

# Time Series Regression and Forecasting

---

Juergen Meinecke

Stationarity and Unit Roots

Stationarity and Trends

## Definition

A time series  $Y_t$  is **stationary** if its probability distribution does not change over time. If a time series is not stationary it is called **non-stationary**.

Intuitively, stationarity requires the future to be like the past (in a probabilistic sense)

If we believe the future to be fundamentally different from the past, then using past data to make forecasts for the future may be unreliable

We would like our time series to be stationary

There is one particular threat to stationarity that we will learn to address...

## Definition

A **trend** is a persistent, long-term movement or tendency.

There are two types of trends

- deterministic: non-random functions of time  
(relatively easy to deal with)
- stochastic: random functions of time  
(require econometric testing)

Stochastic trends create serious econometric challenges

Time series with stochastic trends are non-stationary

Simple examples of time series that follow *deterministic* trends are given by the following equations:

linear  $X_t = \beta_0 + \beta_1 t + u_t$

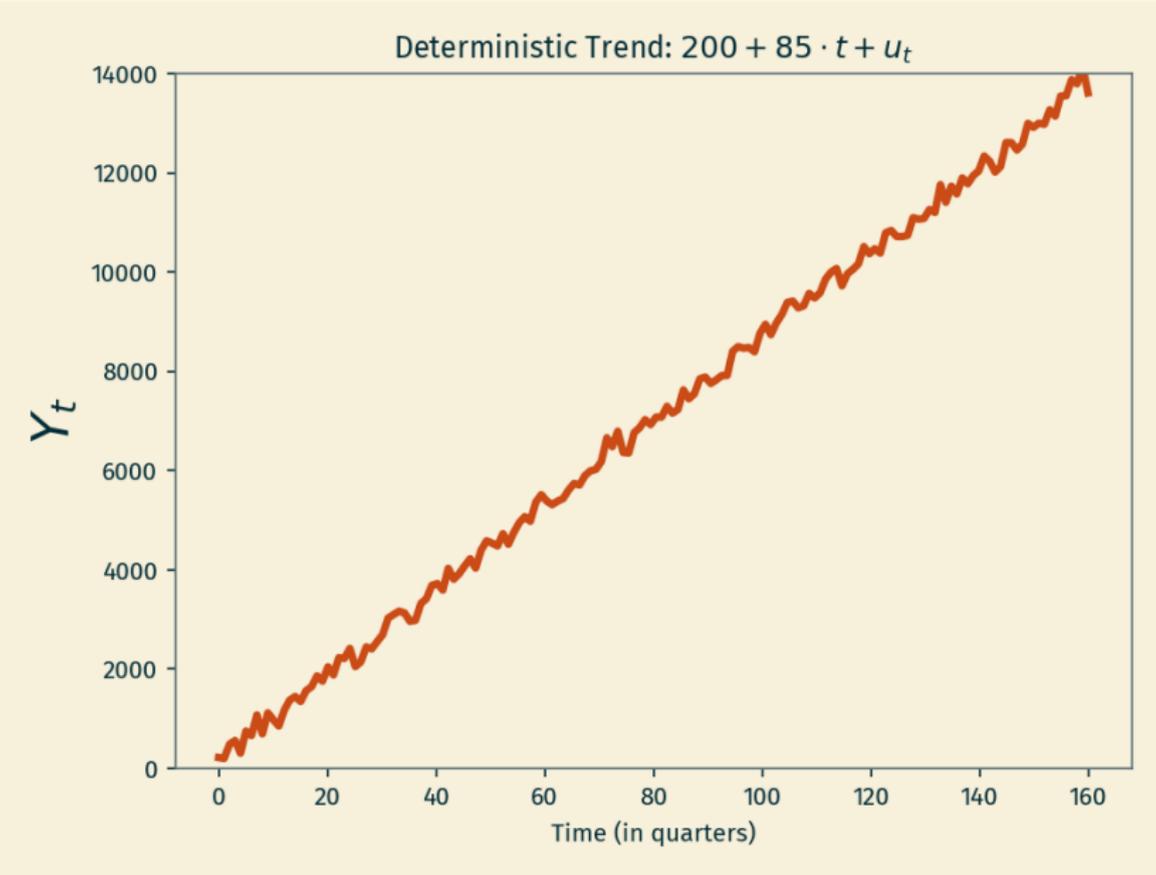
quadratic  $Y_t = \beta_0 + \beta_1 t + \beta_2 t^2 + u_t$

cubic  $Z_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + u_t$

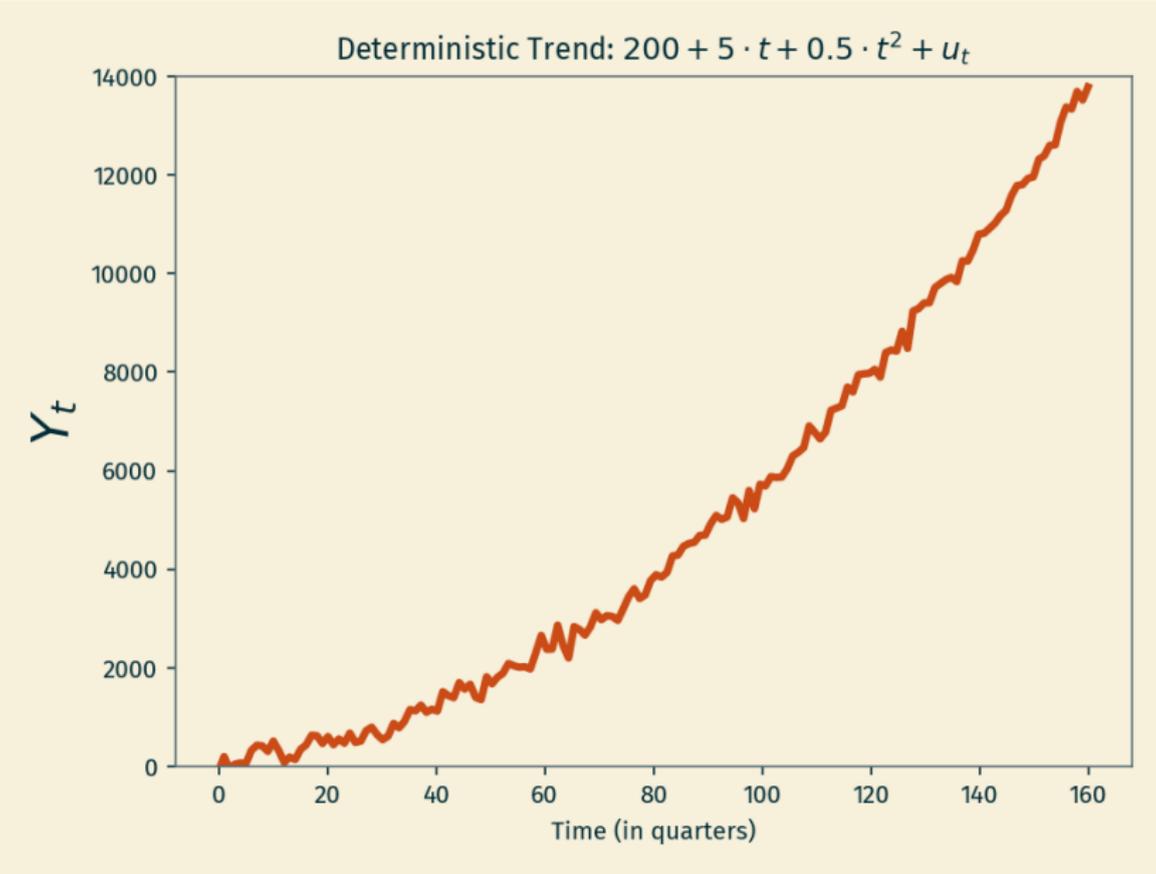
It's easy to account for the presence of deterministic trends:  
Just add a flexible function of time on the rhs of your regressions

In your regression, just include  $t$  and perhaps  $t^2$  as regressors

# Linear deterministic trend



# Linear quadratic trend



Stochastic trends, on the other hand, create bigger problems

Simplest example of a time series that follows stochastic trend:

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + u_t,$$

where

- $\beta_1 = 1$
- $E(u_t | Y_{t-1}, Y_{t-2}, \dots) = 0$

This model is called *random walk with drift*

The coefficient  $\beta_0$  is the drift

When  $\beta_0 = 0$  we get the *random walk*

$$Y_t = Y_{t-1} + u_t$$

Why is it called random walk?

Rewrite it equivalently like so:

$$\Delta Y_t = u_t,$$

where, as usual,  $\Delta Y_t := Y_t - Y_{t-1}$

So period-to-period changes are completely random

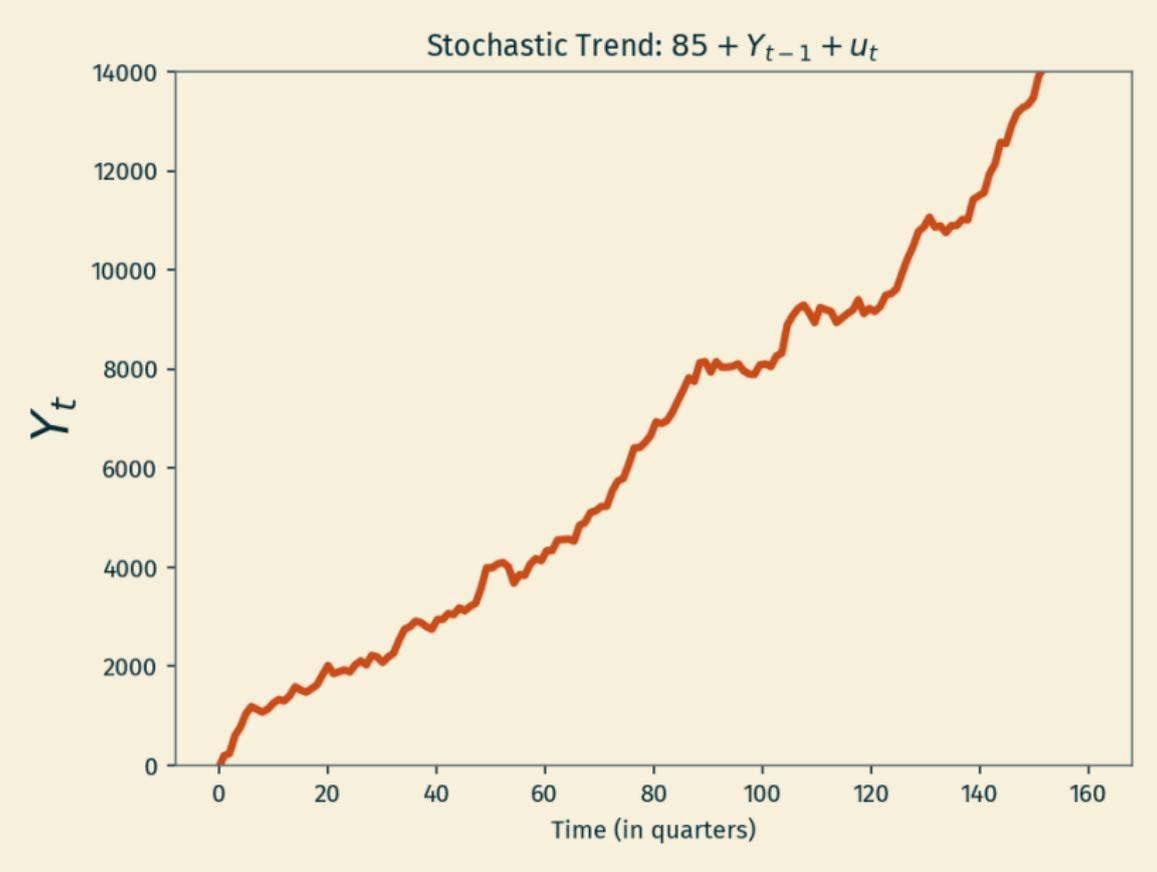
In fact  $E(\Delta Y_t | Y_{t-1}, Y_{t-2}, \dots) = 0$ ,

so the expected period-to-period change given past observations is zero

Past observations don't play a role in predicting current changes in the time series

Time series with this property are also called *martingales*

Here is one computer draw of a random walk with drift



## Two key features of the random walk model

- $Y_{T+1|T} = \beta_0 + Y_T$

The forecast for the next period is today's observation (adjusted for the drift)

- Suppose, for simplicity,  $Y_0 = 0$  then

$$\text{Var}(Y_1) = \text{Var}(Y_0) + \text{Var}(u_1) = \sigma_u^2$$

$$\text{Var}(Y_2) = \text{Var}(Y_1) + \text{Var}(u_2) = \sigma_u^2 + \sigma_u^2 = 2 \cdot \sigma_u^2$$

$$\text{Var}(Y_3) = \text{Var}(Y_2) + \text{Var}(u_3) = 2 \cdot \sigma_u^2 + \sigma_u^2 = 3 \cdot \sigma_u^2$$

$$\text{Var}(Y_4) = \text{Var}(Y_3) + \text{Var}(u_4) = 3 \cdot \sigma_u^2 + \sigma_u^2 = 4 \cdot \sigma_u^2$$

⋮

$$\text{Var}(Y_t) = \text{Var}(Y_{t-1}) + \text{Var}(u_t) = (t-1) \cdot \sigma_u^2 + \sigma_u^2 = t \cdot \sigma_u^2$$

This illustrates that the variance depends on  $t$ , and therefore  $Y_t$  cannot be stationary (why?)

# Time Series Regression and Forecasting

---

Juergen Meinecke

Stationarity and Unit Roots

Unit Roots

In the AR(1) model  $Y_t = \beta_0 + \beta_1 Y_{t-1} + u_t$  in which  $u_t$  is stationary we distinguish two cases:

- if  $|\beta_1| < 1$  then  $Y_t$  will be stationary
- if  $\beta_1 = 1$  then  $Y_t$  will be non-stationary

Aside: economists do not consider the case

- $|\beta_1| > 1$ : explosive (perhaps for periods of hyperinflation?)

When  $Y_t$  is non-stationary, we also sometimes say that it

- has a unit root
- follows a stochastic trend

We use the terms non-stationarity, stochastic trend, and unit root interchangeably

In the more general AR(p) model

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \dots + \beta_p Y_{t-p} + u_t$$

the definition of unit root is a bit more involved

We say that the AR(p) model contains a unit root if at least one of the solutions to the polynomial equation

$$1 - \beta_1 z - \beta_2 z^2 - \dots - \beta_p z^p = 0$$

is equal to 1 in absolute value

We will not prove this and it wouldn't be something you can just check by hand

But you can see that if  $p = 1$  this definition coincides with  $1 - \beta_1 z = 0$  so that the solution  $z = 1/\beta_1$  will be equal to 1 in absolute value only if  $|\beta_1| = 1$

## Unit roots create three problems

### 1. AR coefficients are biased towards zero

- suppose  $Y_t$  contains a unit root but you ignore this
- you estimate the AR(1) model:  $Y_t = \beta_0 + \beta_1 Y_{t-1} + u_t$
- the OLS estimator  $\hat{\beta}_1$  will be biased
- time series econometricians have worked out a simple formula for this bias:  $E[\hat{\beta}_1] = 1 - 5.3/T$
- forecasts will also be biased

### 2. Asymptotic distribution is non-normal

- distribution of the OLS estimator is not close to normal
- cannot simply use the standard errors,  $t$ -statistic,  $p$ -values or confidence intervals that Python spits out in its standard regression output  
(because they all assume the normal distribution)
- in other words: standard statistical inference will not work

3. Spurious regression problem: Because of stochastic trends, any two time series may appear related when they are not

Good example from textbook:

- US unemployment rate and Japanese industrial production
- these two time series aren't meaningfully related
- stylized fact: both contain unit root
- regressing both against each other

using data from 1962 to 1985:

$$\text{US unemployment rate}_t = -2.37 + 2.22 \cdot \ln \text{Japanese IP}_t$$

(0.32)

- between 1962 and 1985 both US unemployment and Japanese industrial production happened to be rising
- this creates a spurious positive correlation between the two
- using data from 1986 to 2017:

$$\text{US unemployment rate}_t = -42.37 - 7.92 \cdot \ln \text{Japanese IP}_t$$

(1.69)

- bottom line: there's a correlation but not a relation!

# Time Series Regression and Forecasting

---

Juergen Meinecke

Stationarity and Unit Roots

Testing for Unit Roots

Because of the three problems just discussed you should not run AR(p) estimations with time series that contain unit roots

Forecasts based on such models will be misleading or wrong

How can we know which time series contain unit roots?

Dickey and Fuller developed a reasonably easy test that should be conducted before you do any time series analysis

It is easiest to illustrate the DF test using an AR(1) model

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + u_t,$$

which is equivalent to

$$\Delta Y_t = \beta_0 + \delta Y_{t-1} + u_t, \quad \delta := \beta_1 - 1$$

To test  $Y_t$  for unit root, we could simply test the coefficient  $\delta$

Formally, we are testing

$$H_0 : \delta = 0 \quad \text{versus} \quad H_1 : \delta < 0$$

Caution: test is one-sided (because we only consider  $\beta_1 \leq 1$ )!

Remember: if  $\delta = 0$  then  $\beta_1 = 1$  (unit root)

To conduct this test, you would simply regress  $\Delta Y_t$  against  $Y_{t-1}$  and check the resulting coefficient estimate  $\hat{\delta}$  for significance

Complication: you cannot simply use the standard error,  $t$ -statistic,  $p$ -value or confidence intervals that Python spits out

The coefficient estimate  $\hat{\delta}$  does not have an approximate normal distribution

So you cannot apply the usual critical value of 1.96 (which comes from the standard normal distribution)

Instead, the distribution is non-normal

Fortunately, Dickey and Fuller have derived the correct critical values for us

In deriving the critical values, DF distinguish between two different AR(1) model specifications:

intercept only

$$\Delta Y_t = \beta_0 + \delta Y_{t-1} + u_t$$

incl. deterministic time trend

$$\Delta Y_t = \beta_0 + \alpha t + \delta Y_{t-1} + u_t$$

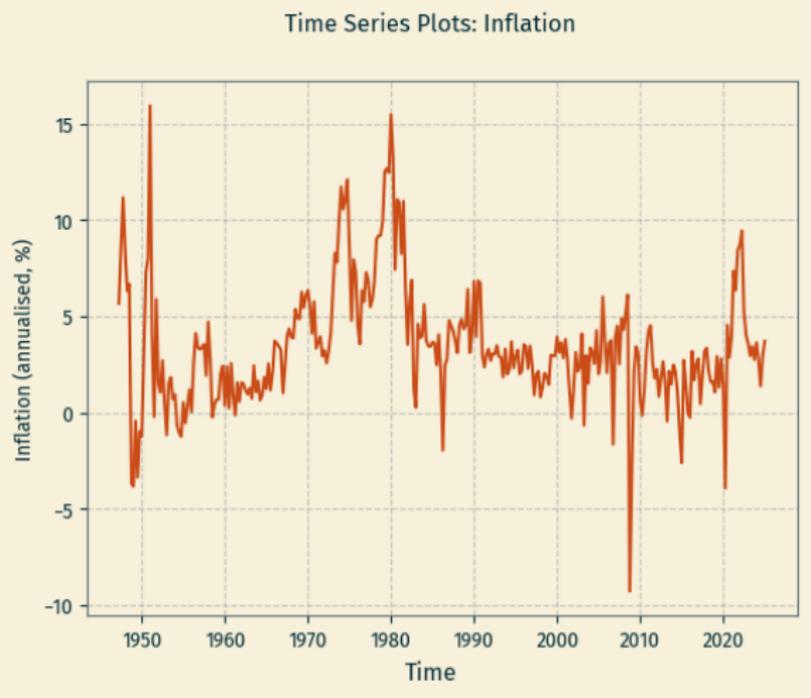
The corresponding critical values (5% significance level) are:

Model specification	Critical value
Intercept only	-2.86
Intercept and time trend	-3.41

For example, in the model specification with intercept only, the null hypothesis of a unit root is rejected if the  $t$ -statistic is less than  $-2.86$

# Example: unit root test for inflation (based on CPIAUCSL)

Does this time series contain a unit root?



(Can't say just from visual inspection)

Let's run the regression that DF suggest

Need to estimate the model  $\Delta \text{infl}_t = \beta_0 + \alpha t + \delta \text{infl}_{t-1} + u_t$

## Python Code (output edited)

```
> df['trend'] = range(1, len(df.infl) + 1)
> unit_root_ar1 = smf.ols('dinfl ~ trend + llinfl', data=df, missing='drop').fit(use_t=False)
> print(unit_root_ar1.summary())
```

deterministic time trend

compare to DF crit value!

```
=====
              coef      std err          z      P>|z|      [0.025      0.975]
-----
Intercept    1.0448      0.295      3.545      0.000      0.467      1.623
trend        -0.0006      0.001     -0.400      0.689     -0.003      0.002
llinfl       -0.2794      0.039     -7.079      0.000     -0.357     -0.202
=====
```

We estimate  $\hat{\delta} = -0.2794$  with a t-stat of  $-7.079$

Caution: Do not use the “robust” option for the DF test!

The  $t$ -statistic is  $-7.079$  which we will compare to the critical value provided by DF

Recall that the DF unit root test is one-sided: the  $t$ -statistic will always be negative; the key question is whether it is sufficiently small to distinguish it from zero

Here we find that  $-7.079 < -3.41$  and conclude that we *reject* the null hypothesis of a unit root (5% significance level)

Here's a convenience function to run the test in one line

## Python Code

```
> from statsmodels.tsa.stattools import adfuller
> adfuller(df.infl.dropna(), maxlag = 0, regression='ct')
(-7.0790892115383155,
 9.351678189617628e-09,
 0,
 311,
 '1%': -3.988178023764844,   '5%': -3.4247019735849102,   '10%': -3.1354051510159615,
 1376.089008314922)
```

p-value

Can compare t-stat to DF critical value or interpret the p-value

Note: the option regression permits the following:

- regression='c': include constant only (default)
- regression='ct': include constant and linear time trend
- regression='ctt': include constant, linear and quadratic trend
- regression='n': include no constant and no trend

Dickey and Fuller also developed a unit root test that is based on the more general AR(p) model

The generalization is straightforward, consider the AR(p) model in first differences

$$\Delta Y_t = \beta_0 + \alpha t + \delta Y_{t-1} + \gamma_1 \Delta Y_{t-1} + \dots + \gamma_p \Delta Y_{t-p} + u_t$$

Depending on whether  $\alpha \neq 0$  or  $\alpha = 0$  you can look at this as a model with or without deterministic time trend

The same critical values from the previous table still apply here

Estimating the above AR(p) model is easy

For example, I am estimating an AR(4) model now...

Let's run the regression that DF suggest

## Python Code (output edited)

```
> # creating four lags of dinfl
> df['l1dinfl'] = df.dinfl.shift(1)
> df['l2dinfl'] = df.dinfl.shift(2)
> df['l3dinfl'] = df.dinfl.shift(3)
> df['l4dinfl'] = df.dinfl.shift(4)

> unit_root_ar4 = smf.ols('dinfl ~ trend + l1infl + l1dinfl + l2dinfl + l3dinfl + l4dinfl',
                          data=df, missing='drop').fit(use_t=False)
> print(unit_root_ar4.summary())
```

```
=====
              coef    std err          z      P>|z|      [0.025      0.975]
-----
Intercept    0.6682    0.300      2.227    0.026     0.080     1.256
trend        2.451e-05  0.001     0.018    0.986    -0.003     0.003
l1infl       -0.2033    0.046    -4.425    0.000    -0.293    -0.113
l1dinfl      -0.1618    0.062    -2.590    0.010    -0.284    -0.039
l2dinfl      -0.1824    0.063    -2.899    0.004    -0.306    -0.059
l3dinfl      0.1316    0.060     2.196    0.028     0.014     0.249
l4dinfl      -0.1016    0.057    -1.788    0.074    -0.213     0.010
=====
```

Caution: do not use the “robust” option for the DF test!

The  $t$ -statistic is  $-4.425$  which we will compare to the critical value provided by DF

Here we find that  $-4.425 < -3.41$  and conclude that we *reject* the null hypothesis of a unit root (5% significance level)

Again using the convenience function to run the test in one line

## Python Code

```
> from statsmodels.tsa.stattools import adfuller
> adfuller(df.infl.dropna(), maxlag = 4, regression='ct')
(-4.425190419784897,
 0.001995606026544201,
 4,
 307,
 '1%': -3.9885651881396162,   '5%': -3.424888419060633,   '10%': -3.135514568313964,
 1330.8744210409302)
```

Compare t-stat to DF critical value or alternatively interpret the p-value