvements/extensions and forward to Joseph; headers and skeleton outline will suffice>*

<u>Other</u>

- An example manual seasonal break correction using a simulation: TSA WDB > Research > Structural Breaks > (Subject: Understanding Seasonal breaks..; Database: Time Series Analysis WDB; Author: Frank Masci; Created: 12/07/2006; Doc Ref: NACT-6RN8QG)