Econometrics 871

LECTURE 1

Overview of course,

difference equations and their solutions

Plan

- Admin and Course Intro
- Intro to time series data
- Difference Equations and their solutions
  - Discussion and Definitions
  - Some practice and examples
- Goal of the lecture:
  - To derive the conditions/rule under which a difference equation has a well-defined solution
  - This will become the conditions under which a univariate stochastic time series process is *stationary*

The Team

Part 1: Time Series Econometrics
6/7 topics

Gideon du Rand (gideondurand@sun.ac.za)
Rm 503 C.G.W Schumann Building

Part 2: Cross Section Econometrics
6 weeks

Cobus Burger (cobusburger@sun.ac.za)
Rm 608 C.G.W. Schumann Building

Assessment

	Dates	Weight in Final Mark
Time Series Assessment **- mid semester test**	12 April 2023	25%
Time Series Research Project	28 April 2023	20%
Cross Section Assessment **- exam**	Date to be announced	25%
Cross Section Research Project	19 June 2023	20%
Graded Exercises	Throughout module	10%

Assessment details

- Read the course outline carefully

Passing requirements:

  - You need a final course mark of 50% to pass the course
  - You must obtain at least 40% in the assessment on ***each*** section

Assessment details

Departmental rules:

Time Series Assessment

- The semester test on Time Series is compulsory
  - Absence only excused with a ***valid*** medical certificate
  - Submitted within 48 hours to **Gideon**
- If you are medically excused from or obtain <40% in the Time Series semester test
  - You *must* write the Time Series sick test (on the *same* day as the Cross Section exam)
  - You *must* obtain at least 40% in the Time Series sick test to pass the course
  - If taken as second opportunity, test mark capped at 50%
- If you obtain 40-49% in the Time Series semester test, you *may* write Time Series sick test
  - The Time Series sick test mark also capped at 50%
- The sick test is the ***final*** opportunity to pass the Time Series component of the course

Assessment details

Departmental rules

Cross Section Assessment

- The exam on Cross Section is compulsory
  - Absence only excused with a ***valid*** medical certificate
  - Submitted within 48 hours to **Carina**
- If you obtain <40% in the Cross Section exam, you fail the course
- If you obtain >40% in the Cross Section exam, but your final course mark is <50%
  - You *must* write the supplementary exam and perform adequately to obtain a final mark of at least 50% to pass the course. Note that your final course mark will be capped at 50%.
- If medically excused from exam,
  - there is no cap on sick exam
- The sick exam is the ***final*** opportunity to pass the Cross Section component of the course

Assessment details

Rules on medical excuses from assessment

- If you submit a valid medical certificate that excuses you from assessment on a given day, you cannot write ***any*** assessment on that **day**

- For this course this implies the following:

  If you are medically excused or obtain less than 40% on the Time Series test,

  - you *must* write the sick test *and* the exam on the same day to be eligible to pass the course
  - A(nother) medical excuse for the exam and sick test will imply an immediate incomplete on the course

Time Series Research Project

- Replication and sensitivity analysis of an extant time series study
  - Find a published time series study to replicate
    - From a top 500 journal on https://ideas.repec.org/top/old/0702/top.journals.simple.html
    - **You need to be able to find the same data used**
    - Post a proposal on SUNLearn for my approval and assistance
  - The focus is on technique, understanding and exposition
    - Not a full, novel research proposition with extensive literature review

Resources

- Textbooks:

Enders, Walter (2014) *Applied Econometric Time Series*, 4th Edition

Hendry, David F. and Nielsen, Bent. (2007). *Econometric Modeling: A Likelihood Approach.*

Hamilton, James D. (1992) *Time Series Analysis*.

- Course notes:
  - Kevin Kotze, UCT

https://www.economodel.com/time-series-analysis

Introduction to Time Series Econometrics

Time Series vs Cross Section

Cross Section

Time Series

Time Series vs Cross Section

Cross Section

Std OLS assumptions:

- Linearity in parameters
- No multi-collinearity
- Random Sample from population

- Mean-independent regressors $E[\varepsilon|x]=0$

- Homoskedasticity

Time Series

Time Series vs Cross Section

Cross Section

Std OLS assumptions:

- Linearity in parameters
- No multi-collinearity
- Random Sample from population

- Mean-independent regressors $E[\varepsilon|x]=0$

- Homoskedasticity

Time Series

Std Time Series situation:

- Linearity in parameters
- No multi-collinearity

Time Series vs Cross Section

Cross Section

Std OLS assumptions:

- Linearity in parameters
- No multi-collinearity
- Random Sample from population same variables over different individuals with different histories
- Mean-independent regressors E[ε|x]=0
- Homoskedasticity

Time Series

Std Time Series situation:

- Linearity in parameters
- No multi-collinearity
- Sample is **fundamentally** non-random: same variables over *time,* "same histories"

Time Series vs Cross Section

Cross Section

Std OLS assumptions:

- Linearity in parameters
- No multi-collinearity
- Random Sample from population
- Mean-independent regressors $E[\varepsilon|x]=0$
- Homoskedasticity

Time Series

Std Time Series situation:

- Linearity in parameters
- No multi-collinearity
- Sample is **fundamentally** non-random: same variables over time
- Regressors: at best predetermined, mostly simultaneous

Time Series vs Cross Section

Cross Section

Std OLS assumptions:

- Linearity in parameters
- No multi-colinearity
- Random Sample from population
- Mean-independent regressors E[ε|x]=0
- Homoskedasticity

Time Series

Std Time Series situation:

- Linearity in parameters
- No multi-collinearity
- Sample is **fundamentally** non-random: same variables over time
- Regressors: at best predetermined, mostly simultaneous
- Persistence / Auto-correlation is always a concern

But wait… there is more!

Time Series vs Cross Section

Cross Section

Std OLS assumptions:

- Linearity in parameters
- No multi-colinearity
- Random Sample from population
- Mean-independent regressors $E[\varepsilon|x]=0$
- Homoskedasticity

Time Series

Std Time Series situation:

- Linearity in parameters
- No multi-collinearity
- Sample is non-random:
  same variables over time
  - Simplest scenario: drawn from the same distribution over time
    (stationary and ergodic)
- Regressors: predetermined or simultaneous
- Auto-correlation is always a concern

Time Series vs Cross Section

Cross Section

Std OLS assumptions:

- Linearity in parameters
- No multi-colinearity
- Random Sample from population
- Mean-independent regressors E[ε|x]=0
- Homoskedasticity

Time Series

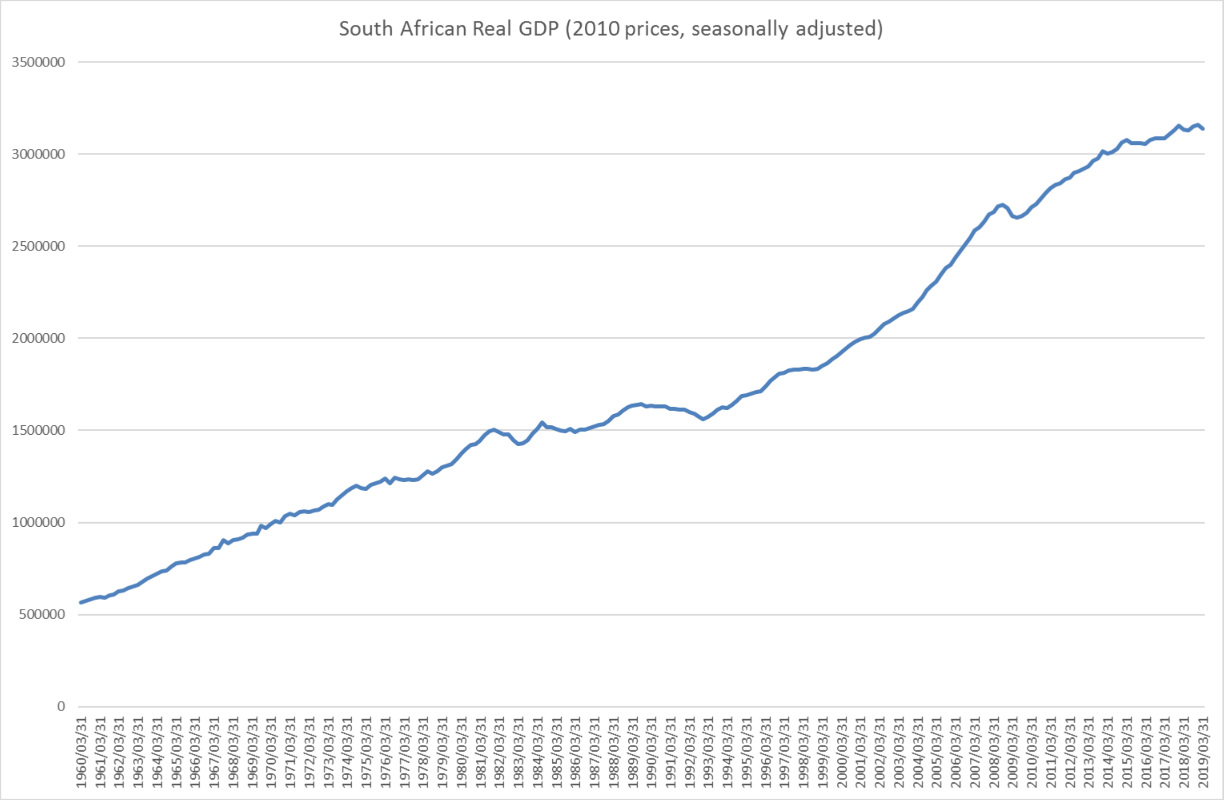
Std Time Series situation:

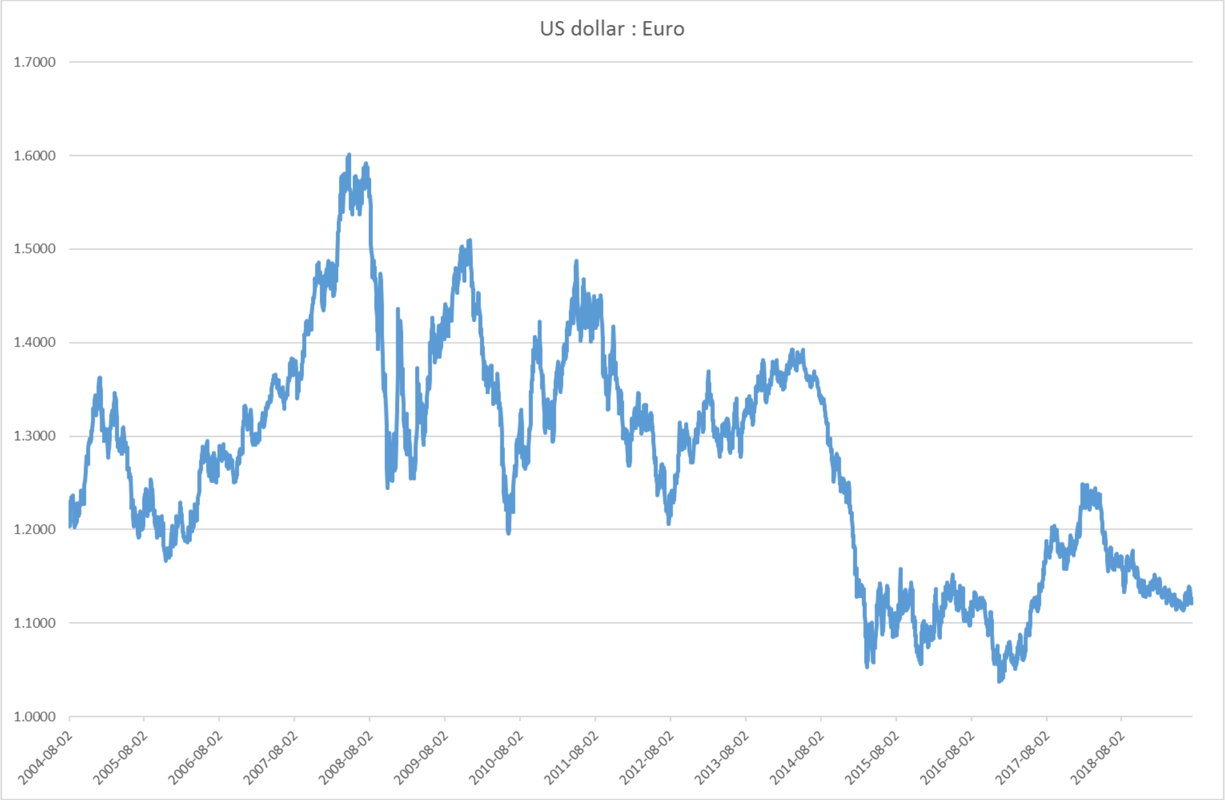
- Linearity in parameters
- No multi-collinearity
- Sample is non-random:
  same variables over time
  - Simplest scenario: drawn from the same distribution over time
    (stationary and ergodic)
  - Common case: non-stationary
- Regressors: predetermined or simultaneous
- Auto-correlation is always a concern

Stylized Facts about economic time series

Stylized Facts about economic time series

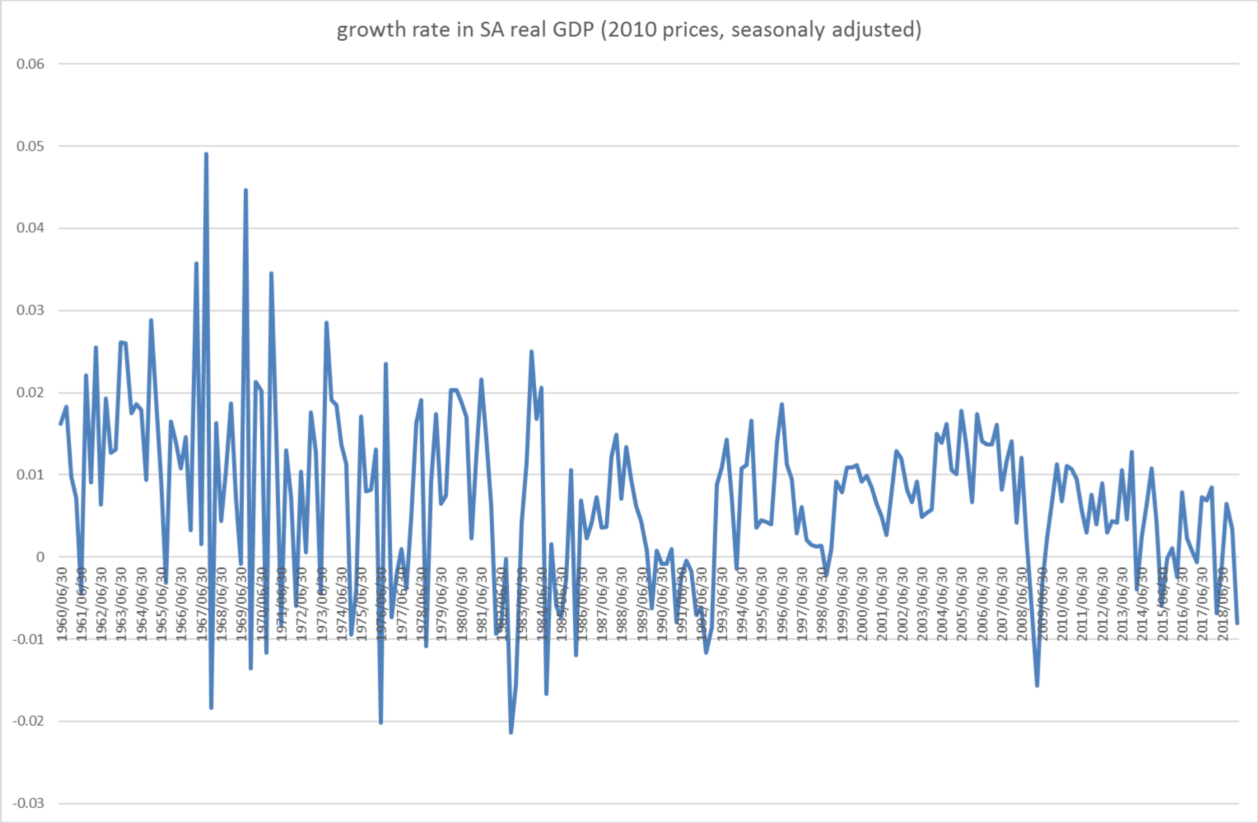
1. persistence, with or without an apparent trend

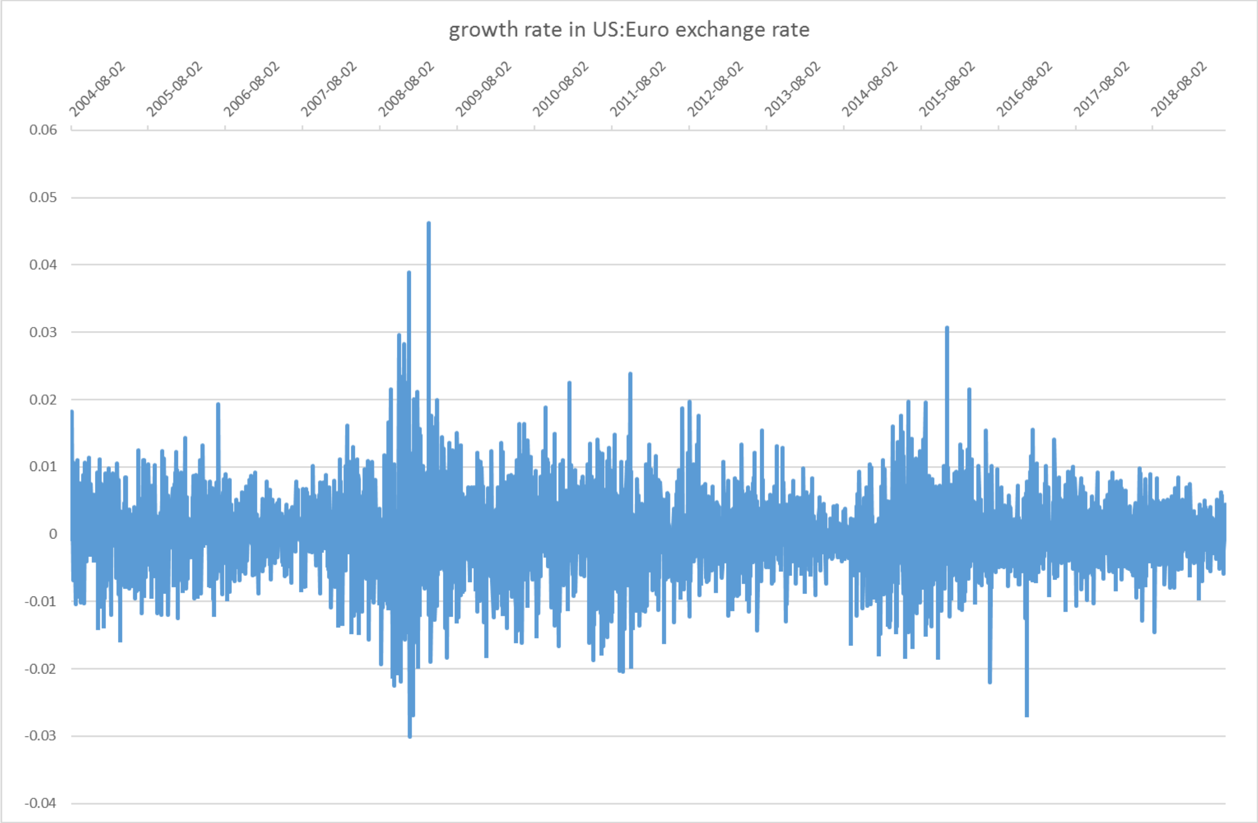




Stylized Facts about economic time series

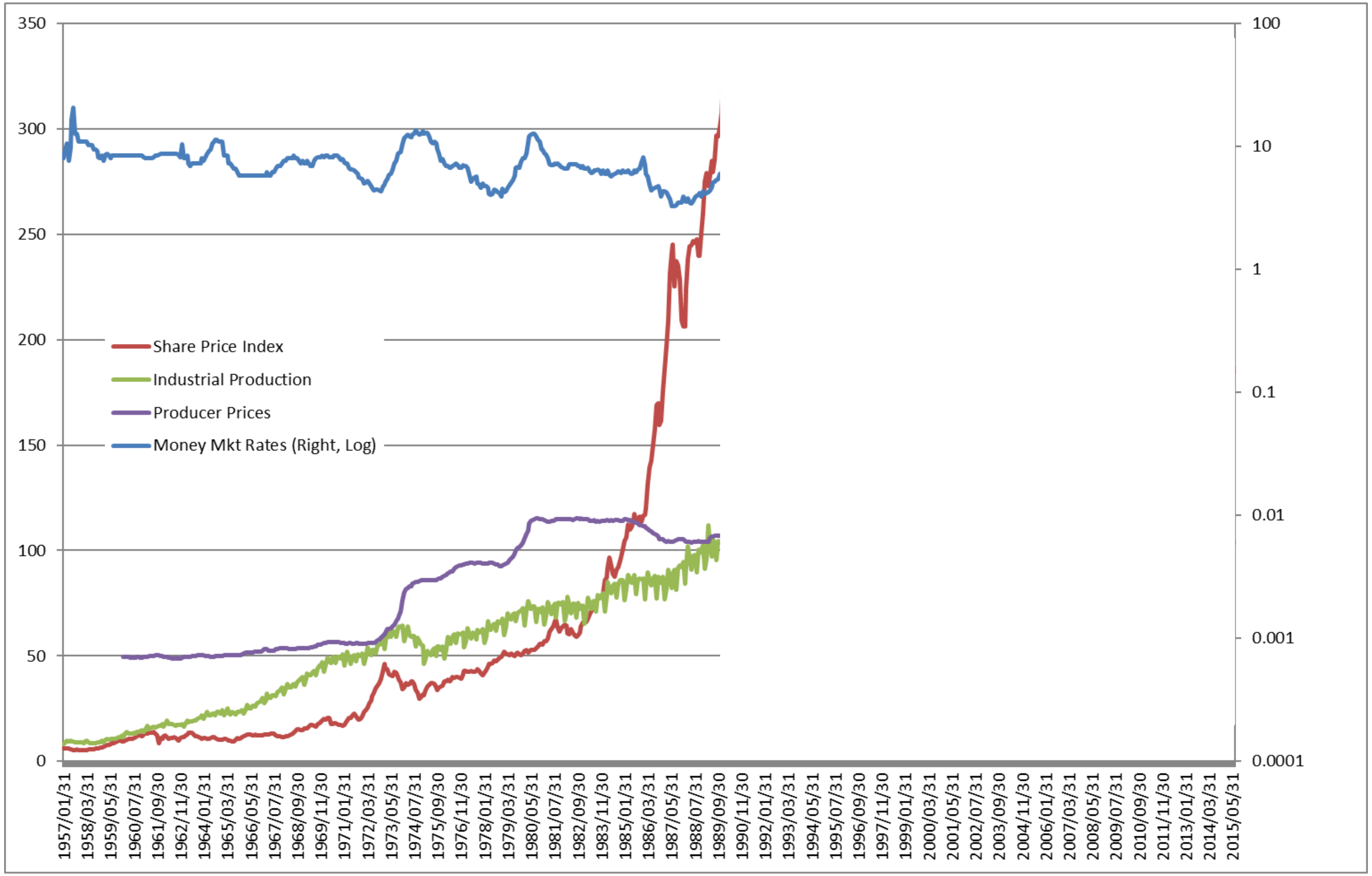
2. the growth rates of persistent series tends to be far less persistent





Stylized Facts about economic time series

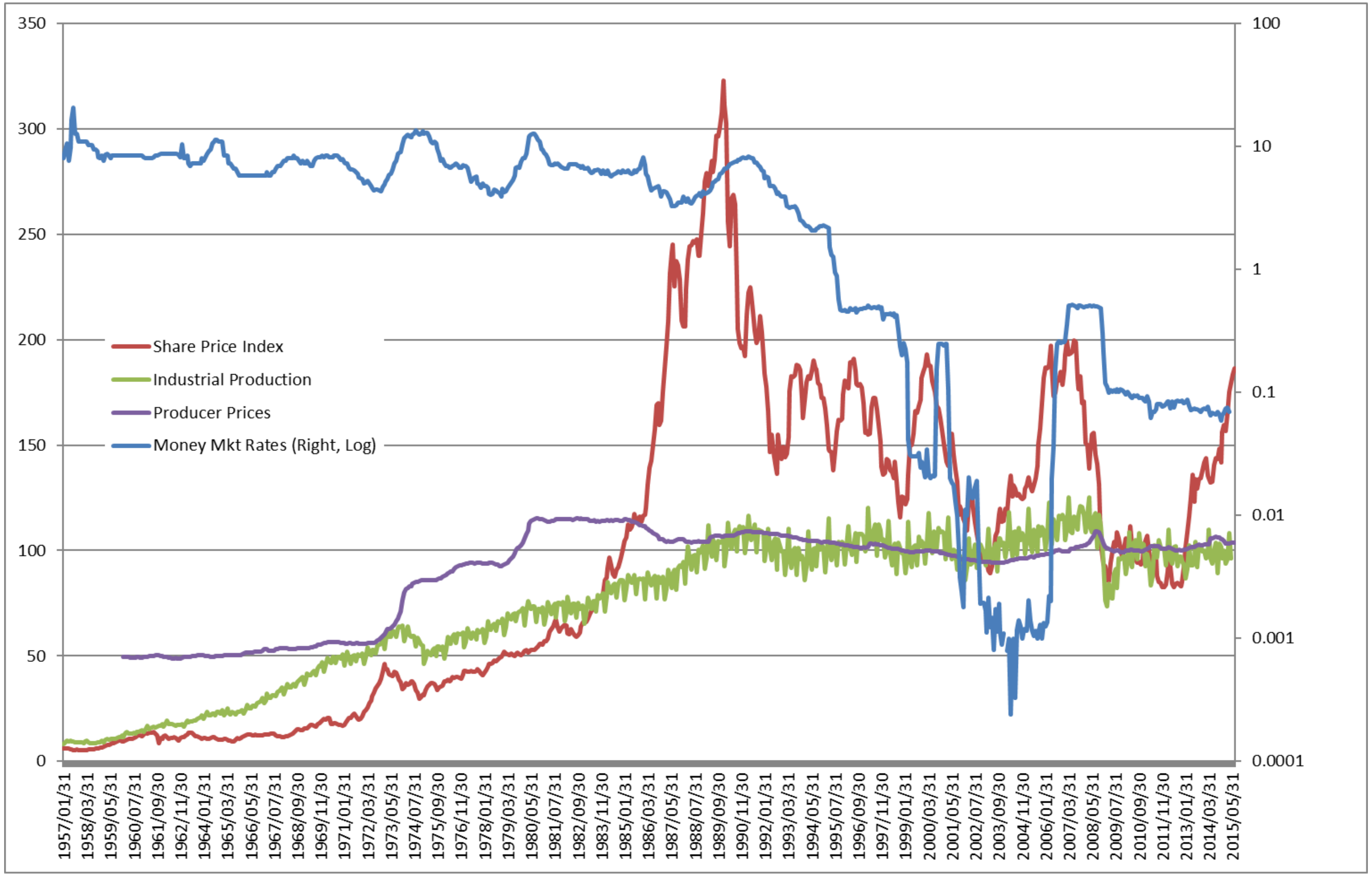
3. apparently systematic behaviour can change without warning (structural breaks)



Data on Japan, source: fred.stlouisfed.org

Stylized Facts about economic time series

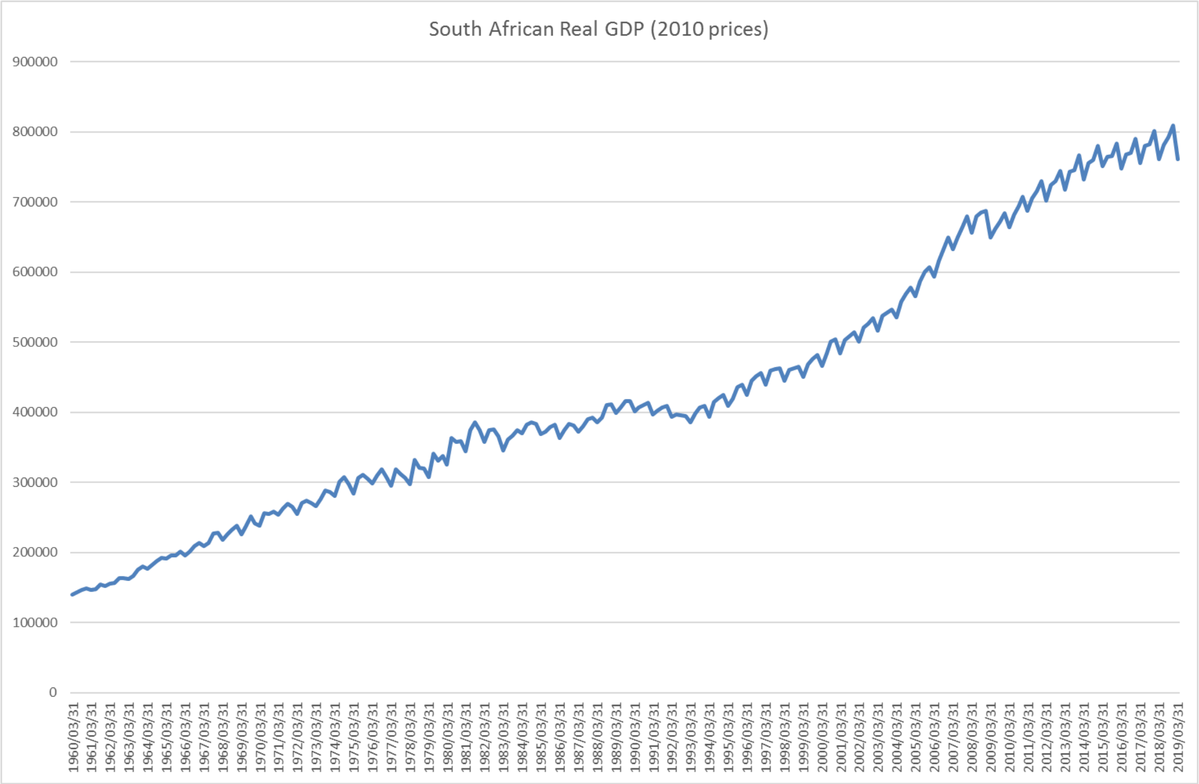
3. apparently systematic behaviour can change without warning (structural breaks)

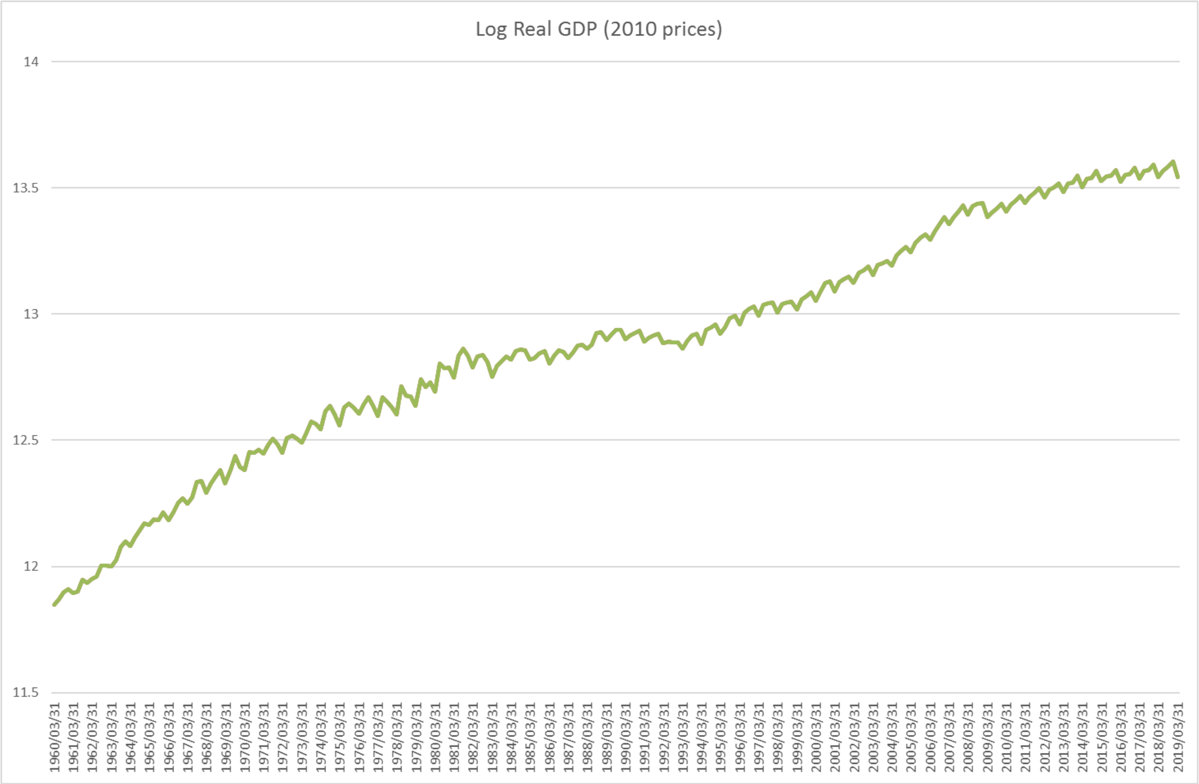


Data on Japan, source: fred.stlouisfed.org

Stylized Facts about economic time series

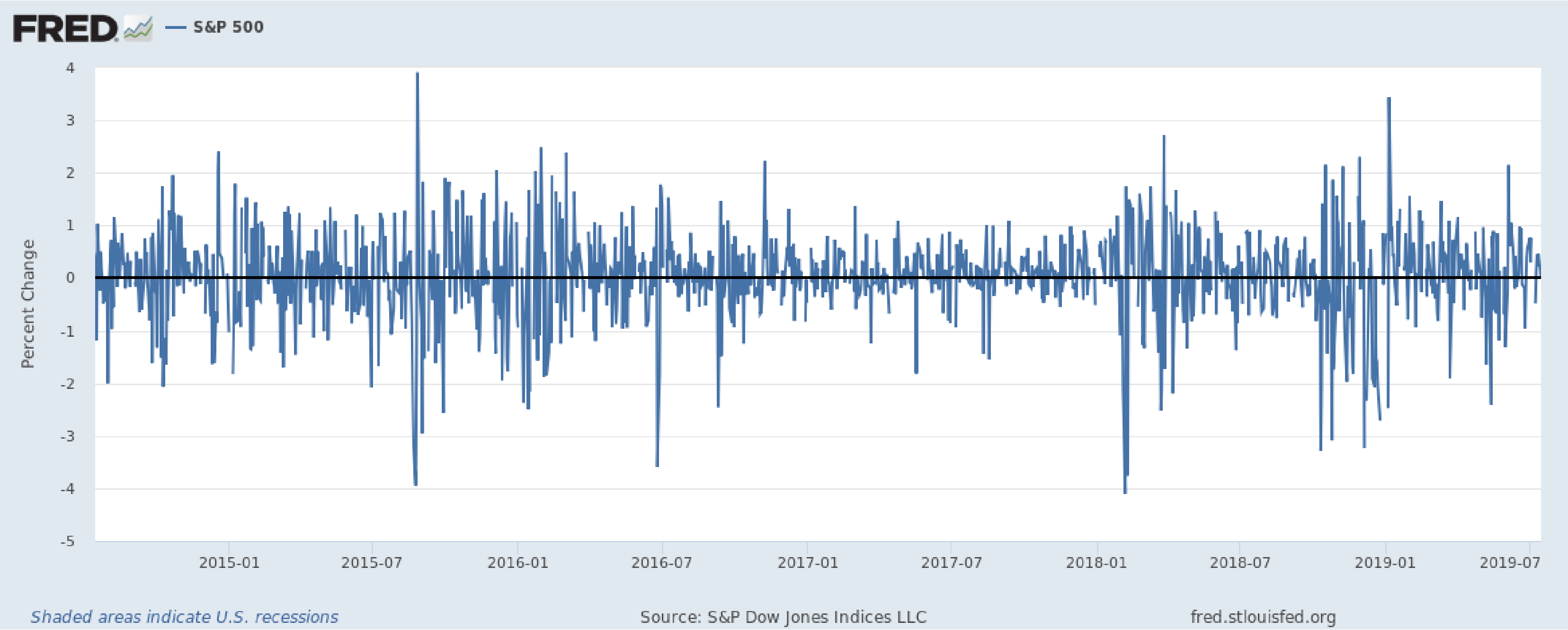
4. Seasonality is a fundamental feature, but it is not “sinusoidal”, and scaling matters





Stylized Facts about economic time series

5. Some series seems to have constant means, but periods of high/low volatility

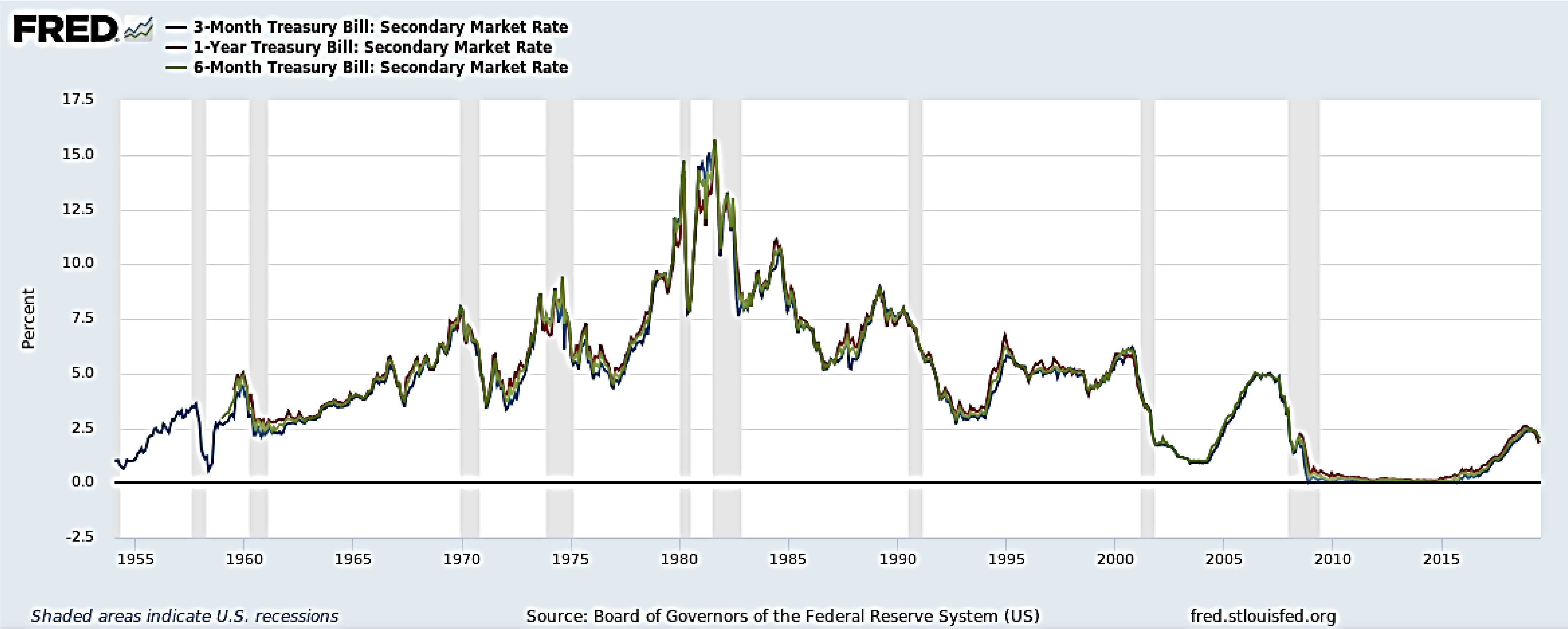


Stylized Facts about economic time series

6. Related economic series tend to co-move systematically

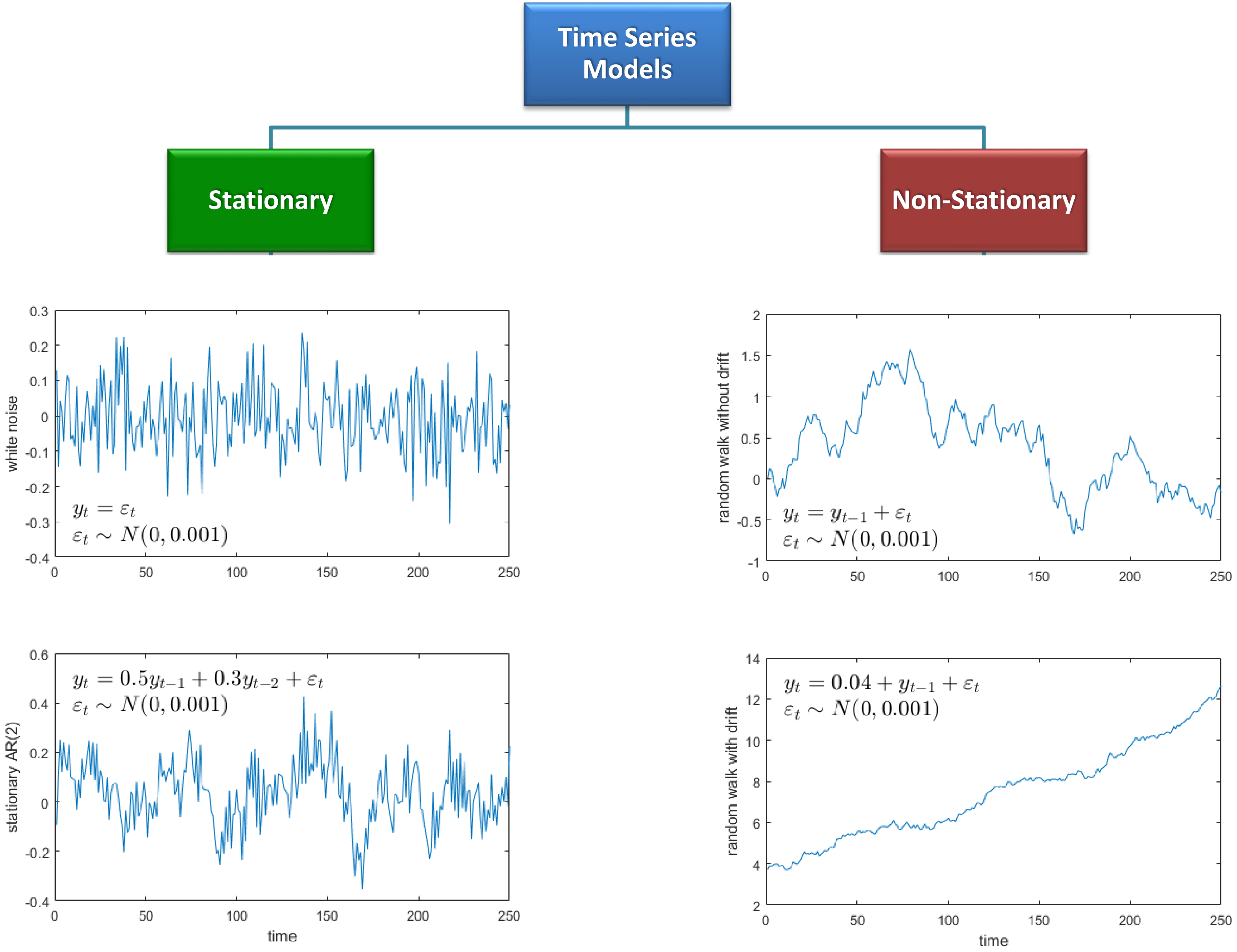
Stylized Facts about economic time series

7. Co-movements reflect relationships with broader economic forces

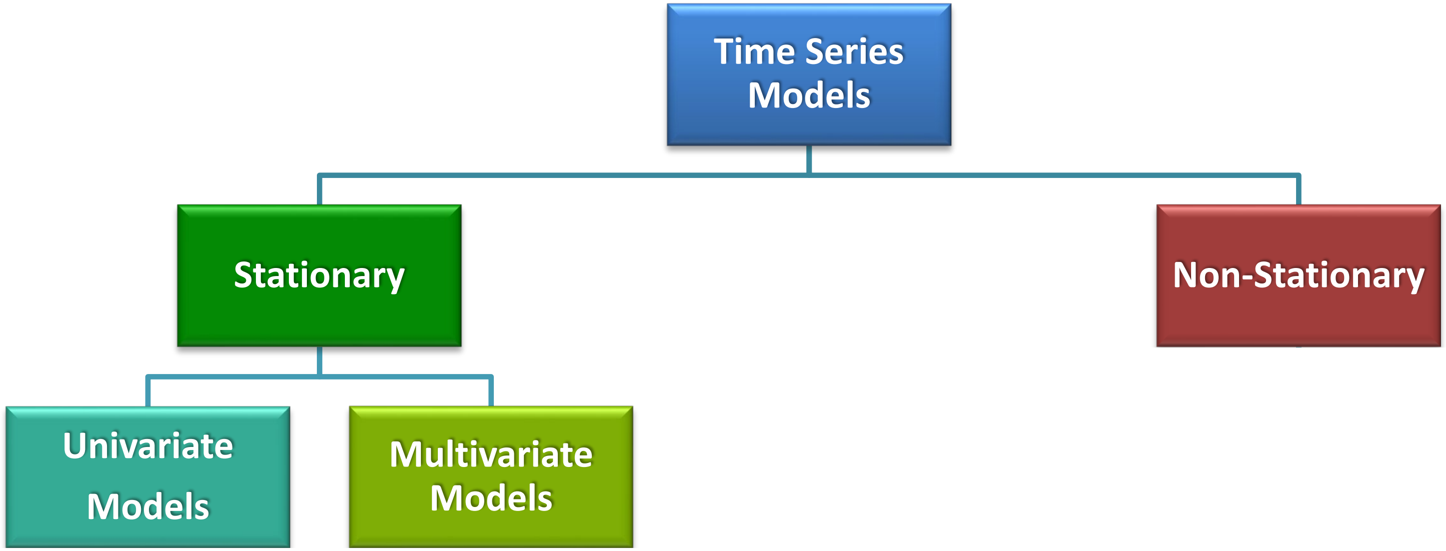


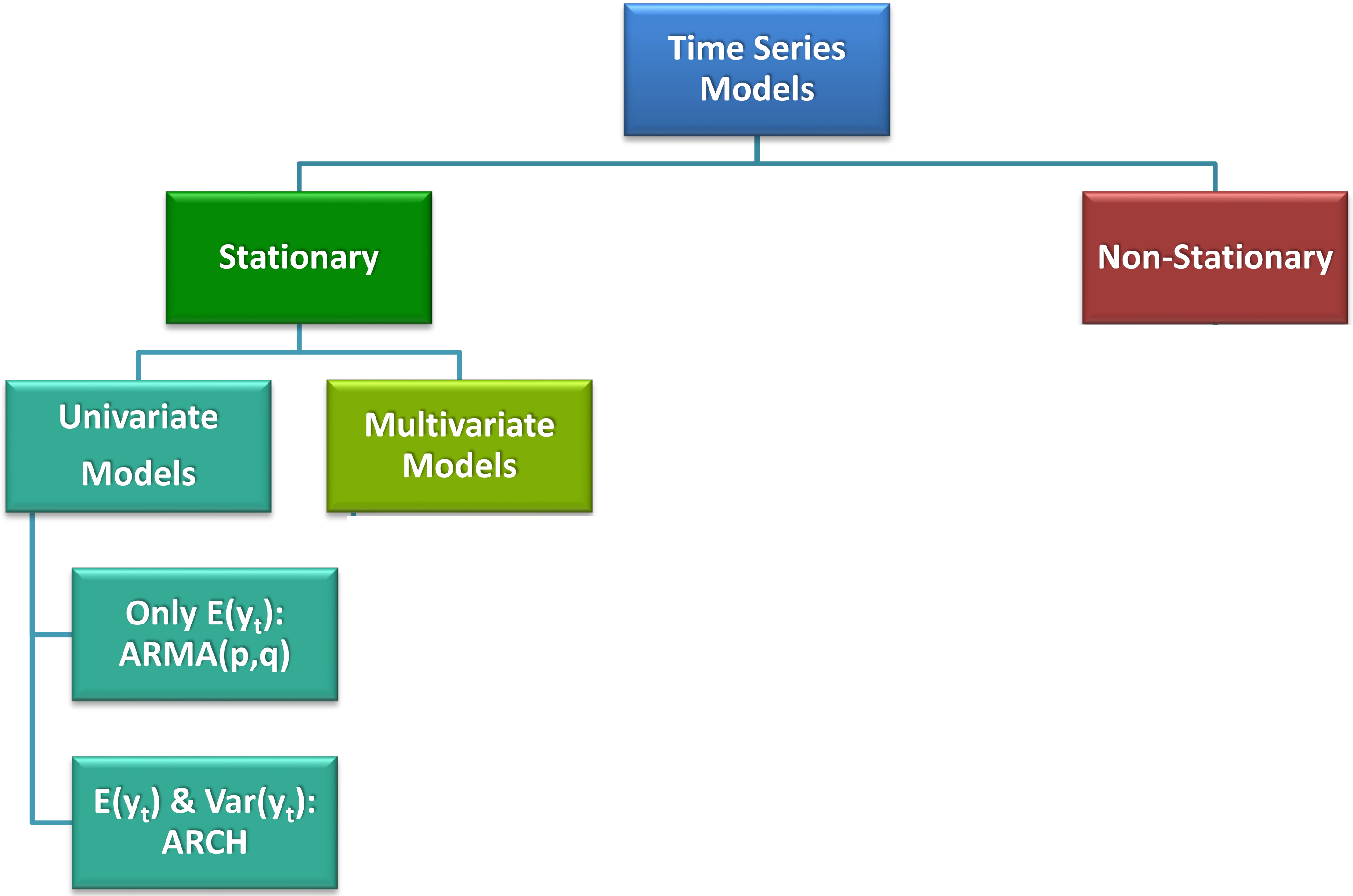
Overview of Time Series Models

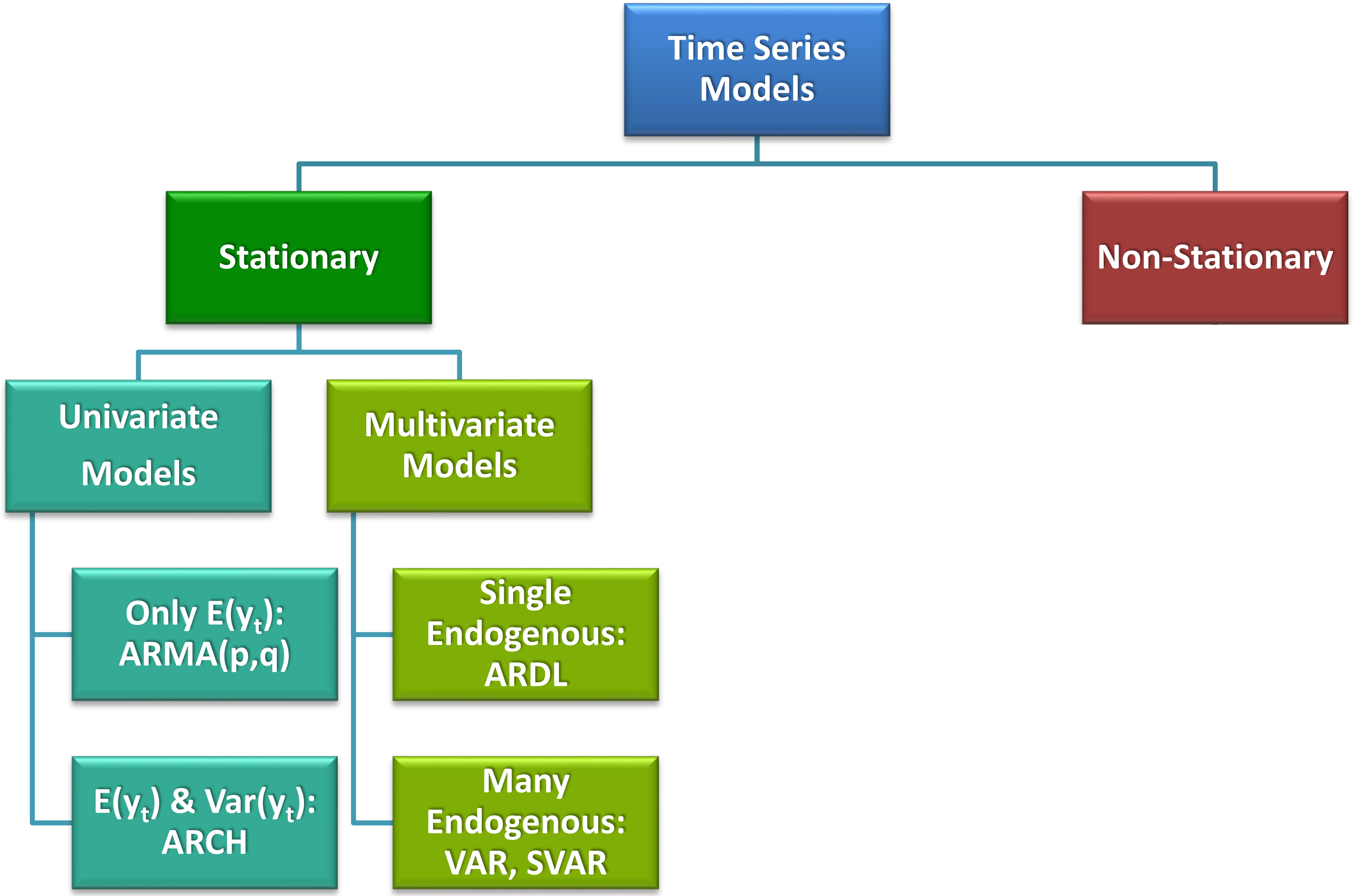
Time Series Models

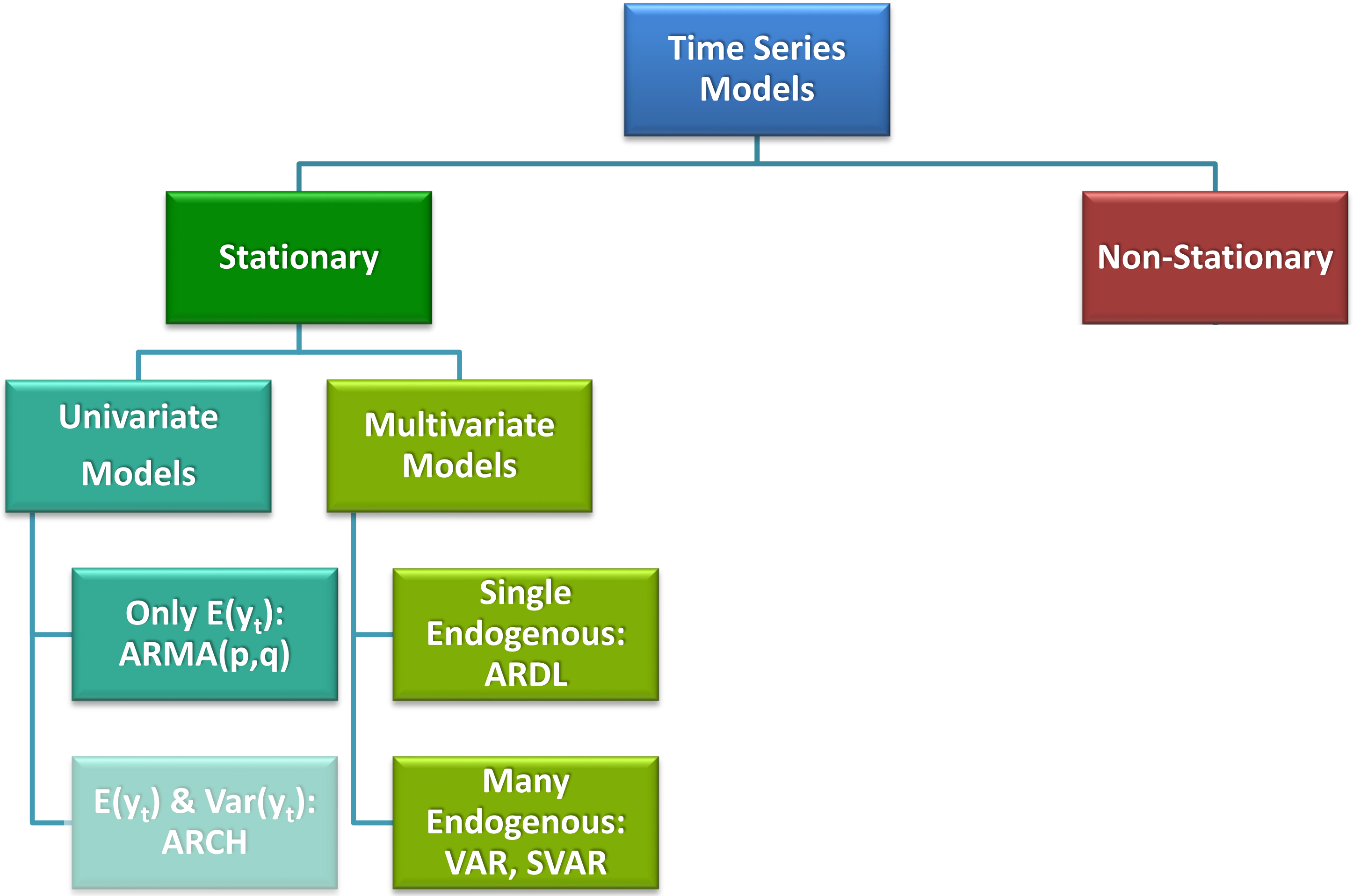


Time Series Models
Stationary
Non-Stationary
white noise
$y_t = \varepsilon_t$
$\varepsilon_t \sim N(0, 0.001)$
stationary AR(2)
$y_t = 0.5y_{t-1} + 0.3y_{t-2} + \varepsilon_t$
$\varepsilon_t \sim N(0, 0.001)$
random walk without drift
$y_t = y_{t-1} + \varepsilon_t$
$\varepsilon_t \sim N(0, 0.001)$
random walk with drift
$y_t = 0.04 + y_{t-1} + \varepsilon_t$
$\varepsilon_t \sim N(0, 0.001)$
time
time

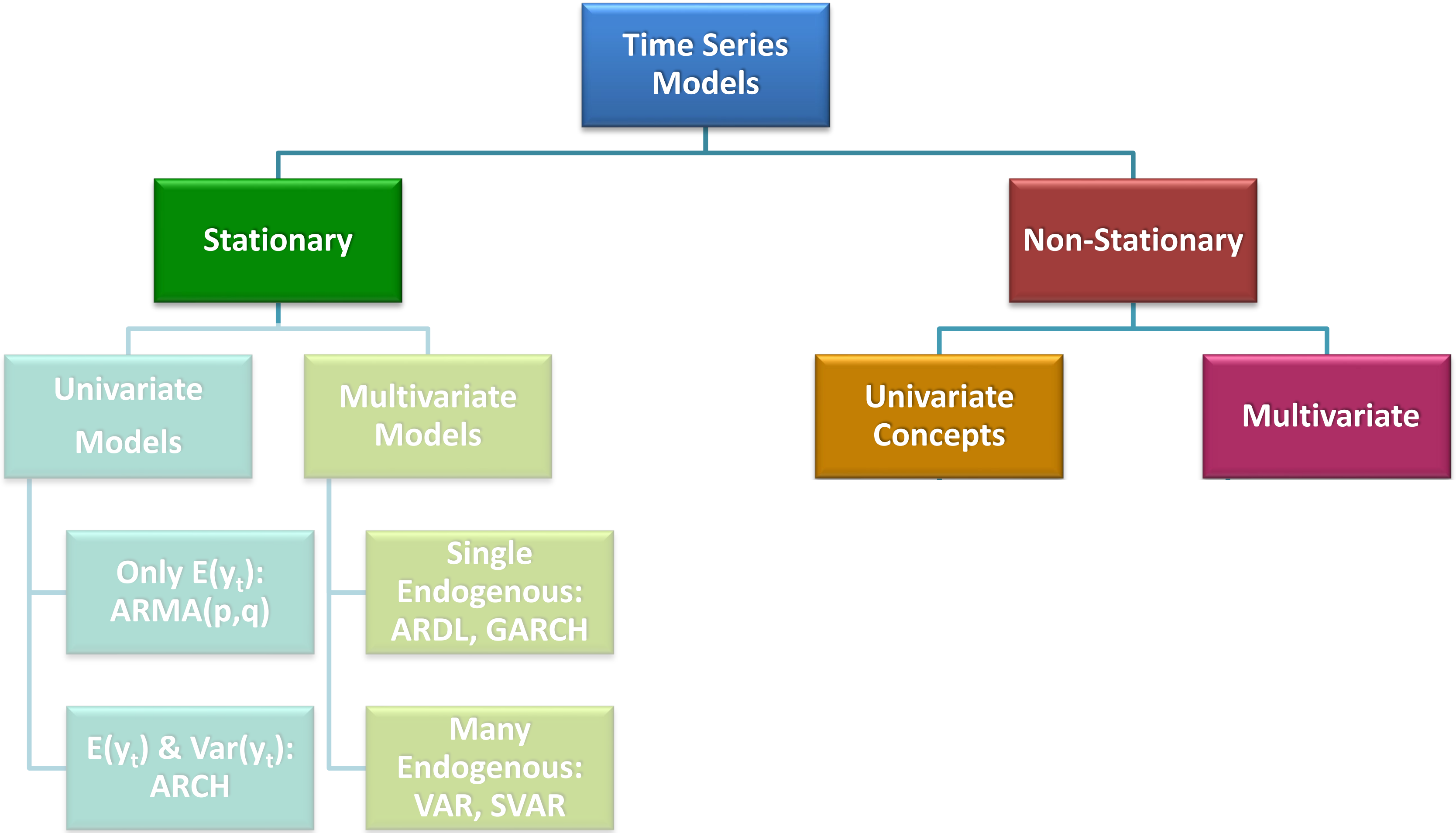

Time Series Models
Stationary
Non-Stationary
Univariate Models
Multivariate Models


Time Series Models
Stationary
Non-Stationary
Univariate Models
Multivariate Models
Only $E(y_t)$: ARMA(p,q)
$E(y_t)$ & $Var(y_t)$: ARCH

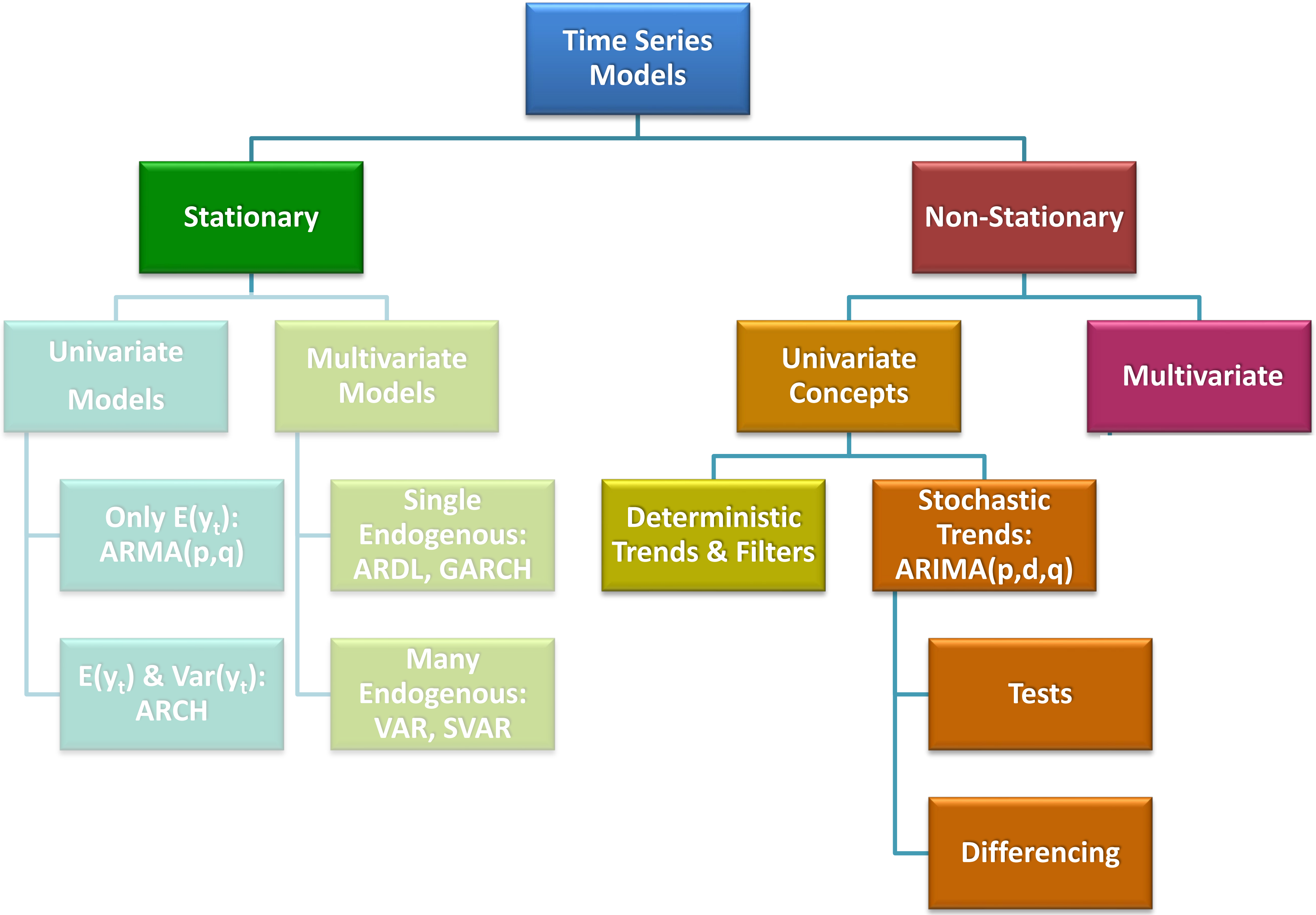

Time Series Models
Stationary
Non-Stationary
Univariate Models
Multivariate Models
Only $E(y_t)$: ARMA(p,q)
$E(y_t)$ & $Var(y_t)$: ARCH
Single Endogenous: ARDL
Many Endogenous: VAR, SVAR

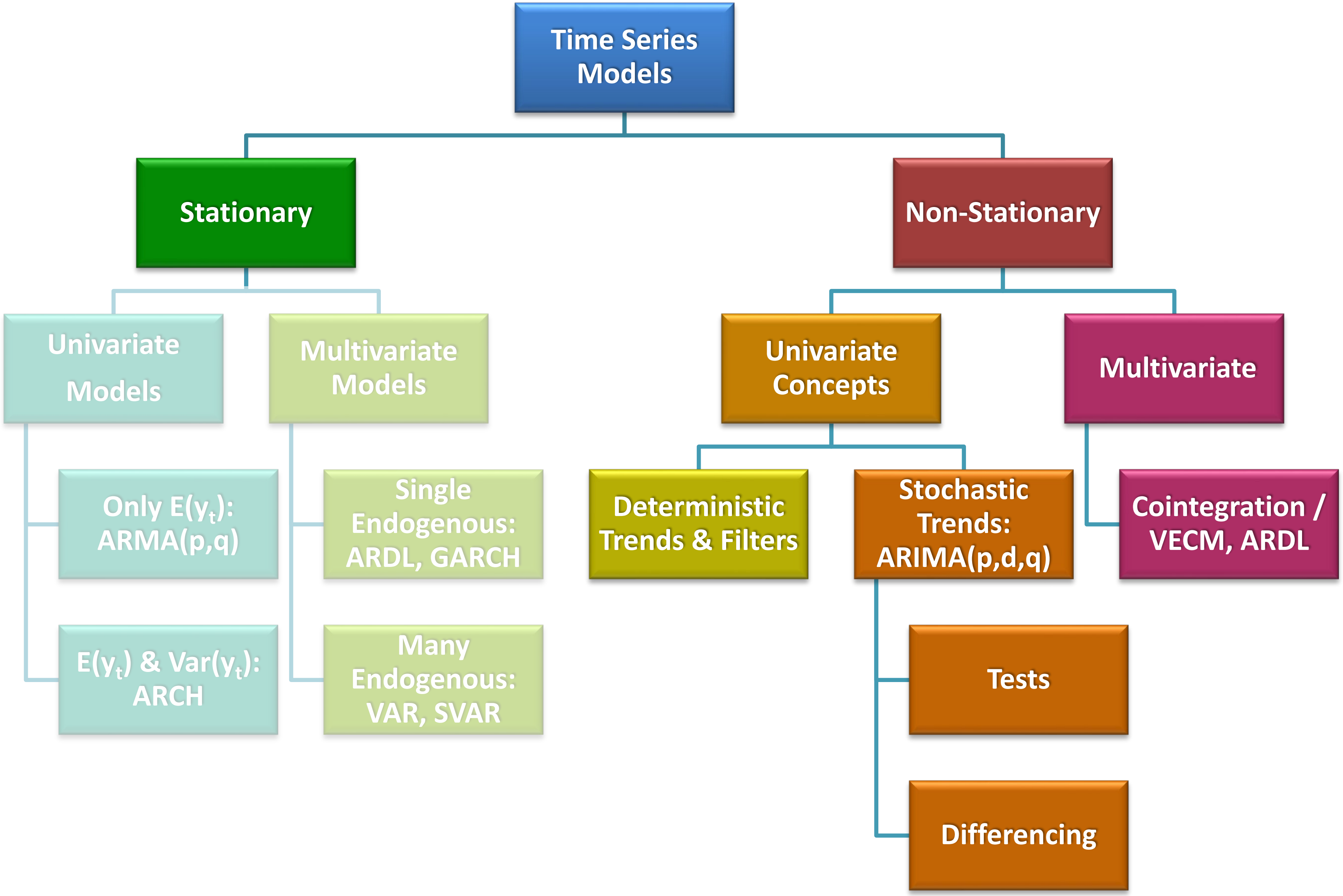


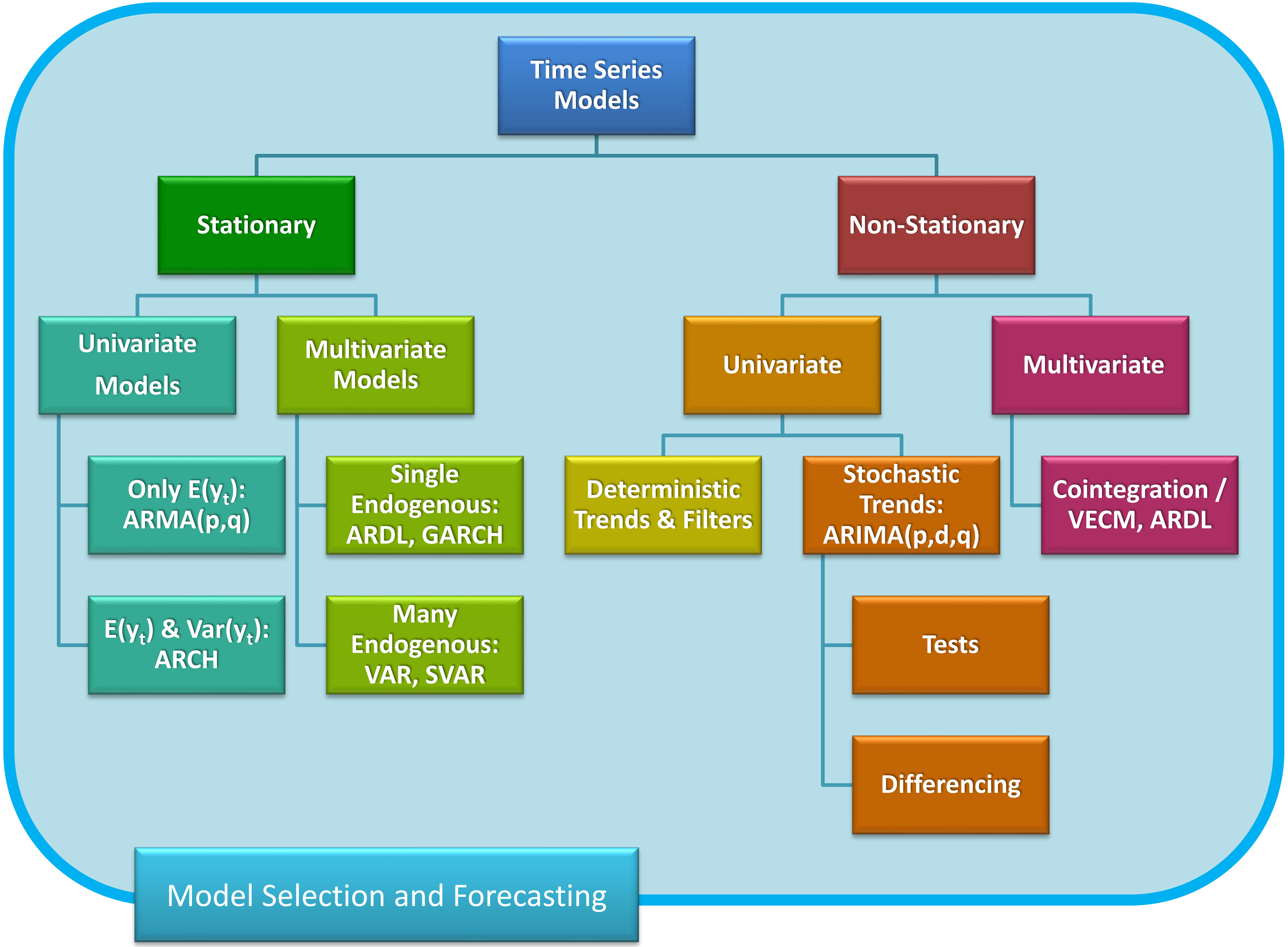
Time Series Models
Stationary
Non-Stationary
Univariate Models
Multivariate Models
Only $E(y_t)$: ARMA(p,q)
$E(y_t)$ & $Var(y_t)$: ARCH
Single Endogenous: ARDL
Many Endogenous: VAR, SVAR



Time Series Models
Stationary
Non-Stationary
Univariate Models
Multivariate Models
Univariate Concepts
Multivariate
Only $E(y_t)$: ARMA(p,q)
$E(y_t)$ & $Var(y_t)$: ARCH
Single Endogenous: ARDL, GARCH
Many Endogenous: VAR, SVAR


Time Series Models
Stationary
Non-Stationary
Univariate Models
Multivariate Models
Only $E(y_t)$: ARMA(p,q)
$E(y_t)$ & $Var(y_t)$: ARCH
Single Endogenous: ARDL, GARCH
Many Endogenous: VAR, SVAR
Univariate Concepts
Multivariate
Deterministic Trends & Filters
Stochastic Trends: ARIMA(p,d,q)
Tests
Differencing


Time Series Models
Stationary
Non-Stationary
Univariate Models
Multivariate Models
Only $E(y_t)$: ARMA(p,q)
$E(y_t)$ & $Var(y_t)$: ARCH
Single Endogenous: ARDL, GARCH
Many Endogenous: VAR, SVAR
Univariate Concepts
Multivariate
Deterministic Trends & Filters
Stochastic Trends: ARIMA(p,d,q)
Tests
Differencing
Cointegration / VECM, ARDL



Time Series Models
Stationary
Non-Stationary
Univariate Models
Multivariate Models
Univariate
Multivariate
Only $E(y_t)$: ARMA(p,q)
$E(y_t)$ & $Var(y_t)$: ARCH
Single Endogenous: ARDL, GARCH
Many Endogenous: VAR, SVAR
Deterministic Trends & Filters
Stochastic Trends: ARIMA(p,d,q)
Tests
Differencing
Cointegration / VECM, ARDL
Model Selection and Forecasting

Today: Difference equations

- In order to effectively and accurately use time series methods, one must be able to manipulate and understand the fundamental mathematical features of difference equations
- Consider the Random Walk hypothesis on stock prices in an efficient market
  - Predictable movements in price imply arbitrage (profit) opportunities
  - Thus an efficient market predicts that stock prices must follow a random walk:

$$y_{t+1} = y_t + \varepsilon_{t+1} \text{ with } \mathrm{E}[\varepsilon_{t+1}] = 0$$

  - We can test this hypothesis with the general difference equation model:

$$y_{t+1} = a_0 + a_1 y_t + \varepsilon_{t+1}$$

    - The Random walk hypothesis is $a_0 = 0$ and $a_1 = 1$
  - A key issues that we need to determine when a difference equation has a solution, as this determines the properties of the estimators

More generally

- Key feature of time series is persistence:
  - persistent process $\equiv$ auto-correlated process
  - high value today $\Rightarrow$ high likelihood of high value tomorrow
  - y_t depends on y_{t-1}, y_{t-2},…
  - Goal of time series analysis is to find a valid model of the equation of motion that describes the data generating process of some variable of interest – this is expressed as a difference equation
  - "Time-series econometrics is concerned with the estimation of difference equations with stochastic components" - Enders
- Any macro/dynamic problem
  - Data observed in discrete intervals
  - Difference equations are the discrete analogues of differential equations

Warning | Suggestion

- Work is HEAVILY dependent on definitions, terminology
  - Interrelated, sometimes poorly named
  - When you encounter a definition, focus on understanding what it means, and learn the name(s) used for it
- Definitions must be used precisely
  - They contain only what they say
  - They need not conform to “common sense”
- Some irregularity in some places:
  - Time series terminology is sometimes less precise than mathematical definitions – unsettled arguments
  - A valuable skill is to carefully internalize what each new author means with a specific concept
- I’m going to attempt to be careful but won’t manage to be careful enough.
  - **Stop me** and get me to be more careful as soon as you get lost/miss something, especially if it is a definition

n^{th} order Linear Difference Equations

- General form

$$y_t = a_1 y_{t-1} + a_2 y_{t-2} + \cdots + a_n y_{t-n} + x_t$$

  – Where x_t is called a "forcing process" and can be anything for now
    - deterministic and/or stochastic
    - For intuition: think of x_t as a normally distributed error with some mean

- Where's the "difference" then?
  – definition of the difference operator: $\Delta y_t = y_t - y_{t-1}$

n^{th} order Linear Difference Equations

- In general form:

$$y_t = a_1 y_{t-1} + a_2 y_{t-2} + \cdots + a_n y_{t-n} + x_t$$

- Subtract y_{t-1} from both sides:

$$y_t - y_{t-1} = a_1 y_{t-1} - y_{t-1} + a_2 y_{t-2} + \cdots + a_n y_{t-n} + x_t$$

$$\Delta y_t = (a_1 - 1) y_{t-1} + a_2 y_{t-2} + \cdots + a_n y_{t-n} + x_t$$

n^{th} order Linear Difference Equations

- In general form:

$$y_t = a_1 y_{t-1} + a_2 y_{t-2} + \cdots + a_n y_{t-n} + x_t$$

- Subtract y_{t-1} from both sides:

$$y_t - y_{t-1} = a_1 y_{t-1} - y_{t-1} + a_2 y_{t-2} + \cdots + a_n y_{t-n} + x_t$$

$$\underbrace{\Delta y_t}_{\text{The change from one period to the next}} \underset{\text{is}}{=} \underbrace{(1 - a_1) y_{t-1} + a_2 y_{t-2} + \cdots + a_n y_{t-n}}_{\text{this linear function the level of past values}} + \underbrace{x_t}_{\text{and other things that happen in this period}}$$

Discrete vs Continuous Time models

- In Honours Macro, you encountered models in both discrete and continuous time.
  - How are they related?
  - Consider the relationship between the difference operator and the differential operator

$$\Delta y_t = \frac{y_t - y_{t-1}}{1}$$

Discrete vs Continuous Time models

- In Honours Macro, you encountered models in both discrete and continuous time.
  - How are they related?
  - Consider the relationship between the difference operator and the differential operator

$$\Delta y_t = \frac{y_t - y_{t-1}}{1}$$

Size of the change in y

Size of the time step in which the change occurs

  - Now let the size of time step be a variable h

Discrete vs Continuous Time models

– Now let the size of time step be a variable h

$$\frac{y_t - y_{t-h}}{h}$$

Size of the change in y

Size of the time step in which the change occurs

– What is the limit?

$$\lim_{h \to 0} \frac{y_t - y_{t-h}}{h}$$

Discrete vs Continuous Time models

– Now let the size of time step be a variable h

$$\frac{y_t - y_{t-h}}{h}$$

Size of the change in y

Size of the time step in which the change occurs

$$\lim_{h \to 0} \frac{y_t - y_{t-h}}{h} = \frac{\partial y_t}{\partial t} = \dot{y}_t$$

– The time derivative of y_t in continuous time

- The instantaneous rate of change

n^{th} order Linear Difference Equations

- What restrictions are we imposing?
  - I.e. what are we not looking at?

n^{th} order Linear Difference Equations

- **Linearity**
  - Strong restriction, but remains majority of the core toolkit
  - Used for estimation of linearized approximations non-linear models
    - around an equilibrium
- Non-linear approaches
  - Quadratic approximations becoming more feasible (cross variables, not cross periods)
    - Higher order?
  - Non-linearity usually takes different forms
    - Asymmetric responses to positive/negative shocks
    - Different behaviours at different ranges of variables

Linear Difference Equations

- Constant Parameters
  - With respect to time, endogenous and forcing variables
    - Empirical Advances: Time varying coefficients, Regime shifts
    - Fall under “Non-linear models” – Advanced course
- Forcing Process
  - Current and/or Lagged values of **other** variables
  - Deterministic functions of time
  - Fully specified stochastic processes
    - (self contained, not function of anything else)
  - Different choices lead to different models
  - Keep it simple for now:
    - Forcing process is deterministic or “white noise”

$$y_{t+1} = a_0 + a_1 y_t + x_{t+1}$$

Linear Difference Equations

- Keep it simple for now:
  - Forcing process is either deterministic or "white noise"

- ε_t is **white noise** if:

$$\mathbb{E}(\varepsilon_t) = 0$$

$$\mathbb{E}(\varepsilon_t^2) = \sigma^2$$

$$\mathbb{E}(\varepsilon_t \varepsilon_s) = 0 \;\; \forall \;\; t \neq s$$

Independent of any other stochastic process

Defining the *solution of a difference equation*

- For some difference equation describing how a variable y_t depends on its past and random shocks:
- E.g.

$$y_t = a_1 y_{t-1} + a_2 y_{t-2} + \cdots + a_n y_{t-n} + x_t$$

- We define a solution to be
  - a function that
    - Always takes on a finite value
    - Contains only exogenous parts
    - Satisfies the difference equation for all t

Defining the *solution of a difference equation*

- $y_t = f(.) < \infty$
- Such that
  - $f(.)$ is a function of only exogenous variables
    - Forcing process (ε_t)
    - Time
    - Initial conditions
  -

Defining the *solution of a difference equation*

- $y_t = f(.)$
- Such that
  - $f(.)$ is a function of only
    - Forcing process (ε_t)
    - Time
    - Initial conditions
  - $f(.)$ satisfies the difference equation for all y_t and t

- E.g. $\triangle y_t = 2$ or equivalently: $y_t = y_{t-1} + 2$
  - Solution: $y_s = 2s + c$ for any arbitrary c
  - Check: does this work for y_{t-1}? Replace s with $t - 1$

Defining the *solution of a difference equation*

- $y_t = f(.)$
- Such that
  - $f(.)$ is a function of only
    - Forcing process (ε_t)
    - Time
    - Initial conditions
  - $f(.)$ satisfies the difference equation for all y_t and t

$$y_{t-1} = 2(t-1) + c$$

- E.g. $\triangle y_t = 2$ or equivalently: $y_t = y_{t-1} + 2$
  - Solution: $y_s = 2s + c$ for any arbitrary c
  - Check: does this work for y_{t-1}? Replace s with $t - 1$

Defining the *solution of a difference equation*

- $y_t = f(.)$
- Such that
  - $f(.)$ is a function of only
    - Forcing process (ε_t)
    - Time
    - Initial conditions
  - $f(.)$ satisfies the difference equation for all y_t and t

$$y_{t-1} = 2(t-1) + c$$

$$y_t = y_{t-1} + 2$$

- E.g. $\triangle y_t = 2$ or equivalently: $y_t = y_{t-1} + 2$
  - Solution: $y_s = 2s + c$ for any arbitrary c
  - Check: does this work for y_{t-1}? Replace s with $t - 1$

Defining the *solution of a difference equation*

- $y_t = f(.)$
- Such that
  - $f(.)$ is a function of only
    - Forcing process (ε_t)
    - Time
    - Initial conditions
  - $f(.)$ satisfies the difference equation for all y_t and t

$$y_{t-1} = 2(t-1) + c$$

$$\begin{aligned} y_t &= y_{t-1} + 2 \\ &= 2(t-1) + c + 2 \end{aligned}$$

- E.g. $\triangle y_t = 2$ or equivalently: $y_t = y_{t-1} + 2$
  - Solution: $y_s = 2s + c$ for any arbitrary c
  - Check: does this work for y_{t-1}? Replace s with $t-1$

Defining the *solution of a difference equation*

- $y_t = f(.)$
- Such that
  - $f(.)$ is a function of only
    - Forcing process (ε_t)
    - Time
    - Initial conditions
  - $f(.)$ satisfies the difference equation for all y_t and t

$$y_{t-1} = 2(t-1) + c$$

$$\begin{aligned} y_t &= y_{t-1} + 2 \\ &= 2(t-1) + c + 2 \\ &= 2t + c \end{aligned}$$

- E.g. $\triangle y_t = 2$ or equivalently: $y_t = y_{t-1} + 2$
  - Solution: $y_s = 2s + c$ for any arbitrary c
  - Check: does this work for y_{t-1}? Replace s with $t - 1$

Solution Methods

Option 1: **Iterative**

- Intuitive
- Cumbersome – but worth doing once or twice

- With initial value
  - Direct; either backward* or forward
  - n^{th} order difference equation requires n initial conditions
- Without initial value
  - Requirements for existence of solution
  - Uniqueness – “Transversality Conditions”
- Application to a first order linear difference equation

Solution by backwards iteration

Consider the first order difference equation that characterizes some process:

$$y_t = a_0 + a_1 y_{t-1} + \varepsilon_t$$

Since it fully describes the process of interest, it must hold in all periods:

$$y_{t-1} = a_0 + a_1 y_{t-2} + \varepsilon_{t-1}$$

Plugging in for y_{t-1} in the original equation:

$$\begin{aligned} y_t &= a_0 + a_1(a_0 + a_1 y_{t-2} + \varepsilon_{t-1}) + \varepsilon_t \\ &= a_0 + a_1 a_0 + a_1^2 y_{t-2} + \varepsilon_t + a_1 \varepsilon_{t-1} \end{aligned}$$

Solution by backwards iteration

Similarly:

$$y_{t-2} = a_0 + a_1 y_{t-3} + \varepsilon_{t-2}$$

Plugging in for y_{t-2}:

$$\begin{aligned} y_t &= a_0 + a_1(a_0 + a_1 y_{t-2} + \varepsilon_{t-1}) + \varepsilon_t \\ &= a_0 + a_1 a_0 + a_1^2 y_{t-2} + \varepsilon_t + a_1 \varepsilon_{t-1} \\ &= a_0 + a_1 a_0 + a_1^2(a_0 + a_1 y_{t-3} + \varepsilon_{t-2}) + \varepsilon_t + a_1 \varepsilon_{t-1} \\ &= a_0 + a_1 a_0 + a_1^2 a_0 + a_1^3 y_{t-3} + \varepsilon_t + a_1 \varepsilon_{t-1} + a_1^2 \varepsilon_{t-2} \end{aligned}$$

Solution by backwards iteration

Recognize patterns:

$$y_t = a_0 + a_1 a_0 + a_1^2 a_0 + a_1^3 y_{t-3} + \varepsilon_t + a_1 \varepsilon_{t-1} + a_1^2 \varepsilon_{t-2}$$

$$= (1 + a_1 + a_1^2) a_0 + 1\varepsilon_t + a_1 \varepsilon_{t-1} + a_1^2 \varepsilon_{t-2} + a_1^3 y_{t-3}$$

$$= a_0 \sum_{i=0}^{2} a_1^i + \sum_{i=0}^{2} a_1^i \, \varepsilon_{t-i} + a_1^3 y_{t-3}$$

Solution by backwards iteration

Initial conditions

- suppose we know the world started at $t = 0$ at some exogenously given y_0, then we can continue with the backwards iteration until we get to the first period:

$$y_t = a_0 \sum_{i=0}^{t-1} a_1^i + \sum_{i=0}^{t-1} a_1^i \, \varepsilon_{t-i} + a_1^t y_0$$

Solution by backwards iteration

$$y_t = a_0 \sum_{i=0}^{t-1} a_1^i + \sum_{i=0}^{t-1} a_1^i \, \varepsilon_{t-i} + a_1^t y_0$$

Convince yourself that this is a valid solution:

- the RHS is only a function of exogenous variables
- is finite for all finite parameters/conditions.
- Homework: go test that it satisfies the original difference equation

We can compute the value of y_s at any point in time for a given sequence of ε_t .

Solution by backwards iteration

$$y_t = a_0 \sum_{i=0}^{t-1} a_1^i + \sum_{i=0}^{t-1} a_1^i \, \varepsilon_{t-i} + a_1^t y_0$$

Go one further period into the past:

$$y_t = a_0 \sum_{i=0}^{t} a_1^i + \sum_{i=0}^{t} a_1^i \, \varepsilon_{t-i} + a_1^{t+1} y_{-1}$$

And then m more periods:

$$y_t = a_0 \sum_{i=0}^{t+m} a_1^i + \sum_{i=0}^{t+m} a_1^i \, \varepsilon_{t-i} + a_1^{t+m+1} y_{-m-1}$$

Solution by backwards iteration

$$y_t = a_0 \sum_{i=0}^{t+m} a_1^i + \sum_{i=0}^{t+m} a_1^i \, \varepsilon_{t-i} + a_1^{t+m+1} y_{-m-1}$$

Now consider the situation as $m \to \infty$.
When will the above equation be **defined**? (i.e. give a **finite** answer)

- We always assume $|\varepsilon_t| < \infty$, and $\left|\lim_{m\to\infty} y_{-m-1}\right| < \infty$. (why?)

Solution by backwards iteration

$$y_t = a_0 \sum_{i=0}^{t+m} a_1^i + \sum_{i=0}^{t+m} a_1^i \, \varepsilon_{t-i} + a_1^{t+m+1} y_{-m-1}$$

Now consider the situation as $m \to \infty$.
When will the above equation be defined? (i.e. give a finite answer)

- We always assume $|\varepsilon_t| < \infty$, and $\left|\lim_{m\to\infty} y_{-m-1}\right| < \infty$. (why?)
- $\left|\lim_{m\to\infty} a_0 \sum_{i=0}^{t+m} a_1^i\right| < \infty$ requires?
- $\left|\lim_{m\to\infty} \sum_{i=0}^{t+m} a_1^i \, \varepsilon_{t-i}\right| < \infty$ requires?
- $\left|\lim_{m\to\infty} a_1^{t+m+1} y_{-m-1}\right| < \infty$ requires?

Solution by backwards iteration

$$y_t = a_0 \sum_{i=0}^{t+m} a_1^i + \sum_{i=0}^{t+m} a_1^i \, \varepsilon_{t-i} + a_1^{t+m+1} y_{-m-1}$$

Now consider the situation as $m \to \infty$.
When will the above equation be defined? (i.e. give a finite answer)

- We always assume $|\varepsilon_t| < \infty$, and $\left|\lim_{m\to\infty} y_{-m-1}\right| < \infty$. (why?)
- $\left|\lim_{m\to\infty} a_0 \sum_{i=0}^{t+m} a_1^i\right| < \infty$ requires? $|a_1| < 1$
- $\left|\lim_{m\to\infty} \sum_{i=0}^{t+m} a_1^i \, \varepsilon_{t-i}\right| < \infty$ requires? $|a_1| < 1$
- $\left|\lim_{m\to\infty} a_1^{t+m+1} y_{-m-1}\right| < \infty$ requires? $|a_1| \leq 1$

Convergent geometric series

- If $|a_1| < 1$, the first term forms a **convergent geometric series**

$$\sum_{i=0}^{m} a_1^i = 1 + a_1 + a_1^2 + \cdots + a_1^m$$

$$\lim_{m\to\infty} \sum_{i=0}^{m} a_1^i = 1 + a_1 + a_1^2 + \cdots$$

- Suppose the limit x exists:

$$x = 1 + a_1 + a_1^2 + \cdots$$

$$a_1 x = \quad a_1 + a_1^2 + \cdots$$

$$x - a_1 x = 1$$

$$x = \frac{1}{1 - a_1}$$

Solution by backwards iteration

$$y_t = a_0 \sum_{i=0}^{t+m} a_1^i + \sum_{i=0}^{t+m} a_1^i \, \varepsilon_{t-i} + a_1^{t+m+1} y_{-m-1}$$

If $|a_1| < 1$

$$\lim_{m\to\infty} a_0 \sum_{i=0}^{t+m} a_1^i = \frac{a_0}{1 - a_1}$$

$$\lim_{m\to\infty} \sum_{i=0}^{t+m} a_1^i \, \varepsilon_{t-i} = \sum_{i=0}^{\infty} a_1^i \, \varepsilon_{t-i} < \infty$$

$$\lim_{m\to\infty} a_1^{t+m+1} y_{-m-1} = 0$$

Solution by backwards iteration

We conclude that, if and only if (iff) $|a_1| < 1$, the following is a solution to the difference equation:

$$y_t = \frac{a_0}{1 - a_1} + \sum_{i=0}^{\infty} a_1^i \, \varepsilon_{t-i}$$

Final proof: show that this solution satisfies the original difference equation. (homework)

Note: this is *a* solution, but it is not unique:

$$y_t = A a_1^t + \frac{a_0}{1 - a_1} + \sum_{i=0}^{\infty} a_1^i \, \varepsilon_{t-i}$$

Is also a solution for *any* constant A. To pin down a unique solution we need either an initial or terminal condition, called a *transversality* condition. This forms a key component of infinite horizon models.

Solution Methods

Option 1: Iterative

Option 2: **General Solution Method**

- Algorithm
  - not easily teachable but learnable (and lookup-able)
- Method of undetermined coefficients

Solution Methods

Option 1: Iterative

Option 2: **General Solution Method**

- Algorithm not easily teachable but learnable (and lookup-able)
- Applicable to linear difference equation of any order
  - Iterative version quickly becomes too messy

Why do we want to solve them?

- Different time series behaviour is generated by different difference equations.
- One part of estimating a model is to “know” what type of behaviour is typically observed from each type of difference equation (different coefficient values, different orders, different forcing processes) so that we can use the correct model in the estimation.
- The solutions tell us exactly how the variable will evolve from any starting point
  - Will the process return to its mean or wander about or explode?
  - If it is mean reverting, does it do so directly or in a oscillatory pattern?

Lag operator

- Process for finding General Solutions
  - cumbersome,
  - requires experience for different forcing processes
- Stability properties (our key interest)
  - Usually only a function of homogenous part
- Lag operator allows a much simpler way to find a general rule for determining the stability properties of a process

Lag operator

Definition:

- Let the lag operator be denoted L
- The definition of the lag operator is simple:

$$Ly_t = y_{t-1}$$

So:

$$Ly_{t-1} = y_{t-2}$$
$$Ly_{t-1} = L(Ly_t) = L^2 y_t$$

Lag operator

- Properties:
  - Lag of a constant
  - Associative – higher order lags
  - Distributive
  - Negative power
  - Zero power

Lag operator

- Properties:
  - Lag of a constant

$$Lc = c$$

  - Associative – higher order lags
  - Distributive
  - Negative power
  - Zero power

Lag operator

- Properties:
  - Lag of a constant

$$Lc = c$$

  - Associative – higher order lags

$$L^i L^j y_t = L^i \left(L^j y_t\right) = L^i y_{t-j} = y_{t-j-i} = L^{i+j} y_t$$

  - Distributive
  - Negative power
  - Zero power

Lag operator

- Properties:
  - Lag of a constant

$$Lc = c$$

  - Associative – higher order lags

$$L^i L^j y_t = L^i \left(L^j y_t\right) = L^i y_{t-j} = y_{t-j-i} = L^{i+j} y_t$$

  - Distributive

$$\left(L^i + L^j\right) y_t = L^i y_t + L^j y_t = y_{t-i} + y_{t-j}$$

  - Negative power

  - Zero power

Lag operator

- Properties:
  - Lag of a constant

$$Lc = c$$

  - Associative – higher order lags

$$L^i L^j y_t = L^i \left(L^j y_t\right) = L^i y_{t-j} = y_{t-j-i} = L^{i+j} y_t$$

  - Distributive

$$\left(L^i + L^j\right) y_t = L^i y_t + L^j y_t = y_{t-i} + y_{t-j}$$

  - Negative power

$$L^{-j} = y_{t-(-j)} = y_{t+j}$$

  - Zero power

Lag operator

- Properties:
  - Lag of a constant

$$Lc = c$$

  - Associative – higher order lags

$$L^i L^j y_t = L^i \left(L^j y_t\right) = L^i y_{t-j} = y_{t-j-i} = L^{i+j} y_t$$

  - Distributive

$$\left(L^i + L^j\right) y_t = L^i y_t + L^j y_t = y_{t-i} + y_{t-j}$$

  - Negative power

$$L^{-j} = y_{t-(-j)} = y_{t+j}$$

  - Zero power

$$L^0 y_t = y_t$$

Using Lag Operator to rewrite n^{th} order Linear Difference Equation

- Consider a typical n^{th} order Linear Difference Equation

$$y_t = a_0 + a_1 y_{t-1} + a_2 y_{t-2} + \cdots + a_n y_{t-n} + \varepsilon_t$$

- Using the properties of the lag operator:

Using Lag Operator to rewrite n^{th} order Linear Difference Equation

- Consider a typical n^{th} order Linear Difference Equation

$$y_t = a_0 + a_1 y_{t-1} + a_2 y_{t-2} + \cdots + a_n y_{t-n} + \varepsilon_t$$

- Using the properties of the lag operator:

$$y_t = a_0 + a_1 L y_t + a_2 L^2 y_t + \cdots + a_n L^n y_t + \varepsilon_t$$

Using Lag Operator to rewrite n^{th} order Linear Difference Equation

- Consider a typical n^{th} order Linear Difference Equation

$$y_t = a_0 + a_1 y_{t-1} + a_2 y_{t-2} + \cdots + a_n y_{t-n} + \varepsilon_t$$

- Using the properties of the lag operator:

$$y_t = a_0 + a_1 L y_t + a_2 L^2 y_t + \cdots + a_n L^n y_t + \varepsilon_t$$

$$= a_0 + (a_1 L + a_2 L^2 + \cdots + a_n L^n) y_t + \varepsilon_t$$

Using Lag Operator to rewrite n^{th} order Linear Difference Equation

- Consider a typical n^{th} order Linear Difference Equation

$$y_t = a_0 + a_1 y_{t-1} + a_2 y_{t-2} + \cdots + a_n y_{t-n} + \varepsilon_t$$

- Using the properties of the lag operator:

$$y_t = a_0 + a_1 L y_t + a_2 L^2 y_t + \cdots + a_n L^n y_t + \varepsilon_t$$

$$= a_0 + (a_1 L + a_2 L^2 + \cdots + a_n L^n) y_t + \varepsilon_t$$

- Since everything is now in terms of y_t, we can collect terms on the LHS:

Using Lag Operator to rewrite n^{th} order Linear Difference Equation

- Consider a typical n^{th} order Linear Difference Equation

$$y_t = a_0 + a_1 y_{t-1} + a_2 y_{t-2} + \cdots + a_n y_{t-n} + \varepsilon_t$$

- Using the properties of the lag operator:

$$y_t = a_0 + a_1 L y_t + a_2 L^2 y_t + \cdots + a_n L^n y_t + \varepsilon_t$$

$$= a_0 + (a_1 L + a_2 L^2 + \cdots + a_n L^n) y_t + \varepsilon_t$$

- Since everything is now in terms of y_t, we can collect terms on the LHS:

$$(1 - a_1 L - a_2 L^2 - \cdots - a_n L^n) y_t = a_0 + \varepsilon_t$$

Using Lag Operator to rewrite n^{th} order Linear Difference Equation

$$(1 - a_1 L - a_2 L^2 - \cdots - a_n L^n) y_t = a_0 + \varepsilon_t$$

Note:

- the RHS has only exogenous elements
- The LHS has only endogenous elements

Consider $(1 - a_1 L - a_2 L^2 - \cdots - a_n L^n)$

- This is an n^{th} order polynomial in the Lag operator
- It is often called the *inverse characteristic polynomial* of the process
- The *inverse* part is one of those conventions to just accept, and also that not everyone is consistent in this. Once you get practice in this, it will become trivial, but initially it might be confusing

Using Lag Operator to rewrite n^{th} order Linear Difference Equation

- Recalling the definition of a solution of a difference equation, we would like to write this as:

$$y_t = (1 - a_1 L - a_2 L^2 - \cdots - a_n L^n)^{-1}(a_0 + \varepsilon_t)$$

  - Since the RHS now only has parameters, constants and forcing processes as required of solution
  - However, for this inverse to exist, there are some restrictions on the values of the set of coefficients a_i
  - The rule that implies that this inverse exists is what we are after:

$$(1 - a_1 L - \cdots - a_n L^n)^{-1} < \infty \Leftrightarrow |roots(1 - a_1 L - \cdots - a_n L^n)| > 1$$

  - Let's return to the first order case that we solved iteratively to derive the rule

Review: roots of a polynomial

- The roots of an function (in our case a polynomial) are those values of the arguments that yield a zero function value:

- If the function is:

$$f(x) = x^2 - 2x + 4$$

- The roots of f are (all) values x^* such that $f(x^*) = 0$

The solution of a 1st order Linear Difference Equation

- The first order difference equation was:

$$y_t = a_0 + a_1 y_{t-1} + \varepsilon_t$$

- Using the lag operator:

$$y_t = a_0 + a_1 L y_t + \varepsilon_t$$
$$(1 - a_1 L) y_t = a_0 + \varepsilon_t$$

- We know from the iterative approach that the solutions exists when $|a_1| < 1$, so

$$y_t = (1 - a_1 L)^{-1} (a_0 + \varepsilon_t)$$

is well defined

The solution of a 1st order Linear Difference Equation

- The first order difference equation was:

$$y_t = a_0 + a_1 y_{t-1} + \varepsilon_t$$

- Using backwards iteration without an initial condition but with the constraint that $|a_1| < 1$ we showed that a solution was:

$$y_t = \frac{a_0}{1 - a_1} + \sum_{i=0}^{\infty} a_1^i \varepsilon_{t-i}$$

The solution of a 1st order Linear Difference Equation

$$y_t = \frac{a_0}{1 - a_1} + \sum_{i=0}^{\infty} a_1^i \varepsilon_{t-i}$$

- Or one step back, before we used the solution to a convergent geometric series and then the lag operator and its properties:

$$y_t = a_0 \sum_{i=0}^{\infty} a_1^i + \sum_{i=0}^{\infty} a_1^i \varepsilon_{t-i}$$

$$y_t = \sum_{i=0}^{\infty} a_1^i \left(a_0 + L^i \varepsilon_t\right)$$

$$y_t = \sum_{i=0}^{\infty} (a_1 L)^i (a_0 + \varepsilon_t)$$

The solution of a 1st order Linear Difference Equation

- Thus the solution

$$y_t = \sum_{i=0}^{\infty} (a_1 L)^i (a_0 + \varepsilon_t)$$

- Must be equivalent to

$$y_t = (1 - a_1 L)^{-1} (a_0 + \varepsilon_t)$$

- Thus we conclude that

$$(1 - a_1 L)^{-1} = \sum_{i=0}^{\infty} (a_1 L)^i < \infty$$

if and only if $|a_1| < 1$

The solution of a 1st order Linear Difference Equation

- Thus the solution

$$y_t = \sum_{i=0}^{\infty} (a_1 L)^i (a_0 + \varepsilon_t)$$

- Must be equivalent to

$$y_t = (1 - a_1 L)^{-1} (a_0 + \varepsilon_t)$$

- Thus we conclude that

$$(1 - a_1 L)^{-1} = \sum_{i=0}^{\infty} (a_1 L)^i < \infty$$

if and only if $|a_1| < 1$

Pay close attention to this result: When it exists, the inverse of a finite order polynomial in the lag operator is a convergent, *infinite* order lag polynomial

The solution of a 1st order Linear Difference Equation

- $(1 - a_1 L)$ is the inverse characteristic polynomial of the first order difference equation $y_t = a_0 + a_1 y_{t-1} + \varepsilon_t$

- The roots of a polynomial in a real/complex *variable* L are the set of values of L that sets the polynomial equal to zero.

- In this case the L^* that solves:

$$(1 - a_1 L^*) = 0$$
$$L^* = \frac{1}{a_1}$$

- Since we know that this inverse of this *Lag* polynomial exists when $|a_1| < 1$, that implies that $|L^*| > 1$

- **In words:** a difference equation has a solution (backwards, into the past) if and only if all the roots of its inverse characteristic equation are all larger than 1 in absolute value

The solution of a 2nd order Linear Difference Equation

- Consider the second order difference equation

$$y_t = a_0 + a_1 y_{t-1} + a_2 y_{t-2} + \varepsilon_t$$

- The inverse characteristic polynomial is

$$(1 - a_1 L - a_2 L^2)$$

- Can you find its roots?
- If not, can you factor the polynomial?

Yes, you can:

- Consider the basic quadratic equation:

$$ax^2 + bx + c = 0$$

- The solutions to this equation, x_1^* and x_2^*, are given by the quadratic equation:

$$x_1^* = \frac{-b + \sqrt{b^2 - 4ac}}{2a}$$

$$x_2^* = \frac{-b - \sqrt{b^2 - 4ac}}{2a}$$

- Let the factors be: $(1-\alpha_1 L)(1-\alpha_2 L)$ then we obtain: $(1-\alpha_1 L)(1-\alpha_2 L)\, y_t = a_0 + \varepsilon_t$

- Now we have effectively converted a second order process into two possibly distinct first order processes:

- Define $x_t = (1-\alpha_2 L)y_t$, so:
$$(1-\alpha_1 L)x_t = a_0 + \varepsilon_t$$

- For this process to be stable/have a solution we need $|\alpha_1| < 1$ or that $|\mathrm{root}(1-\alpha_1 L)| > 1$

- Define $z_t = (1-\alpha_1 L)y_t$, so:
$$(1-\alpha_2 L)z_t = a_0 + \varepsilon_t$$

- For this process to be stable/have a solution we need $|\alpha_2| < 1$ or that $|\mathrm{root}(1-\alpha_2 L)| > 1$

- So for the original process to be stable/have a solution, both roots of the inverse characteristic polynomial must be larger than 1 in absolute value, or equivalently **outside the unit circle**

Important terminology

- The following statements are equivalent:

The roots of the inverse characteristic polynomial are larger than 1 in absolute value

The roots of the inverse characteristic polynomial are outside the unit circle

- What is this unit circle business?

Real and Complex numbers

- In high school math, one typically only works with real numbers
- More advanced math also requires the use of complex numbers
  - Complex numbers contain the "imaginary number" $\sqrt{-1}$
  - A complex number z is typically written as:

$$z = \alpha + \beta\sqrt{-1}$$

Real and Complex numbers

- Consider the basic quadratic equation:

$$ax^2 + bx + c = 0$$

- The solutions to this equation, x_1^* and x_2^*, are given by the quadratic equation:

$$x_1^* = \frac{-b + \sqrt{b^2 - 4ac}}{2a}$$

$$x_2^* = \frac{-b - \sqrt{b^2 - 4ac}}{2a}$$

Real and Complex numbers

$$x_i^* = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

- If $b^2 - 4ac > 0$
  - x_1^* and x_2^* are real numbers, and distinct $x_1^* \neq x_2^*$
- If $b^2 - 4ac = 0$
  - x_1^* and x_2^* are real numbers, and identical $x_1^* = x_2^*$
- If $b^2 - 4ac < 0$
  - x_1^* and x_2^* are complex numbers, and in come in a distinct pair of *complex conjugate* numbers:

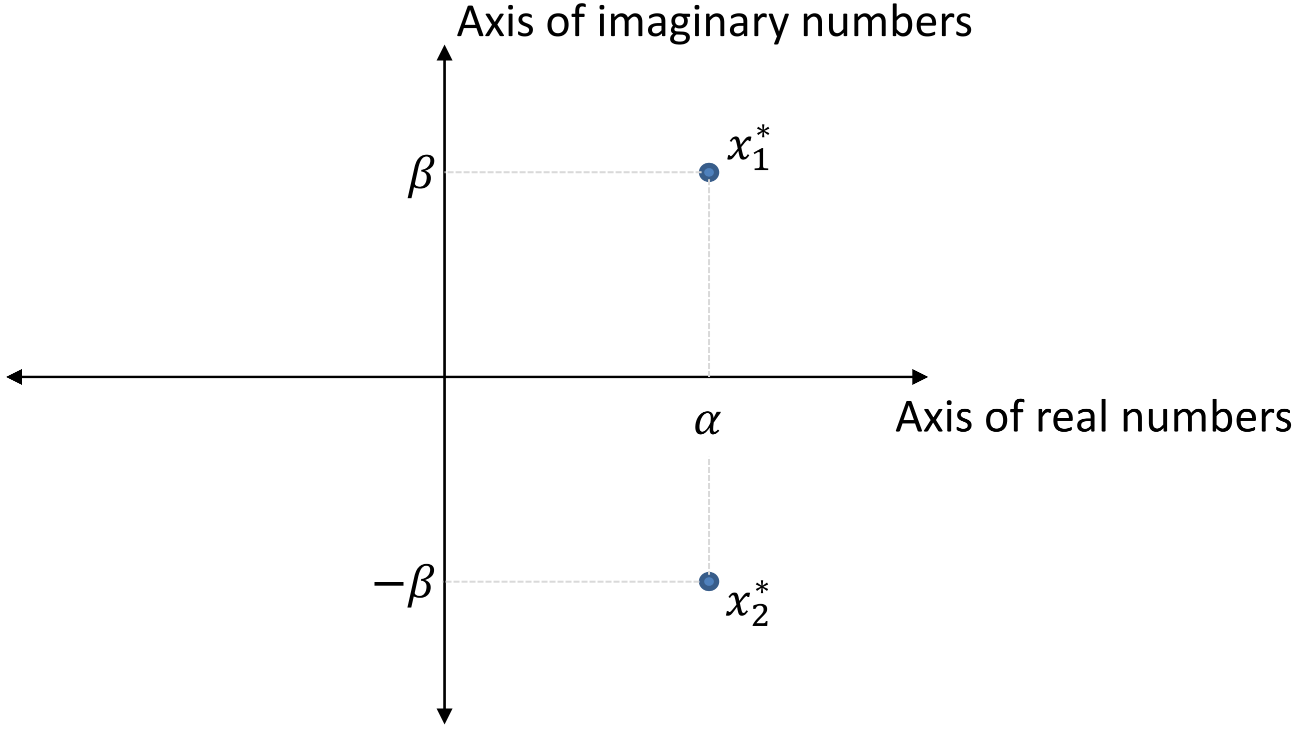
$$x_1^* = \alpha + \beta\sqrt{-1}$$
$$x_2^* = \alpha - \beta\sqrt{-1}$$

Real and Complex numbers

- We can represent a pair of complex conjugates,

$$x_1^* = \alpha + \beta\sqrt{-1}$$
$$x_2^* = \alpha - \beta\sqrt{-1}$$

graphically on the *complex plane*

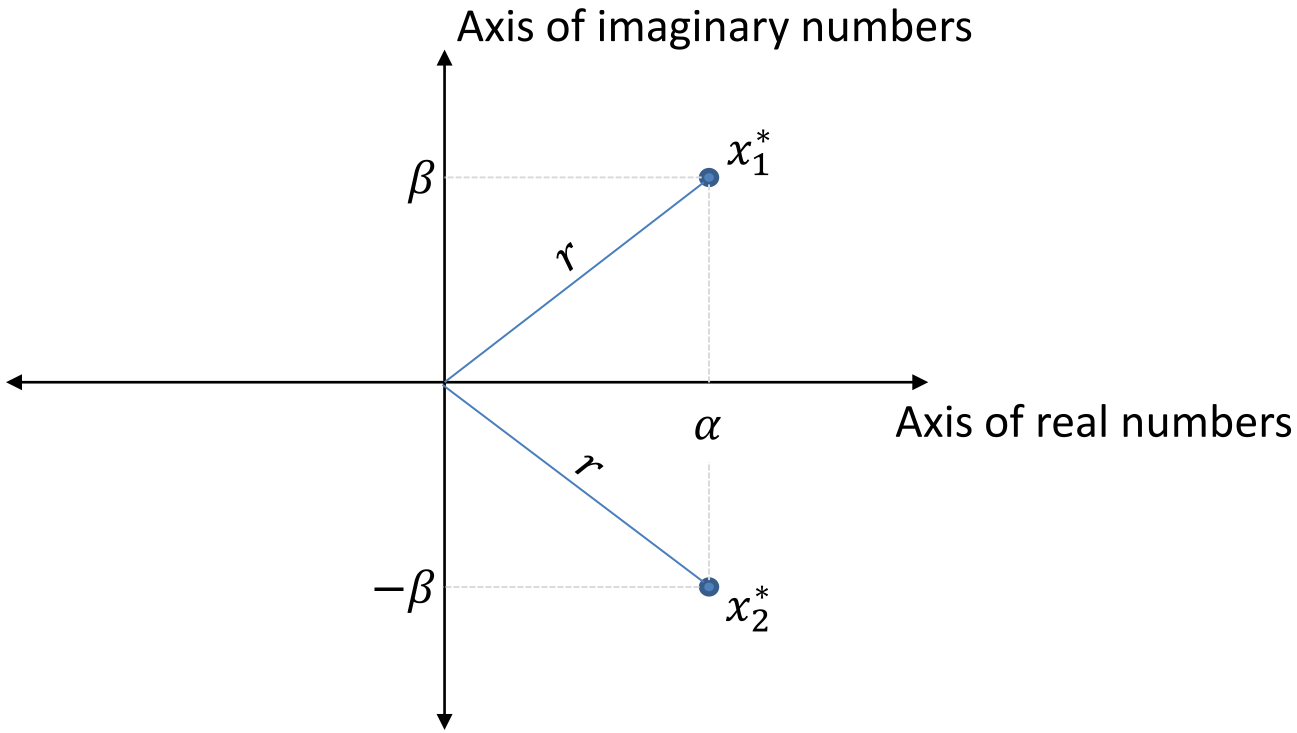


Real and Complex numbers

- The absolute values of a pair of complex numbers is given by

$$|x_1^*| = \sqrt{\alpha^2 + \beta^2} = r$$
$$|x_2^*| = \sqrt{\alpha^2 + (-\beta)^2} = r$$

Thus they are always identical, and equal to the distance to the origin of the complex plane:

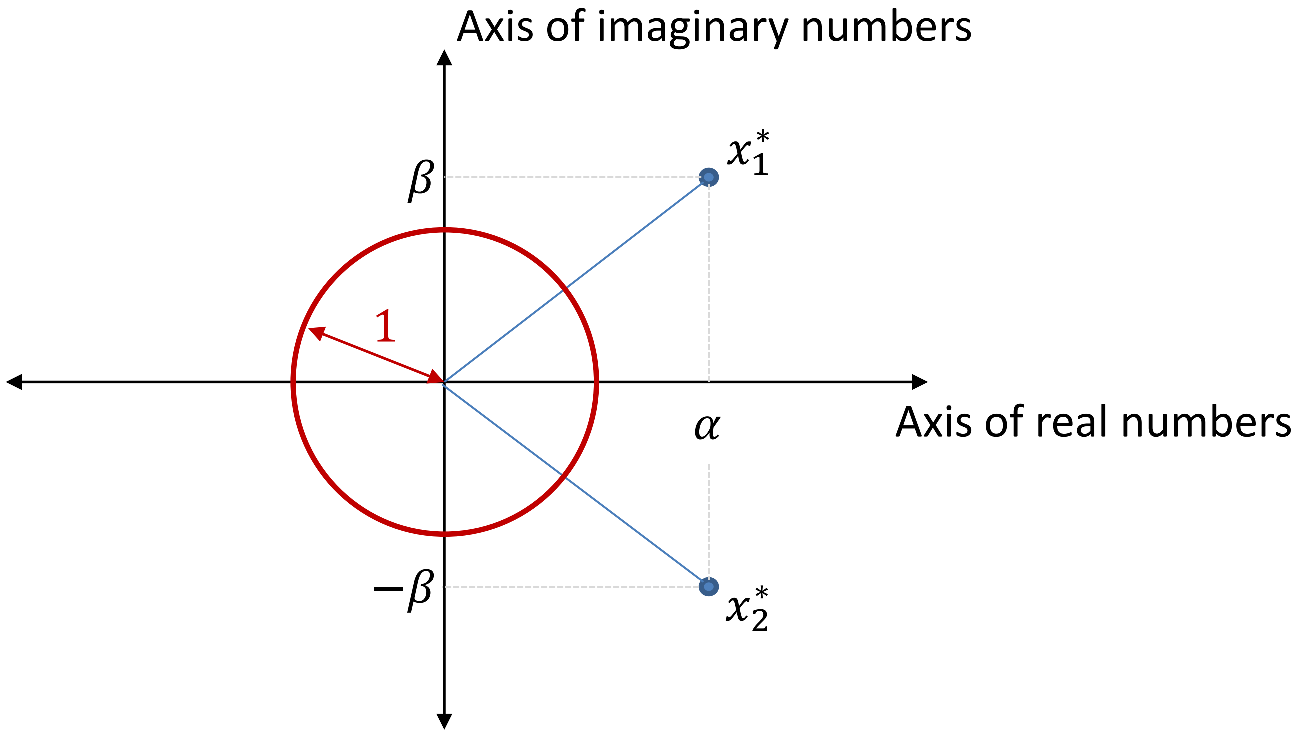


Real and Complex numbers

- If the absolute value of a pair of complex conjugate numbers is larger than 1:

$$|x_1^*| = \sqrt{\alpha^2 + \beta^2} = r > 1$$
$$|x_2^*| = \sqrt{\alpha^2 + (-\beta)^2} = r > 1$$

we say they are "outside the unit circle":



2^{nd} order Linear Difference Equation Behaviour

- In the 1^{st} order difference equation we could predict its behaviour by looking only at a_1.
- It is feasible to do the same for the 2^{nd} order difference equation in terms of a_1 and a_2,
  but it is not intuitive:

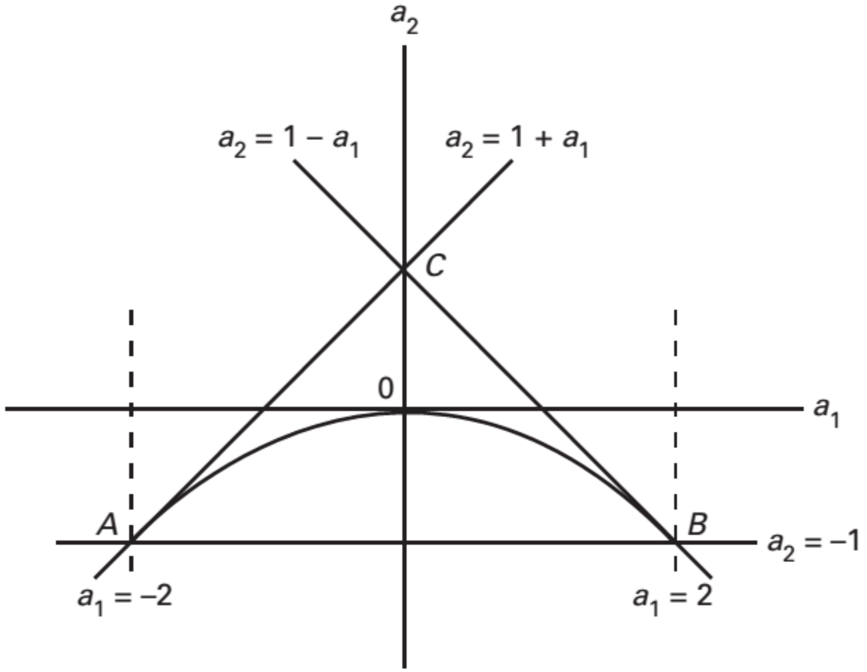


FIGURE 1.5 Characterizing the Stability Conditions

n^{th} order Linear Difference Equation

- A typical n^{th} order Linear Difference Equation

$$y_t = a_0 + a_1 y_{t-1} + a_2 y_{t-2} + \cdots + a_n y_{t-n} + \varepsilon_t$$

- Can be written as:

$$(1 - a_1 L - a_2 L^2 - \cdots - a_n L^n) y_t = a_0 + \varepsilon_t$$

- Which can be factored as:

$$(1 - \alpha_1 L)(1 - \alpha_2 L) \ldots (1 - \alpha_n L) y_t = a_0 + \varepsilon_t$$

- I.e. can be considered as n distinct 1st order difference equations, each of which has to be stable for the nth order difference equation to be stable, hence the rule: all roots of the inverse characteristic polynomial must be outside the unit circle. This will extend directly to Vector Auto Regressions.

Characteristic Polynomial and its inverse:

A typical *n^{th}* order Linear Difference Equation

$$y_t = a_0 + a_1 y_{t-1} + a_2 y_{t-2} + \cdots + a_n y_{t-n} + \varepsilon_t$$

It's *inverse characteristic polynomial*:

$$(1 - a_1 L - a_2 L^2 - \cdots - a_n L^n) = 0$$

Define $\mathrm{z} = L^{-1}$, multiply by $\mathrm{z}^n = L^{-n}$

$$(\mathrm{z}^n - a_1 \mathrm{z}^{n-1} - a_2 z^{n-2} - \cdots - a_n) = 0$$

This is the *characteristic polynomial.*

If all the roots of the ***inverse characteristic polynomial*** are **outside** the unit circle, all the roots of the ***characteristic polynomial*** are ***inside*** the unit circle, then a difference equation is "backwards stable" in the sense that it has a backward solution without an initial condition

Important Cases

- There are several patterns the roots of the inverse characteristic equation may fall in that give different dynamic behaviour:
- Distinct Real Roots
  - All roots outside the unit circle
    - stable, mean reverting, Non-explosive
  - One or more root = 1, others outside the unit circle
    - Random Walk type behaviour
    - Non-mean reverting, Non-explosive
  - One or more root inside unit circle
    - Explosive – homogenous part eventually dominates
- Repeated Roots
  - hump-shaped
- Imaginary Roots
  - oscillating

(Naturally, if e.g. a statistics program is set up to evaluate the roots of the characteristic equation, the rule flips around – inside the unit circle is stable, outside explosive).

Summary

- Our purpose was to derive the rule that tells us when a difference equation has a backward solution without an initial condition
- That rule is:
  - The roots of the inverse characteristic polynomial must be outside the unit circle

  Equivalently:
  - the absolute value of the roots must be larger than 1
- This will become the universal test for *stationarity* in the rest of the course (and in all of time series econometrics)
- Roots for any order polynomial can be trivially computed by any computational program
- Hence we have the fundamental rule that distinguishes different time series models – non-stationary vs stationary models.

For next session:

- Review the work, read through the chapter and practice the examples
- On Friday we will do an introduction to Matlab and simulate some of these processes to get a hands-on feel for the meaning of the stationarity rule