

Time Series Models

Simulated AR Time Series

```
. set obs 500
obs was 0, now 500

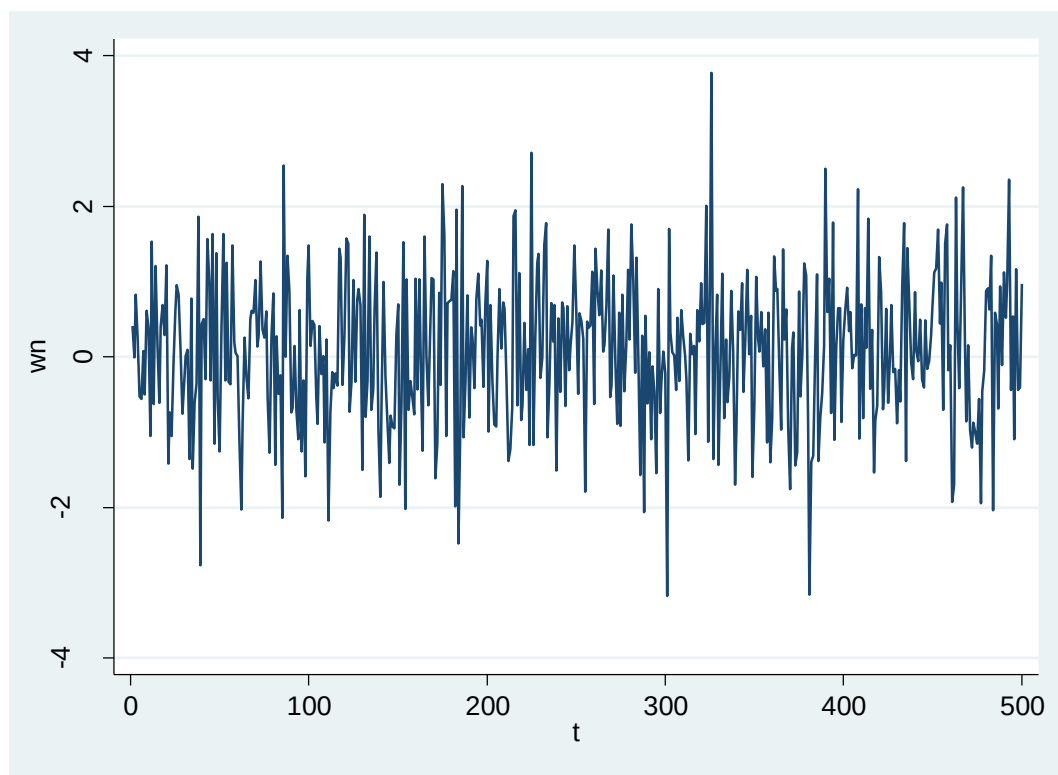
. g t=_n

. tsset t
    time variable: t, 1 to 500
        delta: 1 unit

. set seed 1234

. g wn=rnormal(0,1)

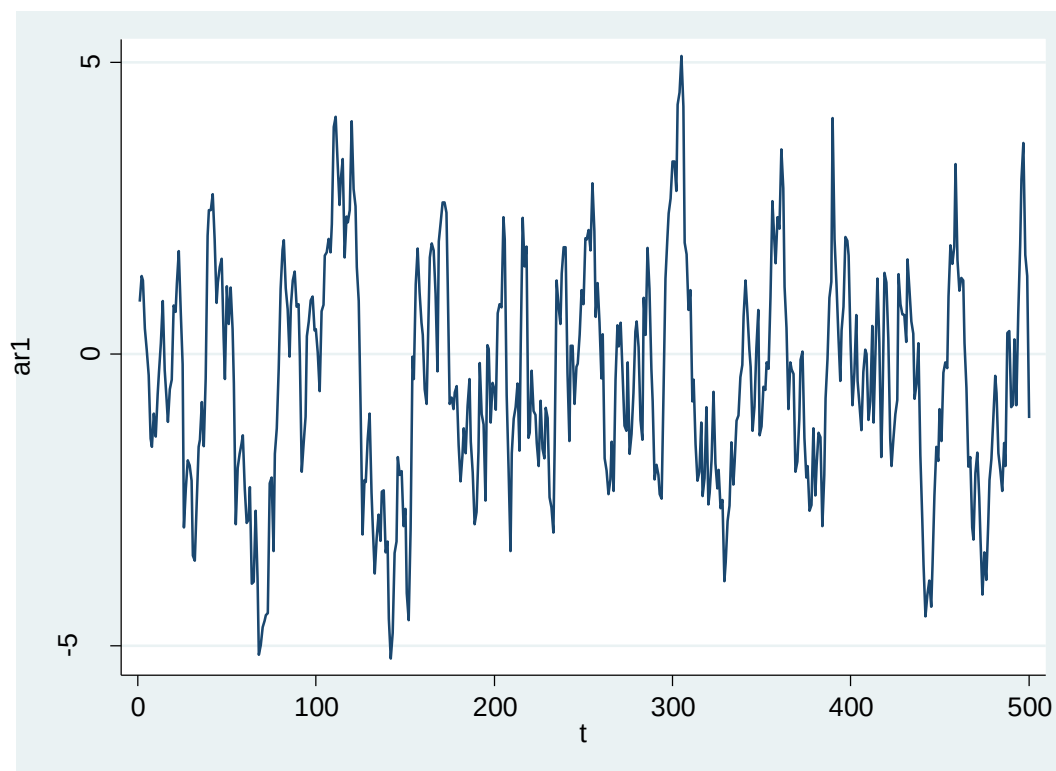
. line wn t
```



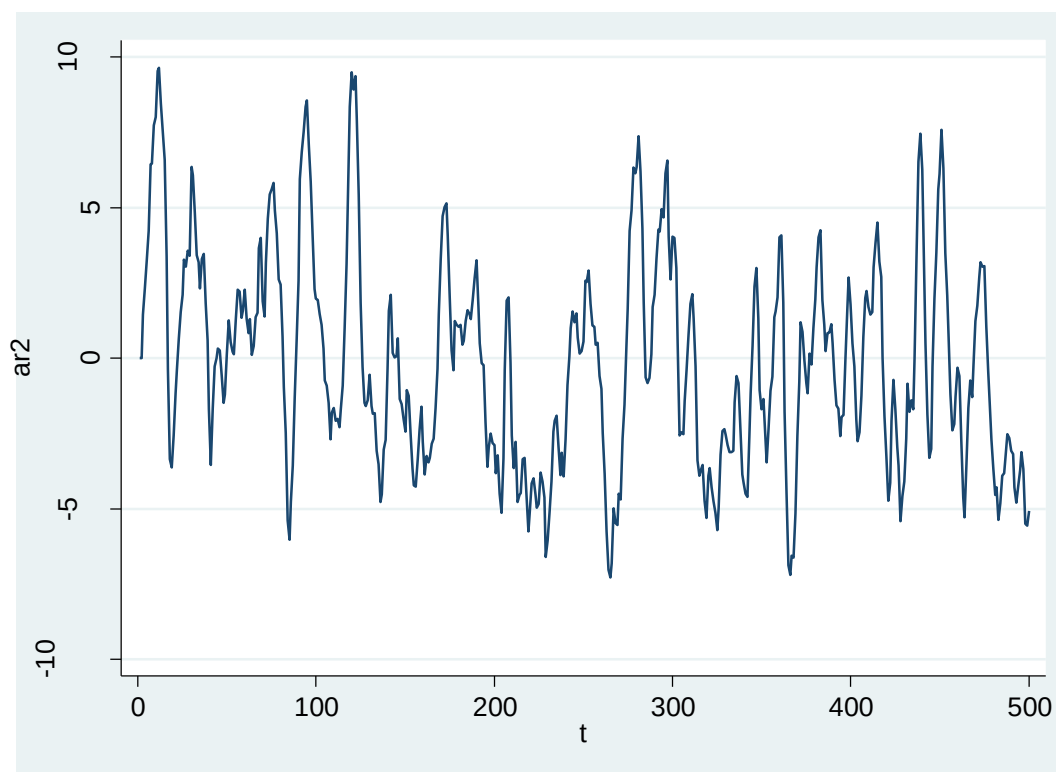
```
. g ar1=0.9 in 1
(499 missing values generated)

. replace ar1=0.9*1.ar1+rnormal(0,1) if t>1
(499 real changes made)

. line ar1 t
```



```
. g ar2=0 if t<3  
(498 missing values generated)  
  
. replace ar2=1.5*1.ar2-0.6*12.ar2+rnormal(0,1) if t>2  
(498 real changes made)  
  
. line ar2 t
```



Random Walk without Drift:

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

where: u_t is white noise error term with mean = 0 and variance = σ^2

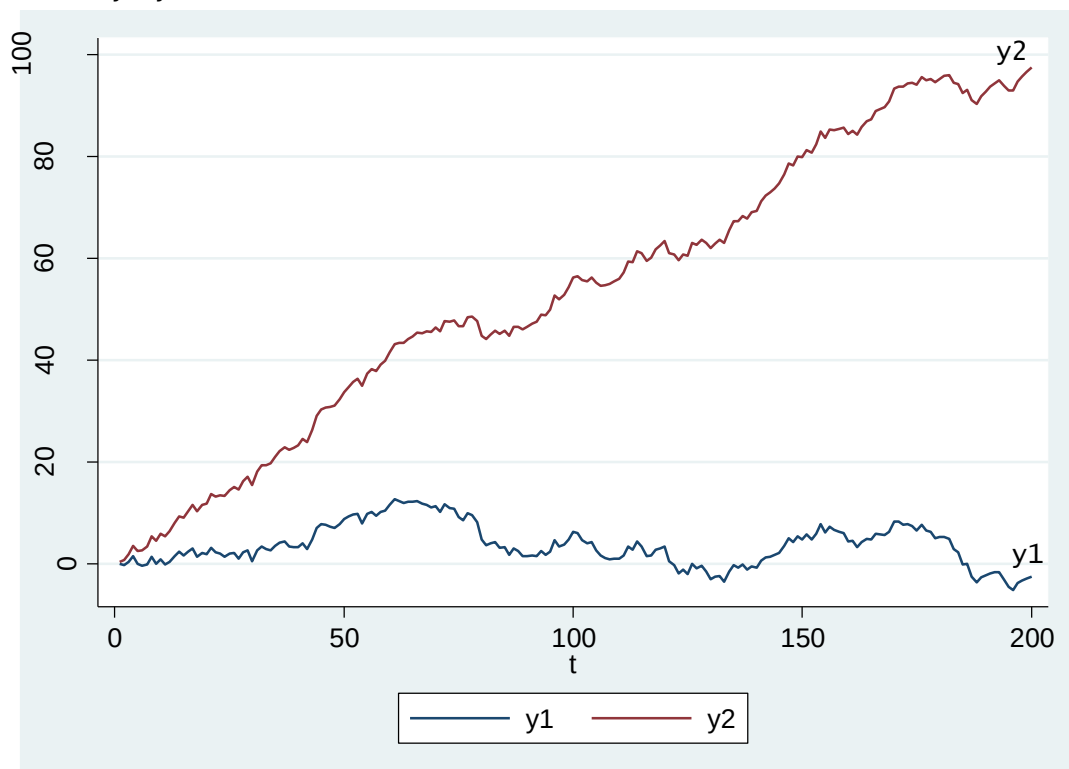
Random Walk with Drift:

$$Y_{2t} = \delta + Y_{2t-1} + u_t$$

```
. set obs 200
obs was 0, now 200

. g t=_n
. tsset t
    time variable: t, 1 to 200
    delta: 1 unit

. set seed 4567
. g u= rnormal(0,1)
. g y1=0 in 1
(199 missing values generated)
. g y2=0.5 in 1
(199 missing values generated)
. replace y1=1.y1 +u if t>1
(199 real changes made)
. replace y2=0.5+1.y2 + u if t>1
(199 real changes made)
. line y1 y2 t
```



Trend Stationary (TS) and Difference Stationary (DS)**Pure random walk**

$$\beta_1 = 0, \beta_2 = 0, \beta_3 = 1$$

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

Random walk with drift

$$\beta_1 \neq 0, \beta_2 = 0, \beta_3 = 1$$

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

Deterministic trend

$$\beta_1 \neq 0, \beta_2 \neq 0, \beta_3 = 0$$

$$Y_{3t} = \beta_1 + \beta_2 t + u_t$$

Random walk with drift and deterministic trend

$$\beta_1 \neq 0, \beta_2 \neq 0, \beta_3 = 1$$

$$Y_{4t} = \beta_1 + \beta_2 t + Y_{4t-1} + u_t$$

Deterministic trend with stationary AR(1) component

$$\beta_1 \neq 0, \beta_2 \neq 0, \beta_3 < 1$$

$$Y_{5t} = \beta_1 + \beta_2 t + \beta_3 Y_{5t-1} + u_t$$

```
. set obs 200
obs was 0, now 200

. g time=_n

. tsset time
    time variable:  time, 1 to 200
                delta:  1 unit

. g y1=0 in 1
(199 missing values generated)

. g y2=0 in 1
(199 missing values generated)

. g y3=0 in 1
(199 missing values generated)

. g y4=0 in 1
(199 missing values generated)

. g y5=0 in 1
(199 missing values generated)

. set seed 7890

. g u = rnormal(0,1)

. g t=_n

. replace y1=l.y1+u if time>1
(199 real changes made)

. replace y2=0.5+l.y2+u if time>1
(199 real changes made)

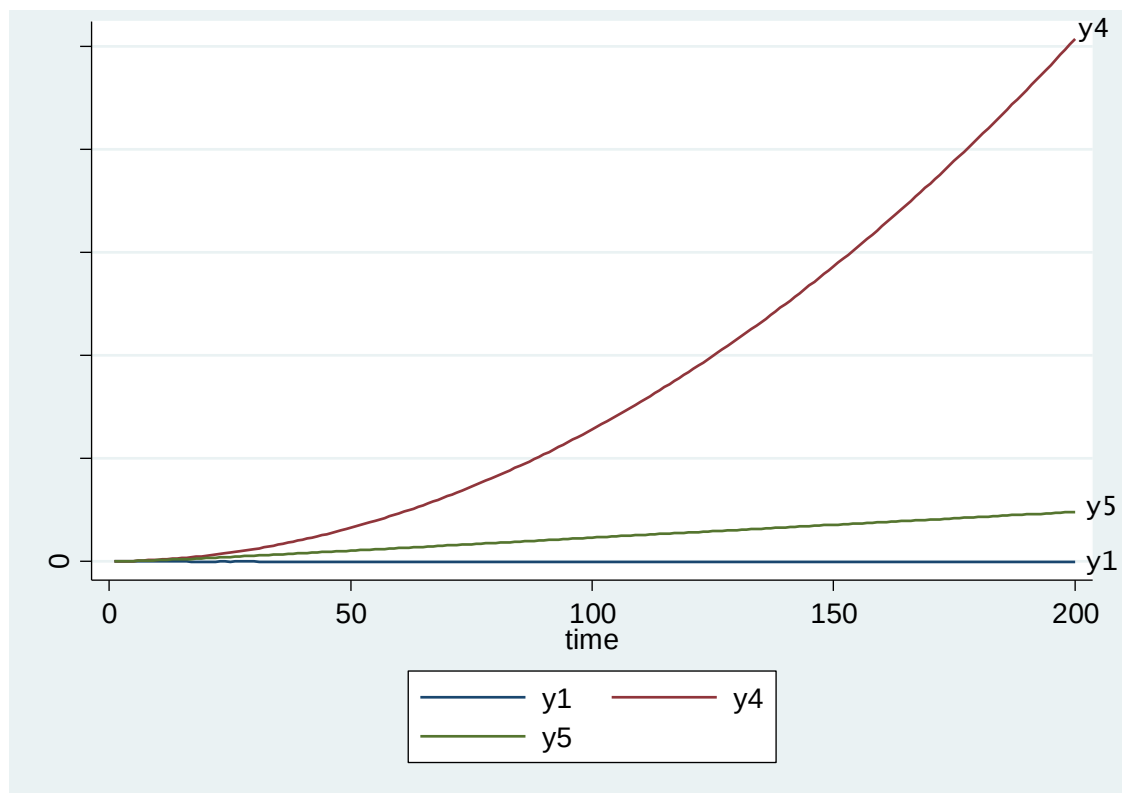
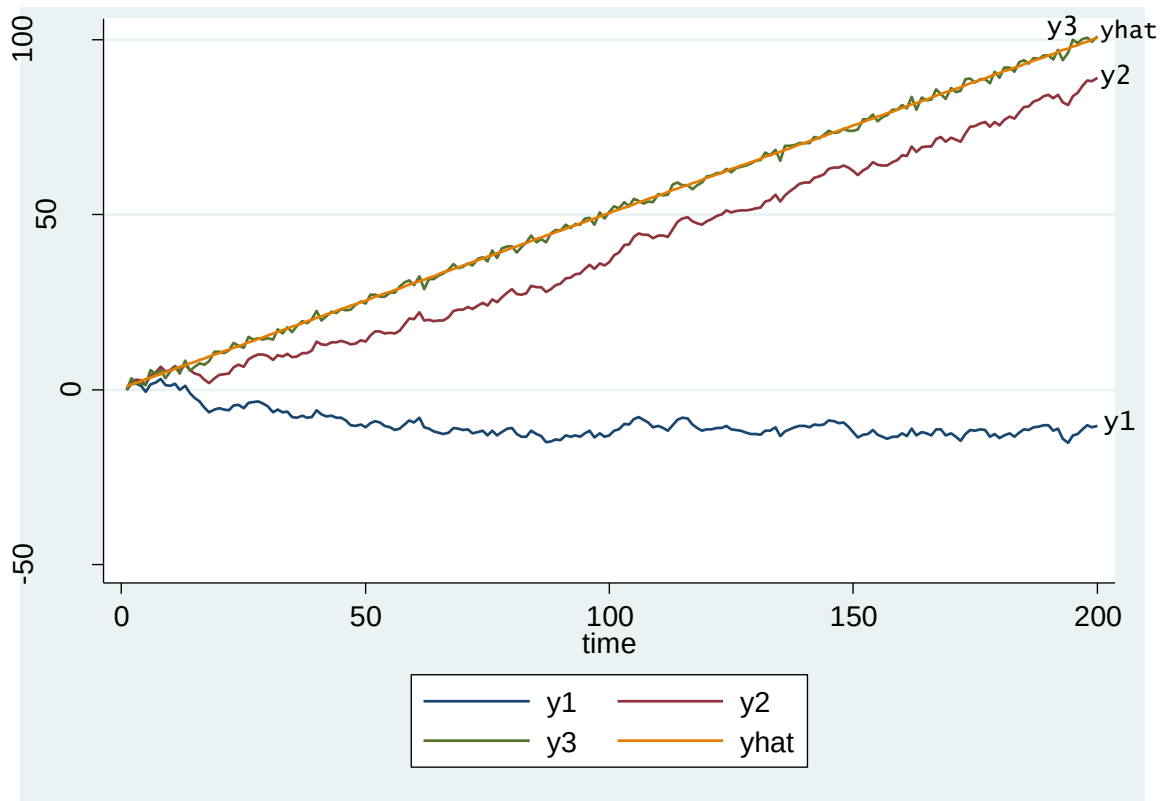
. replace y3=0.5+0.5*t+u if time>1
(199 real changes made)

. replace y4=0.5+0.5*t+l.y4+u if time>1
(199 real changes made)

. replace y5=0.5+0.5*t+0.9*l.y5+u if time>1
(199 real changes made)

. g yhat=0.5+0.5*t
```

```
. line y1 y2 y3 yhat time  
. line y1 y4 y5 time
```



Unit Root Problem**Autoregressive (Stationary) Series**

$$x_{1t} = 0.05 + 0.95 x_{1t-1} + u_{1t}$$

Random Walk with Drift

$$x_{2t} = 0.05 + x_{2t-1} + u_{2t}$$

Explosive Series

$$x_{3t} = 0.05 + 1.05 x_{3t-1} + u_{3t}$$

```
. set obs 500
obs was 0, now 500

. g time=_n

. tsset time
    time variable:  time, 1 to 500
                delta:  1 unit

. g x1=0 in 1
(499 missing values generated)

. g x2=0 in 1
(499 missing values generated)

. g x3=0 in 1
(499 missing values generated)

. set seed 0123

. g u1= rnormal(0,1)

. g u2= rnormal(0,1)

. g u3= rnormal(0,1)

. replace x1=0.05+0.95*_l.x1+u1 if time>1
(499 real changes made)

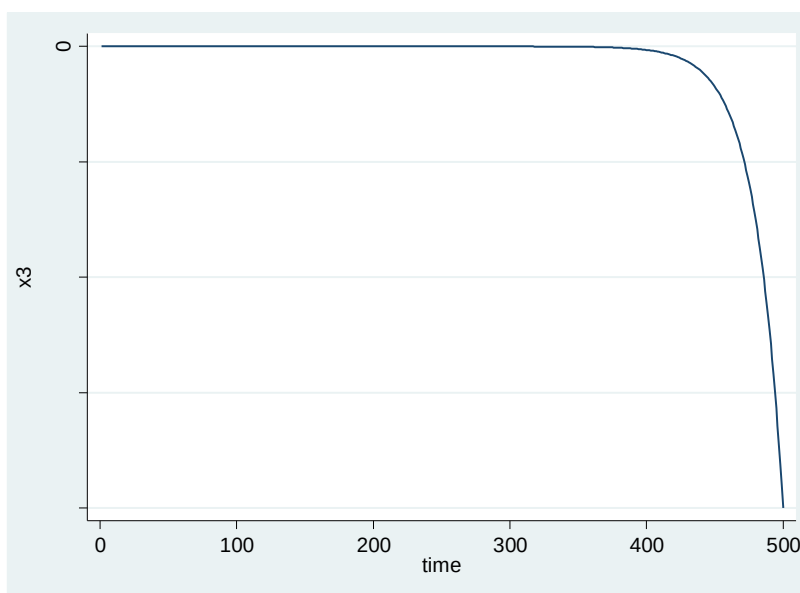
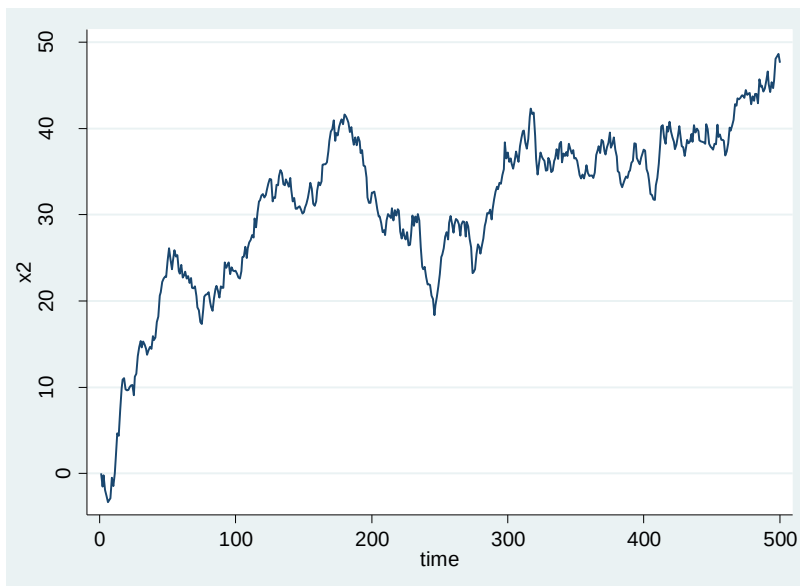
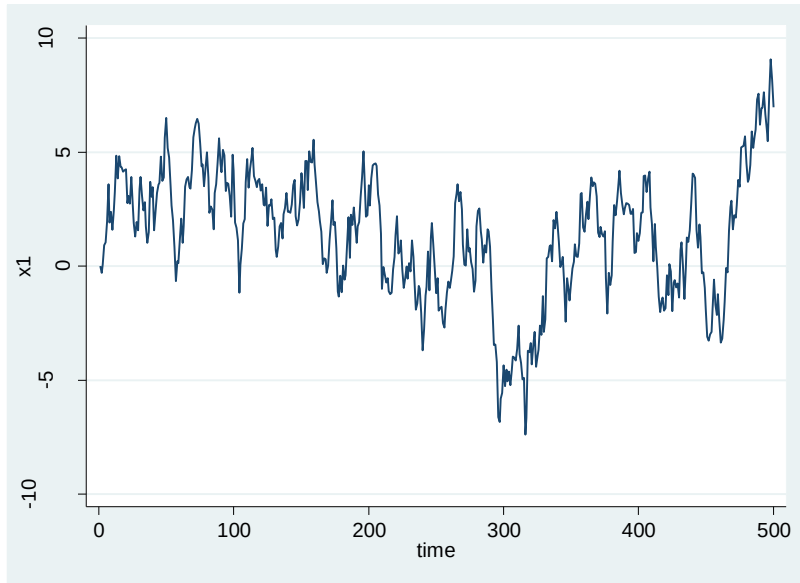
. replace x2=0.05+_l.x2+u2 if time>1
(499 real changes made)

. replace x3=0.05+1.05*_l.x3+u3 if time>1
(499 real changes made)

. line x1 time

. line x2 time

. line x3 time
```



Spurious Regression

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

$$X_t = X_{t-1} + v_t$$

If X_t and Y_t are uncorrelated $I(1)$ processes, regression $Y_t = \beta_1 + \beta_2 X_t + \varepsilon_t$ can lead to spurious problem.

```
. set obs 500
obs was 0, now 500

. g time=_n

. tsset time
    time variable:  time, 1 to 500
              delta: 1 unit

. g y=0 in 1
(499 missing values generated)

. g x=0 in 1
(499 missing values generated)

. set seed 5555

. g u= rnormal(0,1)

. g v= rnormal(0,1)

. replace y=l.y+u if time>1
(499 real changes made)

. replace x=l.x+v if time>1
(499 real changes made)

. regress y x
```

Source	SS	df	MS	Number of obs =	500
Model	32835.1366	1	32835.1366	F(1, 498) =	2267.84
Residual	7210.34126	498	14.4785969	Prob > F =	0.0000
				R-squared =	0.8199
				Adj R-squared =	0.8196
Total	40045.4779	499	80.2514587	Root MSE =	3.8051

y	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
x	-.630954	.0132493	-47.62	0.000	-.6569853	-.6049227
_cons	.4784655	.4633948	1.03	0.302	-.4319844	1.388915

```
. dwstat
```

Durbin-watson d-statistic(2, 500) = .0988012

Unit Root Test

```
. set obs 500
obs was 0, now 500

. g t=_n

. tsset t
    time variable:  t, 1 to 500
        delta: 1 unit

. set seed 5432

. g u=rnormal(0,10)

. g y=0 in 1
(499 missing values generated)

. replace y=1.y+u if t>1
(499 real changes made)
```

1. Test with all terms (intercept, trend, and lags):

```
. dfuller y, trend lags(1) regress
```

Augmented Dickey-Fuller test for unit root Number of obs = 498

	Test Statistic	----- 1% Critical Value	Interpolated Dickey-Fuller 5% Critical Value	----- 10% Critical Value
z(t)	-1.775	-3.980	-3.420	-3.130

Mackinnon approximate p-value for z(t) = 0.7166

D.y	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
y					
L1.	-.0150824	.0084961	-1.78	0.076	-.0317754 .0016106
LD.	-.0325012	.0450527	-0.72	0.471	-.1210196 .0560173
_trend	.0016688	.0035614	0.47	0.640	-.0053285 .0086662
_cons	-.0153098	.9211143	-0.02	0.987	-1.825095 1.794475

The result shows p-value=0.7166 (fail to reject null hypothesis of unit root). Then, check whether trend is significant and find out that p-value=0.640. Thus, there must be no trend. Go to next step:

2. Test with intercept and lags:

```
. dfuller y, lags(1) regress
```

Augmented Dickey-Fuller test for unit root Number of obs = 498

	Test Statistic	----- 1% Critical Value	Interpolated Dickey-Fuller 5% Critical Value	----- 10% Critical Value
z(t)	-1.775	-3.440	-2.870	-2.570

Mackinnon approximate p-value for z(t) = 0.3928

D.y	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
y					
L1.	-.0131265	.0073943	-1.78	0.076	-.0276546 .0014016
LD.	-.033918	.0449156	-0.76	0.451	-.1221668 .0543307
_cons	.3511295	.4863518	0.72	0.471	-.6044389 1.306698

The result shows p-value=0.3928 (fail to reject null hypothesis of unit root). Then, check

whether constant term is significant and find out that $p\text{-value}=0.471$. Thus, there must be no constant term. Go to next step:

3. Test with only lags:

```
. dfuller y, nocon lags(1) regress
```

Augmented Dickey-Fuller test for unit root Number of obs = 498

	Test Statistic	----- 1% Critical Value	Interpolated Dickey-Fuller 5% Critical Value	----- 10% Critical Value
z(t)	-1.623	-2.580	-1.950	-1.620

	D.y	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
	y					
	L1.	-.0109853	.0067702	-1.62	0.105	-.024287 .0023165
	LD.	-.0349845	.0448697	-0.78	0.436	-.1231425 .0531735

The result shows $|Z(t)| = 1.623$ is less than $|5\% \text{ Critical Value}| = 1.950$ then fail to reject null hypothesis of unit root. Therefore, the conclusion is there is unit root problem. The series is nonstationary.

```
. dfuller d.y, nocon lags(1) regress
```

Augmented Dickey-Fuller test for unit root Number of obs = 497

	Test Statistic	----- 1% Critical Value	Interpolated Dickey-Fuller 5% Critical Value	----- 10% Critical Value
z(t)	-15.366	-2.580	-1.950	-1.620

	D2.y	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
	y					
	LD.	-.9954347	.0647826	-15.37	0.000	-1.122717 -.8681519
	LD2.	-.043815	.0449231	-0.98	0.330	-.1320784 .0444484

The result shows $|Z(t)| = 15.366$ is greater than $|5\% \text{ Critical Value}| = 1.950$ then reject null hypothesis of unit root. Therefore, the conclusion is the series is $I(1)$.

Dickey-Fuller GLS Test

```
. dfgls y, maxlag(10)
```

DF-GLS for y Number of obs = 489

[lags]	DF-GLS tau Test Statistic	1% Critical Value	5% Critical Value	10% Critical Value
10	-2.099	-3.480	-2.849	-2.563
9	-1.839	-3.480	-2.852	-2.567
8	-1.949	-3.480	-2.856	-2.570
7	-1.938	-3.480	-2.859	-2.572
6	-1.828	-3.480	-2.862	-2.575
5	-1.746	-3.480	-2.865	-2.578
4	-1.814	-3.480	-2.868	-2.581
3	-1.732	-3.480	-2.871	-2.583
2	-1.919	-3.480	-2.874	-2.586
1	-1.817	-3.480	-2.876	-2.588

Opt Lag (Ng-Perron seq t) = 10 with RMSE 9.753337
 Min SC = 4.618856 at lag 1 with RMSE 9.941964
 Min MAIC = 4.607555 at lag 3 with RMSE 9.887703

```
. dfglr d.y, maxlag(10)
```

```
DF-GLS for D.y
```

```
Number of obs = 488
```

[lags]	DF-GLS tau Test Statistic	1% Critical Value	5% Critical Value	10% Critical Value
10	-5.103	-3.480	-2.849	-2.564
9	-5.327	-3.480	-2.852	-2.567
8	-6.433	-3.480	-2.856	-2.570
7	-6.491	-3.480	-2.859	-2.572
6	-6.995	-3.480	-2.862	-2.575
5	-8.084	-3.480	-2.865	-2.578
4	-9.441	-3.480	-2.868	-2.581
3	-10.421	-3.480	-2.871	-2.583
2	-12.964	-3.480	-2.874	-2.586
1	-14.522	-3.480	-2.876	-2.588

```
Opt Lag (Ng-Perron seq t) = 9 with RMSE 9.871131
Min SC = 4.647202 at lag 1 with RMSE 10.08366
Min MAIC = 5.597676 at lag 10 with RMSE 9.870346
```

Phillips-Perron Test

```
. pperron y, lags(1) regress
```

```
Phillips-Perron test for unit root
```

```
Number of obs = 499
Newey-West lags = 1
```

	Test Statistic	----- 1% Critical Value	Interpolated Dickey-Fuller 5% Critical Value	----- 10% Critical Value
Z(rho)	-6.527	-20.499	-14.000	-11.200
Z(t)	-1.808	-3.440	-2.870	-2.570

```
Mackinnon approximate p-value for Z(t) = 0.3765
```

	y	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
	y					
	L1.	.9864691	.0073583	134.06	0.000	.9720119 1.000926
	_cons	.3541393	.4848813	0.73	0.466	-.5985306 1.306809

```
. pperron d.y, lags(1) regress
```

```
Phillips-Perron test for unit root
```

```
Number of obs = 498
Newey-West lags = 1
```

	Test Statistic	----- 1% Critical Value	Interpolated Dickey-Fuller 5% Critical Value	----- 10% Critical Value
Z(rho)	-518.457	-20.498	-14.000	-11.200
Z(t)	-23.193	-3.440	-2.870	-2.570

```
Mackinnon approximate p-value for Z(t) = 0.0000
```

	D.y	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
	y					
	LD.	-.0404772	.0448604	-0.90	0.367	-.128617 .0476626
	_cons	.0048224	.4464789	0.01	0.991	-.8724007 .8820455

Cointegration Test - Augmented Engle-Granger (AEG) Test

Do file command in generating Cointegrating relationship:

```

set obs 500
g t=_n
tsset t
mat C=(1, 0.7\0.7, 1)
corr2data e1 e2, n(500) mean(0 0) sds(4 5) corr(C)
g x=0 in 1
g y=0 in 1
g dx=0 in 1
g dy=0 in 1
forvalue i=2(1)500 {
    replace dy=0.1+0.2*1.dy+0.3*1.dx-0.5*(1.y-0.1-0.5*1.x)+e1 in `i'
    replace dx=0.2+0.3*1.dy+0.2*1.dx+0.5*(1.y-0.1-0.5*1.x)+e2 in `i'
    replace y=1.y+dy in `i'
    replace x=1.x+dx in `i'
}

```

Unit Root Test to determine Order of Integration of the series

```
. dfuller x, lag(1) trend reg
```

Augmented Dickey-Fuller test for unit root Number of obs = 498

	Test Statistic	----- 1% Critical Value	Interpolated Dickey-Fuller 5% Critical Value	----- 10% Critical Value
Z(t)	-0.351	-3.980	-3.420	-3.130

Mackinnon approximate p-value for Z(t) = 0.9883

D.x	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
x						
L1.	-.0011524	.0032832	-0.35	0.726	-.0076031	.0052983
LD.	.5052605	.0392238	12.88	0.000	.4281944	.5823266
_trend	.0021787	.001759	1.24	0.216	-.0012772	.0056347
_cons	-.3239806	.5129904	-0.63	0.528	-1.331893	.6839315

```
. dfuller x, lag(1) reg
```

Augmented Dickey-Fuller test for unit root Number of obs = 498

	Test Statistic	----- 1% Critical Value	Interpolated Dickey-Fuller 5% Critical Value	----- 10% Critical Value
Z(t)	-0.593	-3.440	-2.870	-2.570

Mackinnon approximate p-value for Z(t) = 0.8725

D.x	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
x						
L1.	-.0019149	.0032267	-0.59	0.553	-.0082546	.0044248
LD.	.5109265	.0389772	13.11	0.000	.4343453	.5875076
_cons	.2299327	.2514875	0.91	0.361	-.2641818	.7240473

```
. dfuller x, nocon lag(1) reg
```

Augmented Dickey-Fuller test for unit root Number of obs = 498

	Test Statistic	----- 1% Critical Value	Interpolated Dickey-Fuller 5% Critical Value	----- 10% Critical Value
--	-------------------	-------------------------------	--	--------------------------------

$z(t)$	-0.446	-2.580	-1.950	-1.620
--------	--------	--------	--------	--------

D.x	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
x					
L1.	-.0014182	.0031801	-0.45	0.656	-.0076663 .0048299
LD.	.5124481	.0389352	13.16	0.000	.4359498 .5889463

. dfuller d.x, nocon lag(1) reg

Augmented Dickey-Fuller test for unit root Number of obs = 497

		----- Interpolated Dickey-Fuller -----			
Test Statistic		1% Critical Value	5% Critical Value	10% Critical Value	
z(t)		-10.013	-2.580	-1.950	-1.620

D2.x	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]

x					
LD.	-.4420748	.0441502	-10.01	0.000	-.5288197 -.35533
LD2.	-.1003079	.0445811	-2.25	0.025	-.1878994 -.0127163

. dfuller y, lag(1) trend reg

Augmented Dickey-Fuller test for unit root Number of obs = 498

	Test Statistic	----- 1% Critical value	Interpolated Dickey-Fuller 5% Critical value	----- 10% Critical value
z(t)	-0.431	-3.980	-3.420	-3.130

Mackinnon approximate p-value for $z(t)$ = 0.9859

D.y	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
y					
L1.	-.0022674	.0052589	-0.43	0.667	-.0125999 .0080651
LD.	.2101265	.0443571	4.74	0.000	.1229747 .2972783
_trend	.0016933	.0014232	1.19	0.235	-.001103 .0044897
_cons	-.2454884	.4159721	-0.59	0.555	-1.062781 .5718044

. dfuller y, lag(1) reg

Augmented Dickey-Fuller test for unit root Number of obs = 498

	Test Statistic	----- 1% Critical Value	Interpolated Dickey-Fuller 5% Critical Value	----- 10% Critical Value
z(t)	-0.649	-3.440	-2.870	-2.570

Mackinnon approximate p-value for $z(t)$ = 0.8594

D.y	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
y					
L1.	-.0033641	.0051796	-0.65	0.516	-.0135408 .0068127
LD.	.2144937	.0442235	4.85	0.000	.1276048 .3013826
_cons	.1855889	.2044389	0.91	0.364	-.2160861 .5872639

. dfuller y, nocon lag(1) reg

Augmented Dickey-Fuller test for unit root Number of obs = 498

----- Interpolated Dickey-Fuller -----

	Test Statistic	1% Critical Value	5% Critical Value	10% Critical Value
Z(t)	-0.499	-2.580	-1.950	-1.620

D.y	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
y					
L1.	-.002543	.0050992	-0.50	0.618	-.0125617 .0074756
LD.	.215445	.0442032	4.87	0.000	.1285963 .3022936

```
. dfuller d.y, nocon lag(1) reg
```

Augmented Dickey-Fuller test for unit root Number of obs = 497

	Test Statistic	1% Critical Value	5% Critical Value	10% Critical Value
Z(t)	-13.412	-2.580	-1.950	-1.620

D2.y	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
y					
LD.	-.7534887	.0561806	-13.41	0.000	-.8638705 -.6431069
LD2.	-.0452885	.0447585	-1.01	0.312	-.1332286 .0426516

Estimate Cointegrating Equation of the series (y and x) using OLS

```
. reg y x
```

Source	SS	df	MS	Number of obs	=	500
Model	762663.128	1	762663.128	F(1, 498)	=	65375.48
Residual	5809.61273	498	11.665889	Prob > F	=	0.0000
				R-squared	=	0.9924
				Adj R-squared	=	0.9924
Total	768472.741	499	1540.02553	Root MSE	=	3.4155

y	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
x	.5039868	.0019711	255.69	0.000	.500114 .5078595
_cons	.1760628	.15523	1.13	0.257	-.1289236 .4810492

Predict residual (uhat)

```
. predict uhat, r
```

Perform Unit Root Test: whether uhat is stationary.

```
. dfuller uhat, trend lag(1) reg
```

Augmented Dickey-Fuller test for unit root Number of obs = 498

	Test Statistic	1% Critical Value	5% Critical Value	10% Critical Value
Z(t)	-10.415	-3.980	-3.420	-3.130

Mackinnon approximate p-value for Z(t) = 0.0000

D.uhat	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
uhat					
L1.	-.4932905	.0473653	-10.41	0.000	-.5863528 -.4002282
LD.	-.1214639	.0447212	-2.72	0.007	-.2093312 -.0335966
_trend	.0009115	.000955	0.95	0.340	-.0009648 .0027878
_cons	-.2308549	.2754893	-0.84	0.402	-.7721301 .3104202

Estimate Error Correction Mechanism

```
. reg d.y d.x l.uhat
```

Source	SS	df	MS	Number of obs	=	499
Model	6008.99264	2	3004.49632	F(2, 496)	=	335.62
Residual	4440.26935	496	8.95215595	Prob > F	=	0.0000
				R-squared	=	0.5751
				Adj R-squared	=	0.5734
Total	10449.262	498	20.9824538	Root MSE	=	2.992

D.y	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
^x d1.	.6287157	.0243408	25.83	0.000	.5808918	.6765395
uhat l1.	-.6766761	.0455654	-14.85	0.000	-.7662011	-.587151
_cons	-.0517604	.1343186	-0.39	0.700	-.3156641	.2121432