ECON-C4200 - Econometrics II

Lecture 8: Time series I

Otto Toivanen

Learning outcomes

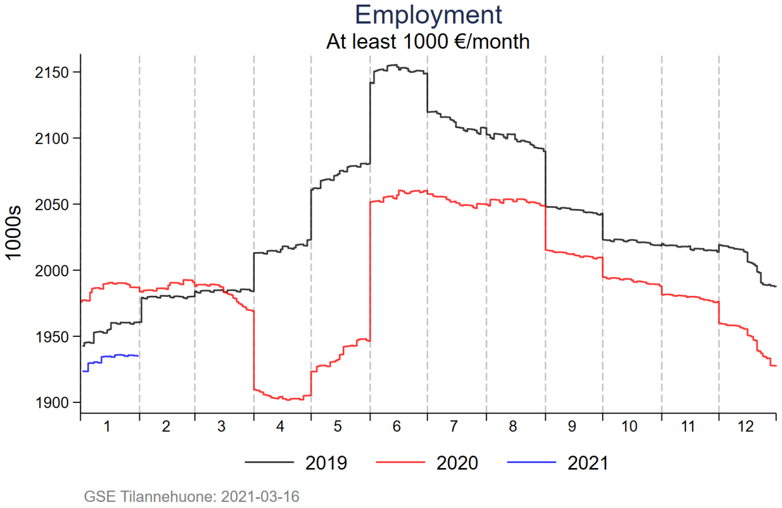
- At the end of lecture 8, you know

1. the definition of a time series
2. what **lags** (=lagged values of a variable) and **differences** are and how to create them
3. what **autocovariance** and **autocorrelation** are
4. what an **autoregressive** model of **order p** is.
5. what a **deterministic** and what a **stochastic** trend is
6. what **stationarity** means and what consequences it has
7. what a **random walk** is and
8. how to test for stationarity

Why separate lectures on time series?

- Time series: values of a (set of) variable(s) Y_t over time.
- Recall OLS assumptions.
- In particular, error terms are assumed to be i.i.d
- This (very) unlikely to hold with a time series: think of the COVID-19 employment shock.

employed in Finland



Time series & economics

- What most lay people consider key economic data are time series:
  1. Price indices (inflation).
  2. Unemployment.
  3. GDP.
  4. Exports and imports.

Any particular reason to study time series?

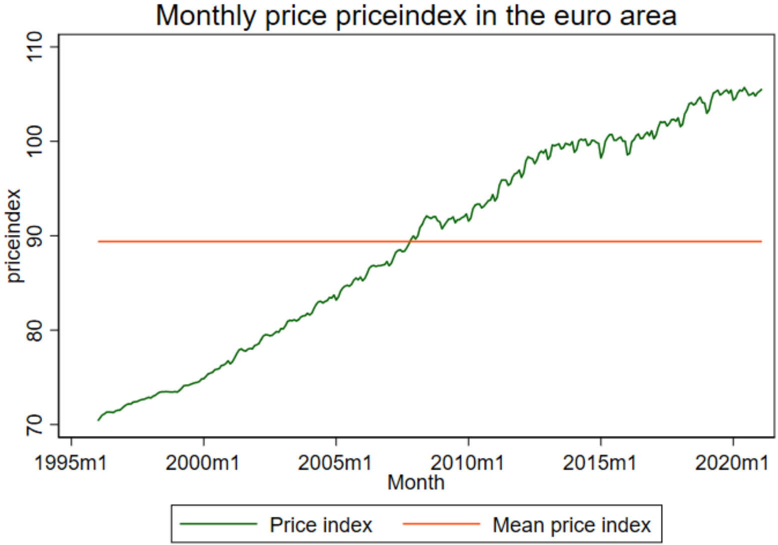
- Forecasting: what is the unemployment rate in 12 months?.
- Dynamic causal effects: Does a change in the central bank interest rate affect inflation 3 months / 12 months ahead?.
- Modeling risks in financial markets (volatility).
- Time series techniques have plenty of uses outside economics (climate modeling, engineering systems, computer science).
- Even if not interested in the time series nature, need to take it into account.

Basic concepts

- Y_t = value of variable Y in period t.
- Time period = time unit you are using.
- Year, quarter, month, week, day, hour, minute, second, ...
- Example: Euro area inflation & unemployment.
- Data set: $\{Y_1, ..., Y_T\}$ are T observations on the time series variable Y.
- We will study time series that are **consecutive**, i.e., there are not breaks (=missing observations) in the series.

unit	coicop	geo	time	priceindex
I15	CP00	EA19	2021m2	105.54
I15	CP00	EA19	2021m1	105.32
I15	CP00	EA19	2020m12	105.15
I15	CP00	EA19	2020m11	104.8
I15	CP00	EA19	2020m10	105.15
I15	CP00	EA19	2020m9	104.96
I15	CP00	EA19	2020m8	104.88
I15	CP00	EA19	2020m7	105.32
I15	CP00	EA19	2020m6	105.69
I15	CP00	EA19	2020m5	105.33

Price level



Monthly price priceindex in the euro area
priceindex
110
100
90
80
70
1995m1
2000m1
2005m1
2010m1
2015m1
2020m1
Month
Price index
Mean price index

Lags, differences, autocorrelation

- Econometric software usually have ready-made operators to produce **differences** and **lags**.
- **First lag**: $LY_t = Y_{t-1}$;
- p^{th} lag: $L^p Y_t = Y_{t-p}$, $p = 1, \ldots$.
- The **first difference**: $\Delta Y_t = Y_t - Y_{t-1}$;
- p^{th} lagged difference $\Delta^p Y_t = Y_{t-p} - Y_{t-p-1}$
- Notice that you "lose" observations from the beginning of the series when you take lags and/or differences.
- Stata time series operators

Lags, Differences, autocorrelation

- Autocovariance and autocorrelation:
  1. $cov(Y_t, Y_{t-\tau}) = \tau^{th}$ autocovariance.
  2. $\rho = \frac{cov(Y_t, Y_{t-\tau})}{var(\sqrt{Y_t})} = \tau^{th}$ autocorrelation.
- These are **population** correlations, i.e., they describe the joint distribution of the population.
- The **sample** autocovariance and autocorrelation are estimates of the population equivalents.

Correlation

```
. pwcorr priceindex L.priceindex L2.priceindex L3.priceindex, sig

             | pricei~x L.pric~x L2.pri~x L3.pri~x
-------------+------------------------------------
  priceindex |   1.0000
             |
             |
L.priceindex |   0.9990   1.0000
             |   0.0000
             |
L2.pricein~x |   0.9979   0.9990   1.0000
             |   0.0000   0.0000
             |
L3.pricein~x |   0.9972   0.9978   0.9990   1.0000
             |   0.0000   0.0000   0.0000
             |
```

Correlation

```
. pwcorr ln_priceindex L.ln_priceindex L2.ln_priceindex L3.ln_priceindex, sig

             | ln_pri~x L.ln_p~x L2.ln_~x L3.ln_~x
-------------+------------------------------------
ln_pricein~x |   1.0000
             |
             |
L.ln_price~x |   0.9991   1.0000
             |   0.0000
             |
L2.ln_pric~x |   0.9981   0.9991   1.0000
             |   0.0000   0.0000
             |
L3.ln_pric~x |   0.9975   0.9981   0.9991   1.0000
             |   0.0000   0.0000   0.0000
             |
```

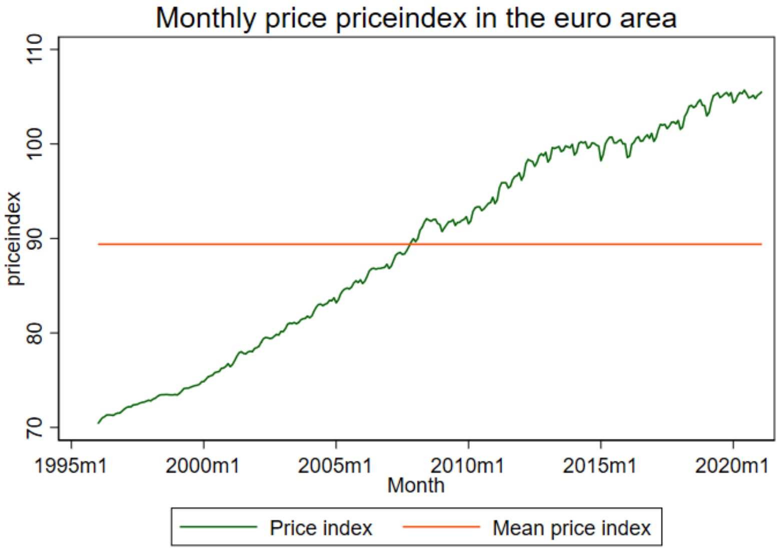
Autocorrelation

- Many economic time series exhibit strong autocorrelation.
- What about (1st) differences?

$$\Delta Y_t = Y_t - Y_{t-1}$$

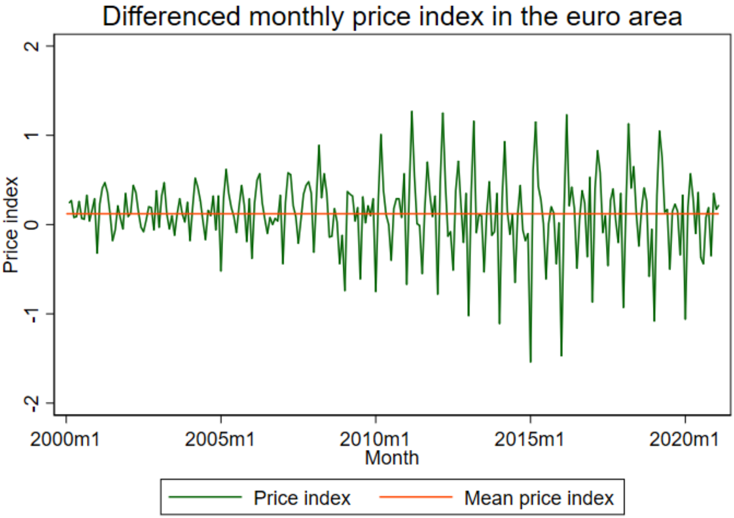
- Note that if one uses logs of Y, then $\Delta \ln Y_t \approx (Y_t - Y_{t-1})/Y_{t-1}$.
- $\rightarrow$ economists often used logs of time series data.
- Illustration using the price level and inflation as examples.

Price level



Monthly price priceindex in the euro area
priceindex
110
100
90
80
70
1995m1
2000m1
2005m1
2010m1
2015m1
2020m1
Month
Price index
Mean price index

Inflation = growth rate of price level



Correlation

```
. pwcorr d1.priceindex d1.L1.priceindex d1.L2.priceindex d1.L3.priceindex, sig
```

	D.pric~x	LD.pri~x	L..pri~x	L..pri~x
D.priceindex	1.0000			
LD.pricein~x	0.0414	1.0000		
	0.5127			
L2D.pricei~x	-0.2087	0.0413	1.0000	
	0.0009	0.5146		
L3D.pricei~x	-0.1312	-0.2093	0.0411	1.0000
	0.0382	0.0009	0.5178	

Correlation

```
. pwcorr d1.ln_priceindex d1.L1.ln_priceindex d1.L2.ln_priceindex d1.L3.ln_priceindex, sig

             | D.ln_p~x LD.ln_~x L..ln_~x L..ln_~x
-------------+------------------------------------
D.ln_price~x |   1.0000
             |
             |
LD.ln_pric~x |   0.0500   1.0000
             |   0.4296
             |
L2D.ln_pri~x |  -0.2001   0.0499   1.0000
             |   0.0014   0.4309
             |
L3D.ln_pri~x |  -0.1269  -0.2005   0.0498   1.0000
             |   0.0450   0.0014   0.4326
             |
```

First order autoregression - AR(1)

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

- let's try on price index and inflation.

AR(1) - levels

```
. reg priceindex L1.priceindex

      Source |       SS           df       MS      Number of obs   =       253
-------------+----------------------------------   F(1, 251)       >  99999.00
       Model |  21526.6321         1  21526.6321   Prob > F        =    0.0000
    Residual |  44.2221747       251  .176183963   R-squared       =    0.9979
-------------+----------------------------------   Adj R-squared   =    0.9979
       Total |  21570.8543       252  85.5986282   Root MSE        =    .41974

------------------------------------------------------------------------------
  priceindex |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
  priceindex |
         L1. |   .9956737   .0028485   349.55   0.000     .9900637    1.001284
             |
       _cons |   .5213354   .2647284     1.97   0.050    -.0000366    1.042707
------------------------------------------------------------------------------
```

AR(1) - differenced

```
. reg d1.priceindex d1.L1.priceindex

      Source |       SS           df       MS      Number of obs   =       252
-------------+----------------------------------   F(1, 250)       =      0.43
       Model |  .076566518         1  .076566518   Prob > F        =    0.5127
    Residual |  44.5378748       250  .178151499   R-squared       =    0.0017
-------------+----------------------------------   Adj R-squared   =   -0.0023
       Total |  44.6144413       251  .177746778   Root MSE        =    .42208

------------------------------------------------------------------------------
D.priceindex |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
  priceindex |
         LD. |   .0414248   .0631882     0.66   0.513    -.0830242    .1658738
             |
       _cons |   .1157865   .0276638     4.19   0.000     .0613027    .1702703
------------------------------------------------------------------------------
```

p^{th} - order autoregression - AR(p)

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

- let's try on price index and inflation.
- let's set p = 4.

AR(4) - levels: price index

```
. reg priceindex L(1/4).priceindex

      Source |       SS           df       MS      Number of obs   =       250
-------------+----------------------------------   F(4, 245)       =  30502.39
       Model |  20621.9269         4  5155.48174   Prob > F        =    0.0000
    Residual |  41.4096457       245  .169018962   R-squared       =    0.9980
-------------+----------------------------------   Adj R-squared   =    0.9980
       Total |  20663.3366       249  82.9852875   Root MSE        =    .41112

------------------------------------------------------------------------------
  priceindex |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
  priceindex |
         L1. |   1.013009   .0634016    15.98   0.000     .8881267     1.13789
         L2. |   -.227738   .0904242    -2.52   0.012     -.405846     -.04963
         L3. |    .083712   .0904537     0.93   0.356    -.0944541     .261878
         L4. |   .1258839   .0632336     1.99   0.048     .0013331    .2504348
             |
       _cons |   .6342373   .2684653     2.36   0.019     .1054429    1.163032
------------------------------------------------------------------------------
```

AR(4) - differenced: "inflation"

```
. reg D1.priceindex L(1/4).D1.priceindex

      Source |       SS           df       MS      Number of obs   =       249
-------------+----------------------------------   F(4, 244)       =     11.16
       Model |  6.89553199         4    1.723883   Prob > F        =    0.0000
    Residual |  37.6940102       244  .154483648   R-squared       =    0.1546
-------------+----------------------------------   Adj R-squared   =    0.1408
       Total |  44.5895422       248  .179796541   Root MSE        =    .39304

------------------------------------------------------------------------------
D.priceindex |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
  priceindex |
         LD. |  -.0120431   .0606884    -0.20   0.843    -.1315831     .107497
        L2D. |  -.2703697   .0602988    -4.48   0.000    -.3891423   -.1515972
        L3D. |  -.1087541    .060321    -1.80   0.073    -.2275705    .0100622
        L4D. |  -.3195227   .0608607    -5.25   0.000    -.4394021   -.1996434
             |
       _cons |   .2061031    .029613     6.96   0.000     .1477733    .2644329
------------------------------------------------------------------------------
```

How many lags?

- Decide through testing.
- F-tests?
- **Bayes information criterion**: $min_p \; BIC(p) = \ln \frac{SSR(p)}{T} + (p+1)\frac{\ln T}{T}$
- **Akaike information criterion**: $min_p \; AIC(p) = \ln \frac{SSR(p)}{T} + (p+1)\frac{2}{T}$
- AIC inconsistent, but yields more lags if (when) $\ln T > 2$.
- Just like in machine learning, BIC and AIC utilize the bias-variance tradeoff.
- Again similar to machine learning, the objective here is prediction (forecasting).

How many lags?

- Through AIC / BIC testing you get a model with possibly biased coefficients, but good forecasts.
- In Problem Set 4 (or 5), you will search for the optimal # of lags using BIC and AIC.

Stationarity

- As we have seen, an autoregression in levels and an autoregression in differences behave very differently.
- This takes us to the concept of **stationarity**.
- Stationarity is an important characteristic of time series.
- It affects regression analysis.
- It affects in particular time series analysis with multiple/explanatory variables.

Stationarity

- **Definition** (for a single time series): A time series Y_t is stationary if its probability distribution does not change over time, that is, if the joint distribution of $(Y_{s+1}, Y_{s+2}, ..., Y_{s+T})$ does not depend on s.
- Otherwise Y_t is **nonstationary**.
- Stationarity requries that in a **probabilistic** sense, the future is like the past.

What is the fuzz about (non)stationarity

- Autoregressive coefficients are biased towards zero.
- It can be shown that with a **random walk** (which we will study more shortly) which is nonstationary, the OLS coefficient of the lagged dependent variable (Y_{t-1})

$$\mathbb{E}[\beta_1] \approx 1 - \frac{5.3}{T}$$

Demonstration with artificial data

Stata code

```
set obs 1000
gen time    = _n
tsset time
gen u      = 3 * invnorm(uniform())
gen y      = u
replace y   = y[_n - 1] + u if time > 1

regr y L.y  if time < T
/* vary T = 26, 51, 151, 1000(+1) */
```

The Monte Carlo data

time	u	y	u+y[t-1]
1	-3.2819	-3.2819	
2	1.1012	-2.1807	-2.1807
3	0.4362	-1.7445	-1.7445
4	0.7973	-0.9471	-0.9471
5	1.4382	0.4911	0.4911
6	-3.7009	-3.2098	-3.2098
7	0.9043	-2.3055	-2.3055
8	-4.6377	-6.9433	-6.9433
9	0.4167	-6.5265	-6.5265
10	3.3998	-3.1267	-3.1267

Results

```
. estimates table t_25 t_50 t_100 t_1000 , b(%7.4f) se(%7.4f) p(%7.4f) stat(N r2 r2_a aic bic)
```

Variable	t_25	t_50	t_100	t_1000
y				
L1.	0.8397	0.9685	0.9917	0.9961
	0.1170	0.0549	0.0136	0.0030
	0.0000	0.0000	0.0000	0.0000
_cons	0.9182	0.8295	-0.2251	0.1034
	0.9938	0.7253	0.2662	0.1035
	0.3651	0.2584	0.3992	0.3181
N	25	50	150	999
r2	0.6915	0.8666	0.9730	0.9909
r2_a	0.6781	0.8638	0.9728	0.9909
aic	133.3537	259.3584	780.5999	5.0e+03
bic	135.7915	263.1825	786.6211	5.0e+03

legend: b/se/p

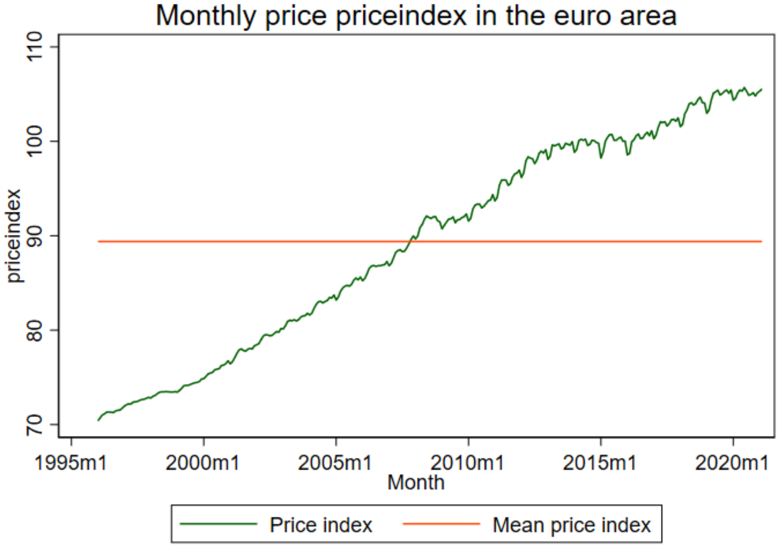
What is the fuzz about (non)stationarity

- In an autoregressive model, the coefficients are biased.
- Also, the t-statistics will have non-normal distributions.
- Maybe more consequentially, you get bad forecasts.

What causes nonstationarity

- Two important sources of nonstationarity:
  1. trends
  2. structural breaks (next lecture)

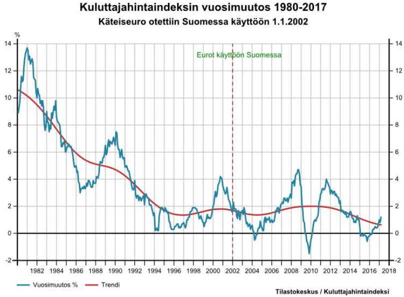
Example of a trend: Price level

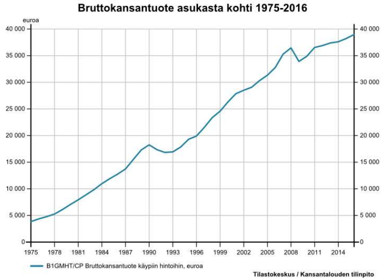


Trends

- So a **trend** is a long-term movement in the data.
- A **deterministic** trend is a non-random function of time, e.g., $Y_t = t$ (i.e., time itself).
- A **stochastic** trend is random and varies over time.
- An important special case of a stochastic trend is a **random walk**.

Examples of trends: Finnish CPI and GDP per capita





2

Random walk

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

- The value of Y_t **today** is **in expectation** the same as what it **actually was yesterday**.
- **Random walk** = today's value of Y_t is equal to where you were yesterday + a (random) step u_t of unknown direction and length.
- More generally, the best prediction p periods into the future is Y_t.

Random walk

- Let's think what happens to the **(expected) value** of a random walk:

period	Y_t	u_t
1	u_1	u_1
2	$u_2 + u_1$	u_2
3	$u_3 + u_2 + u_1$	u_3
.	.	.
.	.	.
l	$\sum_i u_i$	u_l
.	.	.
.	.	.

- The value of a random walk in period t is the sum of the shocks to the series until and including period t.

Random walk

- Let's think what happens to the **variance** of a random walk (assuming $cov(u_t, u_{t-1}) = 0$):

period	$var(Y_t)$	$var(u)$
1	$var(u)$	$var(u)$
2	$var(u) + var(u)$	$var(u)$
3	$var(u) + var(u) + var(u)$	$var(u)$
.	.	.
.	.	.
l	$l \times var(u)$	$var(u)$
.	.	.
.	.	.

- The variance of a random walk in period t is the sum of the period-specific variances.

- Thus the variance of a random walk is increasing over time.

Random walk with drift

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

- The value of Y_t today is in expectation the same as what it actually was yesterday $+\beta_0$.
- Random walk = today's value of Y_t where you were yesterday $+\beta_0 +$ a (random) step u_t of unknown direction and length.
- More generally, the best forecast p periods ahead is

$$\mathbb{E}[Y_{t+p}|Y_t] = \beta_0 p + Y_t$$

Random walk is a special case of AR(1)

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

- If $\beta_1 = 1$, Y_t is nonstationary.
- If $\beta_1 < 1$, Y_t is stationary.
- For AR(p), $p > 1$, similar but more complicated statements apply.

Nonstationarity, unit root & stochastic trend

- If $\beta_1 = 1$, Y_t is nonstationary.
- If $\beta_1 = 1$, Y_t has a **unit root**.
- If $\beta_1 = 1$, Y_t has a **stochastic trend**.
- solving equation $1 - \beta_1 z = 0$ yields the root $z = 1/\beta_1$. If $\beta_1 = 1$, $z = 1$; hence the terminology "unit root".

Testing for a unit root

- First thing to do: Plot the data to visually inspect whether there is a (stochastic) trend.
- More formally, let us study the AR(1) model:

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

- H_0: $\beta_1 = 1$, Y_t is nonstationary / there is a unit root.
- H_1: $\beta_1 < 1$, Y_t is stationary / does not have a unit root.
- BUT: how to test when we know that it does not make sense to directly estimate the above equation unless Y_t is stationary, i.e., unless we know the answer to our question?

Testing for a unit root

- Need a trick:

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

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

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

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

Testing for a unit root: The Dickey-Fuller test (AR(1))

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

- We regress the **first difference** of Y_t on the **lagged level** of Y_t.
- Pay attention to $\delta = \beta_1 - 1$, the coefficient of Y_{t-1}.
- We know that $\beta_1 \leq 1 \rightarrow \delta \leq 0$.
- We want to test the $H_0 : \delta = 0$ against the **one-sided** alternative $H_1 : \delta < 0$.

The augmented Dickey-Fuller test

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

- We can add lagged differences to allow for higher order autocorrelation.
- $H_0 : \delta = 0$ against
- $H_1 : \delta < 0$.

The augmented Dickey-Fuller test

- How do we get the equation for the augmented Dickey-Fuller test? Example with an AR(2):

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + u_t$$

$$Y_t = \beta_0 + (\beta_1 + \beta_2) Y_{t-1} - \beta_2 Y_{t-1} + \beta_2 Y_{t-2} + u_t$$

$$Y_t = \beta_0 + (\beta_1 + \beta_2) Y_{t-1} - \beta_2 [Y_{t-1} - Y_{t-2}] + u_t$$

$$Y_t - Y_{t-1} = \beta_0 - Y_{t-1} + (\beta_1 + \beta_2) Y_{t-1} - \beta_2)[Y_{t-1} - Y_{t-2}] + u_t$$

$$\Delta Y_t = \beta_0 + (\beta_1 + \beta_2 - 1) Y_{t-1} - \beta_2 \Delta Y_{t-1} + u_t$$

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

- $\beta_1 z + \beta_2 z - 1 = 0$, then $\beta_1 + \beta_2 = 1$ and hence $\delta = 0$, too.

The augmented Dickey-Fuller test

- In the AR(1) case, if $\beta_1 = 1$, i.e., there is a unit root,

$$\Delta Y_t = \beta_0 + u_t$$

- In the AR(2) case, if $\beta_1 + \beta_2 = 1$, i.e., there is a unit root,

$$\Delta Y_t = \beta_0 + \gamma_1 \Delta Y_{t-1} + u_t$$

The augmented Dickey-Fuller test with deterministic trend

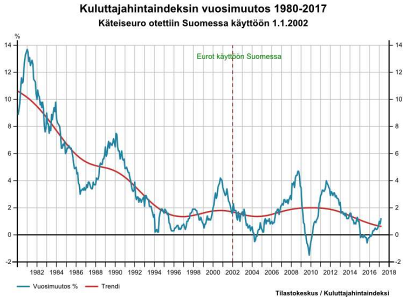
$$\Delta Y_t = \beta_0 + \alpha t + \delta Y_{t-1} + \gamma_1 \Delta Y_{t-1} ... + \gamma_p \Delta Y_{t-p} + u_t$$

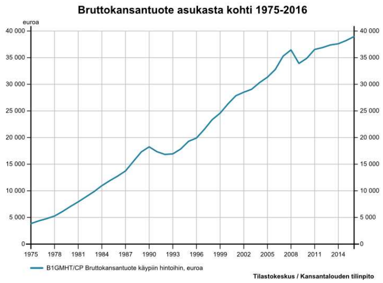
- $H_0 : \delta = 0$ against
- $H_1 : \delta < 0$.

The augmented Dickey-Fuller test with deterministic trend

- What version to use - with or without deterministic trend?
  - If no long-term growth (-/+), then the alternative is that Y_t is stationary round a constant. use the intercept-only version.
  - If the time-series seems to exhibit long-term growth, then the alternative is that Y_t is stationary round a trend. $\rightarrow$ use the intercept plus deterministic trend version.
  - Think back to Finnish CPI and GDP per capita - series.

Examples of trends: Finnish CPI and GDP per capita





2

The augmented Dickey-Fuller test

- The Dickey-Fuller test is **one-sided**.
- Why are we not worried about $\delta > 0$, i.e., $\beta_1 > 1$?
- Need special critical values.

Demonstration #1: Euro-area price index in logs

```
. dfuller lnpind_euro if year > 2000 ,regress

Dickey-Fuller test for unit root                   Number of obs   =       177

                               ---------- Interpolated Dickey-Fuller ---------
                  Test         1% Critical       5% Critical      10% Critical
               Statistic           Value             Value             Value
------------------------------------------------------------------------------
 Z(t)             -1.804            -3.484            -2.885            -2.575
------------------------------------------------------------------------------
MacKinnon approximate p-value for Z(t) = 0.3784

------------------------------------------------------------------------------
D.           |
 lnpind_euro |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
 lnpind_euro |
         L1. |  -.0067908   .0037638    -1.80   0.073    -.0142191    .0006375
             |
       _cons |   .0319235   .0169217     1.89   0.061    -.0014734    .0653205
------------------------------------------------------------------------------
```

3

DF test & critical values for different AR(p)

#lags	ADF	1%	5%	10%
0	-1.804	-3.484	-2.885	-2.575
1	-1.758	-3.484	-2.885	-2.575
2	-1.998	-3.484	-2.885	-2.575
3	-2.109	-3.484	-2.885	-2.575
4	-2.577	-3.484	-2.885	-2.575
5	-2.691	-3.484	-2.885	-2.575
6	-1.933	-3.484	-2.885	-2.575
7	-2.056	-3.484	-2.885	-2.575
8	-2.203	-3.484	-2.885	-2.575

4

The augmented Dickey-Fuller test

- So there is a unit root... what to do?
- As suggested above, transform by differencing.

Demonstration #2: Euro-area inflation = difference in log of price index

- On the LHS, test results and critical values for different AR(p).
- On the RHS, AR(1) test results.

#lags	ADF	1%	5%	10%
0	-12.015	-3.484	-2.885	-2.575
1	-11.006	-3.484	-2.885	-2.575
2	-8.975	-3.484	-2.885	-2.575
3	-9.556	-3.484	-2.885	-2.575
4	-7.861	-3.484	-2.885	-2.575
5	-3.548	-3.484	-2.885	-2.575
6	-3.862	-3.484	-2.885	-2.575
7	-4.149	-3.484	-2.885	-2.575
8	-3.981	-3.484	-2.885	-2.575

```
. dfuller dlnpind_euro if year > 2000, regress

Dickey-Fuller test for unit root                   Number of obs   =       177

                               ---------- Interpolated Dickey-Fuller ---------
                  Test         1% Critical       5% Critical      10% Critical
               Statistic           Value             Value             Value
------------------------------------------------------------------------------
 Z(t)            -12.015            -3.484            -2.885            -2.575
------------------------------------------------------------------------------
MacKinnon approximate p-value for Z(t) = 0.0000

------------------------------------------------------------------------------
D.           |
dlnpind_euro |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
dlnpind_euro |
         L1. |  -.9466037   .0787851   -12.02   0.000    -1.102095   -.7911124
             |
       _cons |   .0013178   .0003365     3.92   0.000     .0006537     .001982
------------------------------------------------------------------------------
```

Some issues in testing for a unit root

- Consider a **break** in the series: A one-time change in the mean of a series (e.g. collapse of Lehman Brothers in 2008). Such a shock would bias conclusions towards a unit root.
- We will consider how to test for a break in the next lecture.
- Consider a large (few large) outlier. The series may then look as if it mean-reverting although it is not. Test results may be biased towards stationarity.

Summary on trends in time series

- The random walk model is the workhorse model for trends in economic time series.
- Always first plot the series.
- Then compute the Dickey-Fuller test (either with or without a trend).
- If Y_t has a unit root, transform data by differencing, i.e., use ΔY_t.
- If Y_t does not have a unit root, move ahead with your analysis.