Composite Seasonal Adjustment

Richard Penny[1]

Statistics New Zealand (StatsNZ)

8 June, 2022

[1] The views expressed in this talk are those of the presenter and not necessarily those of Statistics New Zealand

What is a composite Seasonal Adjustment

This talk motivated by R package `seasonal` V1.9.0 which now can handle composite series

- We have a total that it is the arithmetic combination of some component series
- Do we seasonally adjust the components and add, or add and then seasonally adjust?
  - $C = A + B$. Is $SA(C) = SA(A) + SA(B)$, or $SA(C) = SA(A + B)$?
- $SA(C) = SA(A) + SA(B)$ is termed indirect seasonal adjustment
- $SA(C) = SA(A + B)$ is termed direct seasonal adjustment
- Direct adjustment does not preserve additivity, so often preference is for indirect

Indirect or Direct – Theory

- For additive adjustments the way the components interact is straight-forward

$$\begin{aligned} A + B &= C_A + S_A + I_A + C_B + S_B + I_B \\ &= (C_A + C_B) + (S_A + S_B) + (I_A + I_B) \end{aligned} \tag{1}$$

- Multiplicative components are more complicated which I won't go into, but I have document that covers it.

Direct or Indirect – Practice

- Additive adjustments

$$\begin{aligned} A + B &= C_A + S_A + I_A + C_B + S_B + I_B \\ &= (C_A + C_B) + (S_A + S_B) + (I_A + I_B) \end{aligned} \tag{2}$$

- As the seasonal factors combine nicely it is a case of what the signal to noise ratio is like for indirect versus direct
- Well known that if components are “in-phase” direct adjustment likely be better as the seasonal components reinforce, whereas irregulars not so much (depends on variance in component series)
- Conversely, series “out of phase” better to indirectly adjust as seasonal patterns cancel out, whereas irregular stays about the same
- In all case we can only estimate the components.

Composite Seasonal Adjustment in StatsNZ

- Introduced into StatsNZ in mid 1980's when we migrated from SAS Proc X11 to Statistics Canada's X-11-ARIMA software
- Policy was to choose method based on X-11-ARIMA composite quality statistics, though preference for indirect.
- The quality was based on a smoothness criteria, where $\hat{X}_t$ are the seasonally adjusted values for both outputs and H is the 13-term Henderson filter (Dagum, 1979)
  - $R_1 = \sum_t (\hat{X}_t - \hat{X}_{t-1})^2$
  - $R_2 = \sum_t (\hat{X}_t - H\hat{X}_t)^2 = \sum_t [(I - H)\hat{X}_t]^2$
- We programmed into our s.a. system the ability to graph the direct and indirect s.a. series
  - If there is a significant difference usually a sign something odd is happening
  - Not used as much as it should

Things We Had to Work Out

- How to work from the bottom to the top. That is, if $A + B = C$ and $D + E = F$ and $C + F = G$, how do we integrate our decisions at the lower level into the upper level? And the intermediate levels.
- How to handle non-seasonal components
- Where does the indirect trend come from?
  - For that we had to look at the FORTRAN code associated with that module
- Consider various puzzling output and problems

Puzzle #1

- Model consistency for composite balances
  - That is, composite $C = A - B$ e.g. trade balances
- As the balances have negative values in the time series we have to use additive adjustment. And they look additive too
- Components are clearly multiplicative
- Should we worry?
- Puzzle #1a How are we going to implement RegARIMA models in a consistent manner?

Puzzle #2

- The indirect trend should be, ...
- StatsNZ releases trend series. Just D12 (direct) or ITN (indirect)
- In two cases for indirect outputs users loaded s.a. components and used Db to create sums (i.e. not .isa). Did the same for trend (added the component .d12).
- I thought that this “error” would create a less smooth indirect series, but in both cases I was wrong

Puzzle #3

- The case of the fishing industry
- The components are highly seasonal – though some small values
- However total is not – probably fishing boat capacity
- Should we even bother with composite?

Puzzle #4

- More than one way to create a composite
  - That is, $C = A + B$ **and** $C = D + E$
- Can't do it both ways so
  - Pick the most important way to combine?
  - Pick the one that is indirect (if other way prefers direct)?
  - Pick the composite with best quality?

Puzzle #5

- Outliers in different places in components
- We'll identify outliers in component but unlikely to be aligned to the ones in composite
- Generally not a problem as outliers in components primarily not large (e.g. $2.8\sigma^2 - 4\sigma^2$)
- Have had problems with largish outliers in one or two components not being recognised in direct composite. So are we comparing "like with like"?

Puzzle #6

- Difference between component story and composite story
- For a composite series with two components we had decided that direct gave the better adjustment
- One month the components both went down but the composite (direct) went up. Needless to say the client was very worried
- Never got to the bottom of it as anomaly disappeared next month
- Possibilities

  Human Error: A very complicated structure underneath. There were about 4 layers below so someone could have made a mistake implementing decisions

  Technical glitch: There were big partial C17 weights in the components, but not composite

  Methodologically interesting: May inform us as to which composite identifies turning points first

Puzzle #7

- And the perennial, how do we decide which is the better seasonal adjustment?
- And in StatsNZ's case, also the best trend?

Want to Talk Composites?

Richard.Penny@stats.govt.nz