Time Series ARIMA Models

Ani Katchova

Time Series Models Overview

- Time series examples
- White noise, autoregressive (AR), moving average (MA), and ARMA models
- Stationarity, detrending, differencing, and seasonality
- Autocorrelation function (ACF) and partial autocorrelation function (PACF)
- Dickey-Fuller tests
- The Box-Jenkins methodology for ARMA model selection

Time Series ARIMA Models

Time series examples

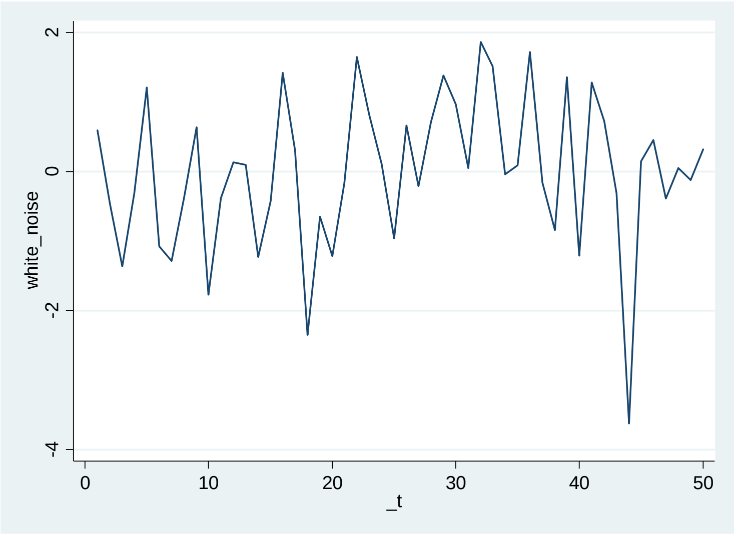
- Modeling relationships using data collected over time – prices, quantities, GDP, etc.
- Forecasting – predicting economic growth.
- Time series involves decomposition into a trend, seasonal, cyclical, and irregular component.

Problems ignoring lags

- Values of y_t are affected by the values of y in the past.
    - For example, the amount of money in your bank account in one month is related to the amount in your account in a previous month.
- Regression without lags fails to account for the relationships through time and overestimates the relationship between the dependent and independent variables.

White noise

- White noise describes the assumption that each element in a series is a random draw from a population with zero mean and constant variance.



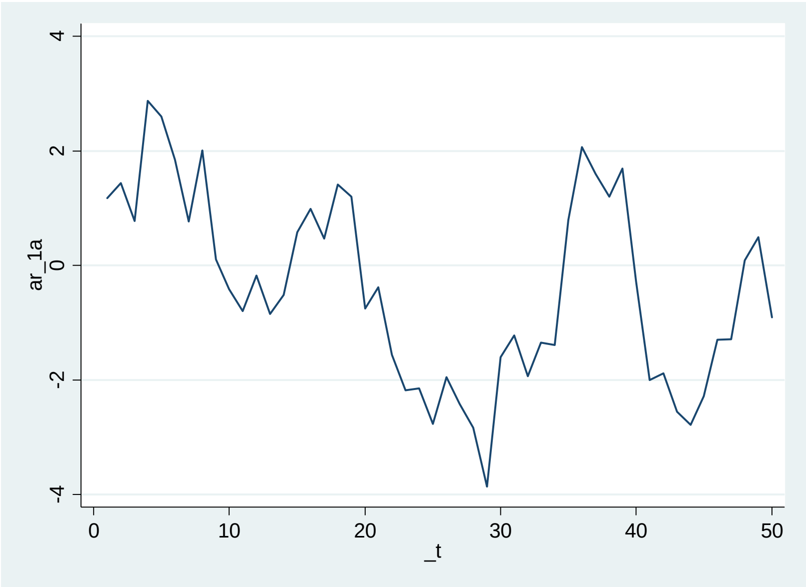
- Autoregressive (AR) and moving average (MA) models correct for violations of this white noise assumption.

Autoregressive (AR) models

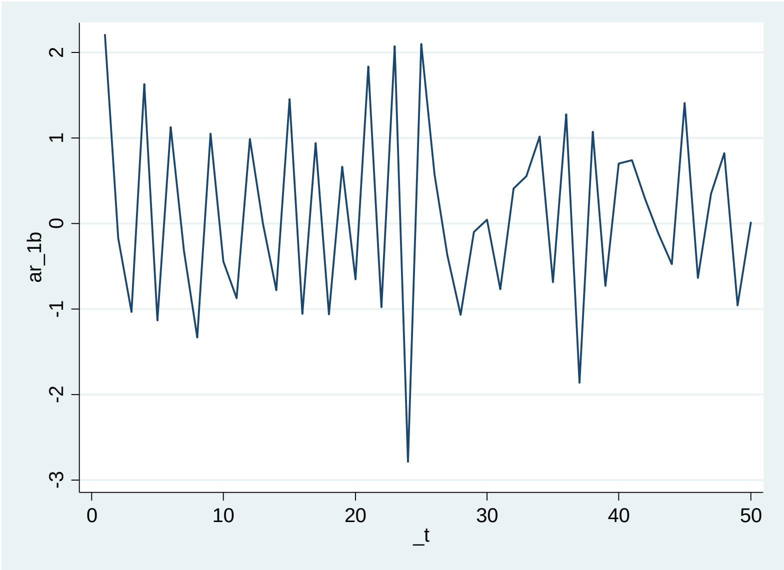
- Autoregressive (AR) models are models in which the value of a variable in one period is related to its values in previous periods.
- AR(p) is an autoregressive model with p lags: $y_t = \mu + \sum_{i=1}^{p} \gamma_i y_{t-i} + \epsilon_t$
  where μ is a constant and γ_p is the coefficient for the lagged variable in time *t-p*.

- AR(1) is expressed as: $y_t = \mu + \gamma y_{t-1} + \epsilon_t = \mu + \gamma(Ly_t) + \epsilon_t$ or $(1 - \gamma L)y_t = \mu + \epsilon_t$

AR(1) with $\gamma = 0.8$



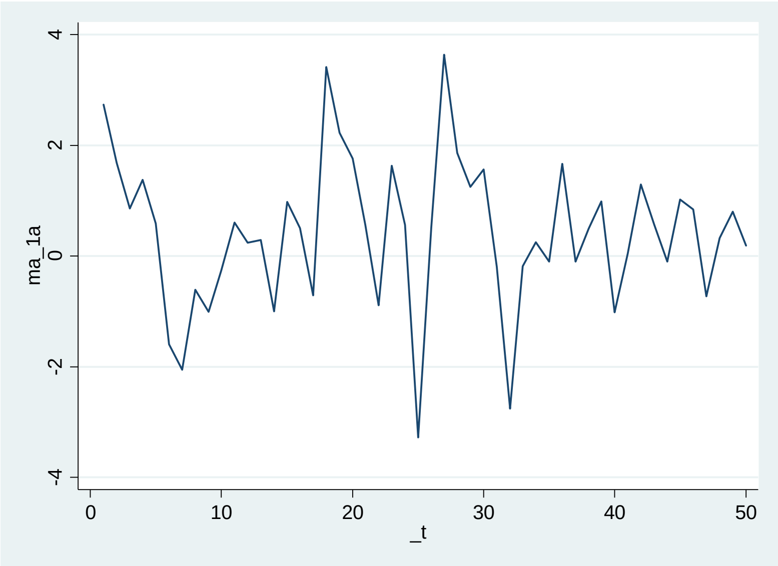
AR(1) with $\gamma = -0.8$



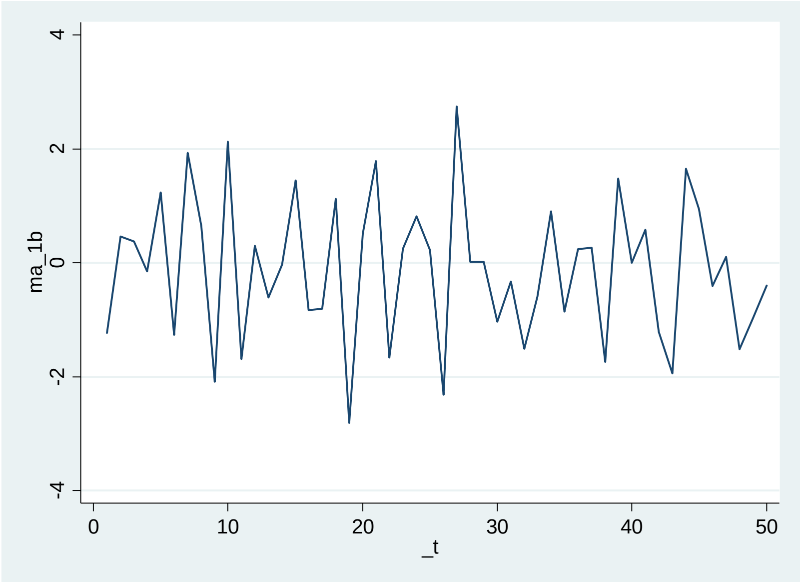
Moving average (MA) models

- Moving average (MA) models account for the possibility of a relationship between a variable and the residuals from previous periods.
- MA(q) is a moving average model with q lags: $y_t = \mu + \epsilon_t + \sum_{i=1}^{q} \theta_i \epsilon_{t-i}$ where θ_q is the coefficient for the lagged error term in time *t*-*q*.
- MA(1) model is expressed as: $y_t = \mu + \epsilon_t + \theta \epsilon_{t-1}$
- Note: SAS (unlike Stata and R), model θ with a reverse sign.

MA(1) with $\theta = 0.7$



MA(1) with $\theta = -0.7$

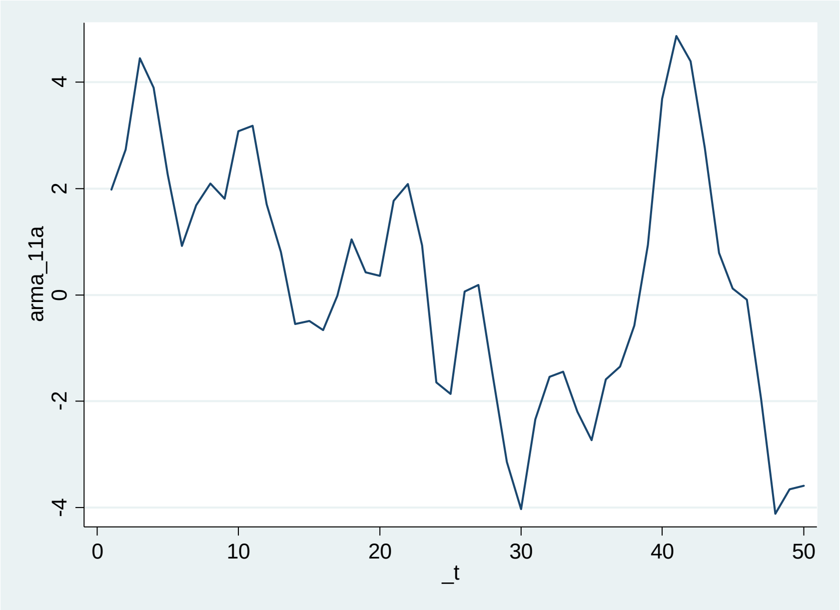


Autoregressive moving average (ARMA) models

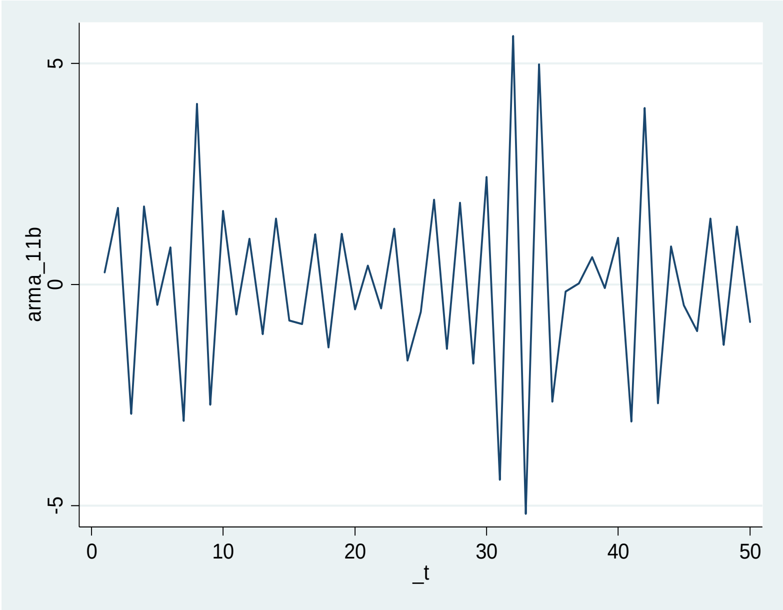
- Autoregressive moving average (ARMA) models combine both p autoregressive terms and q moving average terms, also called ARMA(p,q).

$$y_t = \mu + \sum_{i=1}^{p} \gamma_i y_{t-i} + \epsilon_t + \sum_{i=1}^{q} \theta_i \epsilon_{t-i}$$

ARMA(1,1) with $\gamma = 0.8$ and $\theta = 0.7$



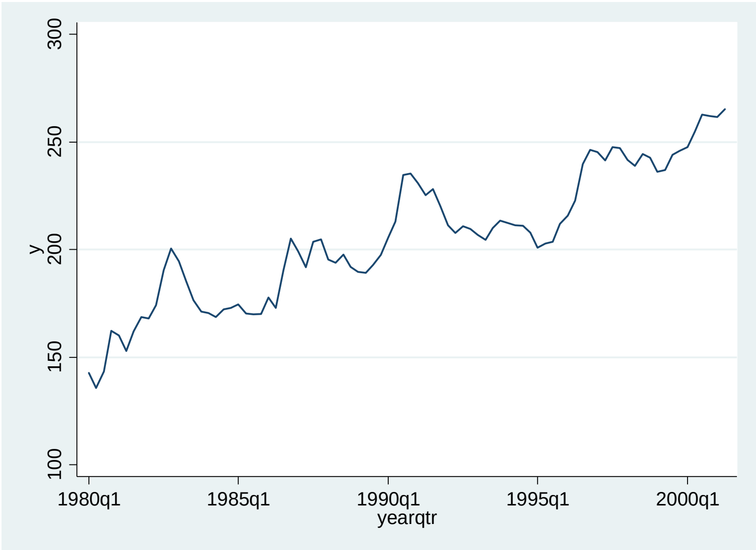
ARMA(1,1) with $\gamma = -0.8$ and $\theta = -0.7$



Stationarity

- Modeling an ARMA(p,q) process requires stationarity.
- A stationary process has a mean and variance that do not change over time and the process does not have trends.
- An AR(1) disturbance process: $u_t = \rho u_{t-1} + \epsilon_t$
- is stationary if $|\rho| < 1$ and ϵ_t is white noise.

- Example of a time-series variable, is it stationary?

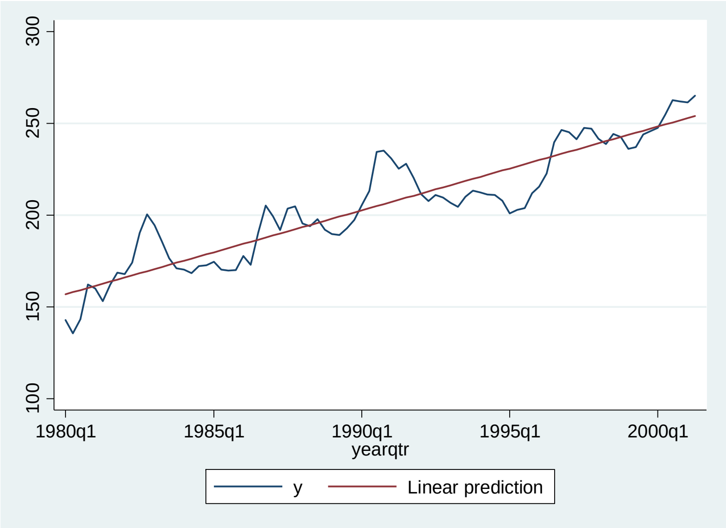


Detrending

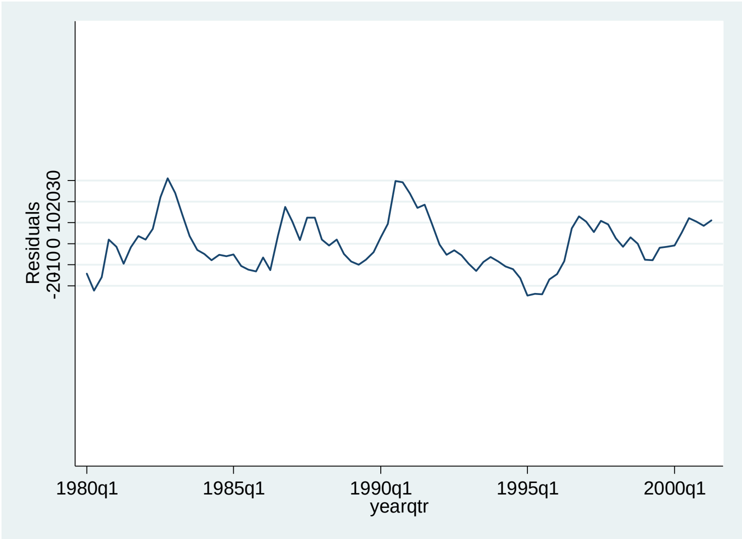
- A variable can be detrended by regressing the variable on a time trend and obtaining the residuals.

$$y_t = \mu + \beta t + \varepsilon_t$$

Variable y_t



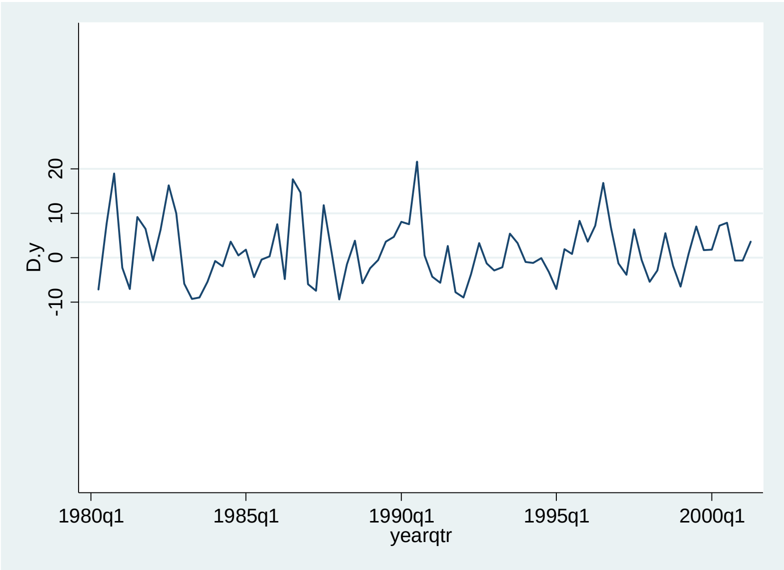
Detrended variable: $\hat{\varepsilon}_t = y_t - \hat{\mu} + \hat{\beta} t$



Differencing

- When a variable y_t is not stationary, a common solution is to use differenced variable: $\Delta y_t = y_t - y_{t-1}$, for first order differences.
- The variable y_t is integrated of order one, denoted $I(1)$, if taking a first difference produces a stationary process.
- ARIMA (p,d,q) denotes an ARMA model with p autoregressive lags, q moving average lags, a and difference in the order of d.

Differenced variable: $\Delta y_t = y_t - y_{t-1}$



Seasonality

- Seasonality is a particular type of autocorrelation pattern where patterns occur every “season,” like monthly, quarterly, etc.
- For example, quarterly data may have the same pattern in the same quarter from one year to the next.
- Seasonality must also be corrected before a time series model can be fitted.

Dickey-Fuller Test for Stationarity

Dickey-Fuller test

- Assume an AR(1) model. The model is non-stationary or a unit root is present if $|\rho| = 1$.

$$y_t = \rho y_{t-1} + e_t$$

$$y_t - y_{t-1} = \rho y_{t-1} - y_{t-1} + e_t$$

$$\Delta y_t = (\rho - 1) y_{t-1} + e_t = \gamma y_{t-1} + e_t$$

- We can estimate the above model and test for the significance of the γ coefficient.
  - If the null hypothesis is not rejected, $\gamma^* = 0$, then y_t is not stationary. Difference the variable and repeat the Dickey-Fuller test to see if the differenced variable is stationary.
  - If the null hypothesis is rejected, $\gamma^* > 0$, then y_t is stationary. Use the variable.
  - Note that non-significance is means stationarity.

Augmented Dickey-Fuller test

- In addition to the model above, a drift μ and additional lags of the dependent variable can be added.

$$\Delta y_t = \mu + \gamma^* y_{t-1} + \sum_{j=1}^{p-1} \phi_j \Delta y_{t-j} + \epsilon_t$$

- The augmented Dickey-Fuller test evaluates the null hypothesis that $\gamma^* = 0$. The model will be non-stationary if $\gamma^* = 0$.

Dickey-Fuller test with a time trend

- The model with a time trend:

$$\Delta y_t = \mu + \beta t + \gamma^* y_{t-1} + \sum_{j=1}^{p-1} \phi_j \Delta y_{t-j} + \epsilon_t$$

- Test the hypothesis that $\beta = 0$ and $\gamma^* = 0$. Again, the model will be non-stationary or will have a unit root present if $\gamma^* = 0$.

Autocorrelation Function (ACF) and Partial Autocorrelation Function (ACF)

Autocorrelation function (ACF)

- ACF is the proportion of the autocovariance of y_t and y_{t-k} to the variance of a dependent variable y_t

$$ACF(k) = \rho_k = \frac{\mathrm{Cov}(y_t, y_{t-k})}{\mathrm{Var}(y_t)}$$

- The autocorrelation function ACF(k) gives the gross correlation between y_t and y_{t-k}.
- For an AR(1) model, the ACF is $ACF(k) = \rho_k = \gamma^k$. We say that this function tails off.

Partial autocorrelation function (PACF)

- PACF is the simple correlation between y_t and y_{t-k} minus the part explained by the intervening lags

$$\rho_k^* = \mathrm{Corr}[y_t - E^*(y_t | y_{t-1}, \dots, y_{t-k+1}), y_{t-k})]$$

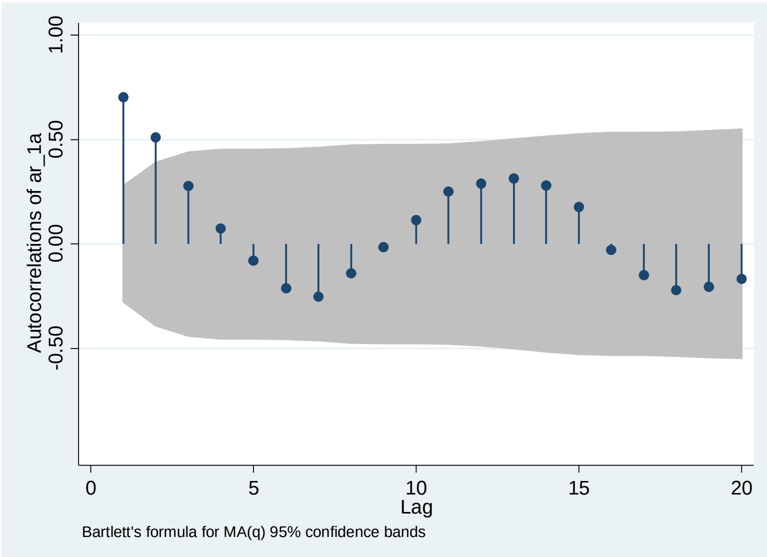
where $E^*(y_t | y_{t-1}, \dots, y_{t-k+1})$ is the minimum mean-squared error predictor of y_t by $y_{t-1}, \dots, y_{t-k+1}$.

- For an AR(1) model, the PACF is γ for the first lag and then cuts off.

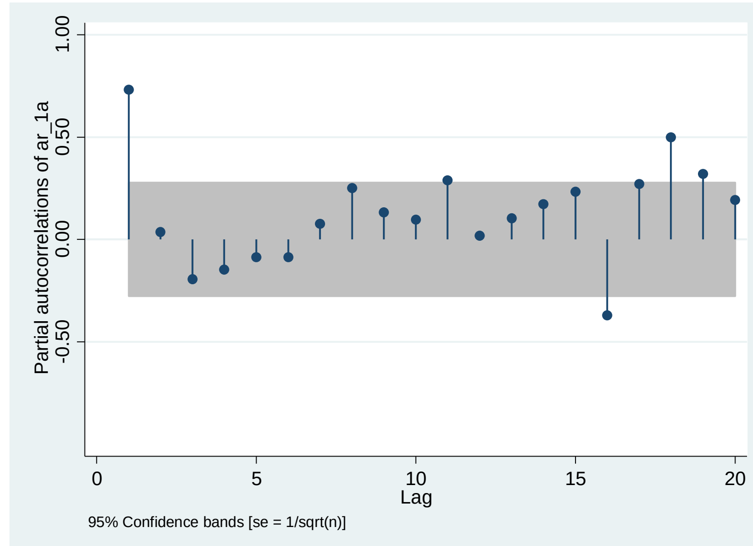
ACF and PACF properties

	AR(p)	MA(q)	ARMA(p,q)
ACF	Tails off	Cuts off after lag q	Tails off
PACF	Cuts off after lag p	Tails off	Tails off

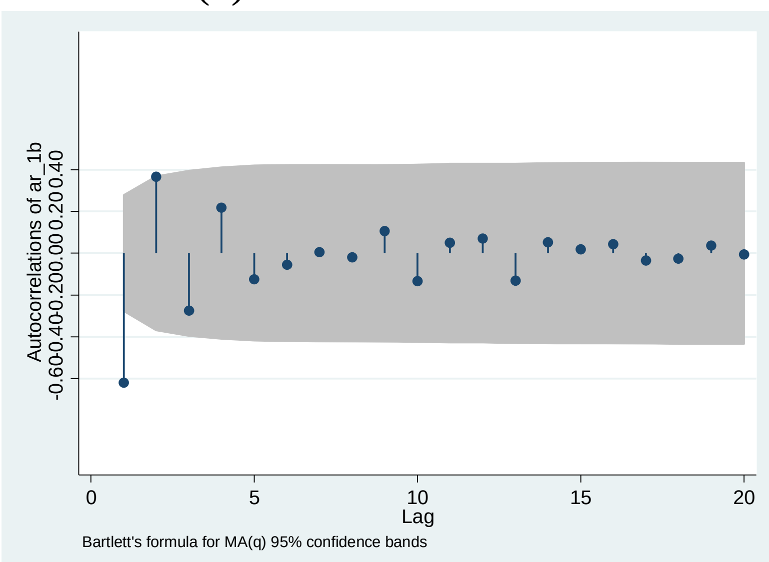
ACF of AR(1) with coefficient 0.8



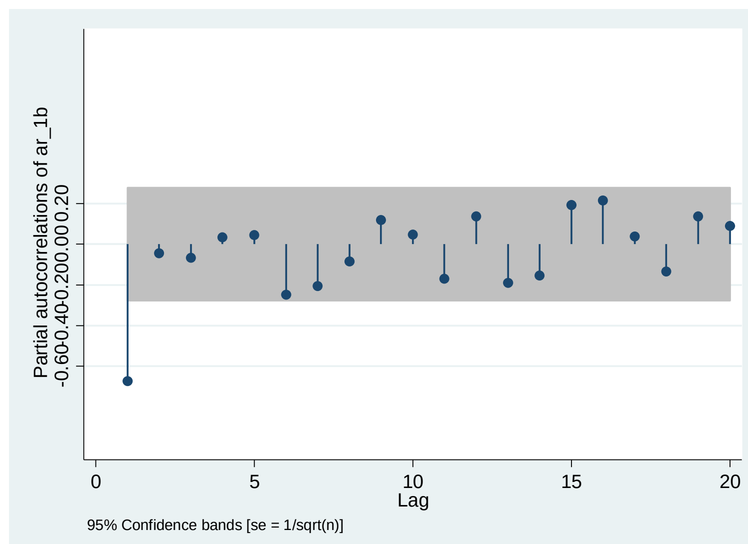
PACF of AR(1) with coefficient of 0.8



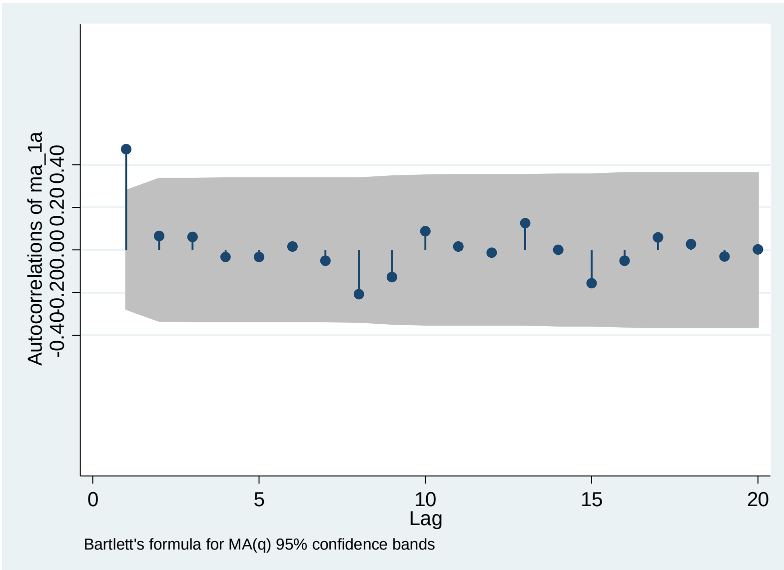
ACF of AR(1) with coefficient -0.8



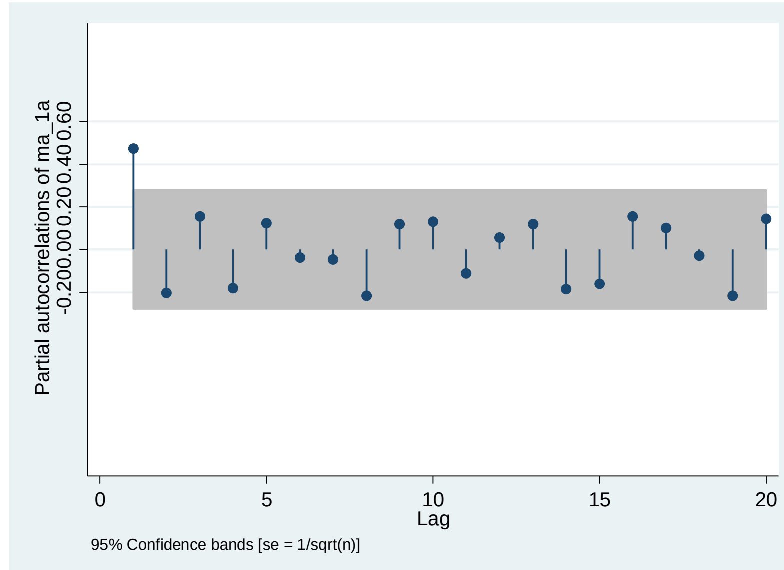
PACF of AR(1) with coefficient of -0.8



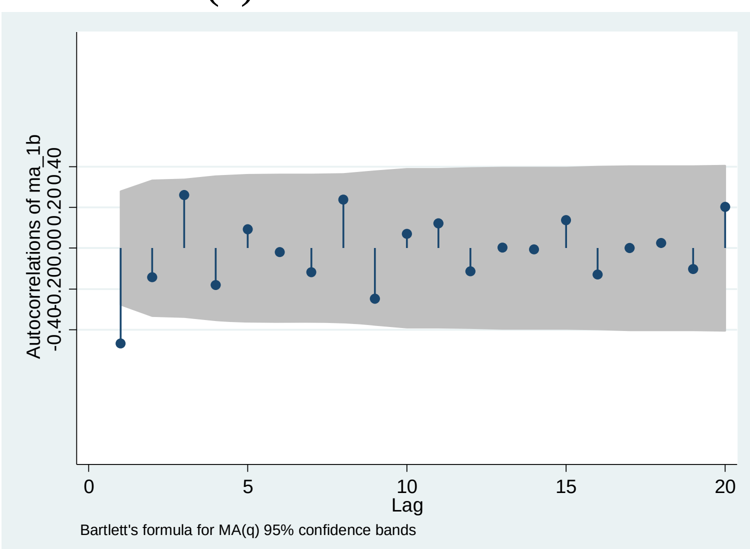
ACF of MA(1) with coefficient of 0.7



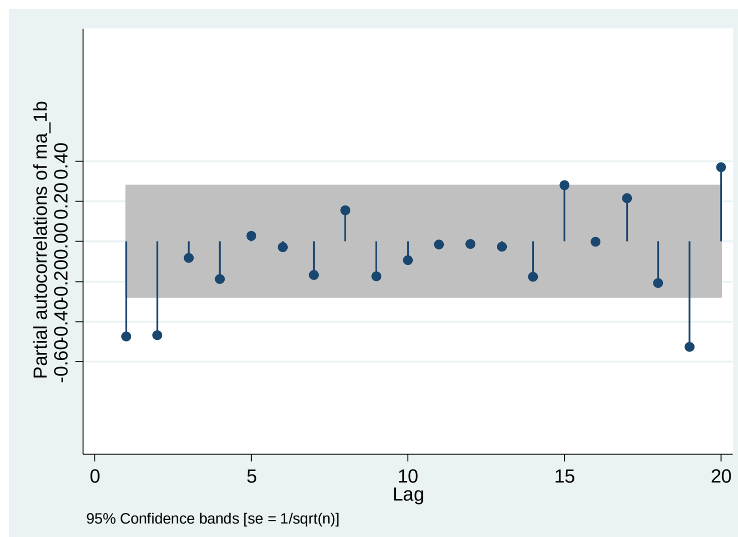
PACF of MA(1) with coefficient of 0.7



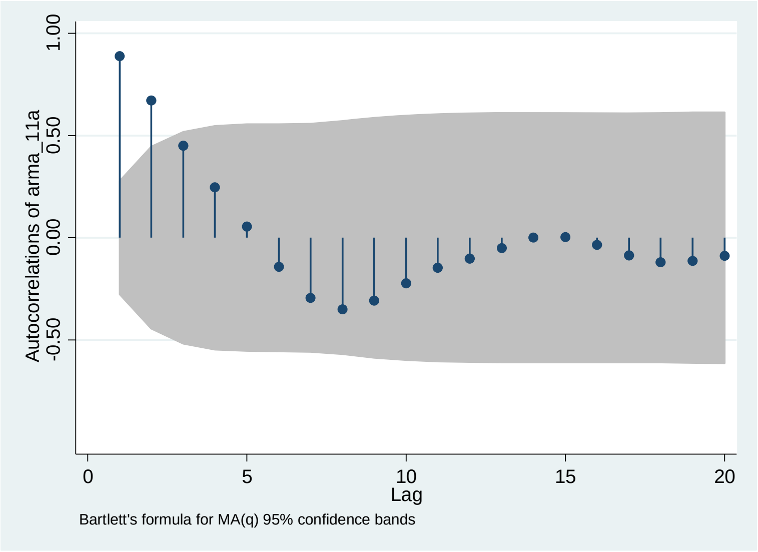
ACF of MA(1) with coefficient of -0.7



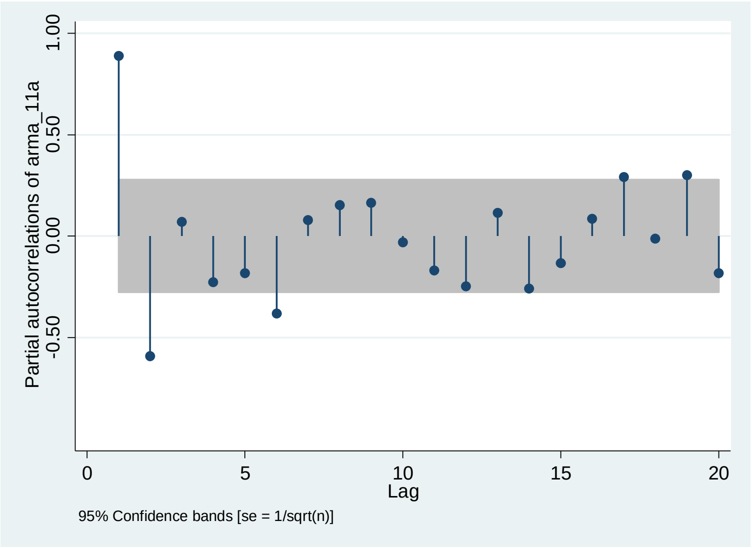
PACF of MA(1) with coefficient of -0.7



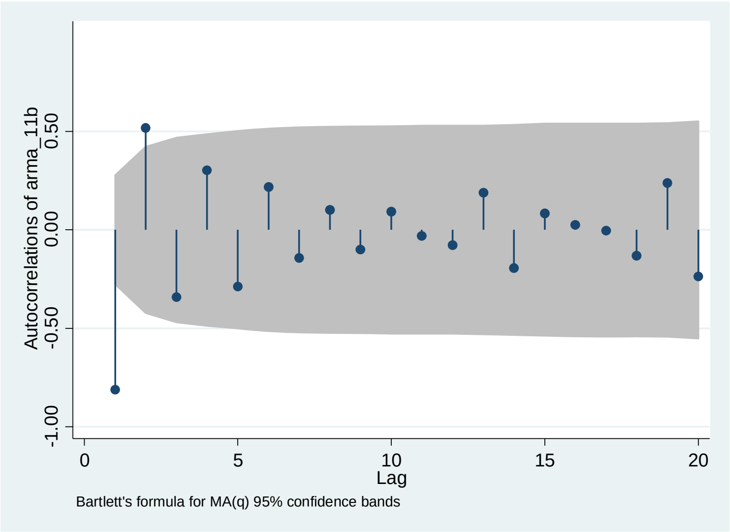
ACF of ARMA(1,1) with coeff 0.8 and 0.7



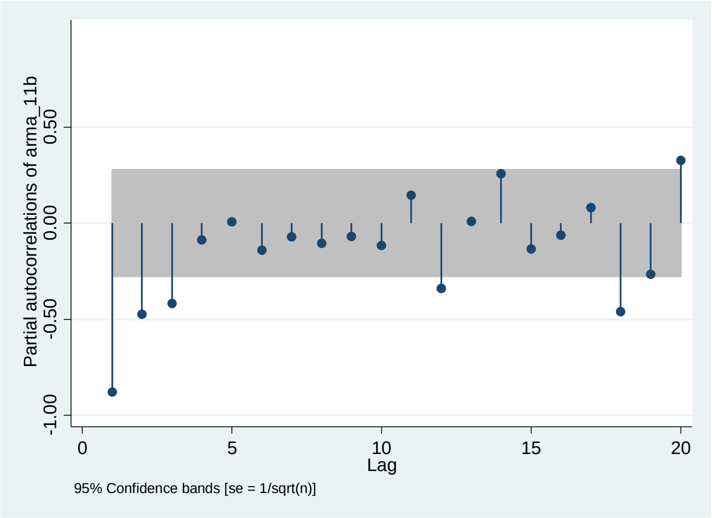
PACF of ARMA(1,1) with coeff 0.8 and 0.7



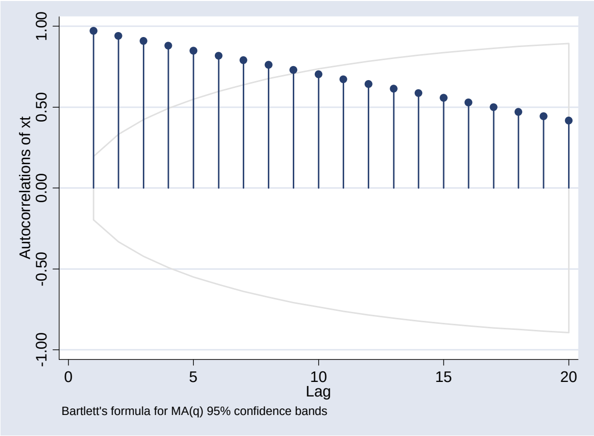
ACF of ARMA(1,1) with coeff -0.8 and -0.7



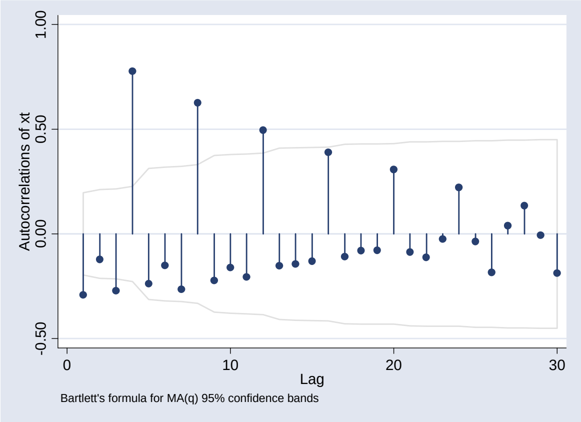
PACF of ARMA(1,1) with coeff -0.8 and -0.7



ACF of non-stationary series - The ACF shows a slow decaying positive ACF.



ACF with seasonal lag (4) – ACF shows spikes every 4 lags.



Diagnostics for ARIMA Models

Testing for white noise

- The Box-Pierce statistic is the following: $Q = T \sum_{k=1}^{P} \rho_k^2$
- The Ljung-Box statistic: $Q' = T(T+2) \sum_{k=1}^{P} \frac{\rho_k^2}{T-k}$

  where ρ_k is the sample autocorrelation at lag k.
- Under the null hypothesis that the series is white noise (data are independently distributed), Q has a limiting χ^2 distribution with p degrees of freedom.

Goodness of fit

- Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) are two measures goodness of fit. They measure the trade-off between model fit and complexity of the model.

$$\text{AIC} = -2\ln(L) + 2k$$

$$\text{BIC} = -2\ln(L) + \ln(N)k$$

where L is the value of the likelihood function evaluated at the parameter estimates, N is the number of observations, and k is the number of estimated parameters.

- A lower AIC or BIC value indicates a better fit (more parsimonious model).

The Box-Jenkins Methodology for ARIMA Model Selection

Identification step

- Examine the time plot of the series.
  - Identify outliers, missing values, and structural breaks in the data.
  - Non-stationary variables may have a pronounced trend or have changing variance.
  - Transform the data if needed. Use logs, differencing, or detrending.
    - Using logs works if the variability of data increases over time.
    - Differencing the data can remove trends. But over-differencing may introduce dependence when none exists.
- Examine the autocorrelation function (ACF) and partial autocorrelation function (PACF).
  - Compare the sample ACF and PACF to those of various theoretical ARMA models. Use properties of ACF and PACF as a guide to estimate plausible models and select appropriate p, d, and q.
  - With empirical data, several models may need to be estimated.
  - Differencing may be needed if there is a slow decay in the ACF.

Estimation step

- Estimate ARMA models and examine the various coefficients.
- The goal is to select a stationary and parsimonious model that has significant coefficients and a good fit.

Diagnostic checking step

- If the model fits well, then the residuals from the model should resemble a while noise process.
  - Check for normality looking at a histogram of the residuals or by using a quantile-quantile (Q-Q) plot.
  - Check for independence by examining the ACF and PACF of the residuals, which should look like a white noise.
  - The Ljung-Box-Pierce statistic performs a test of the magnitude of the autocorrelations of the correlations as a group.
  - Examine goodness of fit using the Akaike Information Criteria (AIC) and Bayesian Information Criteria (BIC). Use most parsimonious model with lowest AIC and/or BIC.