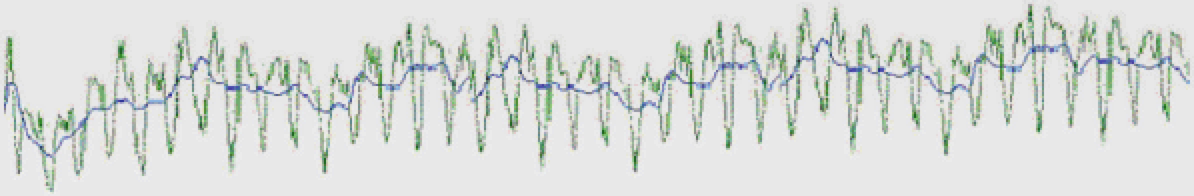
Seasonal Adjustment Tutorial: The Basics

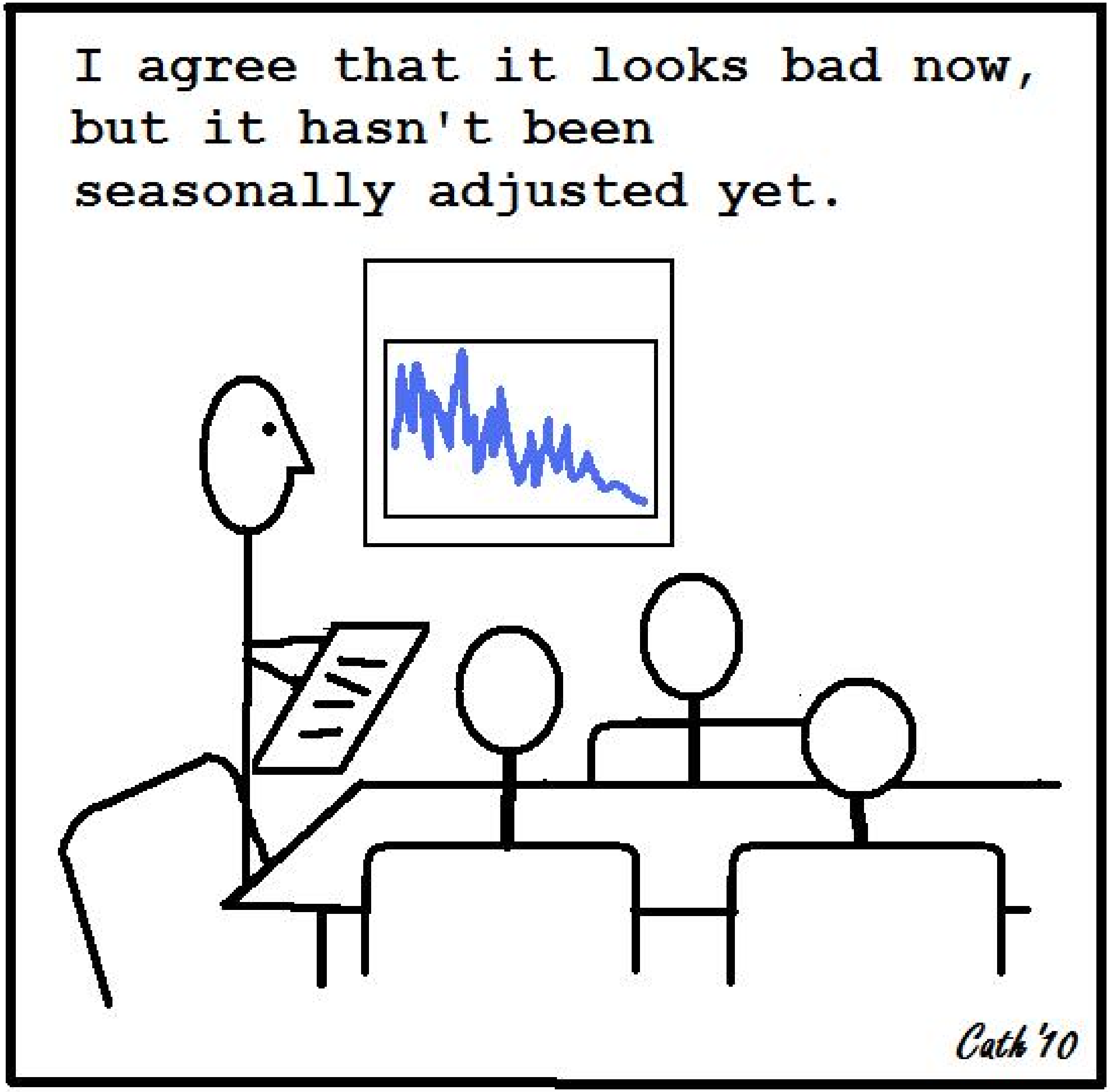
Catherine C.H. Hood

Featuring Brian C. Monsell (retired US Census Bureau, now at BLS)

Catherine Hood
CONSULTING

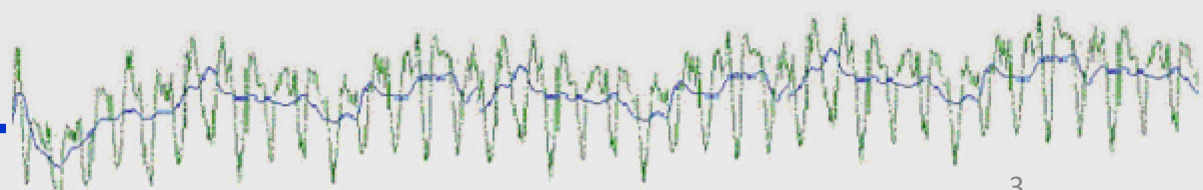


* Random “improv” slides provided by Elijah L. Hood


I agree that it looks bad now, but it hasn't been seasonally adjusted yet.
Cath '10

Time Series Analysis

- A ***time series*** is a set of observations ordered in time.
- The goals of time series analysis
  - Describe the data
  - Summarize the data
  - Fit models to the data
  - Forecast the data



What are we looking for in a time series?

- Important features of a time series, including direction, turning points, cycles, and patterns.
- Consistency between different time series so we can compare
  - Series with different seasonal patterns, or
  - Monthly series to quarterly series or to annual data.

Comparing Series with Different Seasonal Patterns

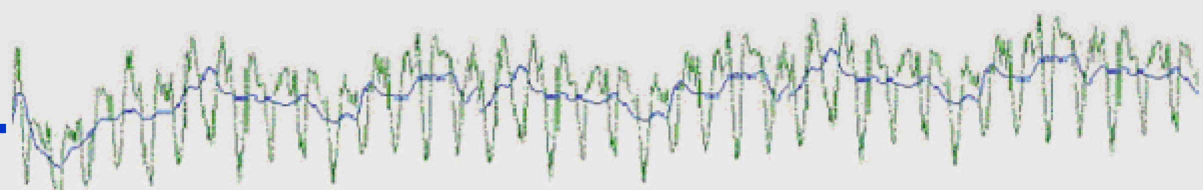
- Example:
  - November to December Month-to-Month Percent Change (from a random year in the past)

    U.S. Midwest Total Housing Starts:

    17091 / 24687 → –31%

    U.S. Department Store Sales:

    117029 / 79395 → 47%

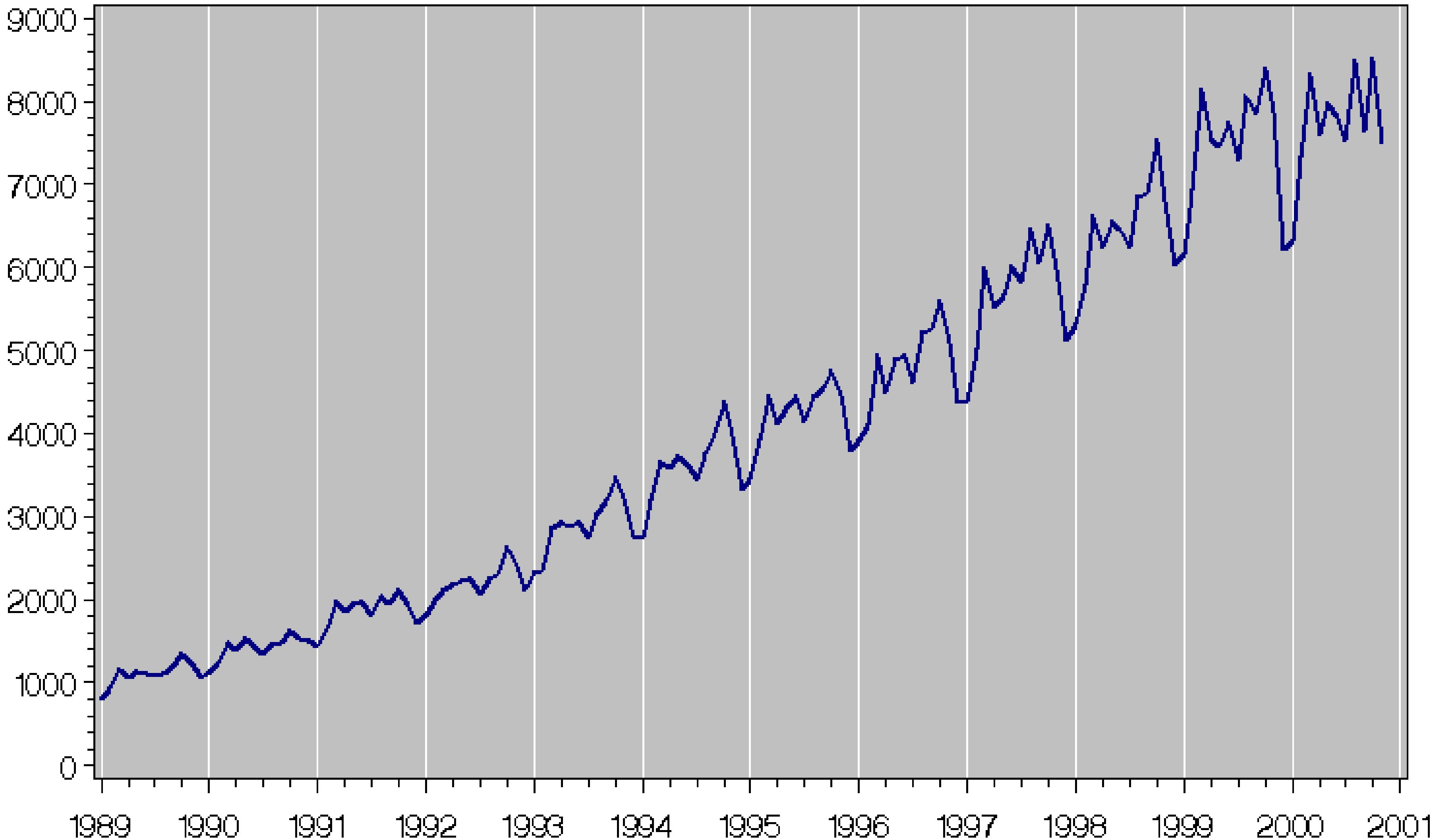


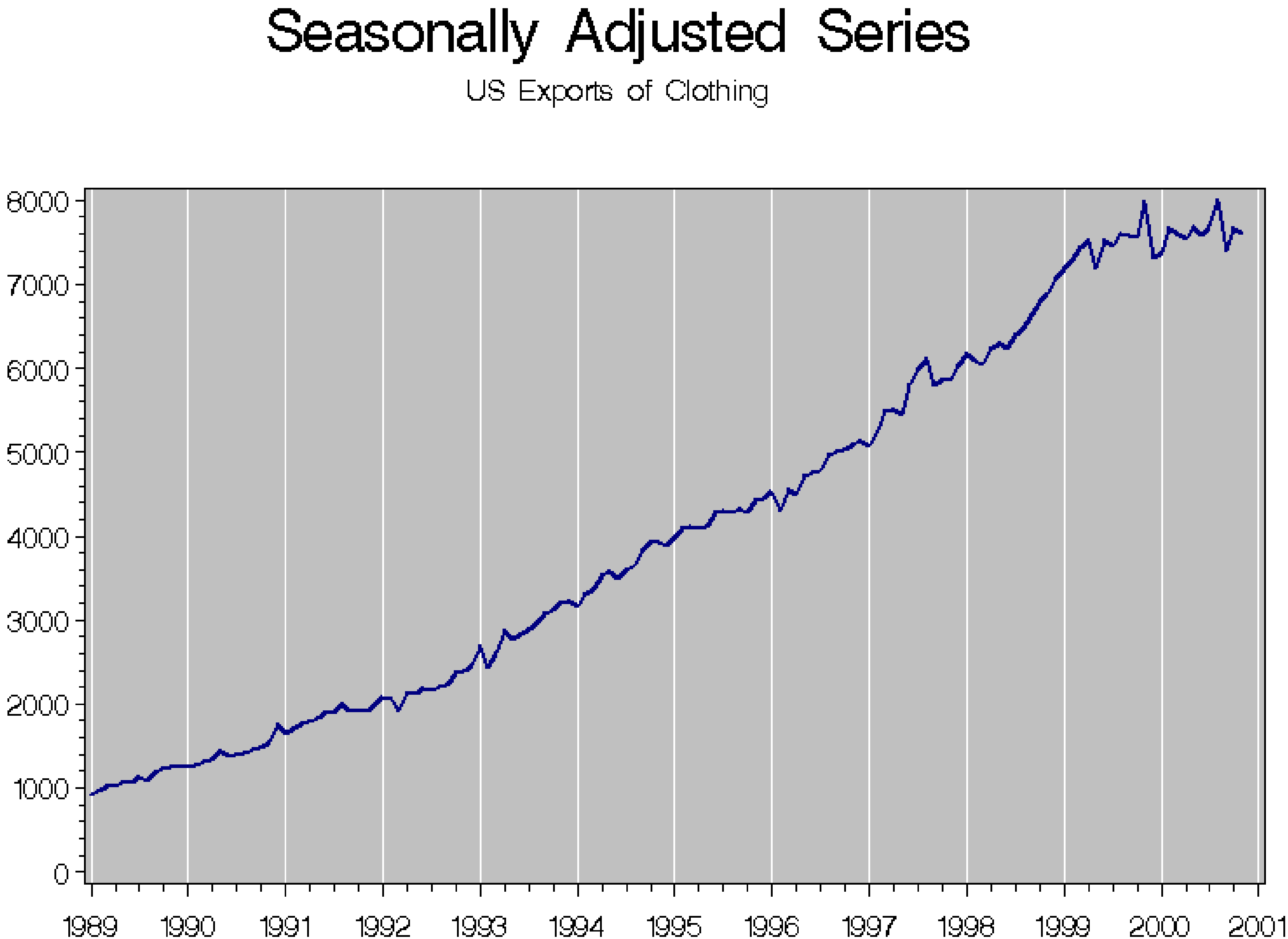
Seasonal Adjustment

- Seasonal movements can make the features we are interested in either difficult or impossible to see.
- The estimation and removal of the seasonal fluctuations from a times series is what we call ***seasonal adjustment***.

Original Series

US Exports of Clothing




Seasonally Adjusted Series
US Exports of Clothing
8000
7000
6000
5000
4000
3000
2000
1000
0
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001

Seasonal Effects

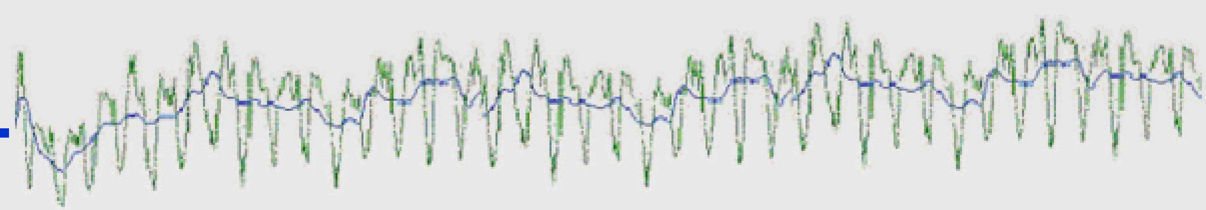
- Reasonably stable in terms of annual timing, direction, and magnitude.
- Possible causes are
  - Natural factors
  - Administrative or legal measures
  - Social/cultural/religious traditions (e.g., fixed holidays, timing of vacations)

Trend-Cycle

- The basic level of the series.
- Reasonably smooth.
- Includes ***cycles*** — cyclical fluctuations longer than a year — if there are any in the series.
- Includes ***turning points*** — places where the series changes from increasing to decreasing, or *vice versa* — if there are any in the series.

Irregular Effects

- Unpredictable in terms of timing, impact, and duration.
- Possible causes
  - Unseasonable weather/natural disasters
  - Strikes
  - Sampling error
  - Nonsampling error

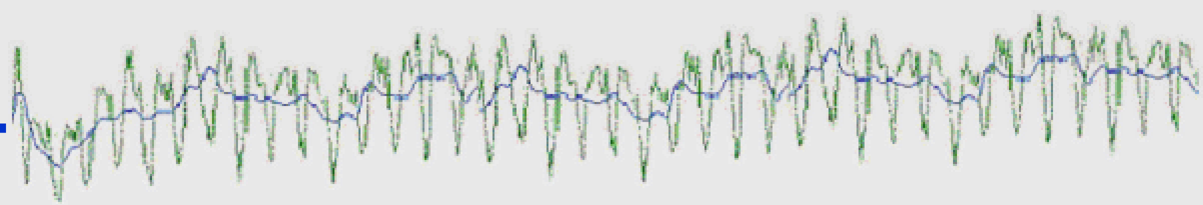


Other Effects

- Trading Day: The number of working or trading days in a period.
- Moving Holidays: Events which occur at regular intervals but not at exactly the same time each year.
- Combined Effects: Trading day and moving holiday effects are persistent, predictable, calendar-related effects, so they are often included with the seasonal effects to form “combined effects.”

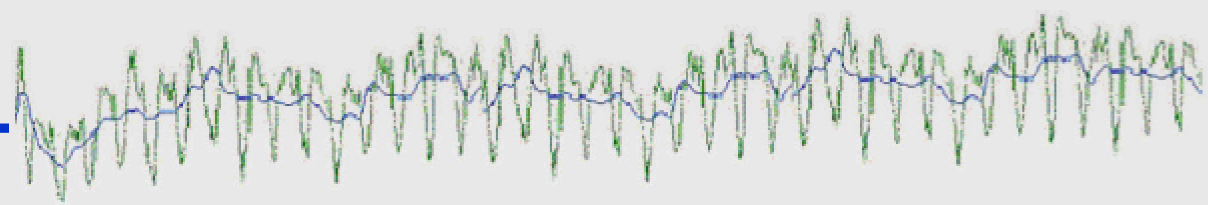
October 2019

S	M	T	W	T	F	S
		1	2	3	4	5
6	7	8	9	10	11	12
13	14	15	16	17	18	19
20	21	22	23	24	25	26
27	28	29	30	31		



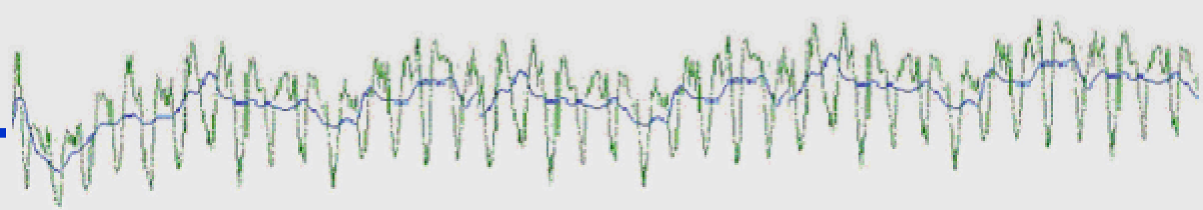
November 2019

S	M	T	W	T	F	S
					1	2
3	4	5	6	7	8	9
10	11	12	13	14	15	16
17	18	19	20	21	22	23
24	25	26	27	28	29	30



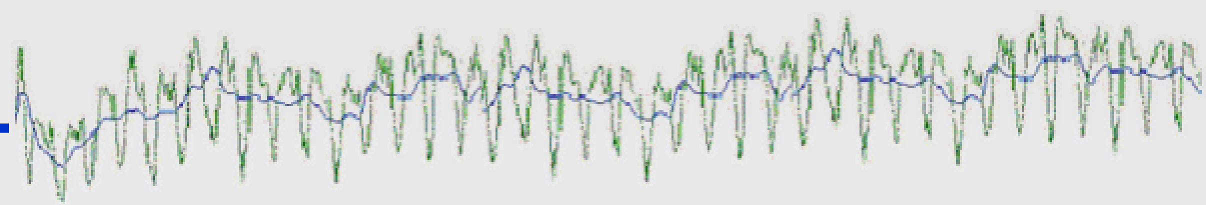
Notation

- Y_t = original series
- C_t = trend-cycle
- I_t = irregular
- S_t = seasonal
- TD_t = trading day
- H_t = moving holiday
- S'_t = combined factors
- A_t = adjusted series



What Seasonal Adjustment Can NOT Do

- The seasonal adjustment process estimates the irregular component, but it does NOT remove the irregular.
- The seasonal adjustment process will NOT remove the turning points or change the direction of the series.



Describing a Time Series

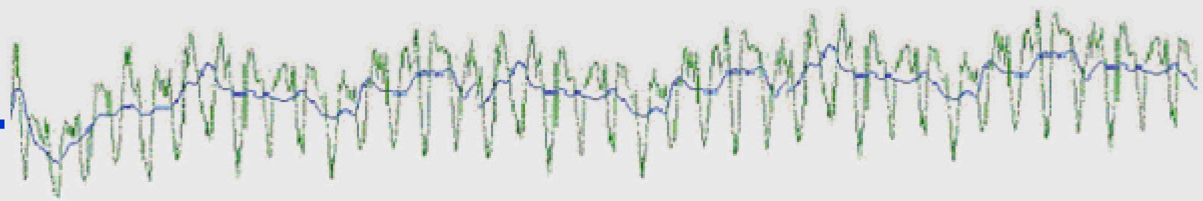
- There is NOT a unique way to represent a series in the time domain.
- Two popular ways to describe time series:
  - Classical Decomposition
  - ARIMA Models

Classical Decomposition

- One method of describing a time series:

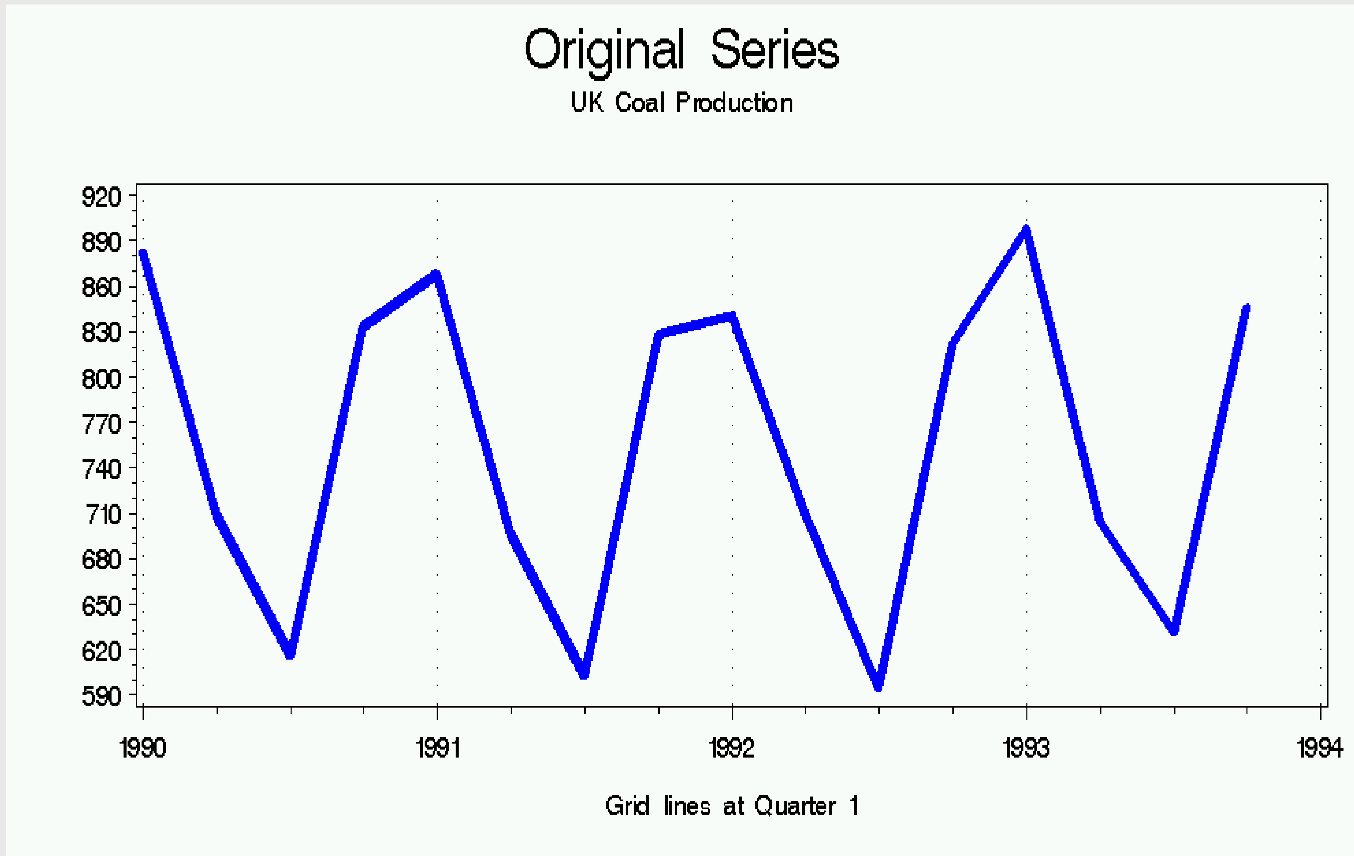
$$Y_t = S_t + C_t + I_t,$$

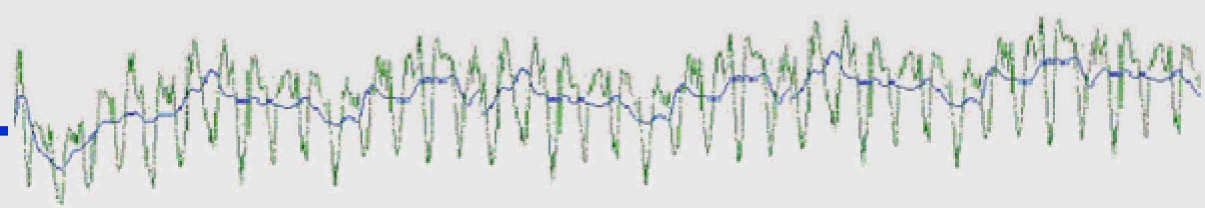
- Two possible estimates:
  - Seasonal adjustment (remove effects of S_t): $A_t = C_t + I_t$
  - Trend-cycle (remove effects S_t and I_t): C_t

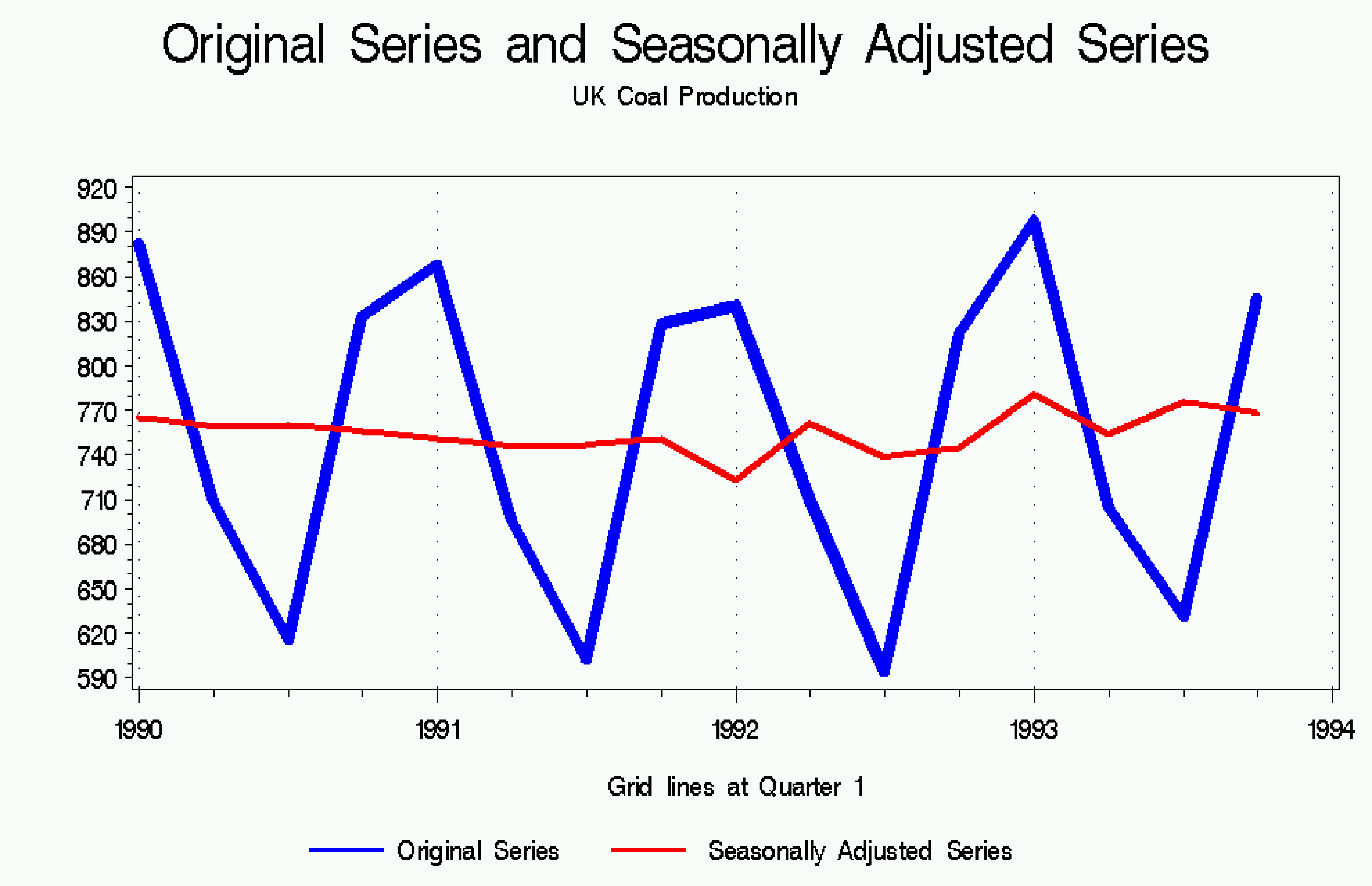


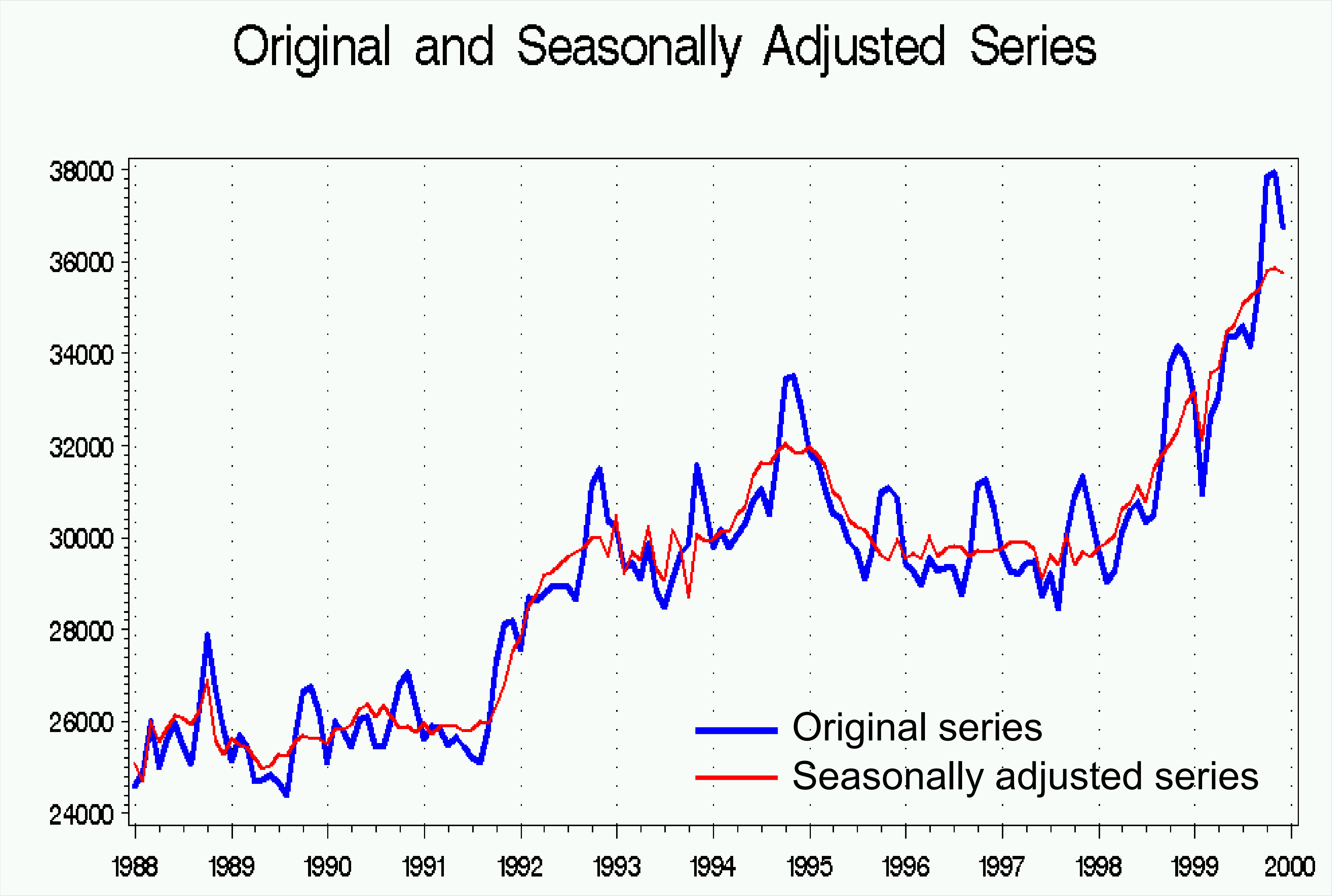
Series #1

1990	1	882
1990	2	709
1990	3	616
1990	4	833
1991	1	868
1991	2	696
1991	3	603
1991	4	828
1992	1	840
1992	2	711
1992	3	594
1992	4	822
1993	1	898
1993	2	704
1993	3	631
1993	4	845






Original Series and Seasonally Adjusted Series
UK Coal Production
920
890
860
830
800
770
740
710
680
650
620
590
1990
1991
1992
1993
1994
Grid lines at Quarter 1
Original Series
Seasonally Adjusted Series


Original and Seasonally Adjusted Series
38000
36000
34000
32000
30000
28000
26000
24000
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
Original series
Seasonally adjusted series

Problem:

- Trend isn’t flat.

Solution:

- Estimate the trend and remove it
- Proceed as before with the detrended data

Problem:

- Trend has cyclical movements.

Solution:

- Local smoothing

Problem:

- Estimating the trend-cycle in the presence of seasonal movements is difficult.
- Estimating seasonal movements is difficult in the presence of a trend-cycle.

Solution:

- Iterate between estimating the trend and seasonal estimation to get successively more refined estimates of the seasonal and trend.

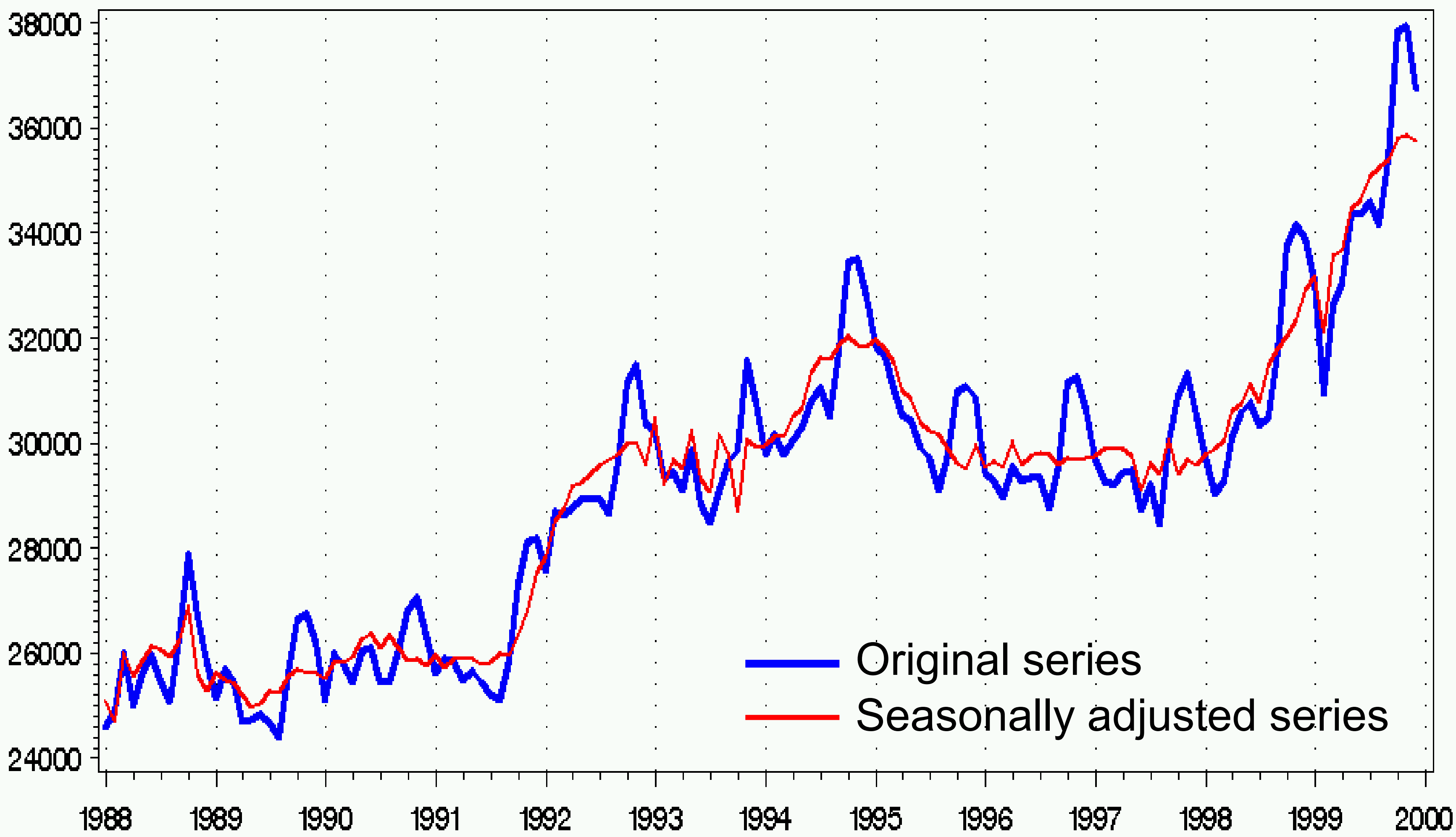
Problem:

- Variation increases as the level increases.

Solution:

- Take the logs of the series.

Original and Seasonally Adjusted Series



Models

Multiplicative model:

$$Y_t = S_t' \times C_t \times I_t,$$

where

$$S_t' = S_t \times TD_t \times H_t$$

$$A_t = C_t \times I_t$$

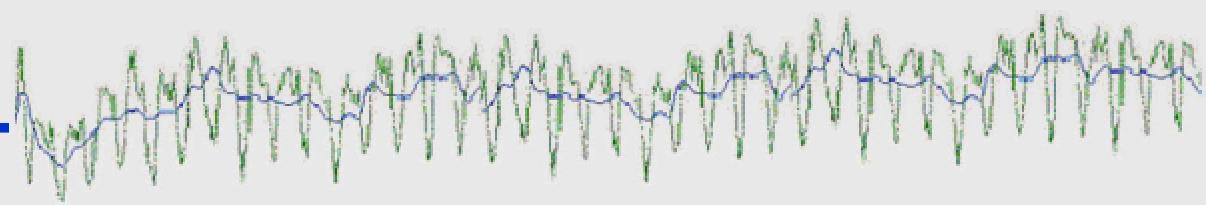
Additive model:

$$Y_t = S_t' + C_t + I_t,$$

where

$$S_t' = S_t + TD_t + H_t$$

$$A_t = C_t + I_t$$

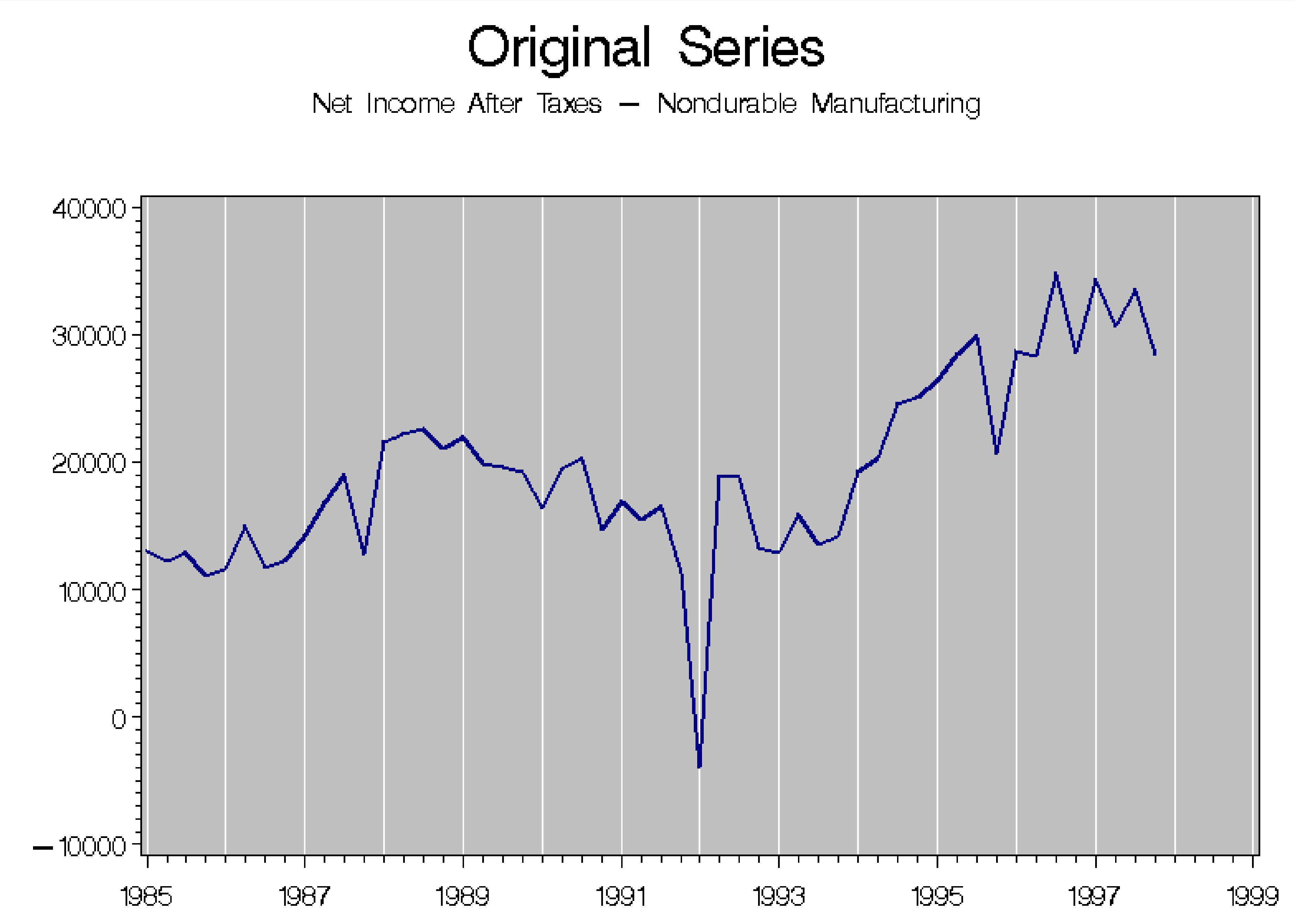


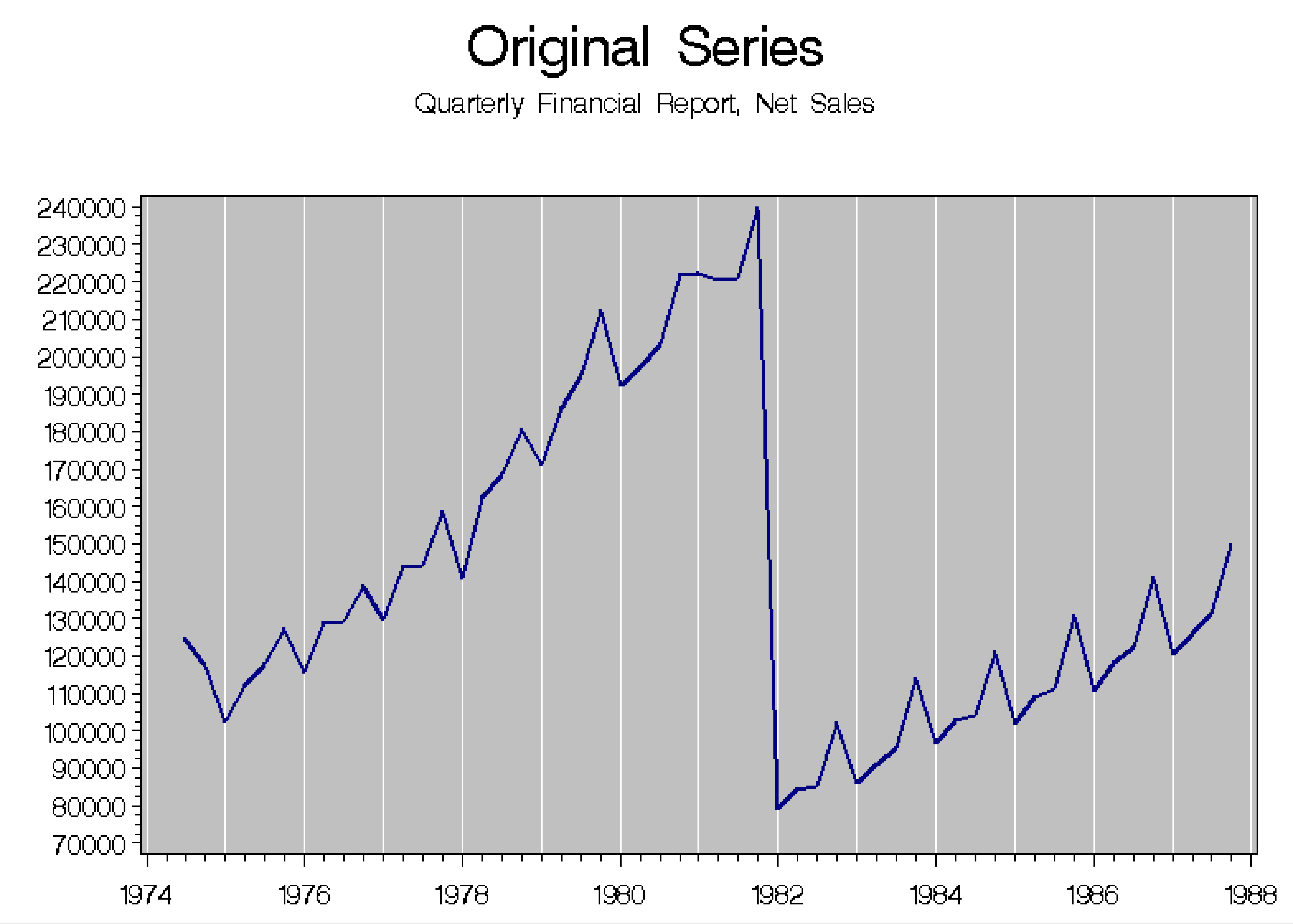
Problem:

- Trading day, moving holidays, and extreme values may be present.

Solution:

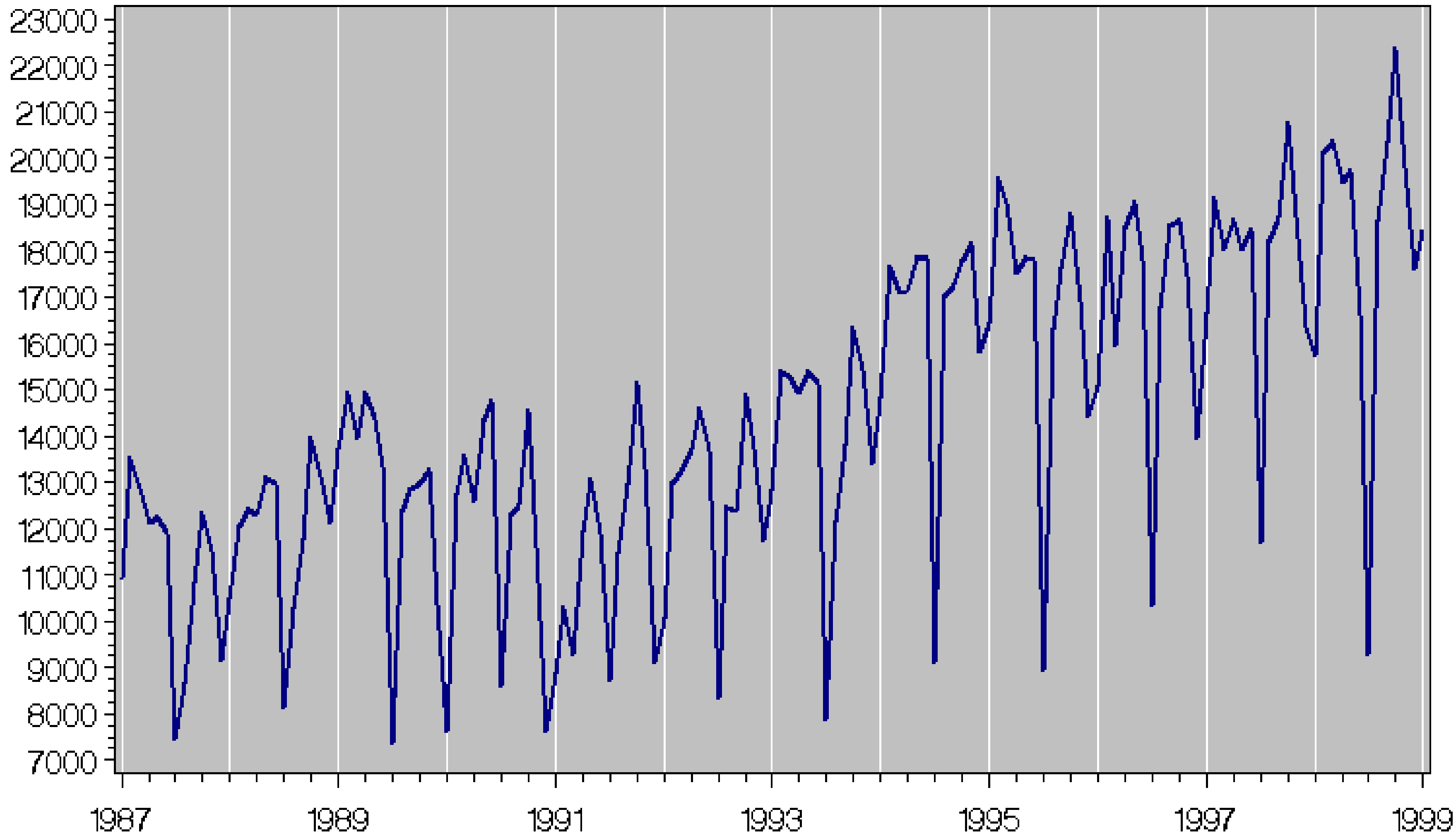
- These effects can be estimated and removed from the series, but they can be difficult to identify and estimate when seasonality and trend are present.


Original Series
Net Income After Taxes — Nondurable Manufacturing
40000
30000
20000
10000
0
−10000
1985
1987
1989
1991
1993
1995
1997
1999


Original Series
Quarterly Financial Report, Net Sales
240000
230000
220000
210000
200000
190000
180000
170000
160000
150000
140000
130000
120000
110000
100000
90000
80000
70000
1974
1976
1978
1980
1982
1984
1986
1988

Original Series

Motor vehicles (U37BVS): Default X12



What would help us eliminate the seasonality?

- Iterative refinement
- Local smoothing
- Robustness against extreme values
- Holiday and Trading Day estimation

ARIMA Models

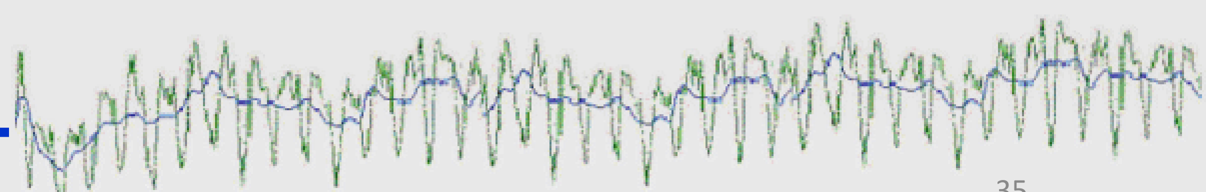
- ARIMA stands for AutoRegressive Integrated Moving Average.
- One way to describe time series.
- Mathematical models of the autocorrelation in a time series.
- Widely used in a variety of fields.
- Popularized by Box and Jenkins (1970).

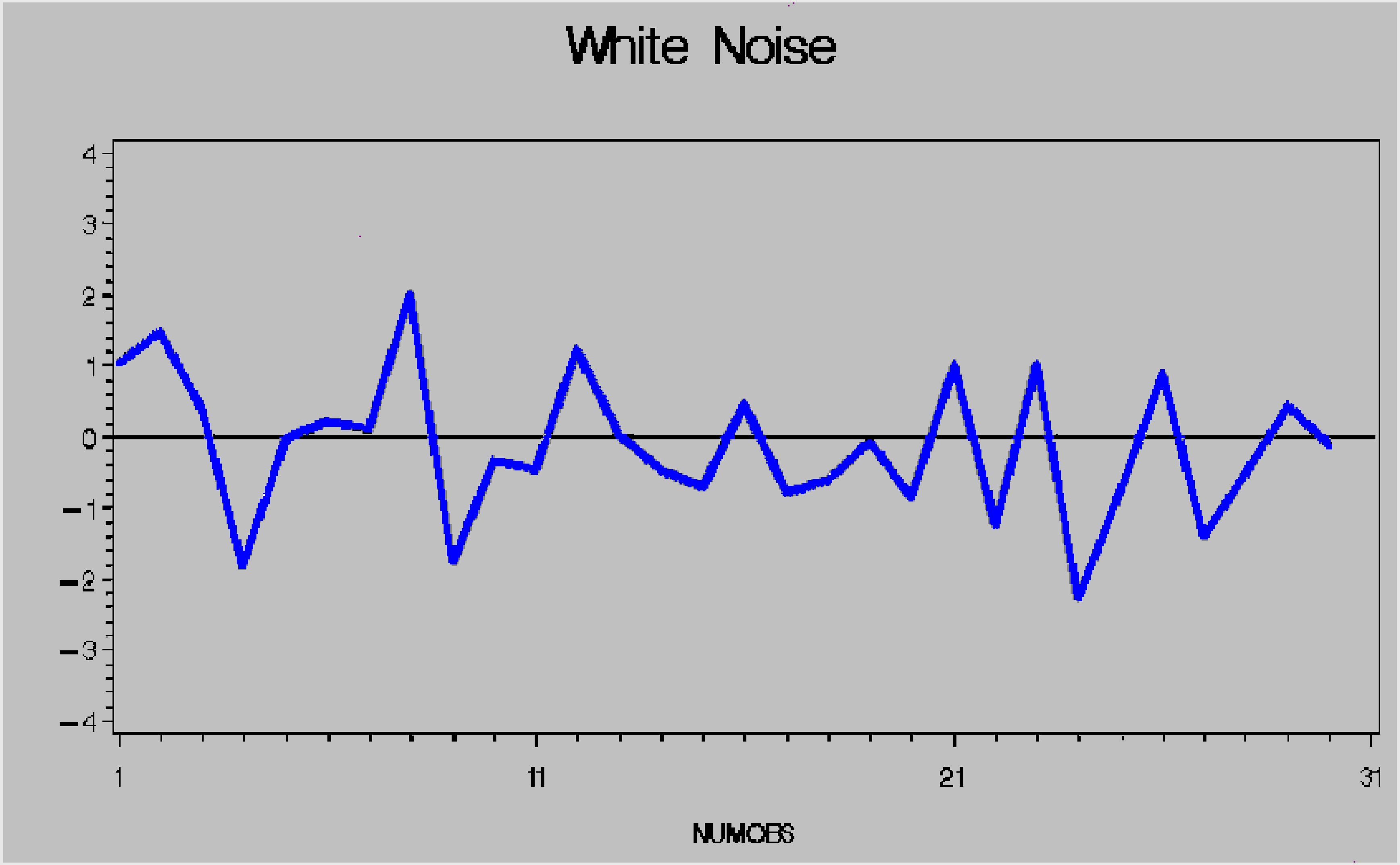
Stochastic Process

- An underlying process +
  random component (white noise)

White Noise

- Random drawings from a fixed distribution, usually assumed to be Normal with mean 0 and variance σ_a^2.
- Notation: a_t




White Noise
4
3
2
1
0
-1
-2
-3
-4
1
11
21
31
NUMOBS

A Related Process

- The current value depends on the previous value times a constant

$$y_t = \varphi\, y_{t-1} + a_t$$

where a_t is white noise and φ is a constant.

- This process is called "autoregressive" – the series is regressed on past values of itself, and because there is one term, it's also called an AR(1) model.

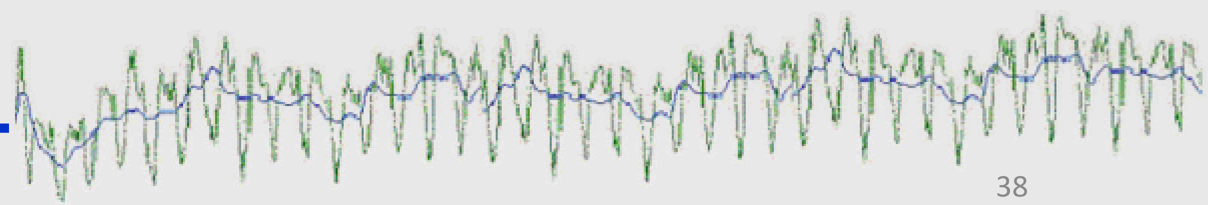
Another Process

- AR(2) – the current value depends on two previous values times constants, plus white noise.

$$y_t = \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + a_t$$

where a_t is white noise and

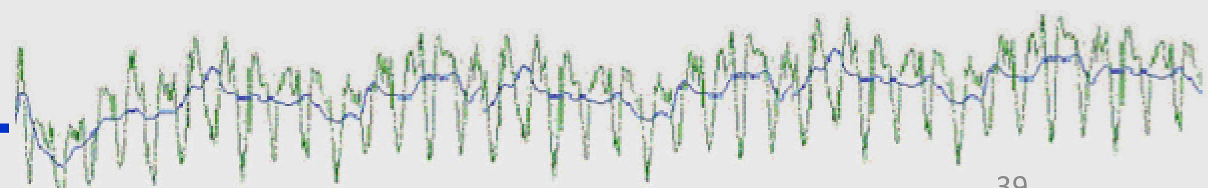
φ_1 and φ_2 are constants



Seasonal Process

- Seasonal models relate the series to past values at seasonal lags.
- For example, for a monthly time series, we could have a seasonal AR process

$$y_t = \Phi\, y_{t-12} + a_t$$



More Complicated Process

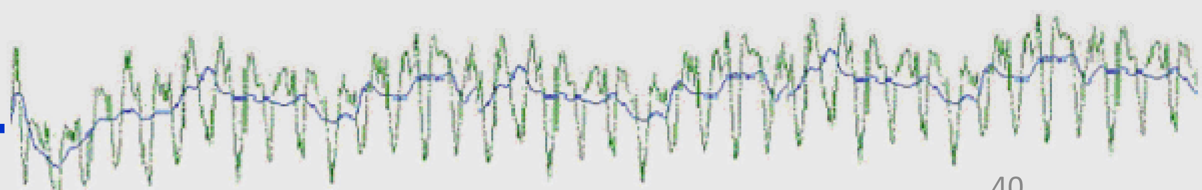
- A series may relate to the past value and the past seasonal value:

$$y_t = \varphi\, y_{t-1} + \Phi\, y_{t-S} + a_t$$

where a_t is white noise,

φ and Φ are constants, and

S is the period (12 for monthly series and 4 for a quarterly series)



Difference

- The AR(1) process with $\varphi = 1$ gives us this equation

$$y_t = y_{t-1} + a_t$$

- Can also think about this as taking the difference between the two points—rewriting the equation as

$$y_t - y_{t-1} = a_t$$

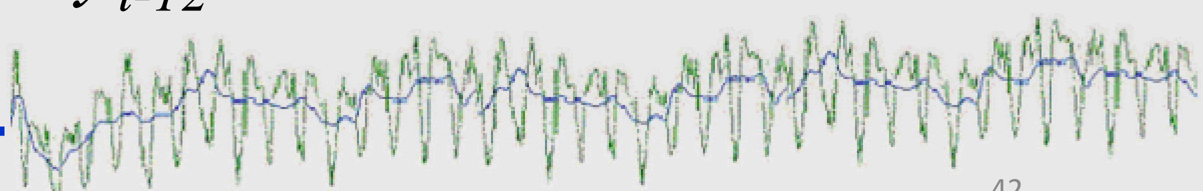
Integrate/Difference

- "I" stands for Integrated, the opposite of differencing
  - First difference – subtracting the previous value from the current value

$$a_t = y_t - y_{t\text{-}1}$$

  - First seasonal difference – subtracting the previous year's value from the current value

$$a_t = y_t - y_{t\text{-}12}$$



Moving Average Process

- The current value depends on lags of the white noise a_t instead of lags of itself

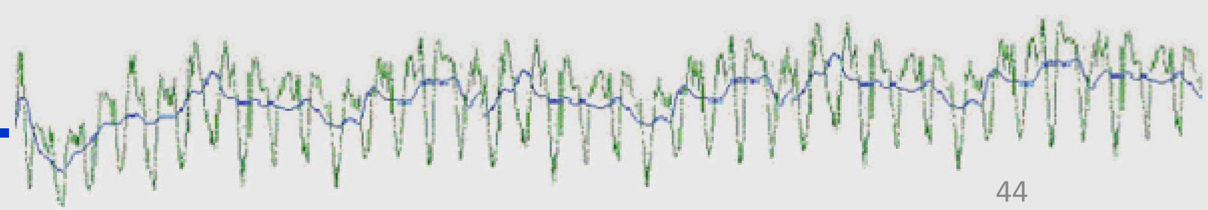
$$y_t = a_t - \theta\, a_{t-1}$$

where a_t is white noise and

θ is a constant

ARIMA Models

- AutoRegressive Integrated Moving Average models
- Usually designated (p d q)(P D Q) where
  - p is the order of the AR model
  - d is the number of differences (integration)
  - q is the order of the MA model
  - P is the order of the seasonal AR model
  - D is the number of seasonal differences
  - Q is the order of the seasonal MA model



ARIMA(0 1 1)

- An MA(1) model for the first differenced series

$$y_t - y_{t-1} = a_t - \theta\, a_{t-1}$$

$$y_t = y_{t-1} + a_t - \theta\, a_{t-1}$$

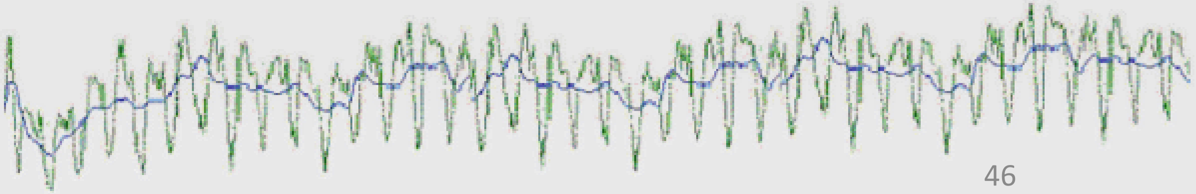
ARIMA(0 1 1)(0 1 1)

- This model combines a differenced series, a seasonally differenced series, an MA(1) model, and a seasonal MA(1) model

$$(y_t - y_{t-1}) - (y_{t-12} - y_{t-13}) = (a_t - \theta a_{t-1}) - (\Theta a_{t-12} - \Theta\theta a_{t-13})$$

- Rewritten:

$$y_t = y_{t-1} + y_{t-12} - y_{t-13} + a_t - \theta a_{t-1} - \Theta a_{t-12} + \Theta\theta a_{t-13}$$



Airline Model

- The most common type of ARIMA model for economic time series is the ARIMA(0 1 1)(0 1 1) model.
- Called the "airline" model because of the Box and Jenkins book.

Scale Model: Fokker F28-4000

- Static Desk Model (does not fly)
- Twin Rolls-Royce Spey Jenkins Engines
- Striking Orange and Grey Design
- Tiny Box Windows
- 1:200 scale

RegARIMA Model

$$\log\left(\frac{Y_t}{D_t}\right) = \boldsymbol{\beta}'\boldsymbol{X}_t + \boldsymbol{Z}_t$$

transformations | ARIMA process

$\boldsymbol{X}_t$ = Regressor for trading day and holiday or calendar effects, additive outliers, temporary changes, level shifts, ramps, and user-defined effects

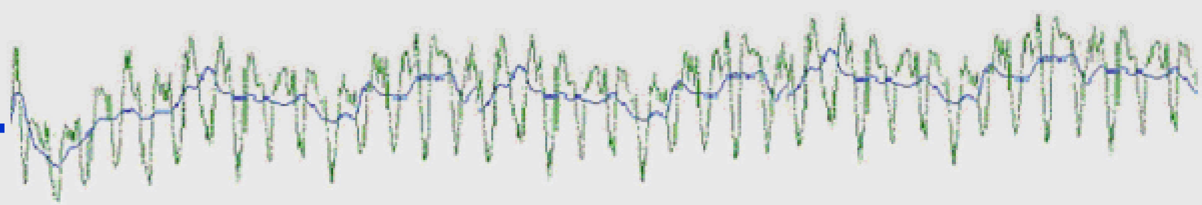
D_t = Leap-year or user-defined prior adjustment

Possible Regression Effects

- Outliers
- Trading day
- Moving holidays
- User-defined regressors

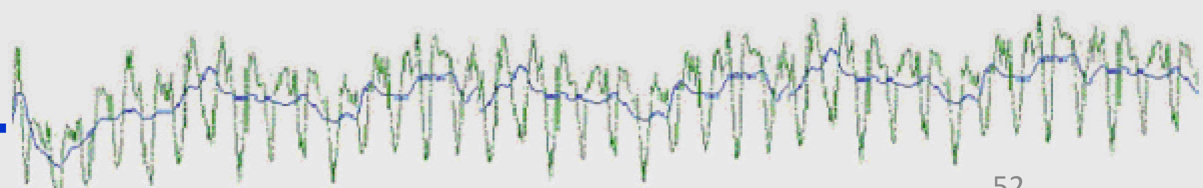
How Do We Estimate the Components and/or Find the Best ARIMA Model?

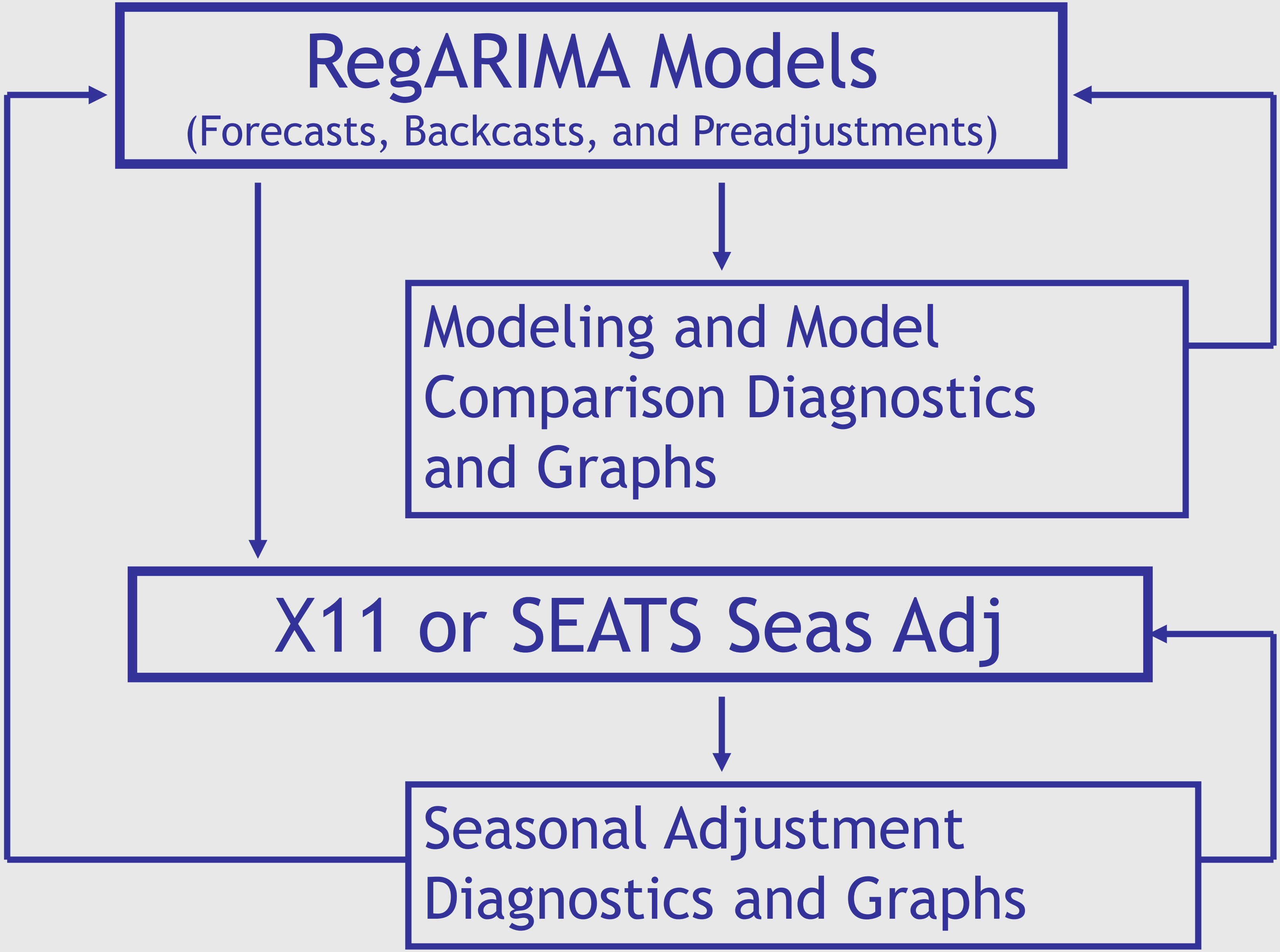
- Seasonal adjustment is normally done with off-the-shelf programs such as:
  - X13-ARIMA-SEATS (US Census Bureau),
  - TRAMO/SEATS (Bank of Spain),
  - Decomp, SABL, STAMP
- The best way to find the ARIMA model is to use an automatic modeling procedure, such as the one in X-13.



Two Pieces of X-13-ARIMA-SEATS

- "X11" or "SEATS"
  - The part of the program that does the seasonal adjustment
- "RegARIMA"
  - The part of the program that prior-adjusts the series before seasonal adjustment

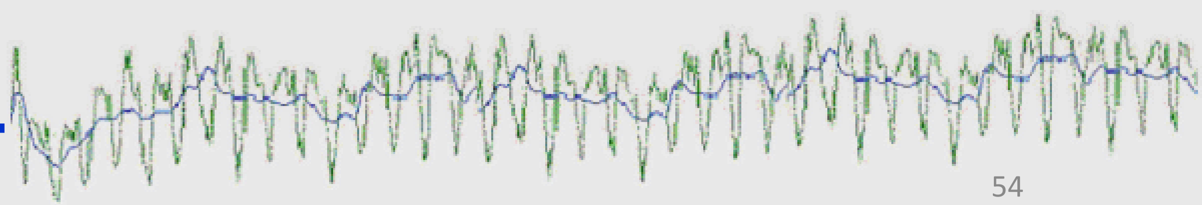



RegARIMA Models
(Forecasts, Backcasts, and Preadjustments)
Modeling and Model Comparison Diagnostics and Graphs
X11 or SEATS Seas Adj
Seasonal Adjustment Diagnostics and Graphs

The X11 Module

- The X11 spec generates a seasonal adjustment using X-11 seasonal adjustment methods.
- The X-11 algorithms rely on set of ***moving average filters***. It this context, a ***filter*** is weighted average where the weights sum to one.

$$y_t = \Sigma w_k x_{t+k}, \quad \Sigma w_k = 1$$



Simple Example

- Sheep population in the UK, 1867-1939
  - Annual data – no seasonality

  From the book *Time Series, 3rd edition* (1990) by Kendall and Ord, Oxford University Press: London

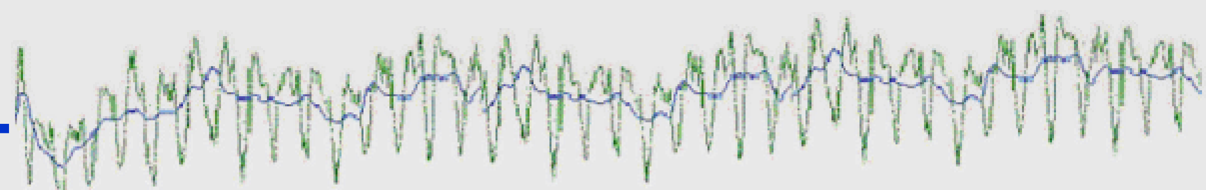
Sheep Population in the UK
2500
2400
2300
2200
2100
2000
1900
1800
1700
1600
1500
1400
1300
1860
1870
1880
1890
1900
1910
1920
1930
1940
YEAR

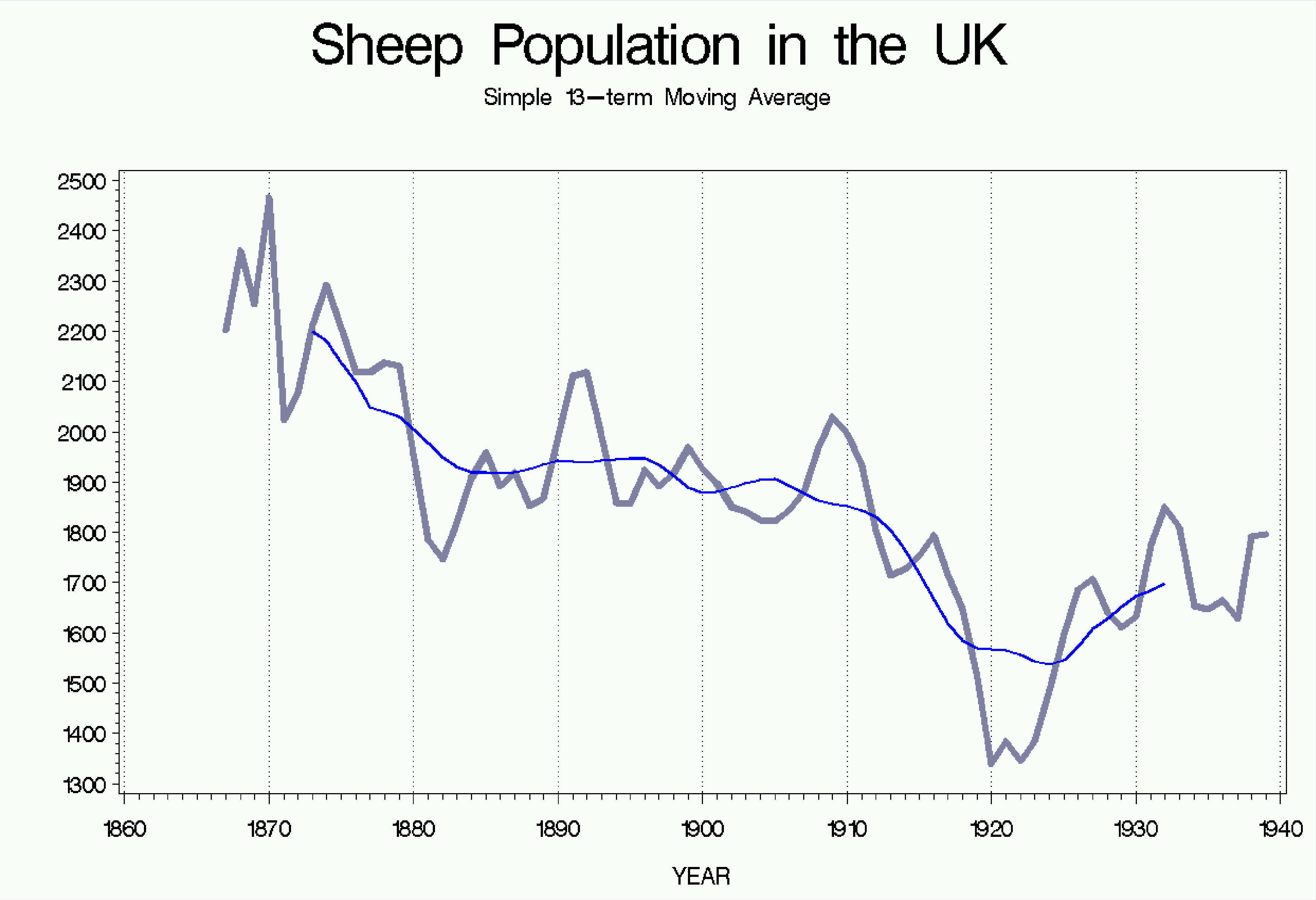
Example: 13-term Moving Average

$$\frac{x_{1867} + x_{1868} + x_{1869} + \ldots + x_{1878} + x_{1879}}{13}$$

So $y_t = (1/13)x_{t-6} + (1/13)x_{t-5} + \ldots + (1/13)x_t$

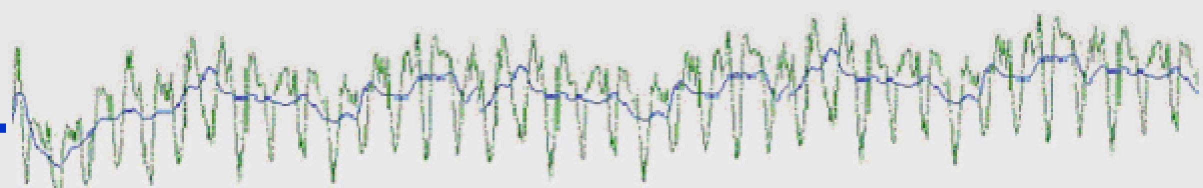
$+ \ldots + (1/13)x_{t+6}$




Sheep Population in the UK
Simple 13-term Moving Average
2500
2400
2300
2200
2100
2000
1900
1800
1700
1600
1500
1400
1300
1860
1870
1880
1890
1900
1910
1920
1930
1940
YEAR

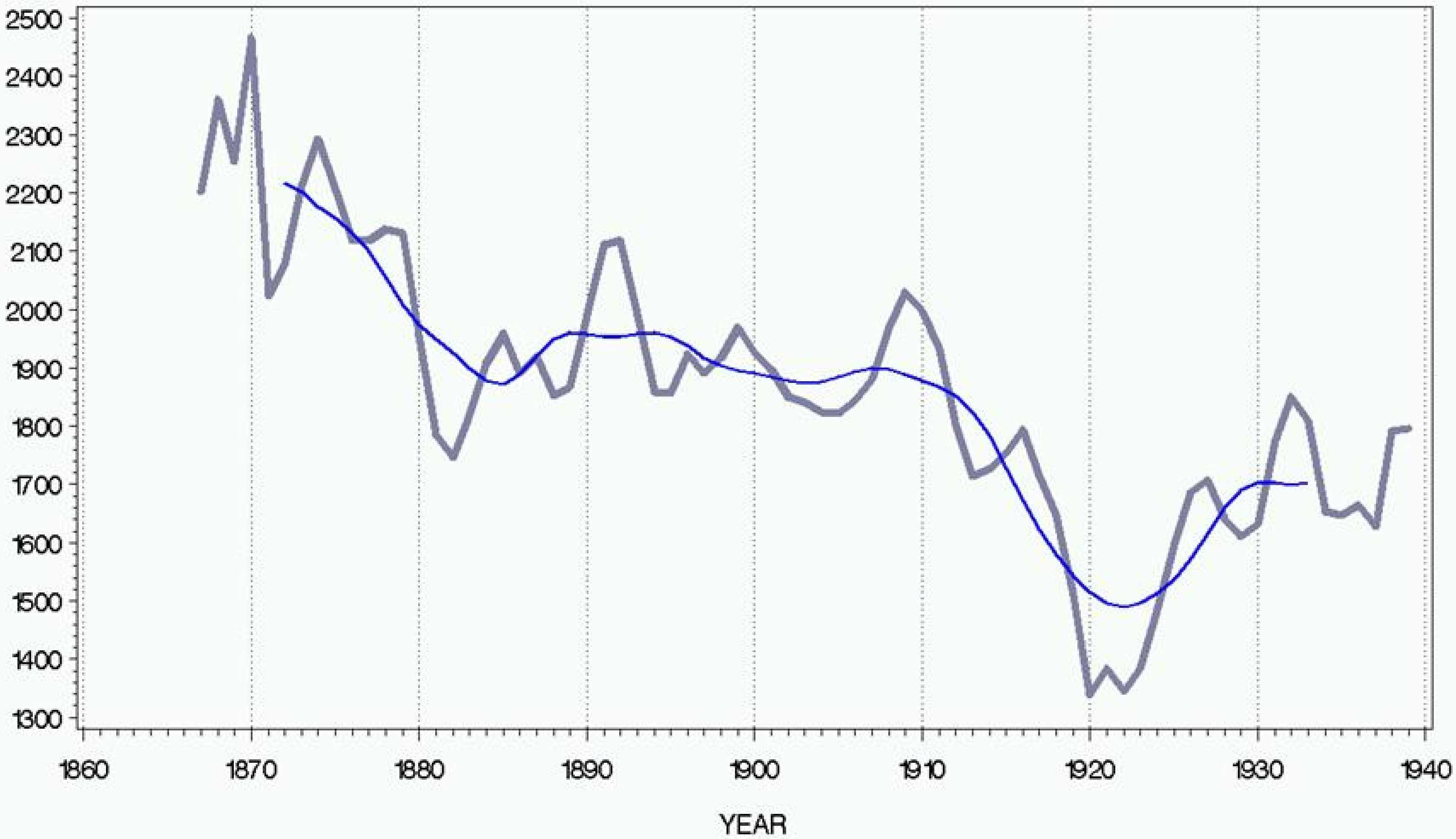
3 by 11 Moving Average

$$\frac{\begin{array}{l} Y_{1867} + Y_{1868} + \ldots + Y_{1877} + \\ \quad Y_{1868} + Y_{1869} + \ldots + Y_{1878} + \\ \quad\quad Y_{1869} + Y_{1870} + \ldots + Y_{1879} \end{array}}{33}$$



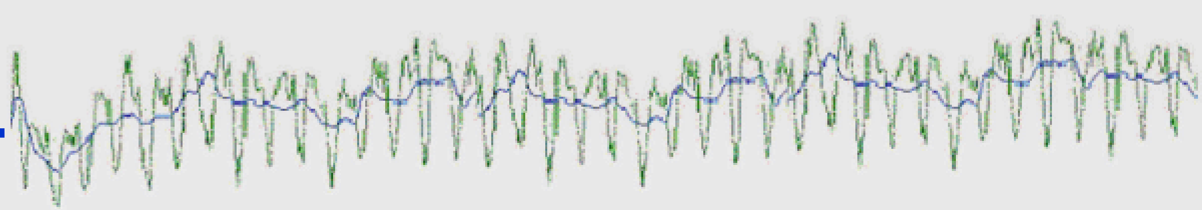
Sheep Population in the UK

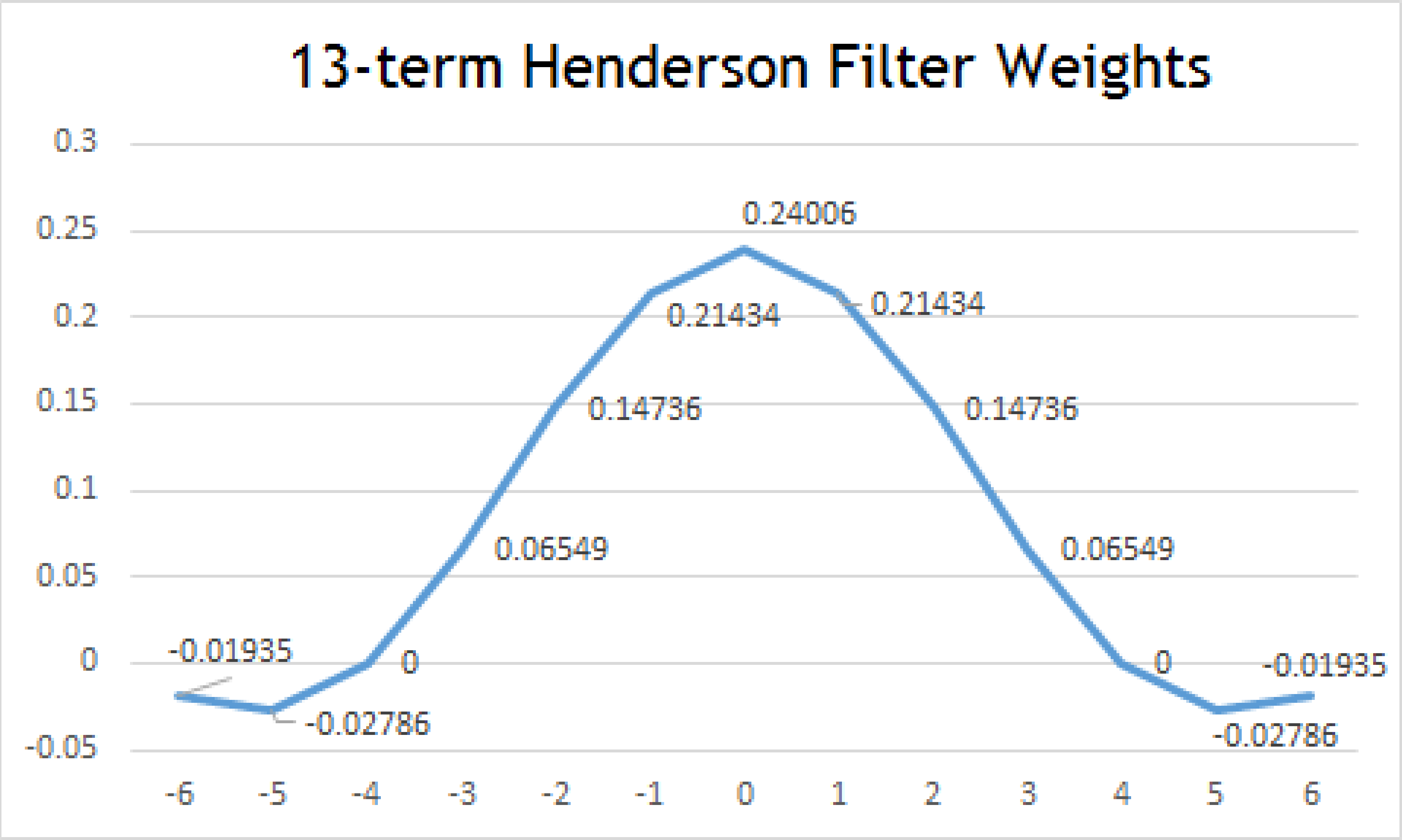
3x11 Moving Average



Henderson Filters, Background

- Derived by Robert Henderson in 1916 for his actuarial work.
- His idea was to develop a set of weights that would follow a cubic polynomial without distorting it.
- Henderson filters work well for economic time series because they don't change the trend or the cycles and yet will smooth out most of the irregular.

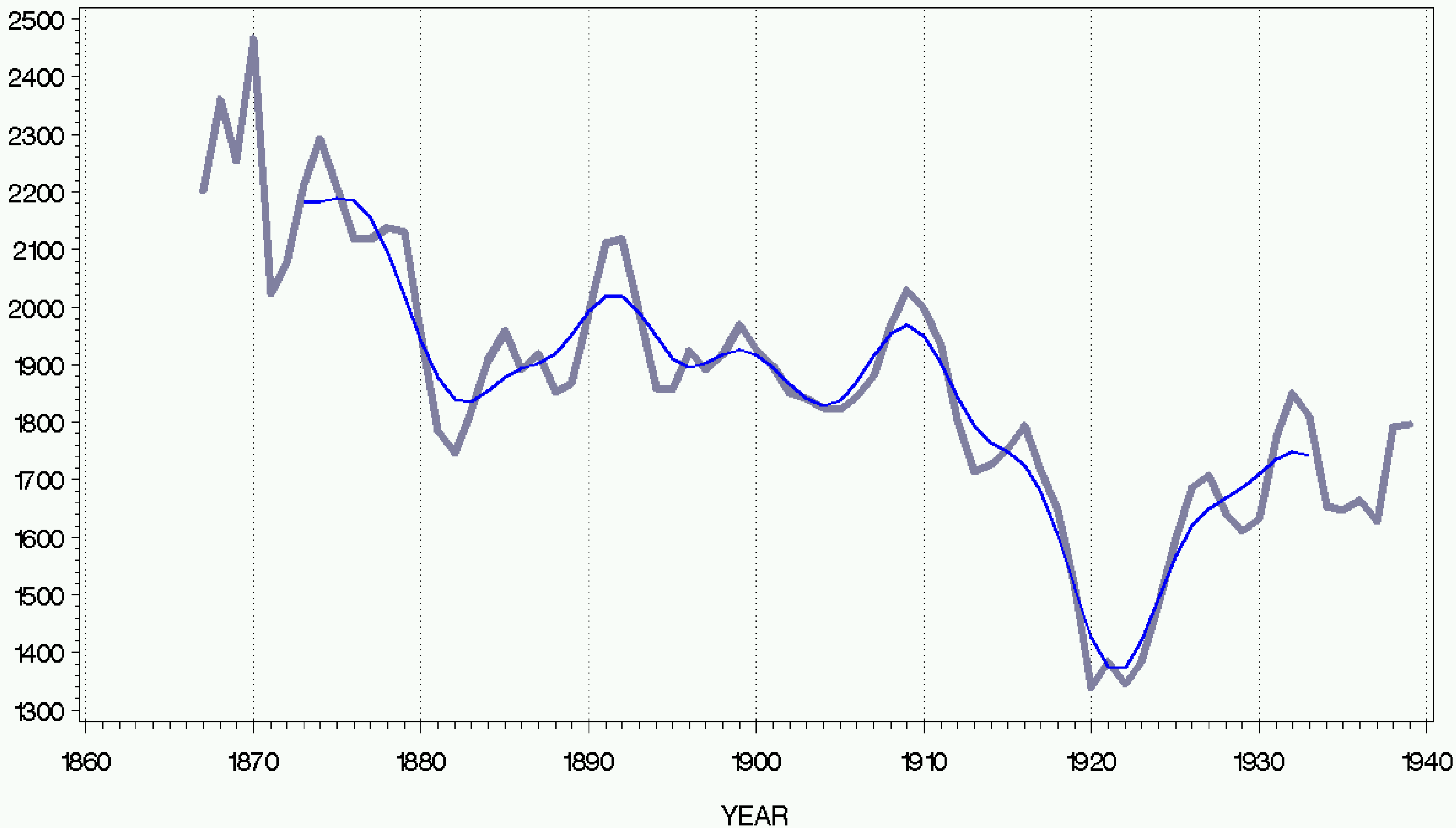


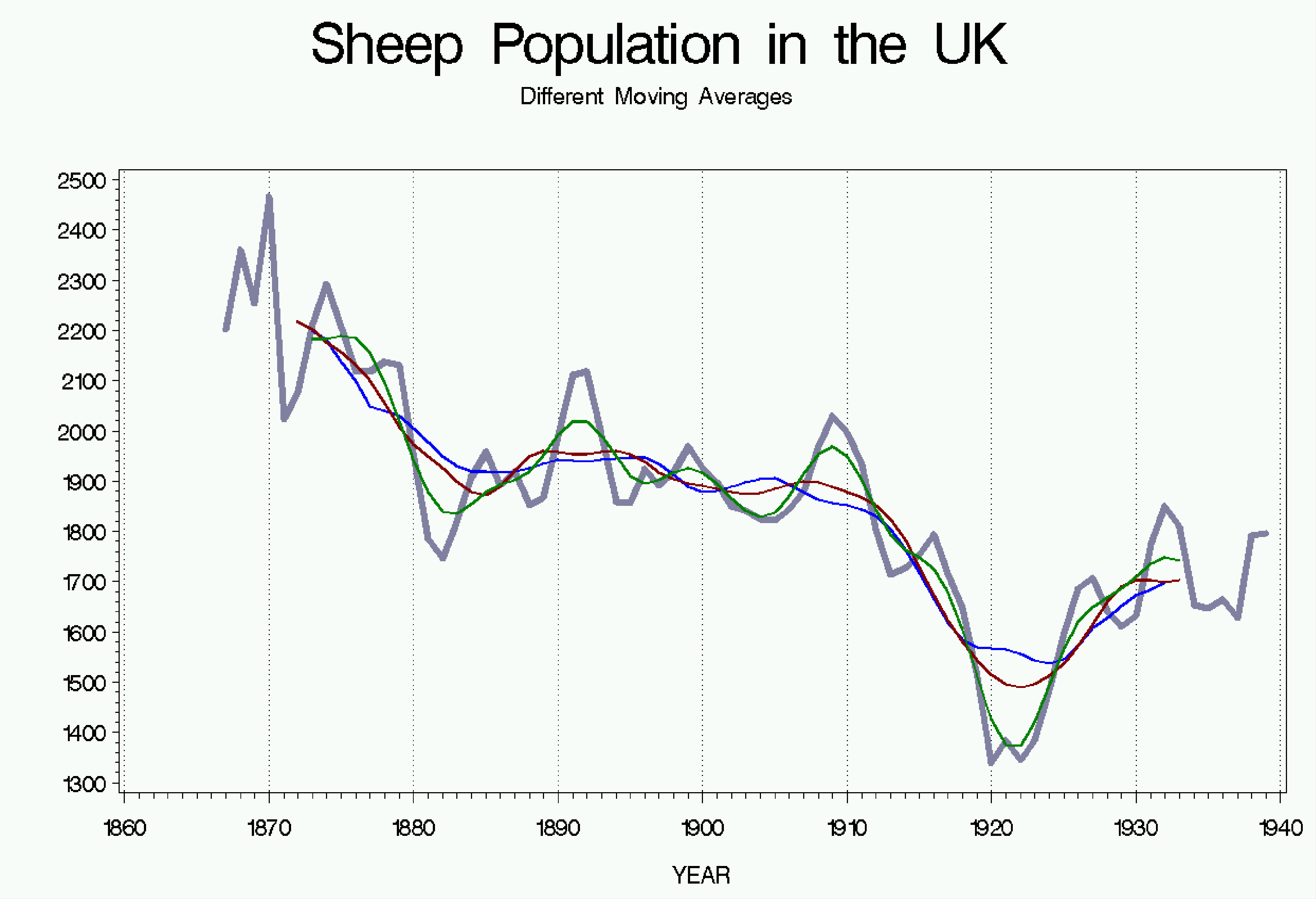


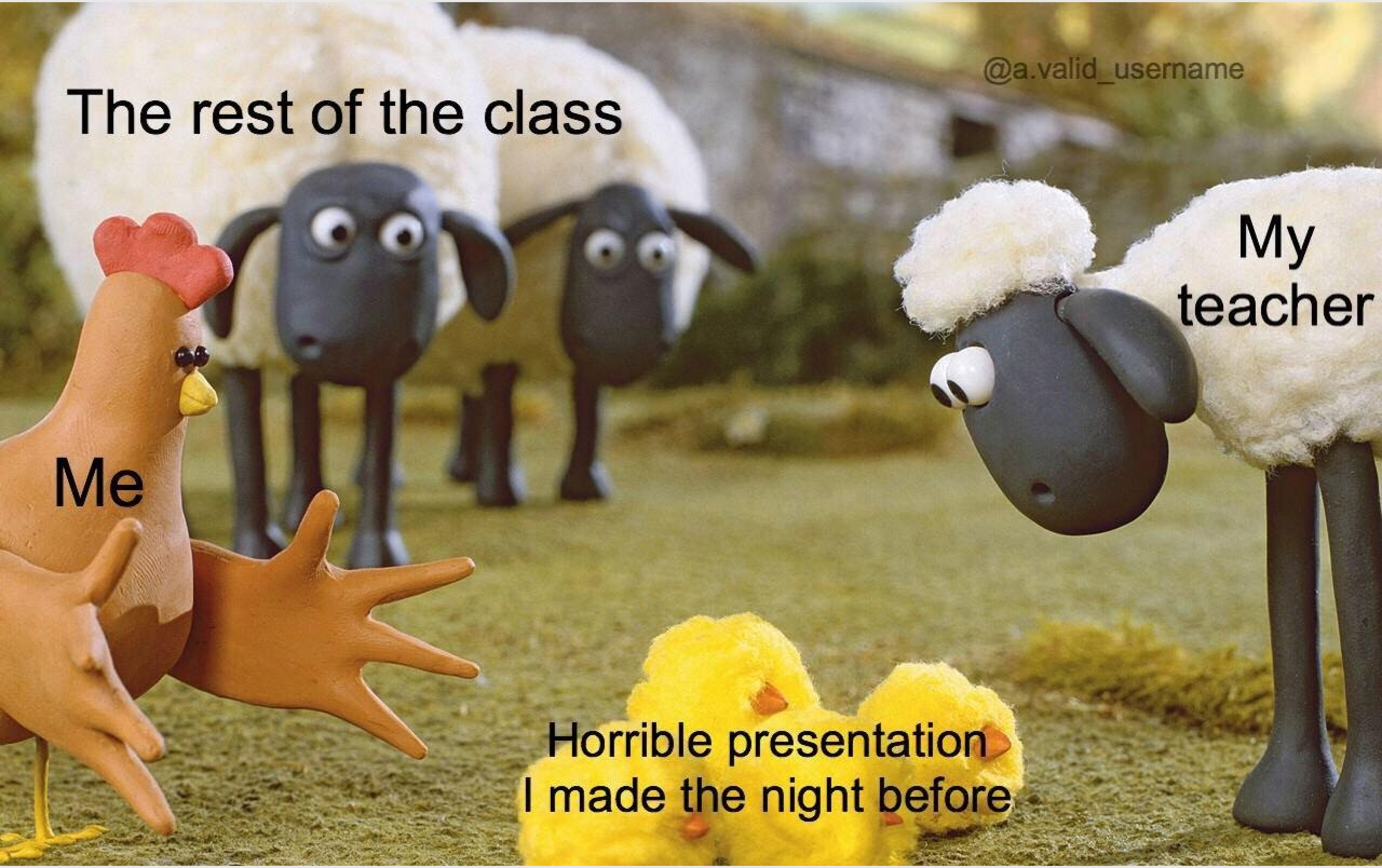
13-term Henderson Filter Weights
0.3
0.25
0.2
0.15
0.1
0.05
0
-0.05
0.24006
0.21434
0.21434
0.14736
0.14736
0.06549
0.06549
-0.01935
0
0
-0.01935
-0.02786
-0.02786
-6
-5
-4
-3
-2
-1
0
1
2
3
4
5
6

Sheep Population in the UK

Henderson–13 Moving Average

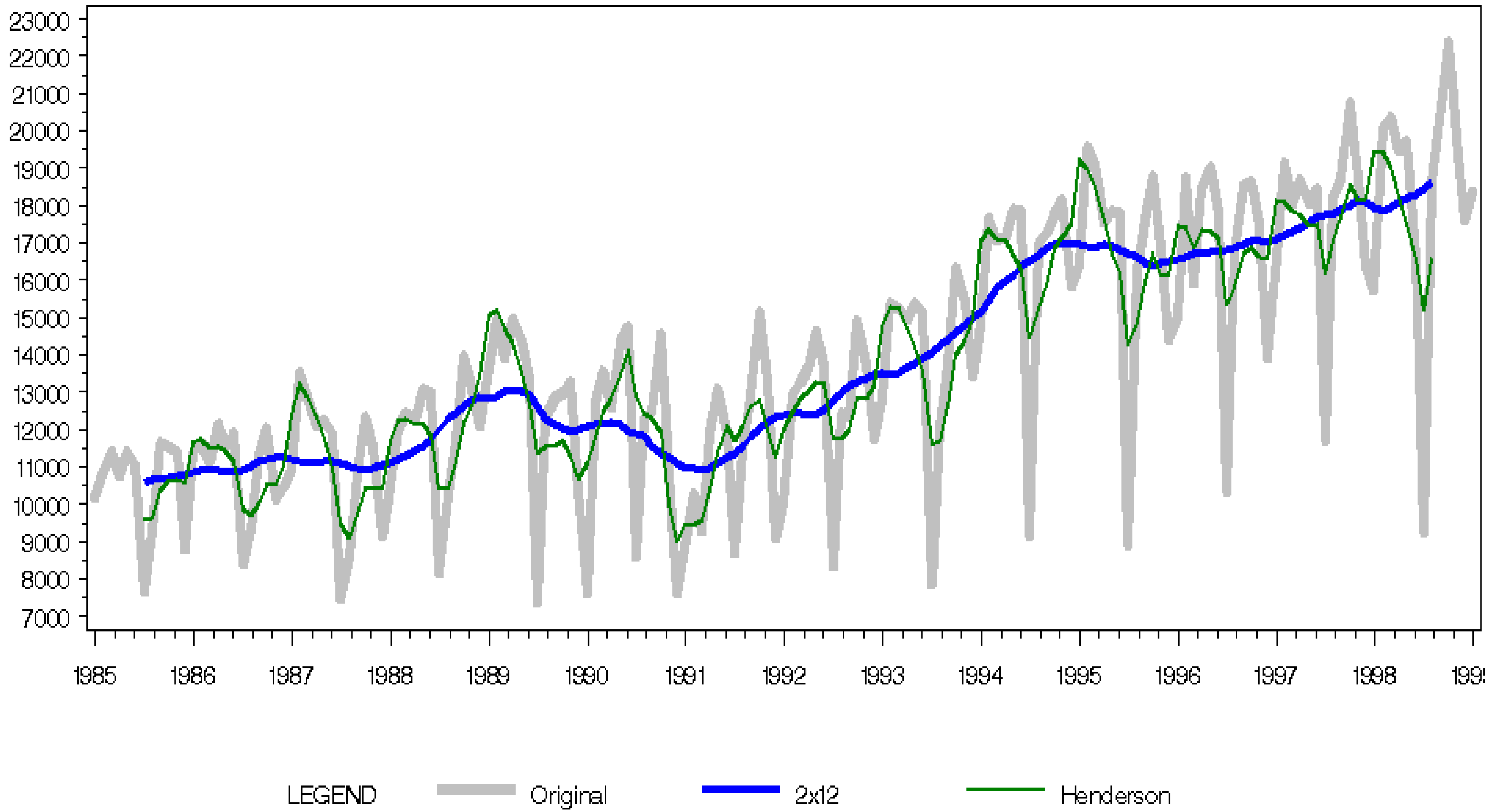



Sheep Population in the UK
Different Moving Averages
2500
2400
2300
2200
2100
2000
1900
1800
1700
1600
1500
1400
1300
1860
1870
1880
1890
1900
1910
1920
1930
1940
YEAR


@a.valid_username
The rest of the class
My teacher
Me
Horrible presentation I made the night before

Motor Vehicles

Different Trend Filters



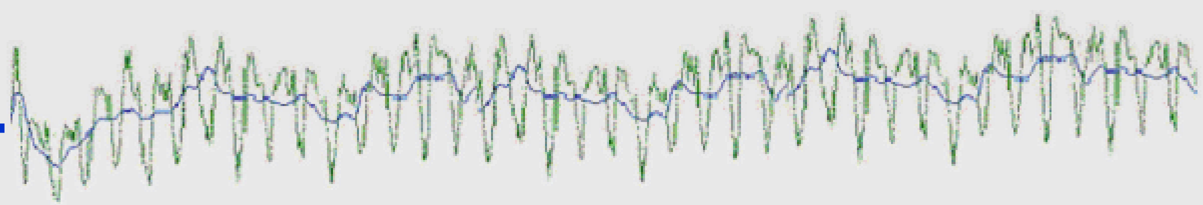
Filters Used by the X-11 Method

- Trend filters
  - 2 x 12 (or 2 x 4) for preliminary trend estimate.
  - Henderson filters for final trend estimate.

Example: 2x4 Trend Filter for a Quarterly Series

Example 2x4 trend filter for 2019 Q1

$$\frac{\begin{array}{l} 2018.3 + 2018.4 + 2019.1 + 2019.2 + \\ \quad\quad\quad\quad 2018.4 + 2019.1 + 2019.2 + 2019.3 \end{array}}{8}$$



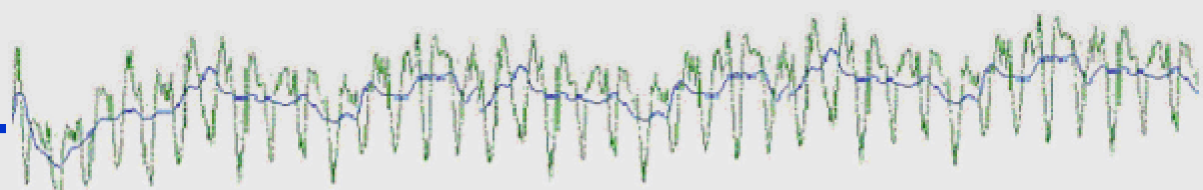
Seasonal Filters

- For seasonal filters, we average values within a month (or quarter)
- Seasonal filters (by default in X-13)
  - 3 x 3 for preliminary seasonal estimate
  - 3 x 3, 3 x 5, or 3 x 9 for final seasonal estimate, chosen by X-13 based on the Global Moving Seasonality Ratio (MSR)

Example: 3x3 Seasonal Filter

3 x 3 filter for January 2015 (or Q1 2015)

$$\frac{\begin{array}{l} 2013.1 + 2014.1 + 2015.1 + \\ \qquad\qquad 2014.1 + 2015.1 + 2016.1 + \\ \qquad\qquad\qquad\qquad 2015.1 + 2016.1 + 2017.1 \end{array}}{9}$$

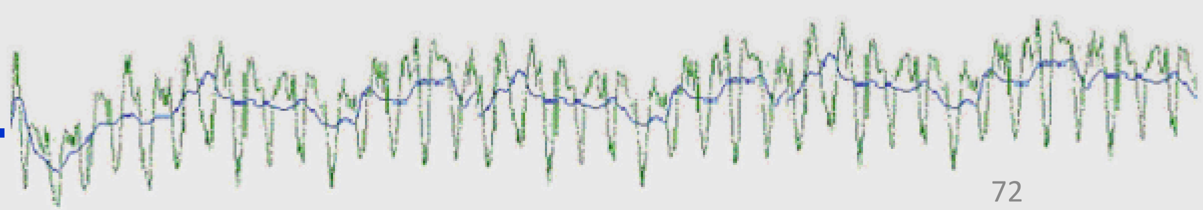


Choosing Seasonal Filters

- 3x5 is most common choice from X-13.
- Use 3x3 filters when seasonal pattern is changing rapidly.
- Use 3x9 filters when seasonal pattern isn't changing or when irregular component is large, because extreme values affect the averages less than with 3x5 or 3x3 filters.

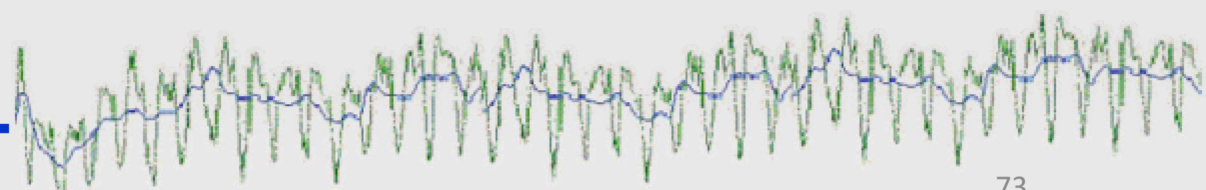
Basic X-11 Algorithm

- Step 1. Estimate the trend.
- Step 2. Detrend the series.
- Step 3. Estimate the seasonal.
- Repeat Steps 1-3
- Estimate the final trend and the final irregular
- **Repeat the entire procedure twice**



X-11 Iterations and Tables

- Part A: Prior Adjustments (regARIMA models) before the core X-11 Procedures
- Part B: Preliminary Estimation of Seasonal, Trend, and Extreme Values
- Part C: Another Estimation of Seasonal and Trend, plus Final Estimation of Extreme Values
- Part D: Final Estimation of Components



Common Table Codes

Raw or prior-adjusted series	B1
Weights for irregular	C17
Seasonal estimation	D10
Seasonally adjusted series	D11
Trend	D12
Combined factors	D16

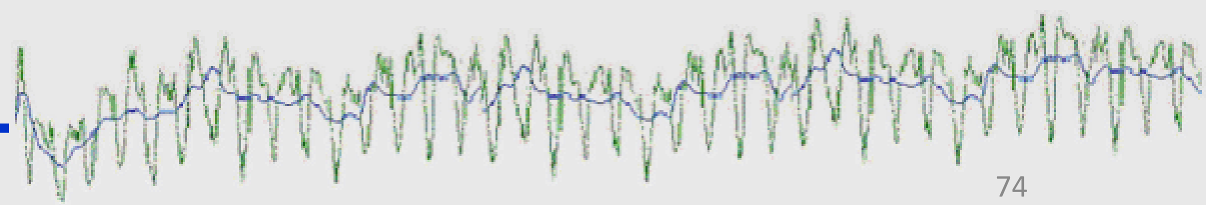


Table with Seats

- Antique Spanish Design and Hand-Stitched Embroidery

- Seasonal Warmth and Fancy Colors

- Comfy Seats

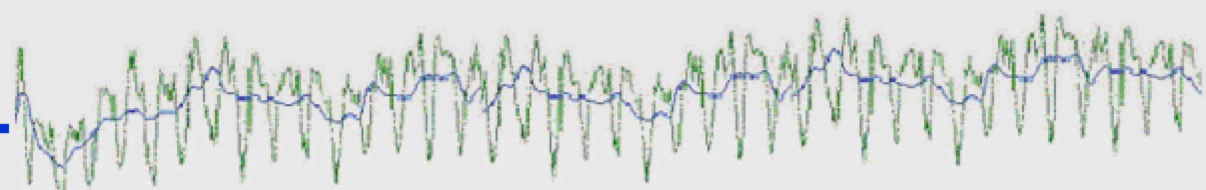
- Wood… probably

The SEATS Module

- The SEATS spec produces seasonal adjustment using a ARIMA-model-based (AMB) method based on SEATS, the program developed by Agustin Maravall at the Bank of Spain.
- Component estimates are formed by
  - Fitting an ARIMA model to the series,
  - This model, plus assumptions, determines models for the components, and then
  - Signal extraction techniques to produce component estimates and mean square errors (MSEs).

Advantages of AMB Adjustment (from Bill Bell, US Census Bureau)

- Flexible approach given wide range of models and parameter values.
- Model selection can be guided by accepted statistical principals.
- Filters are tailored to individual series through parameter estimation, and are “optimal” given
  - True model is used (the bigger worry), and
  - Decomposition assumptions are correct.



More Advantages of AMB (from Bill Bell, US Census Bureau)

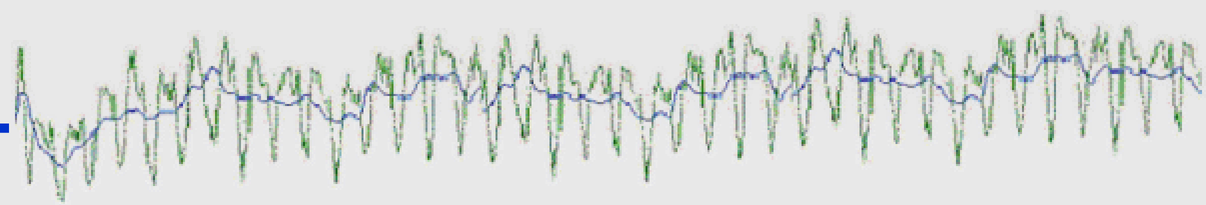
- Signal extraction calculations provide error variances of component estimates (MSEs are based on the model).
- Approach easily extends (in principle) to accommodate a sampling error component. (Work on this by Richard Tiller at the BLS.)

Advantages of SEATS over X11

- The SEATS procedure produces variances (and therefore also confidence intervals) for the various components of the seasonal adjustment.
- It is possible to decompose the trend-cycle into a long-term trend and a cycle component.
- Studies have shown that SEATS works well to provide stable and accurate adjustments of series with a large irregular component.

Advantages of X11 over SEATS

- X11 will work well for shorter series (less than five or six years).
- SEATS can possibly add seasonality to the seasonal adjustment of a nonseasonal series, so it is important to look at diagnostics, and especially diagnostics for residual seasonality.



Bottom Line

- Many series have seasonal adjustments from the X11 module and the SEATS module that are practically identical.
- Diagnostics are important.

Why Seasonal Adjustment?

- Seasonal oscillations can make it difficult to compare time series.
- Large seasonal oscillations can also obscure smaller movements that may be important.
- A seasonally adjusted series makes is easier to see turning points.

Advantages of X13-ARIMA-SEATS

- X-13 combines two of the most useful seasonal adjustment programs into one program with one set of diagnostics.
- X-13 estimates the trend and seasonal component without one getting in the way of the other, and also estimates trading day effects, holiday effects, and outliers.
- X-13 has diagnostics for the regARIMA model and the seasonal adjustment.
- X-13 is able to forecast series.