Advanced time-series analysis (University of Lund, Economic History Department)

30 Jan-3 February and 26-30 March 2012

Lecture 5a Testing for structural consistency

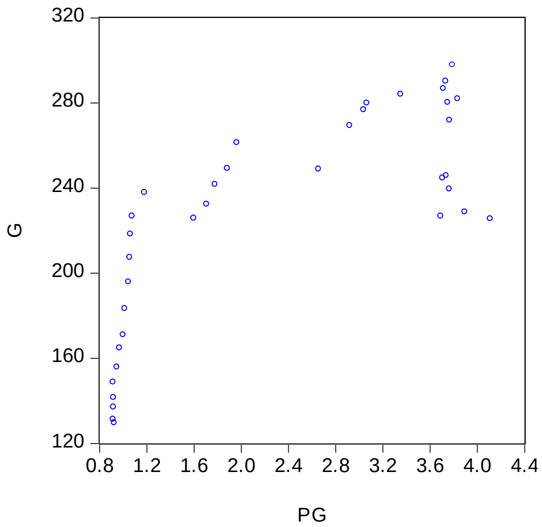
5.a.a. Structural break

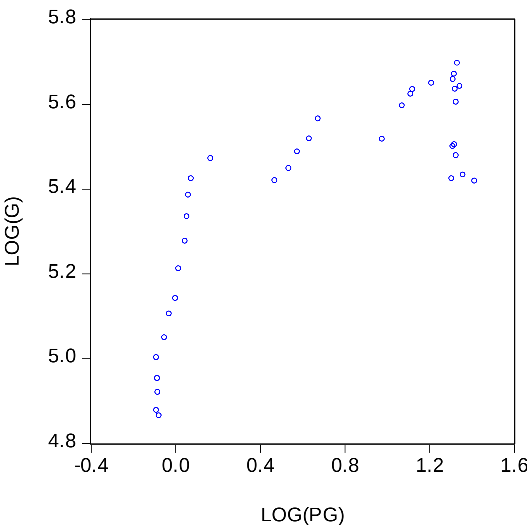
Normally we have an important assumption regarding our model: we require the parameters be stable in the sample. This means the constancy of the parameters.

$y_t = \beta_0 + \beta_1 x_t + u_t$ for all t=1,2,...,t

Let us see a case when it is not true:

Greene (2003) uses the relationship between gasoline price (PG) and total gasoline consumption in the USA (G) as an example of this problem:





Now, you can see that the coefficient stability condition is not fulfilled: the structural parameter β_1 is unstable. The consequences of structural break are numerous:

1. First, since you have misspecified your model your estimate will be biased. In this particular case, for example, you would be most likely to underestimate the own-price elasticity of gasoline consumption for the pre-break period and overestimate it afterwards.
2. Since your model is incorrect it will not fit the model well, so your regression is going to have larger residual variance, causing your standard errors to increase: this may lead to erroneous outcomes of hypothesis tests.

Let us take an example for this latter:

In Table 5_1 of Green we test for a unit root in the level of the log of real GDP:

Null Hypothesis: LNRGDP is stationary
Exogenous: Constant, Linear Trend
Bandwidth: 11 (Newey-West automatic) using Bartlett kernel

		LM-Stat.
Kwiatkowski-Phillips-Schmidt-Shin test statistic		0.291269
Asymptotic critical values*:	1% level	0.216000
	5% level	0.146000
	10% level	0.119000

*Kwiatkowski-Phillips-Schmidt-Shin (1992, Table 1)

Residual variance (no correction)	0.001509
HAC corrected variance (Bartlett kernel)	0.014233

Now, let us assume that we have reason to think that there was structural break around the first quarter of 1966. Would incorporating this knowledge alter our results?

Null Hypothesis: LNRGDP is stationary
Exogenous: Constant, Linear Trend
Bandwidth: 9 (Newey-West automatic) using Bartlett kernel

		LM-Stat.
Kwiatkowski-Phillips-Schmidt-Shin test statistic		0.051138
Asymptotic critical values*:	1% level	0.216000
	5% level	0.146000
	10% level	0.119000

*Kwiatkowski-Phillips-Schmidt-Shin (1992, Table 1)

Residual variance (no correction)	0.000508
HAC corrected variance (Bartlett kernel)	0.002995

The results from the KPSS test are now different and we may believe that the log of real GDP was trend stationary after all.

5.a.b What to do about it?

Let us take now the simple case: you know where your structural break is. Then you have two options.

1. You use sub-samples of your data: this is exactly what we did in case of the log real GDP of the USA: we rerun the test on the subsample between 1966Q1 to 2000Q4. The obvious drawback is that thereby you reduce the available number of observations, and this is what you normally do not want.
2. You modify your specification so that it directly capture the structural break. For this you should know what triggers the break. In the simplest case, some external factor causes a change in the structure of the economic phenomenon at a discrete point in time. In this case would simply use dummy and interaction variables. So:
   $y_t = \beta_0 + \beta_1 x_t + \beta_2 D_t^{break} + \beta_3 \left(x_t D_t^{break} \right) + u_t$ where D_t^{break} is a dummy taking unit value from the period when the break starts and zero otherwise.

   In the more difficult case it is a value of x which triggers the break. This can be handled in a very simple and a more difficult way as well, but we know restrict ourselves to the simpler one. Our adapted specification is:

$y_t = \beta_0 + \beta_1 x_t + \beta_2 \cdot I(x_t \geq c) + \beta_3 \left(x_t \cdot I\left(x_t \geq c \right) \right) + u_t$ where I(x≥c) is an indicator variable (a dummy) which takes unit value if in period t x is larger or equal to c. Parameter c is the threshold value. Now, of course, you either know c or not. In the latter case you should estimate it somehow. One was could be to try several values of c and search for the one which minimizes the residual sum of squared.

Let us see the above solutions for the problem of gasoline consumption:

Dependent Variable: LNG
Method: Least Squares
Date: 02/01/12 Time: 14:37
Sample: 1960 1995
Included observations: 36

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	5.166909	0.036849	140.2196	0.0000
LNPG	0.335383	0.040959	8.188191	0.0000

R-squared	0.663521	Mean dependent var	5.392989
Adjusted R-squared	0.653625	S.D. dependent var	0.248779
S.E. of regression	0.146416	Akaike info criterion	-0.950781
Sum squared resid	0.728877	Schwarz criterion	-0.862807
Log likelihood	19.11405	Hannan-Quinn criter.	-0.920076
F-statistic	67.04648	Durbin-Watson stat	0.215398
Prob(F-statistic)	0.000000		

Now we have a historical knowledge that the break point was in 1973. So we modify our specification:

Dependent Variable: LNG
Method: Least Squares
Date: 02/01/12 Time: 14:39
Sample: 1960 1995
Included observations: 36

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	5.178593	0.022135	233.9578	0.0000
LNPG	2.951069	0.341097	8.651703	0.0000
@YEAR>1972	0.270306	0.054596	4.951017	0.0000
LNPG*(@YEAR>1972)	-2.858351	0.343972	-8.309828	0.0000

R-squared	0.914984	Mean dependent var	5.392989
Adjusted R-squared	0.907013	S.D. dependent var	0.248779
S.E. of regression	0.075862	Akaike info criterion	-2.215360
Sum squared resid	0.184162	Schwarz criterion	-2.039414
Log likelihood	43.87648	Hannan-Quinn criter.	-2.153950
F-statistic	114.7993	Durbin-Watson stat	0.394112
Prob(F-statistic)	0.000000		

Now you can see that improving your specification not only led to a better fit, but also it solved some of the first-order autocorrelation problem (do not be surprised: autocorrelated residuals signify a misspecification). What it did not solve however is the problem of simultaneity, but that is not our concern for now.

If you were improving your model by, say, including the log of per capita disposable income, in real terms, you would obtain even better estimates:

Dependent Variable: LNG
Method: Least Squares
Date: 02/01/12 Time: 14:51
Sample: 1960 1995
Included observations: 36

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-6.099511	0.594583	-10.25846	0.0000
LNPG	0.644050	0.155941	4.130100	0.0003
@YEAR>1972	0.112322	0.017704	6.344615	0.0000
LNPG*(@YEAR>1972)	-0.789991	0.146890	-5.378100	0.0000
LOG(Y)	1.264950	0.066685	18.96916	0.0000

R-squared	0.993257	Mean dependent var	5.392989
Adjusted R-squared	0.992387	S.D. dependent var	0.248779
S.E. of regression	0.021707	Akaike info criterion	-4.694088
Sum squared resid	0.014607	Schwarz criterion	-4.474154
Log likelihood	89.49358	Hannan-Quinn criter.	-4.617325
F-statistic	1141.525	Durbin-Watson stat	1.085922
Prob(F-statistic)	0.000000		

The message is that before 1973 total gasoline consumption grew together with prices while afterwards the relationship became negative: 0.644-0.790=-0.146. You can even use a Wald test to look if this coefficient indeed significant is:

Wald Test:
Equation: Untitled

Test Statistic	Value	df	Probability
t-statistic	-8.163689	31	0.0000
F-statistic	66.64582	(1, 31)	0.0000
Chi-square	66.64582	1	0.0000

Null Hypothesis: C(2)+C(4)=0
Null Hypothesis Summary:

Normalized Restriction (= 0)	Value	Std. Err.
C(2) + C(4)	-0.145941	0.017877

Restrictions are linear in coefficients.

Indeed, it is.

5.a.c How to test for structural break?

The method for testing depends on your knowledge about the break.
Let us know assume that you know when the break happened. In this case you can use a Chow-test. The original Chow test is an F-test and is based on the same idea as any F-test in regression analysis: it is going to compare sum of squared residuals.

the original Chow-test:
The idea is simple. Let us run a regression on the whole sample:

$y_t = \beta_0 + \beta_1 x_t + u_t$ from which the sum of squared residual is: $SSR_T = \sum_{t=0}^{T} u_t^2$.

Now let us split our sample around the break point (t_b) and estimate the same specification on both halfs:

$y_{1t} = \beta_{10} + \beta_{11} x_{1t} + u_{1t}$ $SSR_1 = \sum_{t=0}^{t_b} u_{1t}^2$ and $y_{2t} = \beta_{20} + \beta_{21} x_{2t} + u_{2t}$ $SSR_2 = \sum_{t=t_b}^{T} u_{2t}^2$

The null-hypothesis of the test is that: $\beta_{10} = \beta_{20}$ and $\beta_{11} = \beta_{21}$. This implies that it would be indifferent for the goodness of the model if we were estimating the relationship on the two sub-samples or on the whole sample.

If this is true then $SSR_T = SST_1 + SST_2$.

The test statistics is: $\frac{\left(SSR_T - (SST_1 + SST_2)\right)/k}{(SST_1 + SST_2)/\left(n_1 + n_2 - 2k\right)} \sim F(k, n_1 + n_2 - 2k)$

where n_1 and n_2 are the length of the two subsamples and k is the number of parameters, which is now 2.

the simplfied Chow-test:

You just run the regression like we did in the previous example.

$y_t = \beta_0 + \beta_1 x_t + \beta_2 D_t^{break} + \beta_3 \left(x_t D_t^{break}\right) + u_t$ where the null-hypothesis is that $\beta_2 = \beta_3 = 0$. The two tests are equivalent.

Let us see an example for the problem with gasoline consumption and prices:

Chow Breakpoint Test: 1972
Null Hypothesis: No breaks at specified breakpoints
Varying regressors: All equation variables
Equation Sample: 1960 1995

F-statistic	47.47189	Prob. F(2,32)	0.0000
Log likelihood ratio	49.60830	Prob. Chi-Square(2)	0.0000
Wald Statistic	94.94378	Prob. Chi-Square(2)	0.0000

Theoretically, the test above can be applied to multiple break points.

Now we take the case when you do not know the breakpoint:

Quandt-Andrews test:

The idea is simple: Let us run a Chow test for all possible breakpoints, and choose the one where the test statistics is the highest (that is, the break is the most likely to have occurred).

But of course you cannot start with observation one, since your first sub-sample would not have enough observations. So, instead you have to leave a certain share of your observations at the start and the end of your sample untested. This “trimming” parameter is set to 5-15% of your sample size.

Let us look at this test in practice:

Quandt-Andrews unknown breakpoint test
Null Hypothesis: No breakpoints within 15% trimmed data
Varying regressors: All equation variables
Equation Sample: 1960 1995
Test Sample: 1966 1990
Number of breaks compared: 25

Statistic	Value	Prob.
Maximum LR F-statistic (1972)	47.47189	0.0000
Maximum Wald F-statistic (1972)	94.94378	0.0000
Exp LR F-statistic	21.64455	1.0000
Exp Wald F-statistic	45.22448	1.0000
Ave LR F-statistic	16.53134	0.0011
Ave Wald F-statistic	33.06268	0.0011

Note: probabilities calculated using Hansen's (1997) method

Now, you find that the breakpoint is the most likely to have been in 1972, which nicely coincides with our guess.

the CUSUM test:

This method is based on recursive residuals from recursive regressions. The basic idea is that if the your parameters does not change over time, then using your model estimated on the first t number of periods should be able to forecast your y in t+1 without a systematic bias.

Let us define the following recursive regression:

$y_t = \beta_{0t} + \beta_{1t} x_t + u_t$, where β_{0t} and β_{1t} are the coefficients estimated from the regression on the sample 1,…,t. Now the recursive forecast is:

$f_{t+1} = \beta_{0t} + \beta_{1t} x_{t+1}$ or $f_t = \beta_{0t-1} + \beta_{1t-1} x_t$. Obviously the recursive residual for period t is:

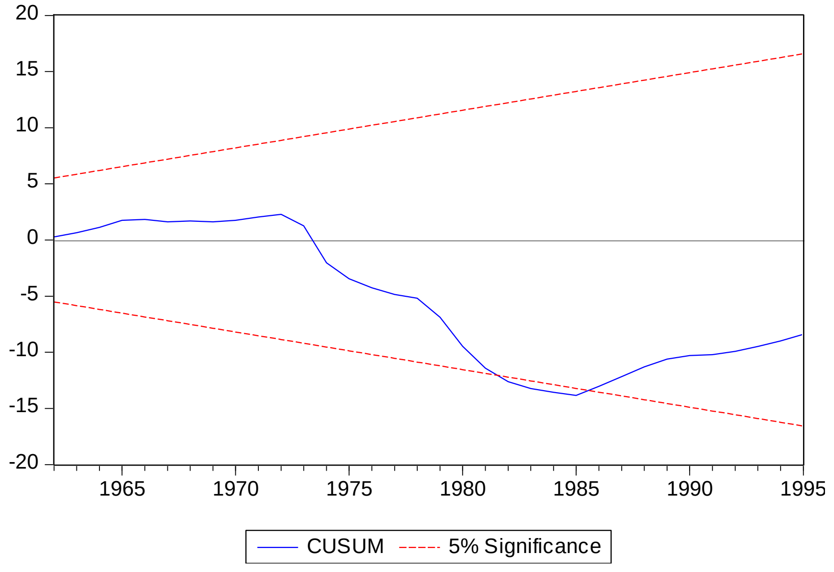
$w_t = f_t - y_t$. Of course, we again cannot start with observation 1 since we also need to have a minimum number of observations for the first regression. So we start at observation K.

Now we actually assume that $E(w_t) = \frac{1}{T-K+1}\sum_{t=K}^{T} w_t = 0$.

The test statistics is the following:

$CUSUM(t) = \sum_{t=K}^{T} \frac{w_t}{\hat{\sigma}_t}$, where $\hat{\sigma}_t^2 = \frac{1}{T-K}\sum_{t=K}^{T}\left(w_t - E(w_t)\right)$

The CUSUM test has critical values depending on t, so you should look for possible points where its test statistics crosses the boundary of critical values either upward or downward (systematic under or overestimation…).


20
15
10
5
0
-5
-10
-15
-20
1965
1970
1975
1980
1985
1990
1995
CUSUM
5% Significance