In the EViews workfile sp500.wf1, you can find two series. *div* is dividends paid out on the S&P 500 index, whereas e is earnings on the S&P 500 index. Both series are in logs. The annual sample period is from 1871 to 2014.

a) Run the following regression:

$$div_t = \alpha + \beta t + \gamma t^2 + \varepsilon_t$$

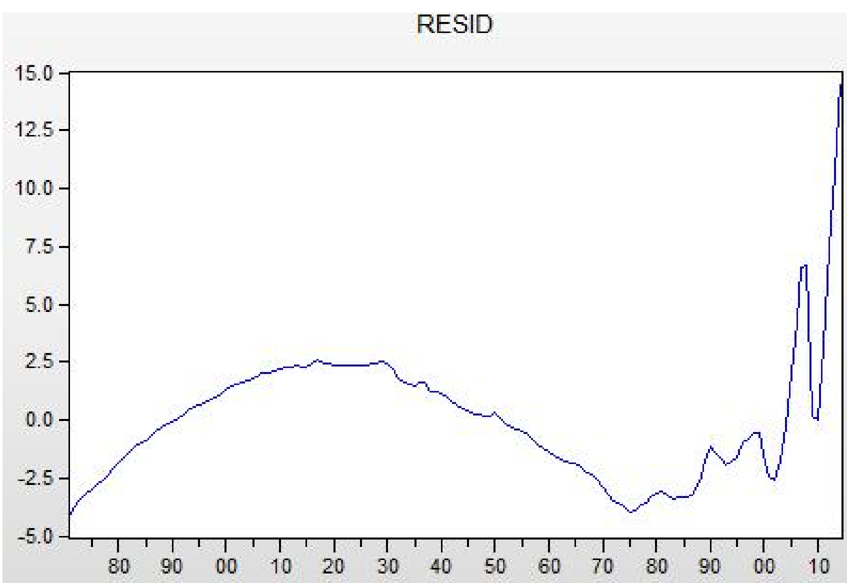
where t is a linear time trend. Based on a plot of the residuals from the regression, explain whether it is likely that *div* is trend stationary.

In the command window: ls div c @trend @trend^2

Dependent Variable: DIV
Method: Least Squares
Date: 02/25/16 Time: 15:02
Sample: 1871 2014
Included observations: 144

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	4.296684	0.682250	6.297812	0.0000
@TREND	-0.267008	0.022046	-12.11146	0.0000
@TREND^2	0.002876	0.000149	19.27364	0.0000

R-squared	0.881362	Mean dependent var	4.876321
Adjusted R-squared	0.879680	S.D. dependent var	7.976978
S.E. of regression	2.766992	Akaike info criterion	4.894012
Sum squared resid	1079.530	Schwarz criterion	4.955883
Log likelihood	-349.3688	Hannan-Quinn criter.	4.919153
F-statistic	523.7471	Durbin-Watson stat	0.112612
Prob(F-statistic)	0.000000		



The plot of residuals, do not show that there is an overall trend during all the period, but if we look the priod from the 1870 to 1920 and also from 1975 to 2014, we can guess there is time trends in the regression.

---

b) Run the regressions:

$$\Delta div_t - \alpha + \rho div_{t-1} + \sum_{i=1}^{2} \beta_i \Delta div_{t-i} + \varepsilon_t$$
$$\Delta^2 div_t = \alpha + \rho \Delta div_{t-1} + \beta \Delta^2 div_{t-1} + \varepsilon_t$$

In both cases, test the null hypothesis that $\rho = 0$. Explain whether div is I(0) or I(1).

The first regression is an AR(3) model. In the first regression, the null hypo says that there is a unit root in the regression. We test it with running the regression: In the command window: ls d(div) c div(-1) d(div(-1)) d(div(-2))

Dependent Variable: D(DIV)
Method: Least Squares
Date: 02/25/16 Time: 15:09
Sample (adjusted): 1874 2014
Included observations: 141 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.022523	0.072475	-0.310762	0.7565
DIV(-1)	0.052704	0.010224	5.155066	0.0000
D(DIV(-1))	0.581634	0.083584	6.958632	0.0000
D(DIV(-2))	-0.412789	0.086847	-4.753042	0.0000

R-squared	0.468067	Mean dependent var	0.277376
Adjusted R-squared	0.456419	S.D. dependent var	0.984837
S.E. of regression	0.726100	Akaike info criterion	2.225701
Sum squared resid	72.22935	Schwarz criterion	2.309354
Log likelihood	-152.9119	Hannan-Quinn criter.	2.259695
F-statistic	40.18382	Durbin-Watson stat	2.056217
Prob(F-statistic)	0.000000		

The p-value (0) and t-stat (5.16) of the ρ, shows that the coef of d_{t-1} is significant. (0.05), The augmented Dickey-Fuller test is based on the t-statistic: ADF = ρ / se(ρ) = 0.05 / 0.01 = 5.15. we compare the value with table 8.1:

Table 8.1 1% and 5% critical values for Dickey–Fuller tests

	Without trend		With trend	
Sample size	1%	5%	1%	5%
$T = 25$	−3.75	−3.00	−4.38	−3.60
$T = 50$	−3.58	−2.93	−4.15	−3.50
$T = 100$	−3.51	−2.89	−4.04	−3.45
$T = 250$	−3.46	−2.88	−3.99	−3.43
$T = 500$	−3.44	−2.87	−3.98	−3.42
$T = \infty$	−3.43	−2.86	−3.96	−3.41

And we strongly reject the null hypo that there is unit root, so in the AR(3) regression, we do not have unit root. And ρ is not zero.

We run the second regression: ls (d(div))^2 c d(div(-1)) (d(div(-1)))^2

Dependent Variable: (D(DIV))^2
Method: Least Squares
Date: 02/25/16 Time: 15:21
Sample (adjusted): 1873 2014
Included observations: 142 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.142474	0.302312	0.471282	0.6382
D(DIV(-1))	2.283922	0.315896	7.229972	0.0000
(D(DIV(-1)))^2	0.371221	0.071760	5.173108	0.0000

R-squared	0.386266	Mean dependent var	1.032646
Adjusted R-squared	0.377435	S.D. dependent var	4.333034
S.E. of regression	3.418885	Akaike info criterion	5.317406
Sum squared resid	1624.739	Schwarz criterion	5.379853
Log likelihood	-374.5358	Hannan-Quinn criter.	5.342782
F-statistic	43.74120	Durbin-Watson stat	2.249607
Prob(F-statistic)	0.000000		

In this case also, the p-value and t-stat suggest us that ρ cannot be zero.

Here from the first regression (AR(3) model), we know that d_t does not have unit root. But we do not know about the AR(1) model of d_t that identifies the I(1) or I(0). We run this regression: ls div c div(-1)

Dependent Variable: DIV
Method: Least Squares
Date: 02/25/16 Time: 15:27
Sample (adjusted): 1872 2014
Included observations: 143 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.032872	0.083531	-0.393527	0.6945
DIV(-1)	1.066210	0.009538	111.7868	0.0000

R-squared	0.988843	Mean dependent var	4.908603
Adjusted R-squared	0.988763	S.D. dependent var	7.995571
S.E. of regression	0.847553	Akaike info criterion	2.520959
Sum squared resid	101.2867	Schwarz criterion	2.562398
Log likelihood	-178.2486	Hannan-Quinn criter.	2.537798
F-statistic	12496.30	Durbin-Watson stat	1.150571
Prob(F-statistic)	0.000000		

We know that $$DF = \frac{\hat{\theta} - 1}{se\left(\hat{\theta}\right)}$$, so DF = (1.07 – 1) / 0.01 = 7. So we strongly reject the null hypo that we have unit root. So the d_t Is I(0).

---

c) Assume that both div and e are I(1). Regress div on a constant and e. Then test whether they are cointegrated.

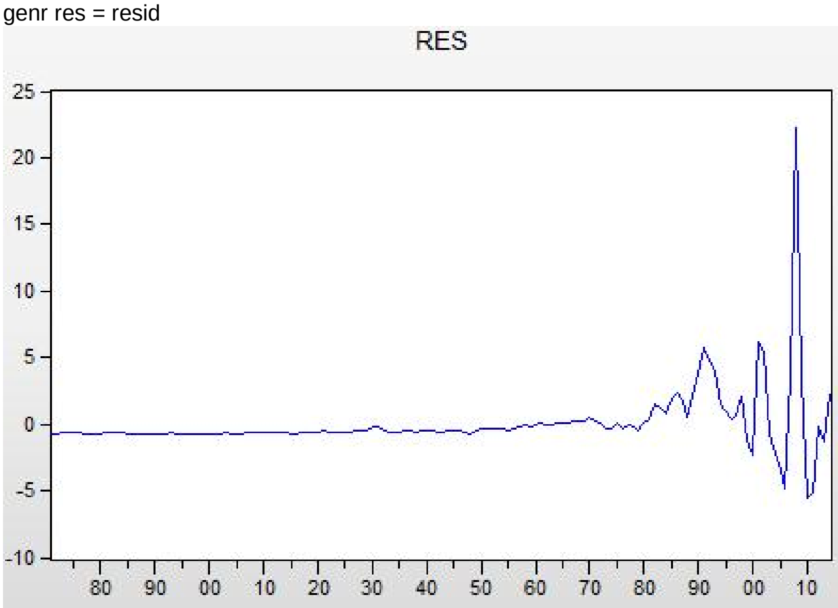
Command: ls div c e

Dependent Variable: DIV
Method: Least Squares
Date: 02/25/16 Time: 15:32
Sample: 1871 2014
Included observations: 144

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.839265	0.230831	3.635841	0.0004
E	0.353983	0.009525	37.16524	0.0000

R-squared	0.906779	Mean dependent var	4.876321
Adjusted R-squared	0.906122	S.D. dependent var	7.976978
S.E. of regression	2.444106	Akaike info criterion	4.639028
Sum squared resid	848.2592	Schwarz criterion	4.680275
Log likelihood	-332.0100	Hannan-Quinn criter.	4.655788
F-statistic	1381.255	Durbin-Watson stat	1.236154
Prob(F-statistic)	0.000000		

Test for cointegration means thata we have to test if the residuals (of regression) are stationary or not. The command:
genr res = resid



The plot shows us that there is not significant trend (it seems that residuals are mean reverting and so they are I(0)). Also it seems that the constant (intercept) is not very significant. But we have to check formally.
Command: ls d(res) c res(-1)
(we use the first lag for starting)

Dependent Variable: D(RES)
Method: Least Squares
Date: 02/25/16 Time: 15:40
Sample (adjusted): 1872 2014
Included observations: 143 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.011392	0.189697	0.060054	0.9522
RES(-1)	-0.619150	0.078149	-7.922668	0.0000

R-squared	0.308039	Mean dependent var	0.021717
Adjusted R-squared	0.303131	S.D. dependent var	2.717331
S.E. of regression	2.268391	Akaike info criterion	4.489906
Sum squared resid	725.5296	Schwarz criterion	4.531345
Log likelihood	-319.0283	Hannan-Quinn criter.	4.506745
F-statistic	62.76867	Durbin-Watson stat	1.697585
Prob(F-statistic)	0.000000		

The p-value and t-stat of res(-1) (slope coefficient) suggest that the result is statistically significant. (-0.62). also R-squared is 0.31 (it is not very strong).DW is 1.7 and so the residuals are highly persistent. Here we have to compare the t-stat with the table 9.2:

Number of variables (incl. Y_t)	Significance level 1%	5%	10%
2	−3.90	−3.34	−3.04

Because we have two variable, we use the first row and strongly reject that regression has a unit root, so two variables are cointegrated.

---

d) Run the following two regressions:

$$\Delta div_t = \alpha + \beta coin_{t-1} + \varepsilon_t$$

$$\Delta e_t = \alpha + \beta coin_{t-1} + \varepsilon_t$$

where *coin* is the cointegrating residual from question c. Carefully explain the implications of the regression results.

We have two equation that are based the difference between d_t and e_t. Because the d_t and e_t are cointegrated (from section c), we can see the residual part as the equilibrium error (here: c_t). It means that c_t measures the deviation from the long run equilibrium. Also we can use it for formulate an error correction model. Here we have two error correction model. But we have to compare them i.e. we can analyze which variable that error corrects d_t, e_t, or both. So we regress these two models:

Dependent Variable: D(DIV)
Method: Least Squares
Date: 02/25/16 Time: 16:01
Sample (adjusted): 1872 2014
Included observations: 143 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.270307	0.068606	3.939980	0.0001
RES(-1)	-0.220606	0.028264	-7.805298	0.0000

R-squared	0.301713	Mean dependent var	0.273986
Adjusted R-squared	0.296760	S.D. dependent var	0.978295
S.E. of regression	0.820392	Akaike info criterion	2.455818
Sum squared resid	94.89901	Schwarz criterion	2.497256
Log likelihood	-173.5910	Hannan-Quinn criter.	2.472656
F-statistic	60.92267	Durbin-Watson stat	0.308137
Prob(F-statistic)	0.000000		

Dependent Variable: D(E)
Method: Least Squares
Date: 02/25/16 Time: 16:02
Sample (adjusted): 1872 2014
Included observations: 143 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.731433	0.538037	1.359448	0.1762
RES(-1)	1.125885	0.221654	5.079461	0.0000

R-squared	0.154681	Mean dependent var	0.712657
Adjusted R-squared	0.148686	S.D. dependent var	6.973080
S.E. of regression	6.433830	Akaike info criterion	6.574904
Sum squared resid	5836.578	Schwarz criterion	6.616343
Log likelihood	-468.1057	Hannan-Quinn criter.	6.591743
F-statistic	25.80093	Durbin-Watson stat	1.605994
Prob(F-statistic)	0.000001		

	Δd_t	Δe_t
c_{t-1}	*-0.22 (-7.81)*	*1.13 (5.08)*

We look at the t-stats values and select the higher value.

The table shows that the error-correction term is a stronger predictor of Δd_t (t-value of -7.81) than of Δe_t (t-value of 5.08). This evidence suggests that when we have deviations from the long-run relationship between dividends and earnings, it is more dividends that adjust such that the cointegration equilibrium is restored.

---

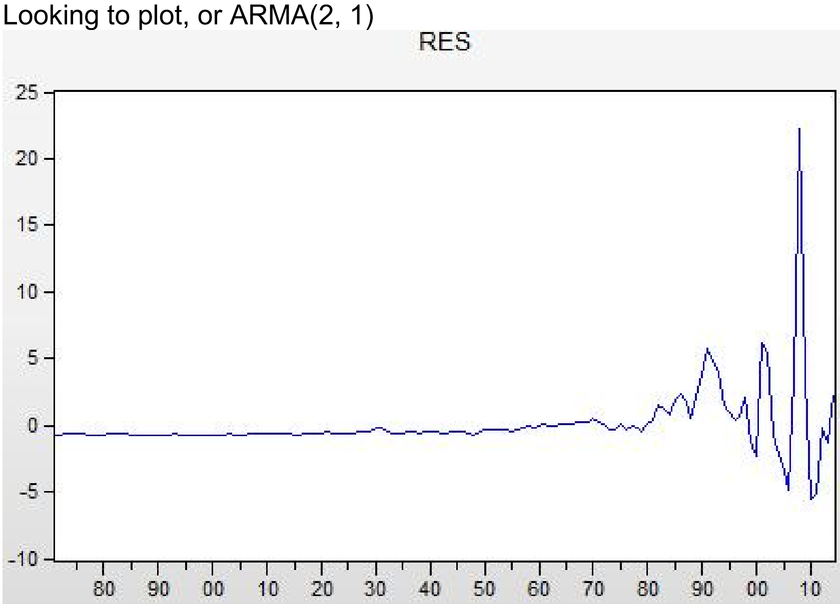
e) Suggest an appropriate ARMA(p,q) model for *coin*. Explain whether *coin* is slow or fast to mean revert.

We know that c_t is I(0). We look plot and correlogram:

Sample: 1871 2014
Included observations: 144

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.378	0.378	21.036	0.000
		2	-0.182	-0.379	25.941	0.000
		3	-0.196	0.058	31.669	0.000
		4	-0.018	-0.011	31.719	0.000
		5	0.100	0.061	33.232	0.000
		6	0.368	0.401	53.827	0.000
		7	0.299	-0.013	67.526	0.000
		8	0.004	0.063	67.529	0.000
		9	-0.013	0.208	67.555	0.000
		10	0.124	0.055	69.951	0.000

We know that The PACF of an AR(p) model cuts off after lag p, also we know that for a MA(q) model the ACF is zero after q lags. So here the suggested p is 2 and q is 1. Looking to plot, or ARMA(2, 1)



We can conclude that at the first steps (1880 to 1970) we have faster in mean reverting. But after that it becomes more and slower to mean reverting. It is close to the definition of the volatility clustering, because when we have a slow mean reverting, then Large. Fluctuations are typically followed by large .fluctuations and small fluctuations are typically followed by small fluctuations. It means that our conditional volatility is time-varying. We use ARCH and GARCH models for showing time varying volatility.

---

f) Use the Q statistic to test for ARCH effects in your suggested model from question e.

We have to estimate the proposed ARMA(2, 1) model.

Dependent Variable: RES
Method: ARMA Maximum Likelihood (BFGS)
Date: 02/25/16 Time: 16:51
Sample: 1873 2014
Included observations: 142
Convergence achieved after 4 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.015350	0.377984	0.040610	0.9677
RES(-1)	0.381917	0.173570	2.200362	0.0295
RES(-2)	-0.332054	0.113041	-2.937471	0.0039
MA(1)	0.175475	0.196186	0.894432	0.3727
SIGMASQ	4.328453	0.268959	16.09339	0.0000

R-squared	0.274544	Mean dependent var	0.009946
Adjusted R-squared	0.253363	S.D. dependent var	2.451295
S.E. of regression	2.118119	Akaike info criterion	4.373730
Sum squared resid	614.6404	Schwarz criterion	4.477809
Log likelihood	-305.5348	Hannan-Quinn criter.	4.416023
F-statistic	12.96171	Durbin-Watson stat	1.980140
Prob(F-statistic)	0.000000		

Inverted MA Roots	-.18

After that we have to use an auxiliary regression for varinaces and residuals.

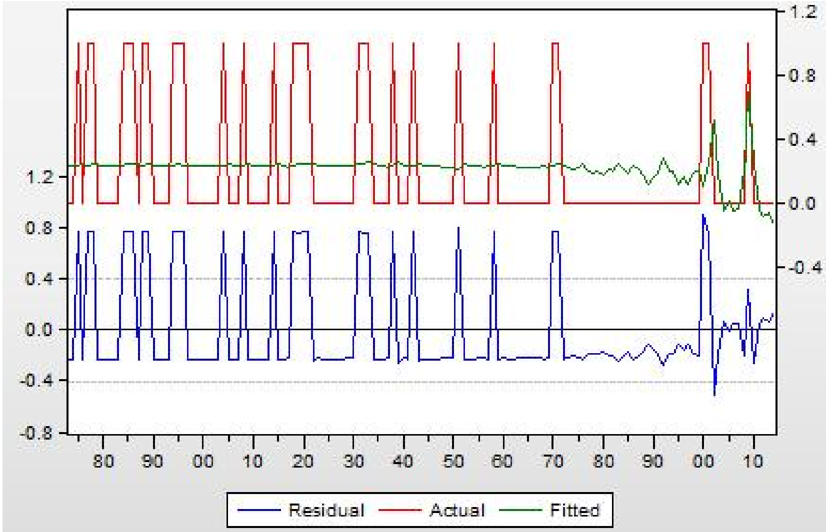
$$\sigma_t^2 = \varpi + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^{q} \beta_i \sigma_{t-i}^2$$

Where p = 2, q = 1.
We use Breusch-Pagan test for heteroscedasticity to test for ARCH effects. We run an auxiliary regression of the squared residuals e^2 on a constant and 2 lagged squared residuals, and 1st lag of variances.
The null hypo is there is no arch effects: Test statistic: TR^2 ~ χ^2(p+q)
TR^2 = 142 * 0.274544 = 38.99 > χ^2(3) = 7.81 so we strongly reject that there is no ARCH effects in the residuals.

Alternatevely we have $c_t = \alpha + \beta_1 c_{t-1} + \beta_2 c_{t-2} + \varepsilon_1 + \beta_3 \varepsilon_{t-1}$ ...
---
g) Run the following command in EViews: genr y − d(div) < 0. This will generate a variable, y, that is equal to 1 if the log change in dividends is negative and 0 otherwise. Then run the regression:

$$y_t = \beta_1 + \beta_2 \Delta div_{t-1} + \beta_3 \Delta e_{t-1} + \varepsilon_t$$

Make a plot of the fitted and actual values of y_t. Comment on the results.
The most important drawback of the linear regression model is that the estimated probabilities may be negative or larger than 100%. As an illustration of this drawback, see plot below.



Another drawback is that linear model by suffers from heteroskedasticity.

---

h) Estimate the following model:

$$P\left(y_t - 1 \mid x_{t-1}\right) - F\left(x_{t-1}'\beta\right)$$

where $x_{t-1} = (1, \Delta div_{t-1}, \Delta e_{t-1})'$, $\beta = (\beta_1, \beta_2, \beta_3)'$, and $F(\cdot)$ is the standard logistic distribution function. Comment on the signs of the estimated coefficients. What is the pseudo R^2 of the model?

$$F\left(x_i'\beta\right) = \frac{\exp\left(x_i'\beta\right)}{1 + \exp\left(x_i'\beta\right)}$$

= exp(b1 + b2*d(div(-1)) + b3*d(e(-1))) / (1 + exp(b1 + b2*d(div(-1)) + b3*d(e(-1))))

We have to estimate the model using ML (Maximum Likelihood) in Eviews. Object – new object – equation – binary – logit: y c d(div(-1)) d(e(-1))

Dependent Variable: Y
Method: ML - Binary Logit (Newton-Raphson / Marquardt steps)
Date: 02/25/16 Time: 17:12
Sample (adjusted): 1873 2014
Included observations: 142 after adjustments
Convergence achieved after 6 iterations
Coefficient covariance computed using observed Hessian

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	-1.217401	0.212827	-5.720153	0.0000
D(DIV(-1))	-0.637443	0.337746	-1.887344	0.0591
D(E(-1))	-0.074814	0.041902	-1.785464	0.0742

McFadden R-squared	0.050639	Mean dependent var	0.211268
S.D. dependent var	0.409653	S.E. of regression	0.403847
Akaike info criterion	1.021294	Sum squared resid	22.66985
Schwarz criterion	1.083741	Log likelihood	-69.51186
Hannan-Quinn criter.	1.046670	Deviance	139.0237
Restr. deviance	146.4393	Restr. log likelihood	-73.21965
LR statistic	7.415570	Avg. log likelihood	-0.489520
Prob(LR statistic)	0.024532		
Obs with Dep=0	112	Total obs	142
Obs with Dep=1	30		

β1 = -1.22, **β**2 = -0.64, **β**3 = -0.07

$$Pseudo\ R^2 = 1 - \frac{1}{1 + 2\left(\log L_1 - \log L_0\right)/N}$$

Where log L1 = -69.51 and log L0 = -73.22, N = 142

---

k) Based on the model in h, compute the marginal effect of a change in Δdiv_{t-1}.

-0.64

Problem 2 (weight=20%).

Consider the data generating process:

$$z_i = (1 + w_i)^{-\gamma} + u_i$$

where the variable w_i is iid $N\left(0, 0.05^2\right)$, the error term u_i is iid $N(0,1)$, and the parameter γ is equal to 25.

a) Write down a program that simulates 10000 observations from the model.

```
'create a workfile using 10000 observation
create sample_2a u 1 10000

rndseed 1
'we have to generate a 0 series first
genr z = 0

'first we have to generate a white noise series ~ N(0, 9)
genr w = 0.05*nrnd
genr u = nrnd
scalar y = 25
```

genr z = (1+w)^(-1)*y+u

---

b) Run the program and use the simulated data to estimate γ using GMM. As moment condition, you should use $E(u_i) = 0$. Comment on the value of the J-statistic.

In the system object: z - (1+w)^(c(1)*(-1)) = 0 @ c

System: UNTITLED
Estimation Method: Generalized Method of Moments
Date: 02/25/16 Time: 17:46
Sample: 1 10000
Included observations: 10000
Total system (balanced) observations 10000
Kernel: Bartlett, Bandwidth: Fixed (12), No prewhitening
Iterate coefficients after one-step weighting matrix
Convergence achieved after: 1 weight matrix, 136 total coef iterations

	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	24.93669	0.062392	399.6788	0.0000

Determinant residual covariance	0.985542
J-statistic	-7.70E-34

Equation: Z - (1+W)^(C(1)*(-1))-(0)
Instruments: C
Observations: 10000

S.E. of regression	0.992794	Sum squared resid	9855.423
Durbin-Watson stat	2.033122		

The number is approximately close to 25 (24.93).
In the overidentified case, we can test the overidentifying restrictions. We can do so by checking whether the sample moments are close to 0.

If we use the optimal weights, the test of overidentifying restrictions is:

$$
\begin{aligned}
J &= Tg\left(\hat{\theta}\right)' W^{opt} g\left(\hat{\theta}\right) \\
&= TQ\left(\hat{\theta}\right)
\end{aligned}
$$

But here we have one moment condition and one parameter. So the model is exactly identified. And for this reason the J-stat is close to zero.
