ECON4150 - Introductory Econometrics

Lecture 15: Introduction to time series

Monique de Haan
(moniqued@econ.uio.no)

Stock and Watson Chapter 14.1-14.6

Lecture outline

- What is time series data
- Estimating a causal effect vs forecasting
- Lags, first differences and growth rates
- Autocorrelation
- Autoregressions
- Auto regressive distributed lag model
- Nonstationarity: stochastic trends
  - random walk with and without drift
  - testing for stochastic trends (Dickey-Fuller test)

What is time series data?

Cross section data is data collected for multiple entities at one point in time

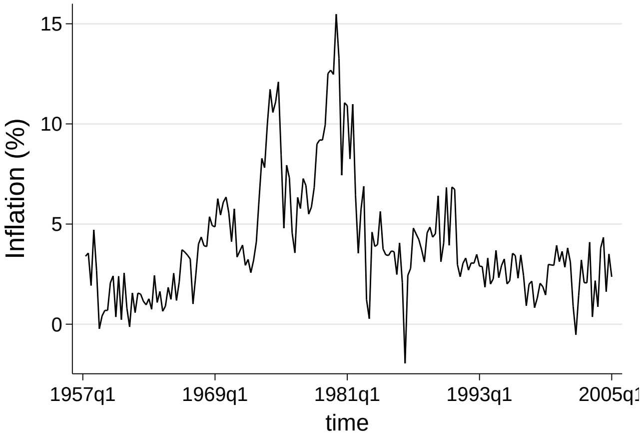
Panel data is data collected for multiple entities at multiple points in time

Time series data is data collected for a single entity at multiple points in time

- Yearly GDP of Norway for a period of 20 years
- Daily NOK/Euro exchange rate for past year
- Quarterly data on the inflation & unemployment rate in the U.S from 1957-2005

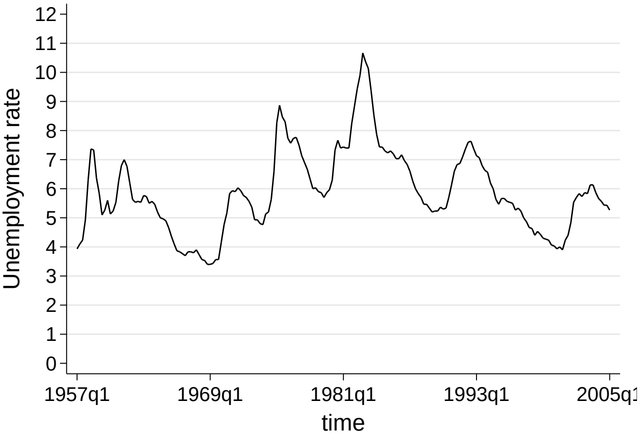
What is time series data?

Quarterly time series data on inflation for the U.S. from 1957 to 2005



What is time series data?

Quarterly time series data on unemployment for the U.S. from 1957-2005



What is time series data?

Quarterly data on inflation & unemployment in the U.S. 2000-2005:

Data Editor (Browse) - [macro_3e.dta]

File Edit View Data Tools

time[199]

	time	inflation	unemployme~e
173	2000q1	3.9386022	4.033333
174	2000q2	3.1231739	3.933333
175	2000q3	3.6388498	4
176	2000q4	2.8415304	3.9
177	2001q1	3.8080787	4.233333
178	2001q2	3.0958201	4.4
179	2001q3	.90159878	4.833333
180	2001q4	-.52568275	5.533333
181	2002q1	1.4252196	5.7
182	2002q2	3.2068807	5.833333
183	2002q3	2.0744651	5.733333
184	2002q4	2.0637621	5.866667
185	2003q1	4.0958529	5.833333
186	2003q2	.36364725	6.133333
187	2003q3	2.1751006	6.133333
188	2003q4	.86675908	5.866667
189	2004q1	3.8057824	5.666667
190	2004q2	4.3359083	5.566667
191	2004q3	1.6227116	5.433333
192	2004q4	3.5051154	5.433333
193	2005q1	2.3660429	5.266667

What is time series data?

- A particular observation Y_t indexed by subscript t
- Total number of observations equals T
- Y_t is current value and value in previous period is Y_{t-1} (first lag)
- In general Y_{t-j} is called jth lag and similarly, Y_{t+j} is the jth future value
- The first difference $Y_t - Y_{t-1}$ is the change in Y from period $t-1$ to period t
- Time series regression models can be used for
  - (1) estimating (dynamic) causal effects;
  - (2) forecasting

Estimating (dynamic) causal effect vs forecasting

- Time series data is often used for forecasting
  - For example next year's economic growth is forecasted based on past and current values of growth & other (lagged) explanatory variables
- Forecasting is quite different from estimating causal effects and is generally based on different assumptions.
- Models that are useful for forecasting need not have a causal interpretation!
  - OLS coefficients need not be unbiased & consistent
- Measures of fit, such as the (adjusted) R^2 or the SER
  - are not very informative when estimating causal effects
  - are informative about the quality of a forecasting model

Logarithms and growth rates

- Suppose that Y_t is some time series, then the rate of growth of Y is

$$\frac{Y_t - Y_{t-1}}{Y_{t-1}} = \frac{\triangle Y_t}{Y_{t-1}}$$

- instead, we often use the “logarithmic growth" or “log-difference":

$$\begin{aligned} \triangle ln Y_t &= ln(Y_t) - ln(Y_{t-1}) \\ &= ln\left(\frac{Y_t}{Y_{t-1}}\right) = ln\left(\frac{Y_t}{Y_{t-1}} + \frac{Y_{t-1}}{Y_{t-1}} - \frac{Y_{t-1}}{Y_{t-1}}\right) \\ &= ln\left(1 + \frac{Y_t - Y_{t-1}}{Y_{t-1}}\right) \approx \frac{\triangle Y_t}{Y_{t-1}} \end{aligned}$$

see also formula (8.16) of S&W

Example: inflation in the U.S.

Annualized rate of inflation:

$$inflation_t \cong 100 \times 4 \times (ln(Y_t) - ln(Y_{t-1}))$$

Stata:

```
tsset time
gen ln_CPI=ln(CPI)
gen ln_CPI_1stlag=ln(L1.CPI)
gen inflation=400*( ln(CPI) - ln(L1.CPI) )
```

time	CPI	ln_CPI	ln_CPI_1stlag	inflation
t	Y_t	$ln(Y_t)$	$ln(Y_{t-1})$	$400 \cdot (ln(Y_t) - ln(Y_{t-1}))$
1957q1	27.77667	3.3241963		
1957q2	28.01333	3.3326806	3.3241963	3.3937128
1957q3	28.26333	3.3415654	3.3326806	3.553894
1957q4	28.4	3.3463891	3.3415654	1.9295103
1958q1	28.73667	3.3581739	3.3463891	4.7138908
1958q2	28.93	3.3648791	3.3581739	2.6821083
1958q3	28.91333	3.3643029	3.3648791	-0.23050418
1958q4	28.94333	3.3653399	3.3643029	0.41480139

Autocorrelation

In time series data, Y_t is typically correlated with Y_{t-j}, this is called **autocorrelation** or **serial correlation**

- The j^{th}autocovariance=$Cov(Y_t, Y_{t-j})$ can be estimated by

$$\widehat{Cov(Y_t, Y_{t-j})} = \frac{1}{T}\sum_{t=j+1}^{T}\left(Y_t - \overline{Y}_{j+1,T}\right)\left(Y_{t-j} - \overline{Y}_{1,T-j}\right)$$

$\overline{Y}_{j+1,T}$ is the sample average of Y computed over observations $t = j+1, ..., T$

$\overline{Y}_{1,T-j}$ is the sample average of Y computed over observations $t = 1, ..., T-j$

- The j^{th}autocorrelation$= \rho_j = \frac{Cov(Y_t, Y_{t-j})}{Var(Y_t)}$ can be estimated by

$$\widehat{\rho}_j = \frac{\widehat{Cov(Y_t, Y_{t-j})}}{\widehat{Var\left(Y_t\right)}}$$

- j start-up observations are lost in constructing these sample statistics
- denominator of j^{th}autocorrelation assumes stationarity of Y_t, which implies (among other things) $Var\left(Y_t\right) = Var\left(Y_{t-j}\right)$

Autocorrelation

First 4 autocorrelations of inflation (inf_t):

```
. corrgram inflation if tin(1960q1,2004q4), noplot lags(4)
```

LAG	AC	PAC	Q	Prob>Q
1	**0.8359**	**0.8361**	**127.89**	**0.0000**
2	**0.7575**	**0.1937**	**233.49**	**0.0000**
3	**0.7598**	**0.3206**	**340.34**	**0.0000**
4	**0.6699**	**-0.1881**	**423.87**	**0.0000**

First 4 autocorrelations of the change in inflation ($\triangle inf_t = inf_t - inf_{t-1}$):

```
. corrgram D.inflation if tin(1960q1,2004q4), noplot lags(4)
```

LAG	AC	PAC	Q	Prob>Q
1	**-0.2618**	**-0.2636**	**12.548**	**0.0004**
2	**-0.2549**	**-0.3497**	**24.507**	**0.0000**
3	**0.2938**	**0.1461**	**40.481**	**0.0000**
4	**-0.0605**	**-0.0220**	**41.162**	**0.0000**

Stationarity

- The denominator of j^{th}autocorrelation assumes stationarity of Y_t
- A time series Y_t is stationary if its probability distribution does not change over time,
  - when the joint distribution of $(Y_{s+1}, ..., Y_{S+T})$ does not depend on s
- Stationarity implies that Y_1 has the same distribution as Y_t for any $t = 1, 2, ...$
- In other words, $\{Y_1, Y_2, ..., Y_T\}$ are identically distributed, however, they are not necessarily independent!
- If a series is nonstationary, then convential hypothesis tests, confidence intervals and forecasts can be unreliable.
- Stationarity says that history is relevant, it is a key requirement for external validity of time series regression.

First order autoregressive model: AR(1)

- Suppose we want to forecast the change in inflation from this quarter to the next
- When predicting the future of a time series a good place start is in the immediate past.
- The first order autoregressive model (AR(1))

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

- Forecast in next period based on AR(1) model:

$$\widehat{Y}_{T+1|T} = \widehat{\beta}_0 + \widehat{\beta}_1 Y_T$$

- Forecast error is the mistake made by the forecast

$$\textit{Forecast error} = Y_{T+1} - \widehat{Y}_{T+1|T}$$

Forecast vs predicted value & forecast error vs residual

- A forecast is not the same as a predicted value
- A forecast error is not the same as a residual

OLS predicted values $\widehat{Y}_t$ and residuals $\widehat{u}_t = Y_t - \widehat{Y}_t$ for $t \leq T$ are “in-sample”:

- They are calculated for the observations in the sample used to estimate the regression.
  - Y_t is observed in the data set used to estimate the regression.

A forecast $\widehat{Y}_{T+j|T}$ and forecast error $Y_{T+j} - \widehat{Y}_{T+j|T}$ for $j \geq 1$ are “out-of-sample”:

- They are calculated for some date beyond the data set used to estimate the regression.
  - Y_{T+j} is not observed in the data set used to estimate the regression.

First order autoregressive model: AR(1)

$$\triangle inflation_t = \beta_0 + \beta_1 \triangle inflation_{t-1} + u_t$$

```
> gen d_inflation=D1.inflation
(2 missing values generated)

1 . regress d_inflation L1.d_inflation if tin(1962q1,2004q4), r

Linear regression                                      Number of obs =        172
                                                          F(  1,   170) =       6.08
                                                          Prob > F      =     0.0146
                                                          R-squared     =     0.0564
                                                          Root MSE      =      1.664
```

d_inflation	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
d_inflation L1.	-.2380471	.0965017	-2.47	0.015	-.4285431	-.047551
_cons	.0171008	.1268849	0.13	0.893	-.2333721	.2675736

```
2 . dis "Adjusted Rsquared = " _result(8)
Adjusted Rsquared = .05082857
```

- $\widehat{\triangle inf}_{2005q1|2004q4} = 0.017 - 0.238 \cdot \triangle inf_{2004q4} = -0.43$
- $Forecast\ error = \triangle inf_{2005q1} - \widehat{\triangle inf}_{2005q1|2004q4} = -1.14 - (-0.43) = -0.71$

Root mean squared forecast error

- Forecasts are uncertain and the Root Mean Squared Forecast Error (RMSFE) is a measure of forecast uncertainty.
- The RMSFE is a measure of the spread of the forecast error distribution.

$$RMSFE = \sqrt{E\left[\left(Y_{T+1} - \widehat{Y}_{T+1|T}\right)^2\right]}$$

- The RMSFE has two sources of error:
  1. The error arising because future values of u_t are unknown
  2. The error in estimating the coefficients β_0 and β_1
- If the sample size is large the first source of error will be larger than the second and $RMSFE \approx \sqrt{Var\left(u_t\right)}$
- $\sqrt{Var\left(u_t\right)}$ can be estimated by the $SER = \frac{1}{T-2}\sum_{t=1}^{T}\widehat{u}_t^2$.

AR(p) model

The *pth* order autoregressive model (AR(p)) is

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

- The AR(p) model uses p lags of Y as regressors
- The number of lags p is called the order or lag length of the autoregression.
- The coefficients generally do not have a causal interpretation.
- We can use t- or F-tests to determine the lag order p
- Or we can determine p using an “information criterion” (more on this later…)

AR(4)

```
> regress d_inflation L1.d_inflation L2.d_inflation L3.d_inflation L4.d_inflation if
>  tin(1962q1,2004q4), r

Linear regression                                      Number of obs =        172
                                                          F(  4,   167) =       7.93
                                                          Prob > F      =     0.0000
                                                          R-squared     =     0.2038
                                                          Root MSE      =     1.5421
```

d_inflation	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
d_inflation						
L1.	-.2579426	.0925934	-2.79	0.006	-.4407471	-.0751381
L2.	-.3220312	.0805465	-4.00	0.000	-.4810518	-.1630106
L3.	.1576089	.0841017	1.87	0.063	-.0084307	.3236484
L4.	-.0302511	.0930471	-0.33	0.746	-.2139512	.153449
_cons	.0224294	.1176344	0.19	0.849	-.2098127	.2546715

```
. dis "Adjusted Rsquared = " _result(8)
Adjusted Rsquared = .18475367
```

- Adjusted R^2 is higher and the RMSE is lower in AR(4) than in AR(1)
- $\widehat{\triangle inf}_{2005q1|2004q4} \cong 0.41$
- $\textit{Forecast error} = \triangle inf_{2005q1} - \widehat{\triangle inf}_{2005q1|2004q4} = -1.14 - (0.4) = -1.55$

Lag length selection

- Is the AR(4) model better that the AR(1) model?

```
. test L2.d_inflation=L3.d_inflation=L4.d_inflation=0

 ( 1)  L2.d_inflation - L3.d_inflation = 0
 ( 2)  L2.d_inflation - L4.d_inflation = 0
 ( 3)  L2.d_inflation = 0

       F(  3,   167) =      6.71
            Prob > F =      0.0003
```

How should we choose the lag length *p*?

- One approach is to start with a model with many lags and to perform a hypothesis test on the final lag
- Delete the final lag if insignificant and perform an hypothesis test on the new final lag,..., continue until all included lags are significant.
- Drawback of this approach is that it can produce too large a model
  - at a 5% significance level: if the true lag length is 5 it will estimate *p* to be 6 in 5% of the time.

Lag length selection

Alternative way to determine the lag length p is to minimize one of the following information criteria:

- **Bayes information criterion (BIC):**

$$BIC(p) = ln\left[\frac{SSR(p)}{T}\right] + (p+1)\frac{ln(T)}{T}$$

- **Akaike information criterion (AIC):**

$$AIC(p) = ln\left[\frac{SSR(p)}{T}\right] + (p+1)\frac{2}{T}$$

- $SSR(p)$ is $\sum_{t=1}^{T} \widehat{u}_t^2$ in an AR(p) model
- T is the number of time periods
- In order to compare the BIC (or AIC) for different p,...
- ....all autoregressions with different lag lengths p should be based on the same number of observations T!

Lag length selection

The BIC and AIC both consist of two terms:

- 1st term $ln\left[\frac{SSR(p)}{T}\right]$: always decreasing in p
  - larger p, better fit
- 2nd term $(p+1)\frac{ln(T)}{T}$ (BIC) or $(p+1)\frac{2}{T}$ (AIC): always increasing in p.
  - This term is a "penalty" for estimating more parameters – and thus increasing the RMSFE.

AIC estimates more lags (larger p) than the BIC for $T > 7.4$,

- the penalty term is smaller for AIC than BIC

In large samples the AIC overestimates p, it is inconsistent.

Lag length selection

```
1 . regress d_inflation L1.d_inflation if tin(1962q1,2004q4), r

  Linear regression                                      Number of obs =        172
                                                            F(  1,   170) =       6.08
                                                            Prob > F      =     0.0146
                                                            R-squared     =     0.0564
                                                            Root MSE      =      1.664
```

d_inflation	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
d_inflation						
L1.	-.2380471	.0965017	-2.47	0.015	-.4285431	-.047551
_cons	.0171008	.1268849	0.13	0.893	-.2333721	.2675736

```
2 . gen BIC_1=ln(e(rss)/e(N))+e(rank)*(ln(e(N))/e(N)) if tin(1962q1,2004q4)
  (21 missing values generated)

3 . gen AIC_1=ln(e(rss)/e(N))+e(rank)*(2/e(N)) if tin(1962q1,2004q4)
  (21 missing values generated)

4 . sum BIC_1 AIC_1
```

Variable	Obs	Mean	Std. Dev.	Min	Max
BIC_1	172	1.066562	0	1.066562	1.066562
AIC_1	172	1.029963	0	1.029963	1.029963

Lag length selection

BIC			AIC		
. sum BIC*			. sum AIC*		
Variable	Obs	Mean	Variable	Obs	Mean
BIC_0	172	1.094665	AIC_0	172	1.076366
BIC_1	172	1.066562	AIC_1	172	1.029963
BIC_2	172	.9549263	AIC_2	172	.9000281
BIC_3	172	.9574141	AIC_3	172	.8842165
BIC_4	172	.9864399	AIC_4	172	.894943

- Optimal lag length according to BIC: $p = 2$
- Optimal lag length according to AIC: $p = 3$

Autoregressive Distributed Lag Model (ADL(p,q))

- Economic theory often suggests other variables that could help to forecast the variable of interest.

- When we add other variables and their lags the result is an

autoregressive distributed lag model ADL(p,q)

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + ... + \beta_p Y_{t-p} + \delta_1 X_{t-1} + ... + \delta_q X_{t-q} + u_t$$

- p is the number of lags of the dependent variable

- q is the number of lags (distributed lags) of the additional predictor X

Autoregressive Distributed Lag Model (ADL(p,q))

- When predicting future changes in inflation economic theory suggests that lagged values of the unemployment rate might be a good predictor
- Short-run Philips curve: negative short run relation between unemployment and inflation
- According to Bays Information Criteria we should include 2 lags of the dependent variable
- In addition we include 2 lags of the unemployment rate
- This gives an ADL(2,2) model

Autoregressive Distributed Lag Model (ADL(p,q))

```
> regress d_inflation L1.d_inflation L2.d_inflation L1.unemployment L2.unemployment if
> tin(1962q1,2004q4), r

Linear regression                                      Number of obs =       172
                                                          F(  4,   167) =     15.41
                                                          Prob > F      =    0.0000
                                                          R-squared     =    0.3514
                                                          Root MSE      =    1.3918

------------------------------------------------------------------------------------
                 |             Robust
     d_inflation |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-----------------+------------------------------------------------------------------
     d_inflation |
             L1. |  -.4685035    .0771115    -6.08   0.000    -.6207425   -.3162645
             L2. |  -.4251441    .0821394    -5.18   0.000    -.5873096   -.2629787
                 |
unemployment_rate|
             L1. |  -2.243865    .4020402    -5.58   0.000    -3.037602   -1.450129
             L2. |   2.044221    .3875693     5.27   0.000     1.279054    2.809388
                 |
           _cons |   1.193032    .4344711     2.75   0.007     .3352683    2.050796
------------------------------------------------------------------------------------

1 . dis "Adjusted Rsquared = " _result(8)
Adjusted Rsquared = .3358903
```

- Adjusted R^2 is higher and the RMSE is lower in ADL(2,2) than in AR(4)
- $\textit{Forecast error} = \triangle inf_{2005q1} - \widehat{\triangle inf}_{2005q1|2004q4} = -1.14 - (0.38) = -1.52$

Granger "causality" test

- Do the included lags of unemployment have useful predictive content conditional on the included lags of the change in inflation?
- The claim that a variable has no predictive content corresponds to the null hypothesis that the coefficients on all lags of the variable are zero.
- The F-statistic of this test is called the Granger causality statistic.
- If the null hypothesis is rejected the variable X is said to *Granger-cause* the dependent variable Y.
- This does not mean that we have estimated the causal effect of X on Y!!
- It means that X is a useful predictor of Y (Granger predictability would be a better term).

Granger "causality" test

The Granger causality test in the ADL(2,2) model:

```
. regress d_inflation L1.d_inflation L2.d_inflation L1.unemployment L2.unemployment if
> tin(1962q1,2004q4), r noheader
```

d_inflation	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
d_inflation						
L1.	-.4685035	.0771115	-6.08	0.000	-.6207425	-.3162645
L2.	-.4251441	.0821394	-5.18	0.000	-.5873096	-.2629787
unemployment_rate						
L1.	-2.243865	.4020402	-5.58	0.000	-3.037602	-1.450129
L2.	2.044221	.3875693	5.27	0.000	1.279054	2.809388
_cons	1.193032	.4344711	2.75	0.007	.3352683	2.050796

```
. test L1.unemployment=L2.unemployment=0

 ( 1)  L.unemployment_rate - L2.unemployment_rate = 0
 ( 2)  L.unemployment_rate = 0

       F(  2,   167) =    16.13
            Prob > F =    0.0000
```

- Null hypothesis that coefficients on the 2 lags of unemployment are zero is rejected at a 1% level.
- Unemployment is a useful predictor for the change in the inflation rate.

Nonstationarity: trends

- A time series Y_t is stationary if its probability distribution does not change over time
- If a time series has a trend, it is nonstationary
- A trend is a persistent long-term movement of a variable over time.
- We consider two types of trends

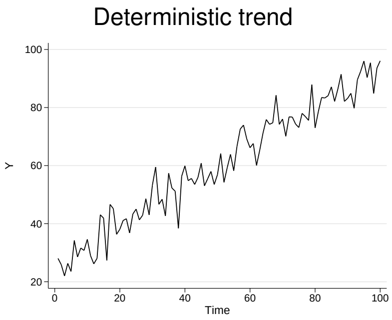
Deterministic trend: $Y_t = \beta_0 + \lambda t + u_1$

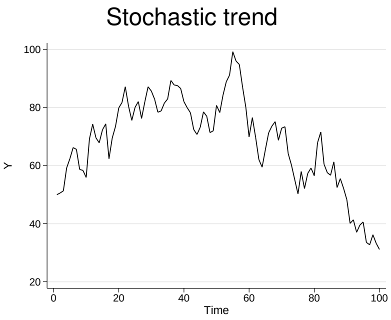
- series is a nonrandom function of time

Stochastic trend: $Y_t = \beta_0 + Y_{t-1} + u_1$

- series is a random function of time

Nonstationarity: trends





Random walk model

- Simplest model of a variable with a stochastic trend is the random walk

$$Y_t = Y_{t-1} + u_t \quad \textit{where } u_t \textit{ is i.i.d.}$$

- The value of the series tomorrow is its value today plus an unpredictable change.
- An extension of the random walk model is the random walk with drift

$$Y_t = \beta_0 + Y_{t-1} + u_t \quad \textit{where } u_t \textit{ is i.i.d.}$$

- β_0 is the "drift" of the random walk, if β_0 is positive Y_t increases on average.
- A random walk is nonstationary: the distribution is not constant over time
- The variance of a random walk increases over time:

$$Var\left(Y_t\right) = Var\left(Y_{t-t}\right) + Var\left(u_t\right)$$

Stochastic trends, autoregressive models and a unit root

- The random walk model is a special case of an AR(1) model with $\beta_1 = 1$
- If Y_t follows and AR(1) with $|\beta_1| < 1$ (and u_t is stationary), Y_t is stationary
- If Y_t follows and AR(p) model

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

Y_t is stationary if its *roots* z are all greater than 1 in absolute value.

- the roots are the values of z that satisfy

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

- In the special case of an AR(1): $1 - \beta_1 z = 0$ gives $z = \frac{1}{\beta_1}$ which is bigger than $|1|$ if $|\beta_1| < 1$

Detecting stochastic trends: Dickey-Fuller test in AR(1) model

- Trends can be detected by informal and formal methods
  - **Informal:** inspect the time series plot
  - **Formal:** Perform the Dickey-Fuller test to test for a stochastic trend.
- Dickey-Fuller test in an AR(1) model:

$$H_0 : \beta_1 = 1 \quad vs \quad H_1 : \beta_1 < 1 \quad in \; Y_t = \beta_0 + \beta_1 Y_{t-1} + u_t$$

- Test can be performed by subtracting Y_{t-1} from both sides of the equation and estimate

$$\triangle Y_t = \beta_0 + \delta Y_{t-1} + \varepsilon_t$$

- We can now test

$$H_0 : \delta = \beta_1 - 1 = 0 \quad vs \quad H_1 : \delta < 0$$

Does U.S. inflation have a stochastic trend?

- DF test for a unit root in U.S. inflation (*Note*: we test for a stochastic trend in $inflation_t$ and not in $\triangle inflation_t$)

```
1 . regress d_inflation L1.inflation if tin(1962q1,2004q4), noheader
```

d_inflation	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
inflation L1.	-.1643553	.0415311	-3.96	0.000	-.2463383	-.0823722
_cons	.7219345	.217599	3.32	0.001	.2923904	1.151479

- Under H_0 , Y_t is nonstationary and the DF-statistic has a nonnormal distribution, we therefore use the following critical values

TABLE 14.5 Large-Sample Critical Values of the Augmented Dickey–Fuller Statistic

Deterministic Regressors	10%	5%	1%
Intercept only	−2.57	−2.86	−3.43
Intercept and time trend	−3.12	−3.41	−3.96

- $DF = -3.96$ this is more negative than -2.86 so we reject the null hypothesis of a stochastic trend at a 5% significance level.

Detecting stochastic trends: Dickey-Fuller test in AR(p) model

- Often an AR(1) model does not capture all the serial correlation in Y_t and we should include more lags & estimate an AR(p) model.
- We can test for a stochastic trend in an AR(p) model by augmenting the DF-regression by lags of $\triangle Y_t$.
- The Augmented Dickey-Fuller test:

$$H_0 : \delta = 0 \qquad vs \quad H_1 : \delta < 0$$

in the regression

$$\triangle Y_t = \beta_0 + \delta Y_{t-1} + \gamma_1 \triangle Y_{t-1} + \gamma_2 \triangle Y_{t-2} + ... + \gamma_p \triangle Y_{t-p} + u_t$$

- *Note:* The DF-statistic should be computed using homoskedasticity-only (nonrobust) standard errors (see footnote 3 in S&W CH14)

Does U.S. inflation have a stochastic trend?

- DF test for a unit root in U.S. inflation – using p = 4 lags (AR(4)-model)

```
1 . regress d_inflation L1.inflation L1.d_inflation L2.d_inflation L3.d_inflation L4.d_in
> flation if tin(1962q1,2004q4), noheader
```

d_inflation	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
inflation						
L1.	-.1134169	.0422344	-2.69	0.008	-.1968029	-.030031
d_inflation						
L1.	-.1864426	.0805144	-2.32	0.022	-.3454068	-.0274783
L2.	-.2563879	.081463	-3.15	0.002	-.4172251	-.0955507
L3.	.1990491	.0793514	2.51	0.013	.0423811	.3557171
L4.	.0099994	.0779921	0.13	0.898	-.1439849	.1639837
_cons	.5068158	.2141807	2.37	0.019	.0839466	.9296851

- $DF = -2.69$ this is less negative than -2.86 so we do not reject the null hypothesis of a stochastic trend at a 5% significance level.

Does U.S. inflation have a stochastic trend?

- Instead of testing the null hypothesis of a stochastic trend against the alternative hypothesis of no trend....
- ...the alternative hypothesis can be that Y_t is stationary around a deterministic trend.
- The Dickey-Fuller regression then includes a deterministic trend

$$\triangle Y_t = \beta_0 + \alpha t + \delta Y_{t-1} + \gamma_1 \triangle Y_{t-1} + \gamma_2 \triangle Y_{t-2} + ... + \gamma_p \triangle Y_{t-p} + u_t$$

- And we have to use the critical values in the second row:

TABLE 14.5 Large-Sample Critical Values of the Augmented Dickey–Fuller Statistic

Deterministic Regressors	10%	5%	1%
Intercept only	−2.57	−2.86	−3.43
Intercept and time trend	−3.12	−3.41	−3.96

Avoiding problems caused by stochastic trends

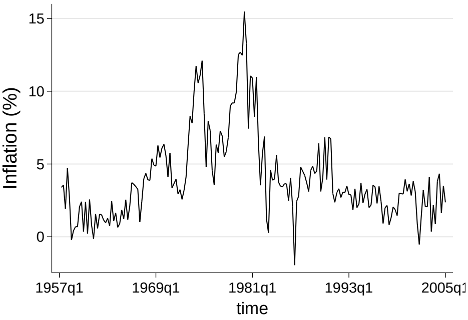
- Best way to deal with a trend is to transform the series such that it does not have a trend.
- If the series Y_t has a stochastic trend, then the first difference of the series $\triangle Y_t$ does not have a stochastic trend.
- For example if Y_t follows a random walk with drift

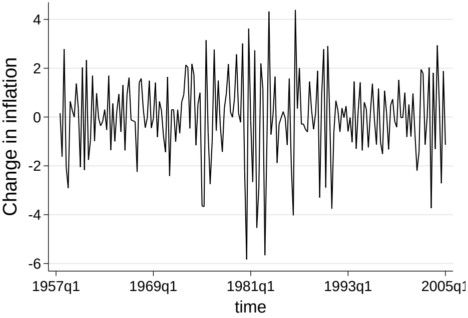
$$Y_t = \beta_0 + Y_{t-1} + u_t$$

the first difference is stationary

$$\triangle Y_t = \beta_0 + u_t$$

- On slide 37 we saw that we did not reject the null hypothesis of a stochastic trend in $\mathit{inflation}_t$.
- This is the reason that in the beginning of the lecture we estimated AR's with $\triangle \mathit{inflation}_t$.


Inflation (%)
15
10
5
0
1957q1
1969q1
1981q1
1993q1
2005q1
time


Change in inflation
4
2
0
-2
-4
-6
1957q1
1969q1
1981q1
1993q1
2005q1
time

Good luck with the exam!

(Don’t forget to bring the book and a calculator!)