European Commission

X-13ARIMA

ESTP Training
17-19 October 2023

A Bit of History

- 1954: X-0: first computerized seasonal adjustment program (X=eXperimental).
- 1965: X-11 (US Bureau of Census)
- 1980: X-11 ARIMA (Statistics Canada)
- 1998: X-12-ARIMA (US Bureau of Census)
- 2006: X-13ARIMA-SEATS (US Bureau of Census)

- 2015: JDemetra+ (Eurostat, ECB, NBB)

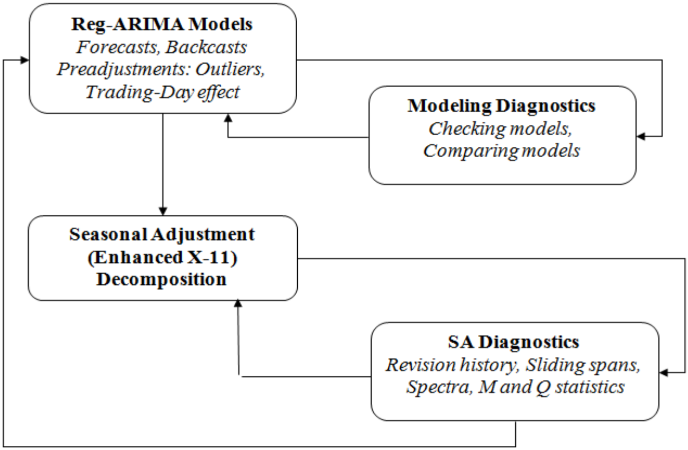
A long history, lots of improvements but the core of the decomposition (moving averages) did not really change since X-11.

References

- USBC website: `https://www.census.gov/srd/www/x13as/`;
  - with lots of papers on seasonal adjustment.
- Ladiray, D., Quenneville, B. (2001), Seasonal Adjustment with the X-11 Method. New York: Springer. Lecture Notes in Statistics, vol. 158, Springer-Verlag;
- Findley, D.F., Monsell, B.C, Bell, W.R., Otto, M.C., Chen, S., (1998), New Capabilities and Methods of the X12ARIMA Seasonal Adjustment Program (with discussion), Journal of Business and Economics Statistics, 16, 127-177.

X-13ARIMA-SEATS

Two building blocks: RegARIMA and X11.



Outline of X-13ARIMA-SEATS

- First cleaning of the series (Outliers, Trading-Day effects), ARIMA modeling and seasonal adjustment with X11;
- ARIMA modeling: to stabilize end points estimations (use of symmetric moving averages on the forecasts);
- An old idea (DeForest, 1877; Macaulay, 1931) used by E.B. Dagum for X11-ARIMA (1975):
- "*However, graduation of the ends of almost any series is necessarily extremely hypothetical unless facts outside the range covered by the graduation are used in obtaining the graduation ..... Though mathematically inelegant, the most desirable procedure in a majority of the cases of graduation is to graduate not only the actual data, but extrapolated data which sometimes may be extremely crude estimates*". Macaulay (1931).

The RegARIMA model

A usual additive model for SA is:

$$X_t = TC_t + S_t + O_t + TD_t + MH_t + I_t,$$

Where all the components are orthogonal and are non-seasonal (except S_t of course).

A SA in 2 steps:

Step 1: $X_t = O_t + TD_t + MH_t + Z_t,$ where Z_t follows an ARIMA model.
Step 2: Decomposition of $Z_t = TC_t + S_t + I_t$

More details

- Reg-Arima Model.
    - Automatic choice of the decomposition model (additive-Multiplicative);
    - Automatic estimation of missing values;
    - Automatic detection and correction of Outliers;
    - Automatic detection and correction of calendar effects (using a National calendar if available);
    - Automatic ARIMA modeling, backcasts and forecasts;

    - Quality diagnostics to check the validity of the model.
- The decomposition with X11.
    - Automatic choice of the moving averages;
    - New and automatic detection and correction of additive outliers;
    - Automatic decomposition in Trend-Cycle, Seasonality and Irregular;

    - Quality diagnostics to check the validity of the decomposition.

X11 basic algorithm

Monthly time series, additive model: $X_t = TC_t + S_t + I_t$.

1. Estimation of Trend-Cycle by 2×12 moving average:
   $TC_t^{(1)} = M_{2\times 12}(X_t)$
2. Estimation of the Seasonal-Irregular component:
   $(S_t + I_t)^{(1)} = X_t - TC_t^{(1)}$
3. Estimation of the Seasonal component by 3×3 moving average over each month: $S_t^{(1)} = M_{3\times 3}\left[(S_t + I_t)^{(1)}\right]$ and normalization: $Snorm_t^{(1)} = M_{2\times 12}\left[S_t^{(1)}\right]$
4. Estimation of the seasonally adjusted series:
   $Xsa_t^{(1)} = X_t - Snorm_t^{(1)} = (TC_t + I_t)^{(1)}$

5. Estimation of Trend-Cycle by (13-term) Henderson moving average:
   $TC_t^{(2)} = H_{13}\left[Xsa_t^{(1)}\right]$
6. Estimation of the Seasonal-Irregular component:
   $(S_t + I_t)^{(2)} = X_t - TC_t^{(2)}$
7. Estimation of the Seasonal component by 3×5 moving average over each month: $S_t^{(2)} = M_{3\times 5}\left[(S_t + I_t)^{(2)}\right]$ and normalization: $Snorm_t^{(2)} = M_{2\times 12}\left[S_t^{(2)}\right]$
8. Estimation of the seasonally adjusted series:
   $Xsa_t^{(2)} = X_t - Snorm_t^{(2)} = (TC_t + I_t)^{(2)}$

Seasonal Adjustment Quality Issues

- No residual seasonality (trading day, etc.)
    - Various tests: Fisher, Ljung-Box, Friedman etc.;
- Adequate stability (e.g., large revisions indicate lack of usefulness)
- Adequate smoothness (for interpretability)

X-12: M and Q-statistics

- Basic idea: to check if X-11 filters are well adapted to the series
- Basic assumptions of X-11 filters: the series is not very noisy and the seasonal factors do not evolve a lot.
- Therefore, basically 2 kind of statistics:
  - on the irregular: randomness (M4), "size" (M1, M2, M3, M5, M6)
  - on the seasonal component: M7 to M11
- Q-statistics are linear combinations of these 11 indicators.

M-statistics

M1	The relative contribution of the irregular over three months span (from Table F 2.B).
M2	The relative contribution of the irregular component to the stationary portion of the variance (from Table F 2.F).
M3	The amount of month to month change in the irregular component as compared to the amount of month to month change in the trend-cycle (from Table F2.H).
M4	The amount of autocorrelation in the irregular as described by the average duration of run (Table F 2.D).
M5	The number of months it takes the change in the trend-cycle to surpass the amount of change in the irregular (from Table F 2.E).
M6	The amount of year to year change in the irregular as compared to the amount of year to year change in the seasonal (from Table F 2.H).
M7	The amount of moving seasonality present relative to the amount of stable seasonality (from Table F 2.I).
M8	The size of the fluctuations in the seasonal component throughout the whole series.
M9	The average linear movement in the seasonal component throughout the whole series.
M10	Same as 8, calculated for recent years only.
M11	Same as 9, calculated for recent years only.

They are all normalized to 1.

Q-statistics

- Linear combinations of the M-Statistics;
  $Q_1 = \frac{10M_1+11M_2+10M_3+8M_4+11M_5+10M_6+18M_7+7M_8+7M_9+4M_{10}+4M_{11}}{100}$
- Q_2: without the M_2 statistics;
- If M_6 does not make sense (no 3x5), weight 0;
- Easy to interpret: good if < 1

Towards a simple Quality Report (QR)

- X13 outputs a lot (too many?) of diagnostics;
- It is important in production time to get an idea of the quality of the adjustment at a glance;
- How to summarize the information? Several basic common sense ideas:
    - Testing for the presence of seasonality, outliers and calendar effects;
    - How important are seasonality and calendar effects?
    - An idea of the overall quality of the adjustment (Q statistics);
    - Testing for the absence of residual seasonality and calendar effects:

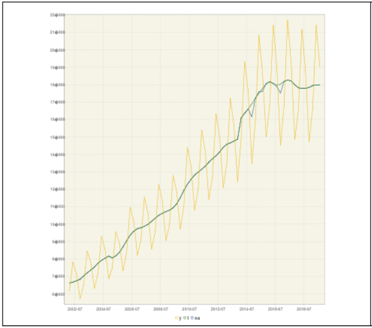
A basic selection of indicators (1)

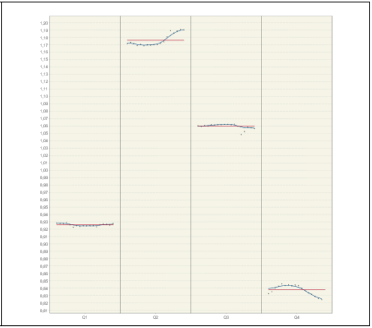
- The graph of the series, the seasonally adjusted series and the trend-cycle;
- The graph of the Seasonal-Irregular component;
- The number of observations and the span of the series;
- The decomposition model;
- The presence/absence of trading-day effects;
- The presence/absence of specific calendar effects?
- The ARIMA model used;

A basic selection of indicators (2)

- The list of detected outliers;
- The Q-statistics;
- The results of the tests for the presence of seasonality in the linearized series;
- The results of the tests for the presence of residual seasonality in the seasonally adjusted series;
- The results of the tests for the presence of residual trading-day effects in the seasonally adjusted series;
- The decomposition of the variance of the stationary part of the series.

An example





	ARIMA
Estimation span: [I-2002 - III-2019], 71 observations	(0,1,1)(0,1,1)
Series has been log-transformed	
No trading days effects	
No Easter effect	

Q	Q2
0,069	0,079

Cycle	Seasonality	Irregular	TDH
1,67	98,31	0,02	0,00

Outliers	Coefficient	T-Stat
LS (I-2005)	-0,0291	-3,85
LS (I-2014)	0,0671	8,81
AO (IV-2014)	-0,0442	-9,94
AO (IV-2016)	-0,0291	-6,53

	Fisher	Lung-Box	Kruskal-Wallis	Friedman	Combined
Seasonality	0,00	0,00	0,02	0,00	Present
Residual Seasonality	0,27	0,01	1,00	0,20	None
Residual TD effect	0,91				