Time Series Analysis

Philip Leifeld

GV903: Advanced Research Methods, Week 8

University of Essex

1. Introduction to Time Series Analysis

What is a Time Series and what can we do with it?

What is a Time Series and what can we do with it?

- Like in the linear regression case, we have a DV and possibly a set of IVs.

What is a Time Series and what can we do with it?

- Like in the linear regression case, we have a DV and possibly a set of IVs.
- But now all the observations have a time index.

What is a Time Series and what can we do with it?

- Like in the linear regression case, we have a DV and possibly a set of IVs.
- But now all the observations have a time index.
- Recorded annually, monthly, quarterly, weekly, daily, hourly or similar.

What is a Time Series and what can we do with it?

- Like in the linear regression case, we have a DV and possibly a set of IVs.
- But now all the observations have a time index.
- Recorded annually, monthly, quarterly, weekly, daily, hourly or similar.
- This has implications for the analysis: observations can depend on earlier observations; they are not a sample (i. e., the second Gauss-Markov assumption is possibly violated).

What is a Time Series and what can we do with it?

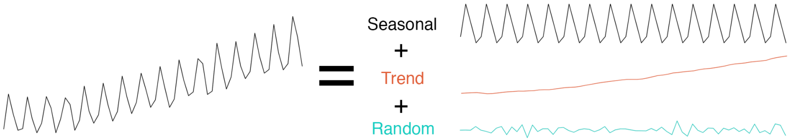
- Like in the linear regression case, we have a DV and possibly a set of IVs.
- But now all the observations have a time index.
- Recorded annually, monthly, quarterly, weekly, daily, hourly or similar.
- This has implications for the analysis: observations can depend on earlier observations; they are not a sample (i. e., the second Gauss-Markov assumption is possibly violated).
- We can visualise and decompose time series, and/or we can model them using regression models. We can also try to predict future values.

Examples of Time Series in Political Science

- Public opinion (on policies, candidates etc).
- Government approval ratings.
- Annual budgets for policies or offices.
- News media attention (via content analysis).
- Changes in POLITY scores etc of a country.
- ...

2. Time Series Decomposition

Time Series Decomposition



What is Time Series Decomposition?

- A time series is split up into three constituent parts of the data-generating process:

What is Time Series Decomposition?

- A time series is split up into three constituent parts of the data-generating process:
  1. A time trend component (can be linear or complex).

What is Time Series Decomposition?

- A time series is split up into three constituent parts of the data-generating process:
  1. A time trend component (can be linear or complex).
  2. A seasonality component (can be additive or multiplicative).

What is Time Series Decomposition?

- A time series is split up into three constituent parts of the data-generating process:
  1. A time trend component (can be linear or complex).
  2. A seasonality component (can be additive or multiplicative).
  3. The remaining variation ("noise" or "residual variation").

What is Time Series Decomposition?

- A time series is split up into three constituent parts of the data-generating process:
  1. A time trend component (can be linear or complex).
  2. A seasonality component (can be additive or multiplicative).
  3. The remaining variation ("noise" or "residual variation").
- We can de-trend a time series and average out the seasonality and then start modelling the remaining variation using other tools, for example time series regression.

What is Time Series Decomposition?

- A time series is split up into three constituent parts of the data-generating process:
  1. A time trend component (can be linear or complex).
  2. A seasonality component (can be additive or multiplicative).
  3. The remaining variation ("noise" or "residual variation").
- We can de-trend a time series and average out the seasonality and then start modelling the remaining variation using other tools, for example time series regression.
- Or we can model time trends and seasonality all within a regression framework (shown later).

What is Time Series Decomposition?

- A time series is split up into three constituent parts of the data-generating process:
  1. A time trend component (can be linear or complex).
  2. A seasonality component (can be additive or multiplicative).
  3. The remaining variation ("noise" or "residual variation").
- We can de-trend a time series and average out the seasonality and then start modelling the remaining variation using other tools, for example time series regression.
- Or we can model time trends and seasonality all within a regression framework (shown later).
- We can also use decomposition for forecasting and anomaly detection.

Working with ts Objects in R

```
library("forecast") # useful for decomposition
library("fpp")      # some time series data
library("Ecdat")    # more time series data
data(AirPassengers)
data(ausbeer)
class(beer)
## [1] "ts"

head(beer)
##      Jan Feb Mar Apr May Jun
## 1991 164 148 152 144 155 125

tail(beer)
##      Mar Apr May Jun Jul Aug
## 1995 152 127 151 130 119 153
```

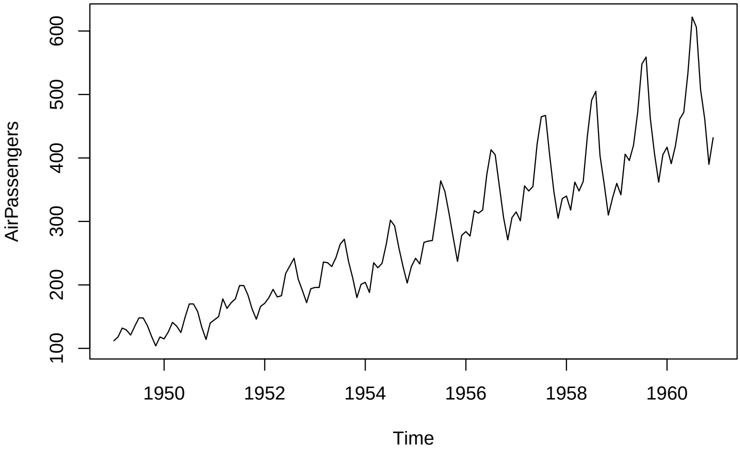
Creating New ts Objects

```
ts(1:10, frequency = 4, start = c(1959, 2))
##      Qtr1 Qtr2 Qtr3 Qtr4
## 1959         1    2    3
## 1960    4    5    6    7
## 1961    8    9   10

rp <- rpois(n = 20, lambda = 10)
ts(rp, frequency = 12, start = 2010)
##      Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec
## 2010  11   7  15   7  10   9   7   9   7   6   6  12
## 2011  13  14   7   7   8  11   9  13
```

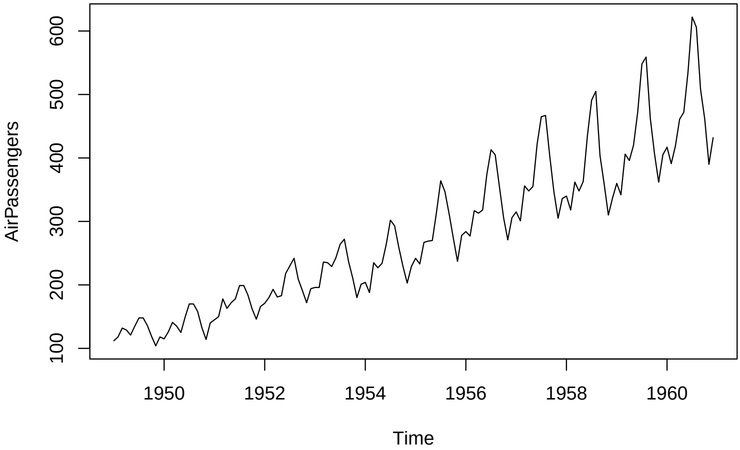
Monthly Airline Passenger Numbers 1949–1960

```r
plot(AirPassengers)
```



Monthly Airline Passenger Numbers 1949–1960

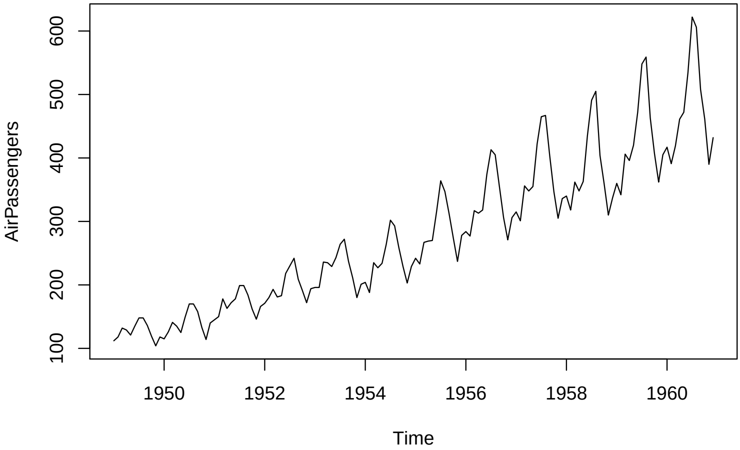
```r
plot(AirPassengers)
```



Note the linear trend in the data!

Monthly Airline Passenger Numbers 1949–1960

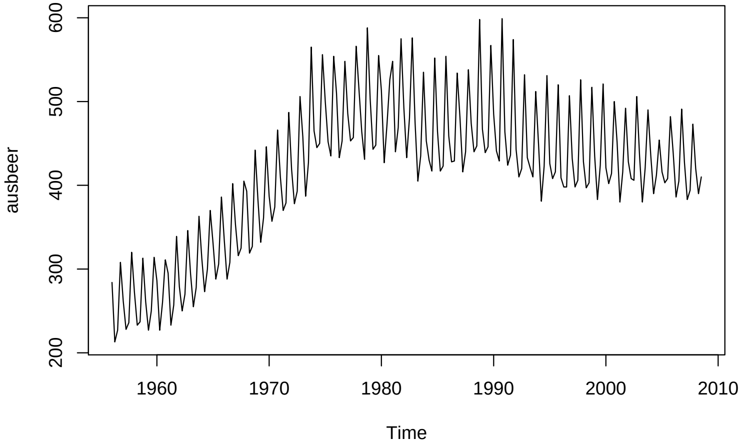
```
plot(AirPassengers)
```



Note the linear trend in the data! Also note the seasonality!

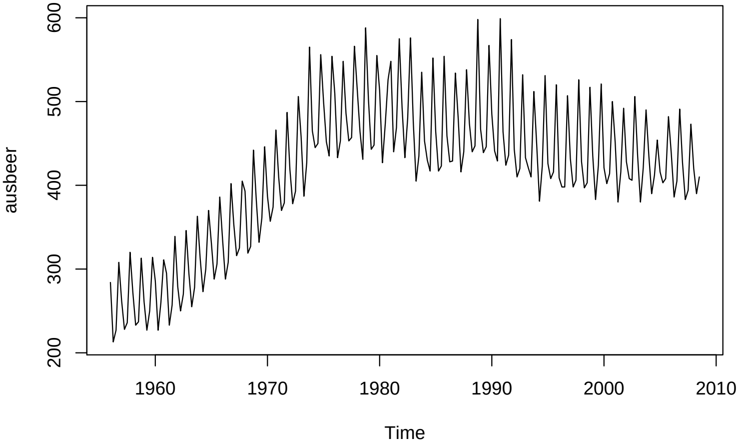
Quarterly Australian Beer Production 1956–2008

```r
plot(ausbeer)
```



Quarterly Australian Beer Production 1956–2008

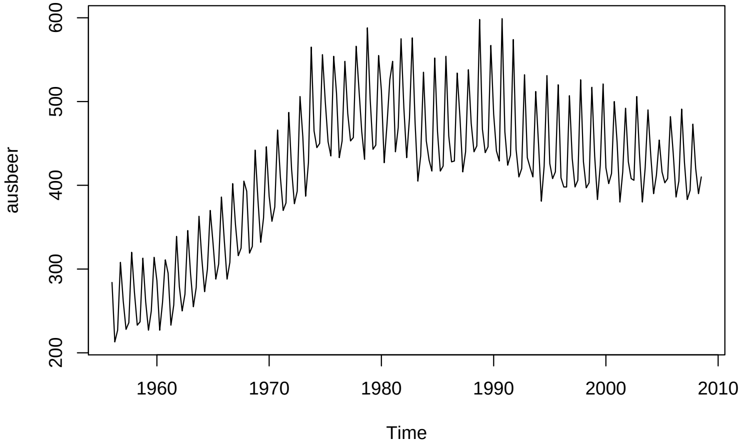
```r
plot(ausbeer)
```



This is an example of a more complex time trend.

Quarterly Australian Beer Production 1956–2008

```r
plot(ausbeer)
```

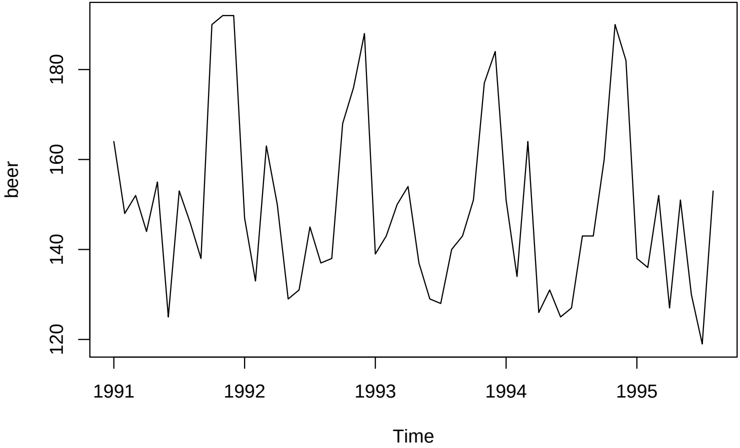


This is an example of a more complex time trend.

There is also seasonality.

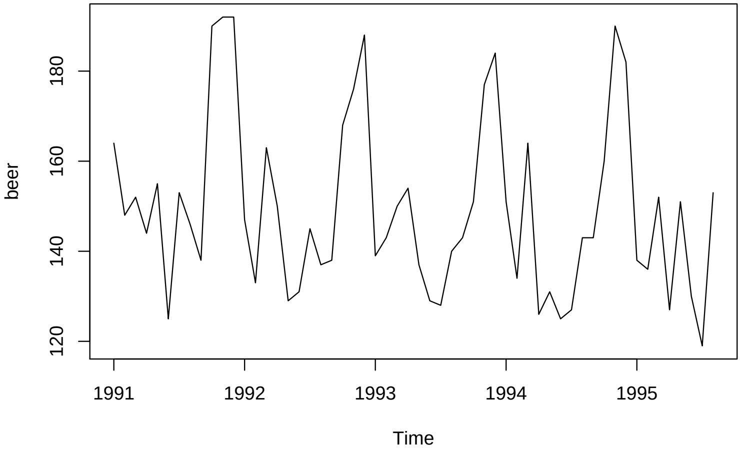
Monthly Australian Beer Production 1991–1995

```
plot(beer)
```



Monthly Australian Beer Production 1991–1995

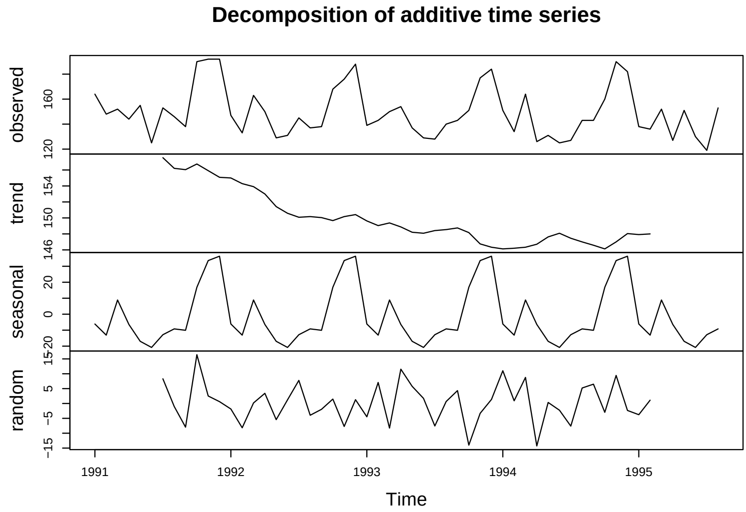
```r
plot(beer)
```



Slightly negative trend and seasonality.

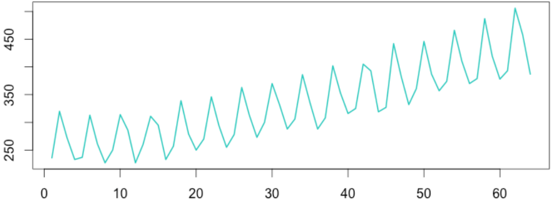
Decomposition

```r
plot(decompose(beer, "additive"))
```

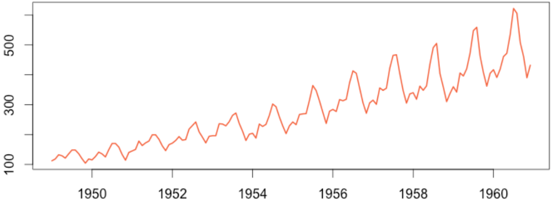


Additive versus Multiplicative Models

Additive Time Series: Seasonality stays constant.

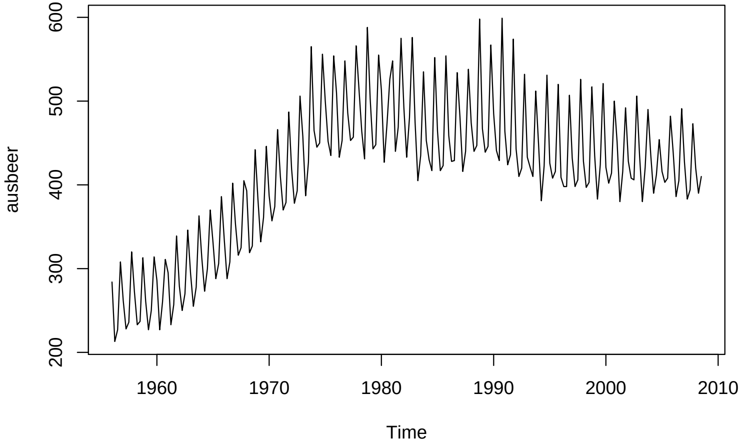


Multiplicative Time Series: Seasonality increases.



Quarterly Australian Beer Production 1956–2008

```r
plot(ausbeer)
```

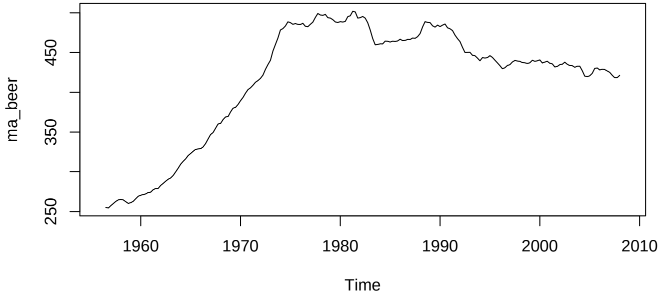


How do we compute the trend and then de-trend the time series?

De-Trending: Computing a Moving Average

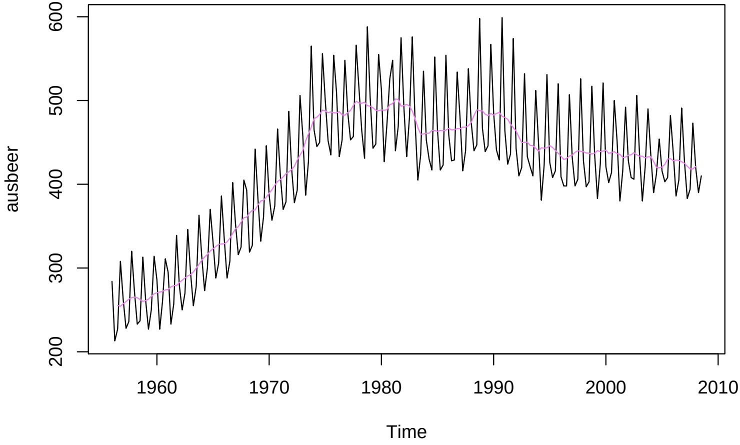
```
ma_beer <- ma(ausbeer, order = 4)
head(ma_beer) # 4 means 2 in each direction
##         Qtr1    Qtr2    Qtr3    Qtr4
## 1956      NA      NA 255.250 254.375
## 1957 257.375 260.000

plot(ma_beer)
```



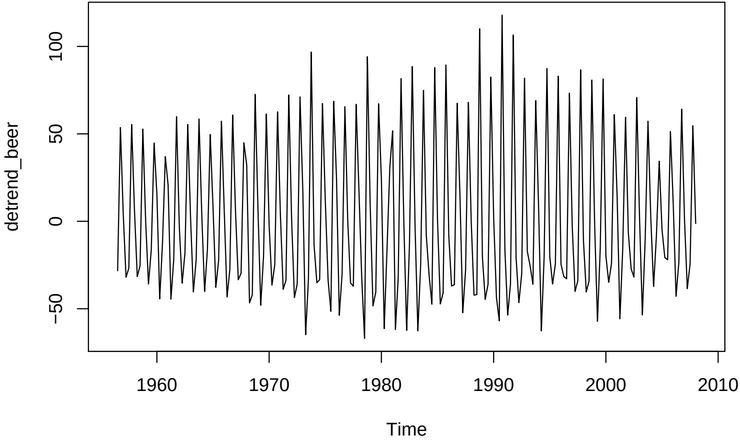
We can also Plot this Together

```r
plot(ausbeer)
lines(ma_beer, col = "violet")
```



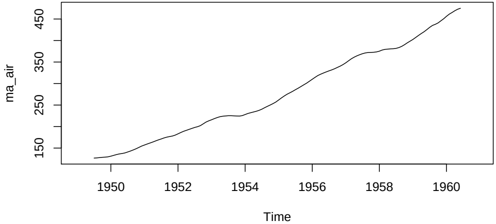
De-Trending: Additive Case

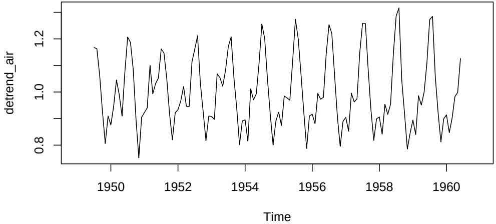
```r
detrend_beer <- ausbeer - ma_beer
plot(detrend_beer)
```



De-Trending: Multiplicative Example

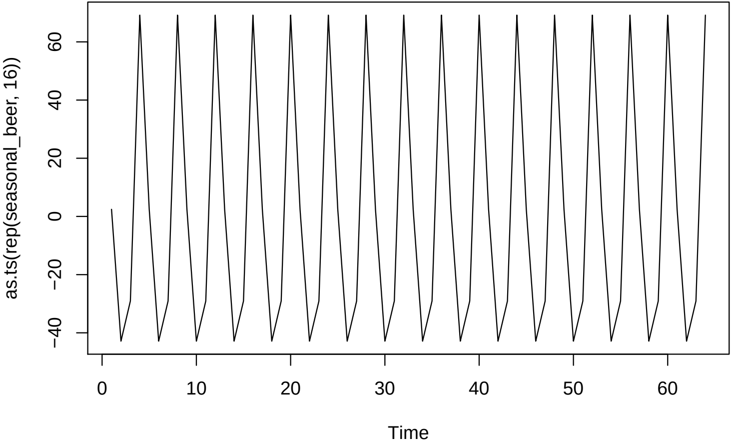
```
ma_air <- ma(AirPassengers, order = 12) # monthly obs
plot(ma_air)
detrend_air <- AirPassengers / ma_air
plot(detrend_air)
```





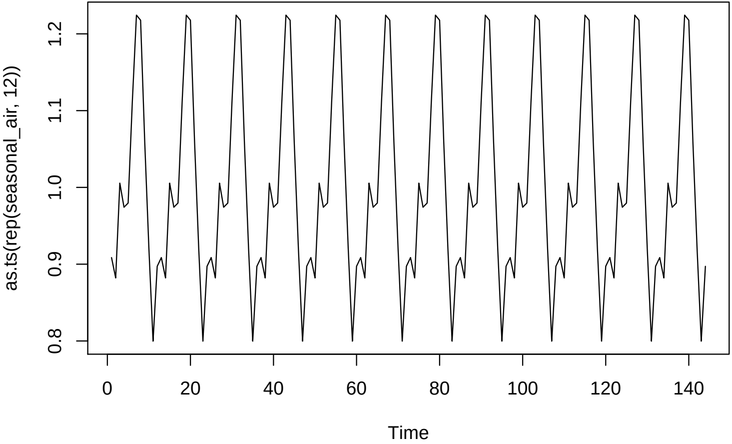
Averaging the Seasonality: Beer Example

```r
mat <- matrix(detrend_beer, ncol = 4, byrow = TRUE)
seasonal_beer = colMeans(mat, na.rm = TRUE)
plot(as.ts(rep(seasonal_beer, 16)))
```



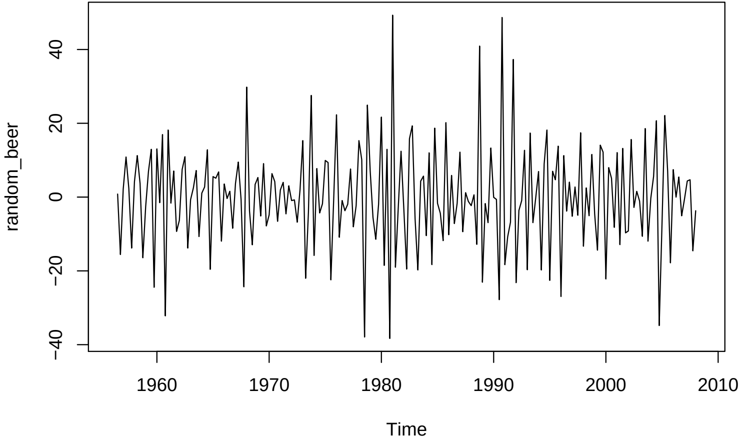
Averaging the Seasonality: Air Passengers

```
m_air = matrix(detrend_air, ncol = 12, byrow = TRUE)
seasonal_air = colMeans(m_air, na.rm = TRUE)
plot(as.ts(rep(seasonal_air, 12)))
```



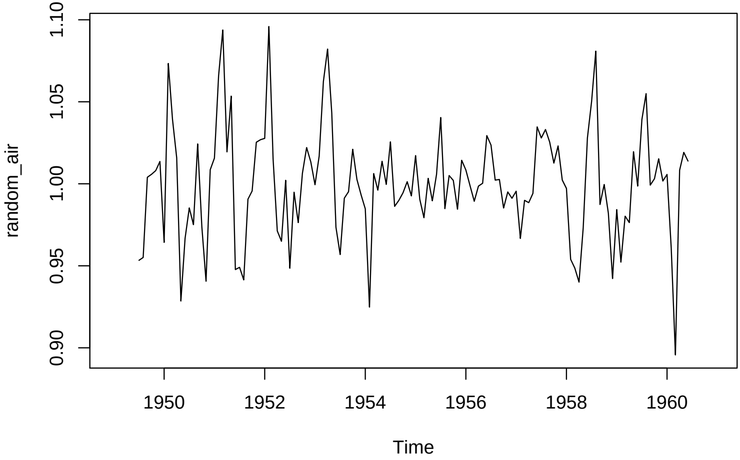
Looking at the Residual Variance (Noise)

```r
random_beer = ausbeer - ma_beer - seasonal_beer
plot(random_beer)
```



Looking at the Residual Variance (Noise)

```
random_air = AirPassengers / (ma_air * seasonal_air)
plot(random_air)
```



Re-Composing Time Series

In principle, we can now put everything back together.

Re-Composing Time Series

In principle, we can now put everything back together.

Additive case:

```
recomposed_beer = ma_beer + seasonal_beer + random_beer
```

Re-Composing Time Series

In principle, we can now put everything back together.

Additive case:

```
recomposed_beer = ma_beer + seasonal_beer + random_beer
```

Multiplicative case:

```
recomposed_air = ma_air * seasonal_air * random_air
```

Re-Composing Time Series

In principle, we can now put everything back together.

Additive case:

```
recomposed_beer = ma_beer + seasonal_beer + random_beer
```

Multiplicative case:

```
recomposed_air = ma_air * seasonal_air * random_air
```

. . . or we can find and remove anomalies (= outliers) in the three components and remove them before putting the time series back together.

Anomaly Detection

- Anomaly detection can be useful to explore the effect of external events on disrupting a time series.

Anomaly Detection

- Anomaly detection can be useful to explore the effect of external events on disrupting a time series.
- Or you can use it to recompose a “cleaner” time series, for example when you are interested in drawing a polarization curve that is affected by outliers.

Anomaly Detection

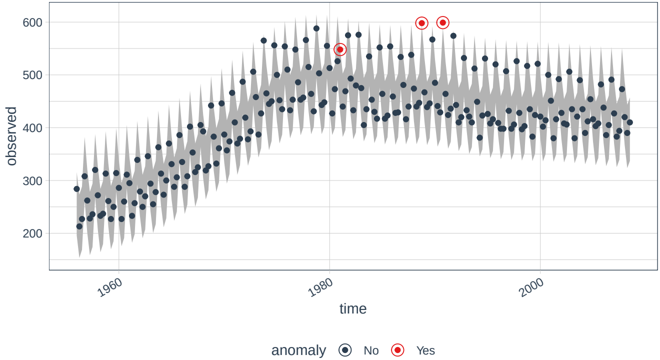
- Anomaly detection can be useful to explore the effect of external events on disrupting a time series.
- Or you can use it to recompose a “cleaner” time series, for example when you are interested in drawing a polarization curve that is affected by outliers.
- There are lots of different methods that can be applied to decomposed time series.

Anomaly Detection

- Anomaly detection can be useful to explore the effect of external events on disrupting a time series.
- Or you can use it to recompose a “cleaner” time series, for example when you are interested in drawing a polarization curve that is affected by outliers.
- There are lots of different methods that can be applied to decomposed time series.
- Just to give you a visual idea, let’s look at an example...

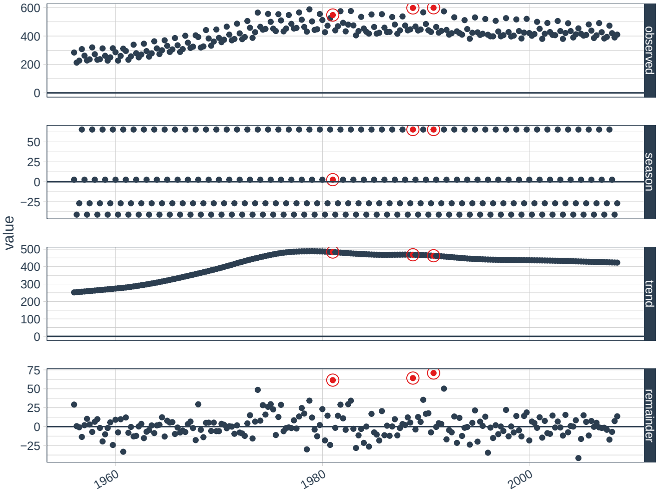
The anomalize Package

```
library("anomalize")    # tidy anomaly detection
library("tibbletime")   # tidy data structures for time
library("tsbox")        # convert time series objects
df_beer <- as_tbl_time(ts_df(ausbeer), index = time)
df_beer %>% time_decompose(value) %>%
  anomalize(remainder) %>% time_recompose() %>%
  plot_anomalies(time_recomposed = TRUE)
```



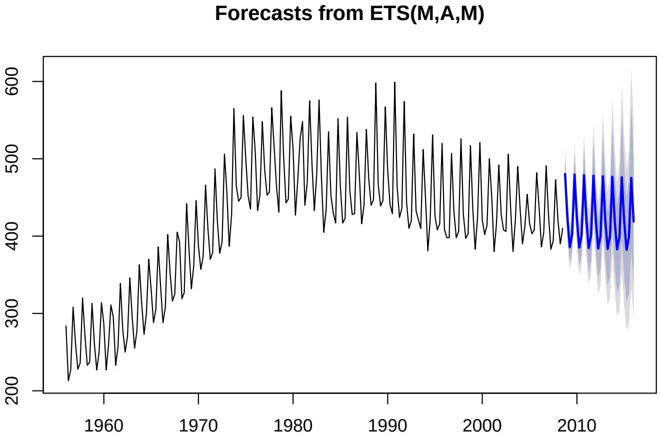
The anomalize Package

```
df_beer %>%
  time_decompose(value) %>%
  anomalize(remainder) %>%
  plot_anomaly_decomposition()
```



Forecasting with the `forecast` Package

```
fit <- ets(ausbeer)
fcast <- forecast(fit, h = 30) # predict 30 periods
plot(fcast)
```



(Multiplicative error; Additive trend; Multiplicative seasonality.)

3. Time Series Regression

Energy Policy Monthly TV News Coverage

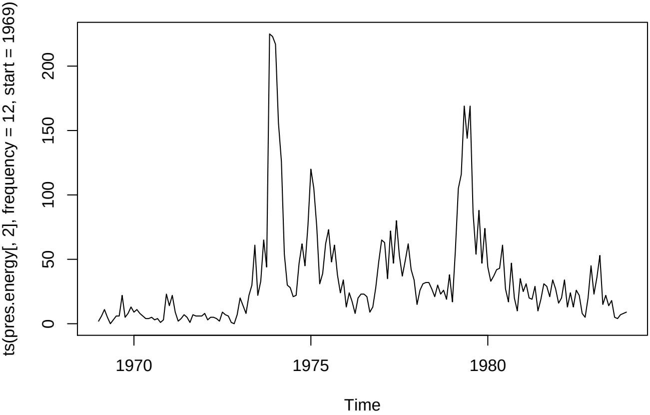
Data and Example from Monogan (2015): `https://doi.org/10.7910/DVN/ARKOTI`

```
pres.energy <- read.csv("PESenergy.csv")
head(pres.energy[, 1:4])
##     Date Energy Unemploy Approval
## 1 Jan-69      2      3.4       59
## 2 Feb-69      6      3.4       60
## 3 Mar-69     11      3.4       64
## 4 Apr-69      5      3.4       60
## 5 May-69      0      3.4       63
## 6 Jun-69      3      3.5       63
```

Energy Policy Monthly TV News Coverage

Data and Example from Monogan (2015): `https://doi.org/10.7910/DVN/ARKOTI`

```
plot(ts(pres.energy[, 2], frequency = 12, start = 1969))
```



Static Time Series Regression Models

```
static.ols <- lm(Energy ~ rmn1173 + grf0175 + grf575 + jec477 + jec1177 +
  jec479 + embargo+ hostages + oilc + Approval + Unemploy, data = pres.energy)
coeftest(static.ols)
##
## t test of coefficients:
##
##               Estimate Std. Error t value  Pr(>|t|)
## (Intercept) 319.74423   46.83584  6.8269 1.511e-10 ***
## rmn1173      78.82610   28.80117  2.7369  0.006870 **
## grf0175      60.79046   26.70059  2.2767  0.024063 *
## grf575       -4.26762   26.53152 -0.1609  0.872404
## jec477       47.03877   26.67596  1.7633  0.079661 .
## jec1177      15.44268   26.37859  0.5854  0.559048
## jec479       72.05191   26.50274  2.7187  0.007243 **
## embargo      96.37602   13.31055  7.2406 1.530e-11 ***
## hostages     -4.52890    7.39449 -0.6125  0.541055
## oilc         -5.87649    1.08478 -5.4172 2.069e-07 ***
## Approval     -1.06930    0.21472 -4.9800 1.567e-06 ***
## Unemploy     -3.70181    1.38610 -2.6707  0.008314 **
## ---
## Signif. codes:
## 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Stationary and Integrated Time Series

- Stationary time series: Any new data point depends only on exogenous IVs, not on the previous data points of the time series. This is what we (erroneously) assumed on the previous slide.

Stationary and Integrated Time Series

- Stationary time series: Any new data point depends only on exogenous IVs, not on the previous data points of the time series. This is what we (erroneously) assumed on the previous slide.
- Non-stationary time series: New data points also depend on previous data points.

Stationary and Integrated Time Series

- Stationary time series: Any new data point depends only on exogenous IVs, not on the previous data points of the time series. This is what we (erroneously) assumed on the previous slide.
- Non-stationary time series: New data points also depend on previous data points.
- Non-stationary time series can be integrated (a.k.a. unit root time series), which is based on a random walk.

Stationary and Integrated Time Series

- Stationary time series: Any new data point depends only on exogenous IVs, not on the previous data points of the time series. This is what we (erroneously) assumed on the previous slide.
- Non-stationary time series: New data points also depend on previous data points.
- Non-stationary time series can be integrated (a.k.a. unit root time series), which is based on a random walk.
- Examples: stock prices; public opinion.

Stationary and Integrated Time Series

- Stationary time series: Any new data point depends only on exogenous IVs, not on the previous data points of the time series. This is what we (erroneously) assumed on the previous slide.
- Non-stationary time series: New data points also depend on previous data points.
- Non-stationary time series can be integrated (a.k.a. unit root time series), which is based on a random walk.
- Examples: stock prices; public opinion.
- From the perspective of current observation y_t, we need to consider the so-called lagged variables y_{t-1}, y_{t-2} etc. as IVs to make sure we have modelled all dependencies.

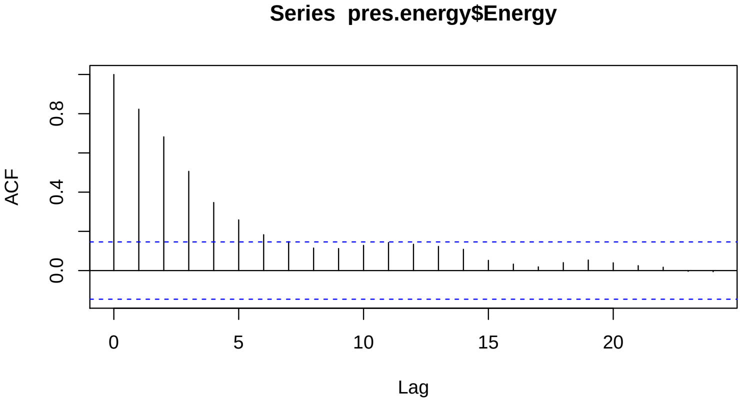
Stationary and Integrated Time Series

- Stationary time series: Any new data point depends only on exogenous IVs, not on the previous data points of the time series. This is what we (erroneously) assumed on the previous slide.
- Non-stationary time series: New data points also depend on previous data points.
- Non-stationary time series can be integrated (a.k.a. unit root time series), which is based on a random walk.
- Examples: stock prices; public opinion.
- From the perspective of current observation y_t, we need to consider the so-called lagged variables y_{t-1}, y_{t-2} etc. as IVs to make sure we have modelled all dependencies.
- If we fail to do this for a non-stationary time series, we get omitted variable bias.

Autocorrelation Function

This shows the correlation between y_t and y_{t-k}.

```r
acf(pres.energy$Energy, lag.max = 24)
```

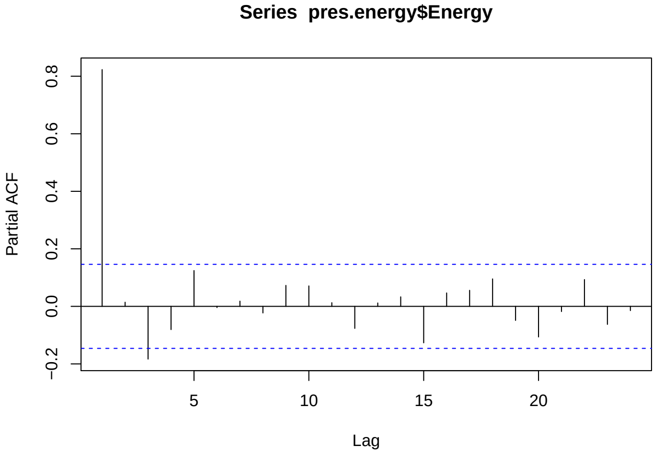


Important because we need to understand the data-generating process to be modelled. (Sometimes, we need to remove seasonality before looking at the ACF.)

Partial Autocorrelation Function

This shows the partial correlation between y_t and y_{t-k}.

```
pacf(pres.energy$Energy, lag.max = 24)
```



Correlation between y_t and y_{t-k} that is unaccounted for by any previous (i. e., closer) lags. A one-period lag is sufficient here.

What should you do then?

What should you do then?

- Use a hypothesis test for the presence of autocorrelation:
  - Durbin–Watson test (`dwtest`).
  - Breusch–Godfrey test (`bgtest`).

What should you do then?

- Use a hypothesis test for the presence of autocorrelation:
  - Durbin–Watson test (`dwtest`).
  - Breusch–Godfrey test (`bgtest`).
- Use linear model, but estimate via FGLS. There are two popular estimation techniques:
  - The Prais–Winsten estimator.
  - The Cochrane–Orcutt estimator.

What should you do then?

- Use a hypothesis test for the presence of autocorrelation:
  - Durbin–Watson test (`dwtest`).
  - Breusch–Godfrey test (`bgtest`).
- Use linear model, but estimate via FGLS. There are two popular estimation techniques:
  - The Prais–Winsten estimator.
  - The Cochrane–Orcutt estimator.
- Include lagged DVs.

What should you do then?

- Use a hypothesis test for the presence of autocorrelation:
  - Durbin–Watson test (`dwtest`).
  - Breusch–Godfrey test (`bgtest`).
- Use linear model, but estimate via FGLS. There are two popular estimation techniques:
  - The Prais–Winsten estimator.
  - The Cochrane–Orcutt estimator.
- Include lagged DVs.
- Estimate first-differences model.

What should you do then?

- Use a hypothesis test for the presence of autocorrelation:
  - Durbin–Watson test (`dwtest`).
  - Breusch–Godfrey test (`bgtest`).
- Use linear model, but estimate via FGLS. There are two popular estimation techniques:
  - The Prais–Winsten estimator.
  - The Cochrane–Orcutt estimator.
- Include lagged DVs.
- Estimate first-differences model.
- Estimate an ARIMA model.

Including Lagged DVs: Autoregressive Models

We call a linear model of a time series a K-order autoregressive model—or AR(K) process—if we include K lagged dependent variables:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \left(\sum_{k=1}^{K} \delta_k y_{t-k} \right) + u$$

Including Lagged DVs: Autoregressive Models

We call a linear model of a time series a K-order autoregressive model—or AR(K) process—if we include K lagged dependent variables:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \left(\sum_{k=1}^{K} \delta_k y_{t-k} \right) + u$$

In most cases (also here), an AR(1) process is sufficient:

$$y_t = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + \ldots + \delta y_{t-1} + u_t$$

Including Lagged DVs: Autoregressive Models

We call a linear model of a time series a K-order autoregressive model—or AR(K) process—if we include K lagged dependent variables:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \left(\sum_{k=1}^{K} \delta_k y_{t-k} \right) + u$$

In most cases (also here), an AR(1) process is sufficient:

$$y_t = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + \ldots + \delta y_{t-1} + u_t$$

Some people in political science claim that lagged DVs are inappropriate because they are correlated with other IVs. However, they *should* be included. Once they are controlled for, the other IVs can be interpreted.

The PACF showed that a lag of 1 is sufficient. . .

```
library("dynlm")
energy.ts <- ts(pres.energy$Energy, frequency = 12, start = 1969)
dynamic.ols <- dynlm(energy.ts ~ pres.energy$rmn1173 +
  pres.energy$grf0175 + pres.energy$grf575 + pres.energy$jec477 +
  pres.energy$jec1177 + pres.energy$jec479 + pres.energy$embargo +
  pres.energy$hostages + pres.energy$oilc + pres.energy$Approval +
  pres.energy$Unemploy + lag(energy.ts, -1))
```

The `dynlm` package can use `ts` objects with the right time structure as a DV.

It also provides convenient model terms for including time trends, seasonality etc. We can also include lags for various IVs if that's what our theory tells us.

Lagged DV in the Energy Example

	Static	Dynamic
(Intercept)	319.74 (46.84)***	62.11 (36.97)
rmn1173	78.83 (28.80)**	171.63 (20.15)***
grf0175	60.79 (26.70)*	51.70 (17.73)**
grf575	−4.27 (26.53)	7.06 (17.62)
jec477	47.04 (26.68)	39.02 (17.71)*
jec1177	15.44 (26.38)	−10.78 (17.59)
jec479	72.05 (26.50)**	28.68 (17.83)
embargo	96.38 (13.31)***	10.54 (10.61)
hostages	−4.53 (7.39)	−2.51 (4.91)
oilc	−5.88 (1.08)***	−1.14 (0.81)
Approval	−1.07 (0.21)***	−0.15 (0.16)
Unemploy	−3.70 (1.39)**	−0.89 (0.97)
lag(energy.ts, -1)		0.74 (0.05)***
R^2	0.59	0.82
Adj. R^2	0.57	0.81
Num. obs.	180	179

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Differencing the Time Series

Instead of using lagged DVs, one could just model the change between any two time points as the DV, rather than the value at any time point.

Differencing the Time Series

Instead of using lagged DVs, one could just model the change between any two time points as the DV, rather than the value at any time point.

This removes any trending and seasonality and thus autocorrelation.

$$y_t - y_{t-1} = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + \ldots + u_t$$

Differencing the Time Series

Instead of using lagged DVs, one could just model the change between any two time points as the DV, rather than the value at any time point.

This removes any trending and seasonality and thus autocorrelation.

$$y_t - y_{t-1} = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + \ldots + u_t$$

If $K = 2$, use second-order differences:

$$\Delta(\Delta y) = (y_t - y_{t-1}) - (y_{t-1} - y_{t-2})$$

Use PACF to determine K.

Differencing the Time Series

Instead of using lagged DVs, one could just model the change between any two time points as the DV, rather than the value at any time point.

This removes any trending and seasonality and thus autocorrelation.

$$y_t - y_{t-1} = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + \ldots + u_t$$

If $K = 2$, use second-order differences:

$$\Delta(\Delta y) = (y_t - y_{t-1}) - (y_{t-1} - y_{t-2})$$

Use PACF to determine K.

Note that differencing will also require having changes on the IV side rather than static IVs.

How can we Choose the Best Specification?

- There are many things we can do to model a time series:

How can we Choose the Best Specification?

- There are many things we can do to model a time series:
  - Lagged DV ($K = 1$ or higher) to account for autocorrelation.

How can we Choose the Best Specification?

- There are many things we can do to model a time series:
  - Lagged DV ($K = 1$ or higher) to account for autocorrelation.
  - Lagged IVs ($K = 1$ or higher) to account for dynamic causality.

How can we Choose the Best Specification?

- There are many things we can do to model a time series:
  - Lagged DV ($K = 1$ or higher) to account for autocorrelation.
  - Lagged IVs ($K = 1$ or higher) to account for dynamic causality.
  - Index variable (e.g., `1:n`) to control for linear time trends.

How can we Choose the Best Specification?

- There are many things we can do to model a time series:
  - Lagged DV ($K = 1$ or higher) to account for autocorrelation.
  - Lagged IVs ($K = 1$ or higher) to account for dynamic causality.
  - Index variable (e.g., `1:n`) to control for linear time trends.
  - Adding dummy variables for months or quarters to control for seasonality.

How can we Choose the Best Specification?

- There are many things we can do to model a time series:
  - Lagged DV ($K = 1$ or higher) to account for autocorrelation.
  - Lagged IVs ($K = 1$ or higher) to account for dynamic causality.
  - Index variable (e.g., `1:n`) to control for linear time trends.
  - Adding dummy variables for months or quarters to control for seasonality.
- How do we choose among them?

How can we Choose the Best Specification?

- There are many things we can do to model a time series:
  - Lagged DV ($K = 1$ or higher) to account for autocorrelation.
  - Lagged IVs ($K = 1$ or higher) to account for dynamic causality.
  - Index variable (e.g., `1:n`) to control for linear time trends.
  - Adding dummy variables for months or quarters to control for seasonality.
- How do we choose among them?
- Easy: start with all of them and then use F tests or LM tests to decide what restrictions are sensible.

How can we Choose the Best Specification?

- There are many things we can do to model a time series:
  - Lagged DV ($K = 1$ or higher) to account for autocorrelation.
  - Lagged IVs ($K = 1$ or higher) to account for dynamic causality.
  - Index variable (e.g., `1:n`) to control for linear time trends.
  - Adding dummy variables for months or quarters to control for seasonality.
- How do we choose among them?
- Easy: start with all of them and then use F tests or LM tests to decide what restrictions are sensible.
- E. g., leave out the lag if the test that compares the model with and without lag is insignificant.

How can we Choose the Best Specification?

- There are many things we can do to model a time series:
  - Lagged DV ($K = 1$ or higher) to account for autocorrelation.
  - Lagged IVs ($K = 1$ or higher) to account for dynamic causality.
  - Index variable (e.g., `1:n`) to control for linear time trends.
  - Adding dummy variables for months or quarters to control for seasonality.
- How do we choose among them?
- Easy: start with all of them and then use F tests or LM tests to decide what restrictions are sensible.
- E. g., leave out the lag if the test that compares the model with and without lag is insignificant.
- Most importantly: Also think about what specifications make theoretical sense!

How can we Choose the Best Specification?

- There are many things we can do to model a time series:
  - Lagged DV ($K = 1$ or higher) to account for autocorrelation.
  - Lagged IVs ($K = 1$ or higher) to account for dynamic causality.
  - Index variable (e.g., `1:n`) to control for linear time trends.
  - Adding dummy variables for months or quarters to control for seasonality.
- How do we choose among them?
- Easy: start with all of them and then use F tests or LM tests to decide what restrictions are sensible.
- E. g., leave out the lag if the test that compares the model with and without lag is insignificant.
- Most importantly: Also think about what specifications make theoretical sense!
- See De Boef and Keele (2008).

More Advanced Time Series Topics

ARIMA models Autoregressive Integrated Moving Average model. This is a combination of all of the above (and some more), where the analyst can decide what components to include (e. g., first- or second-order differencing, moving average, lags, trends. . .).

VAR models Vector Autoregression model. Use this is you have multiple time series that interact with each other.

GARCH models Generalized Autoregressive Conditional Heteroskedasticity model. Deals with volatility. For example, stock market data have volatile and calm periods and thus heteroskedasticity.