Testing for unit roots and structural breaks

Prepared by Ziyodullo Parpiev, PhD

for Regional Summer School

September 21, 2016

Unit Root Tests

- Traditionally, Augmented Dickey–Fuller (ADF) and Phillips–Perron (PP) tests are used to assess the order of integration of the variables.

- Uniform outcomes of both tests are necessary for the final conclusion about the stationarity properties of each series.

- Usually, all variables are tested with a linear trend and/or intercept or none.

Structural Breaks

- A well-known weakness of the ADF and PP unit root tests is their potential confusion of structural breaks in the series as evidence of non-stationarity.

- In other words, they may fail to reject the unit root hypothesis if the series have a structural break.

- In other words, for the series that are found to be $I(1)$, there may be a possibility that they are in fact stationary around the structural break(s), $I(0)$, but are erroneously classified as $I(1)$.

Nonstationarity: Structural Breaks

The coefficients of the model might not be constant over the full sample. Clearly, it is a problem for forecasting if the model describing the historical data differs from the current model – you want the current model for your forecasts!

So we will:

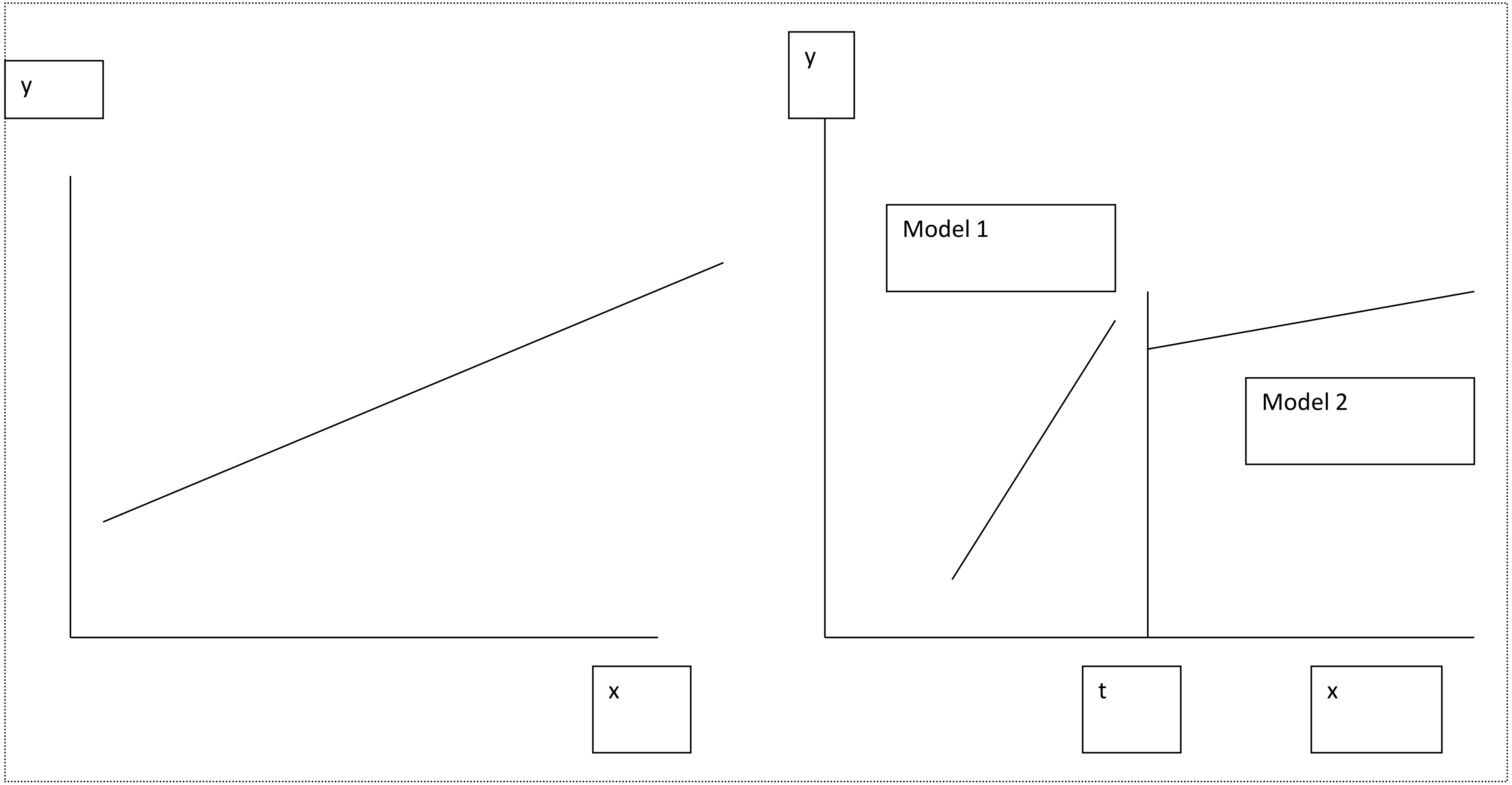
- Go over the way to detect changes in coefficients: tests for a break
- Work through an example: the U.S. labor productivity

Case I: Break date is known

A series of data can often contain a structural break, due to a change in policy or sudden shock to the economy, i.e. 1987 stock market crash. In order to test for a structural break, we often use the Chow test, this is Chow' first test (the second test relates to predictions).

The model in effect uses an F-test to determine whether a single regression is more efficient than two separate regressions involving splitting the data into two sub-samples. This could occur as follows, where in the second case we have a structural break at t.:

Case 1 Case2



Chow Test for Structural Stability

- In the first case we have just a single regression line to fit the data points (scatterplot), it can be expressed as:

$$y_t = \alpha_0 + \alpha_1 x_t + u_t$$

- In the second case, where there is a structural break, we have two separate models, expressed as:

$$y_t = \beta_1 + \beta_2 x_t + u_{1t}$$

$$y_t = \delta_1 + \delta_2 x_t + u_{2t}$$

- This suggests that model 1 applies before the break at time t, then model 2 applies after the structural break. If the parameters in the above models are the same, i.e. , then models 1 and 2 can be expressed as a single model as in case 1, where there is a single regression line. The Chow test basically tests whether the single regression line or the two separate regression lines fit the data best. The stages in running the Chow test are:

1. Firstly run the regression using all the data, before and after the structural break, collect RSS_c.
2. Run two separate regressions on the data before and after the structural break, collecting the RSS in both cases, giving RSS_1 and RSS_2.
3. Using these three values, calculate the test statistic from the following formula:

$$F - \frac{RSS_c - (RSS_1 + RSS_2)/k}{RSS_1 + RSS_2 / n - 2k}$$

4. Find the critical values in the F-test tables, in this case it has F(k,n-2k) degrees of freedom.
5. Conclude, the null hypothesis is that there is no structural break

Chow Test (stages in using test)

- Run the regression using all the observations, before and after the structural break, collect the RSS
- Run 2 separate regressions, one before, RSS(1) and one after, RSS(2) the structural break.
- Calculate the test statistic using the following formulae:

Chow Test

$$F - \frac{RSS_c - (RSS_1 + RSS_2)/k}{RSS_1 + RSS_2 / n - 2k}$$

$$RSS_c - combined\_RSS$$

$$RSS_1 - pre - break\_RSS$$

$$RSS_2 - post - break\_RSS$$

Chow Test

- The final stage of the Chow Test is to compare the test statistic with the critical value from the F-Tables.
- The null hypothesis in this case is structural stability, if we reject the null hypothesis, it means we have a structural break in the data
- We then need to decide how to overcome this break.

Chow Test

- If there is evidence of a structural break, it may mean we need to split the data into 2 samples and run separate regressions.
- Another method to overcome this problem is to use dummy variables (To be covered later in term), the benefit of this approach is that we do not lose any degrees of freedom through a loss of observations.

Chow Test Example

- The following model is regressed using data in quarterly form from 1990 to 2005 (64 observations)for Malaysian stock prices against output (structural break in 1997).

$$s_t = \alpha_0 + \alpha_1 y_t + u_t$$

Chow Test

- The first regression using all the data produced a RSS(c) of 0.56, then 2 regressions were run on a sub-sample of the data from 1990-1997, giving a RSS(1) of 0.23. The final regression was on the sample from 1998 to 2005, producing a RSS(2) of 0.17, n=64, k=2.

$$F - \frac{0.56-(0.23+0.17)/2}{0.23+0.17/64-2*2} = \frac{0.08}{0.00667} = 12$$

Chow Test

- As the critical value for F(2,60) =3.15(5%)
- As 12> 3.15, we reject the null hypothesis of structural stability.
- We conclude that there is a structural break in this model, we need to split the data into 2 sub-samples or use another method to overcome the break.

Problems with Chow Test

- The test may suggest splitting the data, this may mean fewer degrees of freedom
- When should the cut off point be for the test, usually there should be a theoretical basis for this.
- There is the potential for structural instability across the whole data range. It is possible to test every observation for a structural break.

Tests for a break (change) in regression coefficients using dummies

- Dummy Variables are a common way of solving structural breaks, as it does not involve splitting the data.
- These variables consist of 1s and 0s and are often termed ‘on-off’ variables.
- They can be used to determine the importance of policy actions on models and are often used to account for qualitative effects.
- Their coefficients and t-statistics can then be interpreted in the usual way.

Suppose the break is known to have occurred at date τ. Stability of the coefficients can be tested by estimating a <u>fully interacted regression model</u>. In the ADL(1,1) case:

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \delta_1 X_{t-1} + \gamma_0 D_t(\tau) + \gamma_1[D_t(\tau)\times Y_{t-1}] + \gamma_2[D_t(\tau)\times X_{t-1}] + u_t$$

where $D_t(\tau) = 1$ if $t \geq \tau$, and $= 0$ otherwise.

If $\gamma_0 = \gamma_1 = \gamma_2 = 0$, then the coefficients are constant over the full sample.

If at least one of γ_0, γ_1, or γ_2 are nonzero, the regression function changes at date τ.

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \delta_1 X_{t-1}$$

$$+ \gamma_0 D_t(\tau) + \gamma_1[D_t(\tau)\times Y_{t-1}] + \gamma_2[D_t(\tau)\times X_{t-1}] + u_t$$

where $D_t(\tau) = 1$ if $t \geq \tau$, and $= 0$ otherwise

The ***Chow test statistic*** for a break at date τ is the (heteroskedasticity-robust) F-statistic that tests:

$$H_0: \gamma_0 = \gamma_1 = \gamma_2 = 0$$

vs. H_1: at least one of γ_0, γ_1, or γ_2 are nonzero

- Note that you can apply this to a subset of the coefficients, e.g. only the coefficient on X_{t-1}.
- Unfortunately, you often don't have a candidate break date, that is, you don't know τ …

Dummy Variables

- Dummy Variables are a common way of solving structural breaks, as it does not involve splitting the data.
- These variables consist of 1s and 0s and are often termed ‘on-off’ variables.
- They can be used to determine the importance of policy actions on models and are often used to account for qualitative effects.
- Their coefficients and t-statistics can then be interpreted in the usual way.

Case II: The break date is unknown

Why consider this case?

- You might suspect there is a break, but not know when
- You might want to test the null hypothesis of coefficient stability against the general alternative that there has been a break sometime.
- Even if you think you know the break date, if that “knowledge” is based on prior inspection of the series then you have in effect “estimated” the break date. This invalidates the Chow test critical values.

The Quandt Likelihood Ratio (QLR) Statistic
(also called the "sup-Wald" statistic)

The QLR statistic = the maximum Chow statistic

- Let $F(\tau)$ = the Chow test statistic testing the hypothesis of no break at date τ.
- The *QLR* test statistic is the ***maximum*** of all the Chow *F*-statistics, over a range of τ, $\tau_0 \leq \tau \leq \tau_1$:

$$QLR = \max[F(\tau_0), F(\tau_0+1), \ldots, F(\tau_1-1), F(\tau_1)]$$

- A conventional choice for τ_0 and τ_1 are the inner 70% of the sample (exclude the first and last 15%).
- Should you use the usual $F_{q,\infty}$ critical values?

TABLE 14.6 Critical Values of the QLR Statistic with 15% Trimming

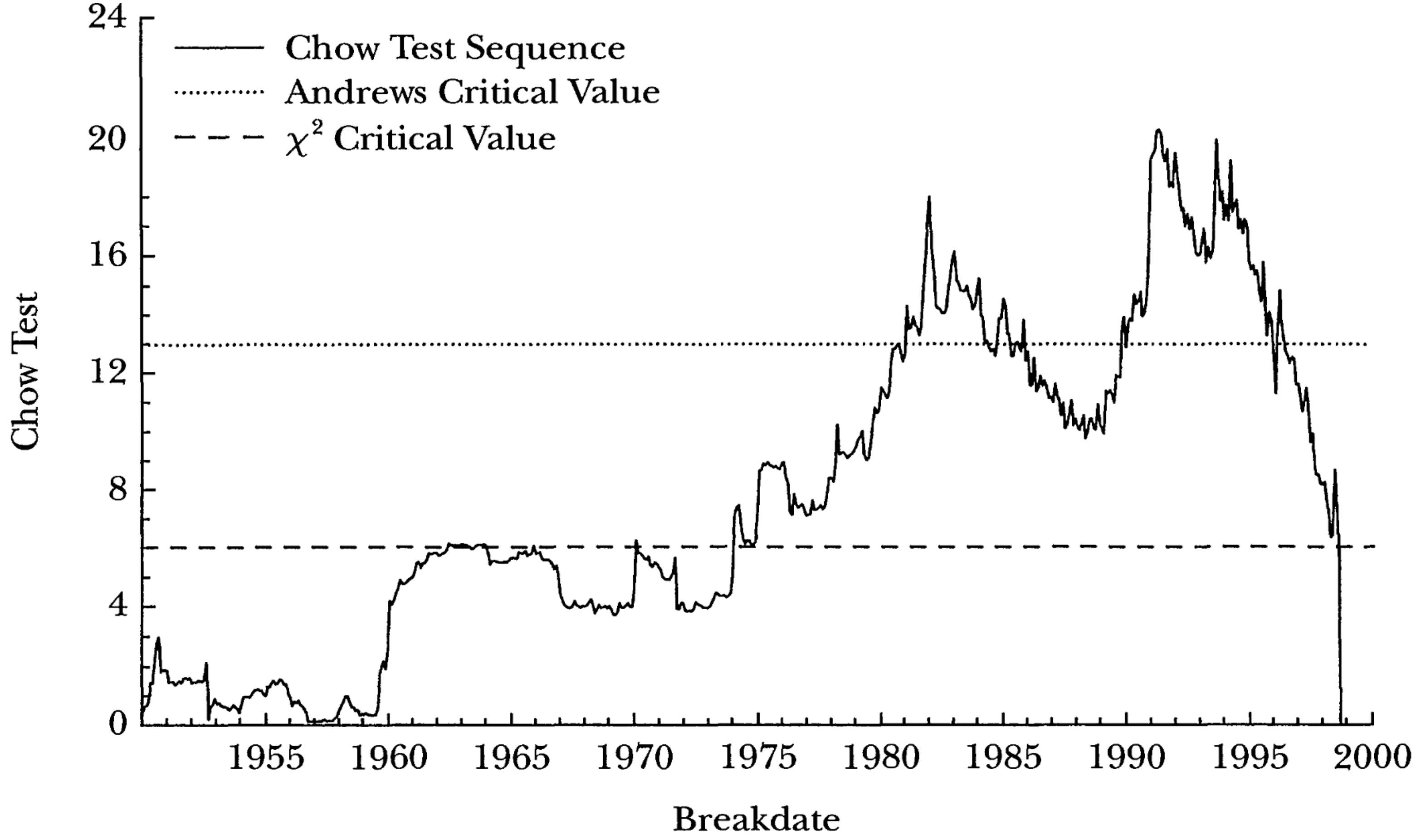
Number of Restrictions (q)	10%	5%	1%
1	7.12	8.68	12.16
2	5.00	5.86	7.78
3	4.09	4.71	6.02
4	3.59	4.09	5.12
5	3.26	3.66	4.53
6	3.02	3.37	4.12
7	2.84	3.15	3.82
8	2.69	2.98	3.57
9	2.58	2.84	3.38
10	2.48	2.71	3.23
11	2.40	2.62	3.09
12	2.33	2.54	2.97
13	2.27	2.46	2.87
14	2.21	2.40	2.78
15	2.16	2.34	2.71
16	2.12	2.29	2.64
17	2.08	2.25	2.58
18	2.05	2.20	2.53
19	2.01	2.17	2.48
20	1.99	2.13	2.43

These critical values apply when $\tau_0 = 0.15T$ and $\tau_1 = 0.85T$ (rounded to the nearest integer), so the F-statistic is computed for all potential break dates in the central 70% of the sample. The number of restrictions q is the number of restrictions tested by each individual F-statistic. Critical values for other trimming percentages are given in Andrews (2003).

Note that these critical values are larger than the $F_{q,\infty}$ critical values – for example, $F_{1,\infty}$ 5% critical value is 3.84.

Figure 1

Testing for Structural Change of Unknown Timing: Chow Test Sequence as a Function of Breakdate



Testing for structural break with unknown date

- Perron (1989) shows that failure to allow for an existing break leads to a bias that reduces the ability to reject a false unit root null hypothesis.
- To overcome this, the author proposes allowing for a known or exogenous structural break in the Augmented Dickey-Fuller (ADF) tests.
- Following this development, many authors, including Zivot and Andrews (1992) and Perron (1997), proposed determining the break point ‘endogenously’ from the data.

Testing for structural break with unknown date

- Enders (2004) argues that Perron-Vogelsang (1992) unit root tests are more appropriate “if the date of the break is uncertain”.

- Shrestha and Chowdhury (2005) argue that, in the case of a structural break, the testing power of the Perron-Vogelsang unit root test is superior to that of the Zivot-Andrews test.

- Applying the unit root tests which allow for the possible presence of the structural break prevents obtaining a test result which is possibly biased towards non-rejection, as suspected by Perron (1989).

- Also, since this procedure can identify the date of the structural break, it facilitates the analysis of whether a structural break on a certain variable is associated with a particular event such as a change in government policy, a currency crisis, war and so forth.

Multiple Structural Breaks

- The Zivot-Andrews and Perron-Vogelsang (1992) unit root tests allow for one structural break, whereas the Clemente-Montanes-Reyes (1998) unit root test allows for two structural breaks in the mean of the series196.

- Clemente *et al* (1998) base their approach on Perron and Vogelsang (1992), allowing for the possibility of having two structural breaks in the mean of the series.

Multiple Structural Breaks

- In these tests, the null hypothesis is that the series has a unit root with structural break(s) against the alternative hypothesis that they are stationary with break(s).

- The advantage of these tests is that they do not require an *a priori* knowledge of the structural break dates.

- Ben-David *et al* (2003) cautions that “just as failure to allow one break can cause non-rejection of the unit root null by the Augmented Dickey –Fuller test, failure to allow for two breaks, if they exist, can cause non-rejection of the unit root null by the tests which only incorporate one break” (Ben-David *et al*, 2003: 304).

Multiple Structural Breaks

- Lumisdaine and Papell (1997) extended the Zivot and Andrews (1992) model to accommodate two structural breaks.

- However, this test was criticized for the absence of the breaks under the null hypothesis of unit root as this could result in a tendency for these tests to suggest evidence of stationarity with breaks (see Glynn *et al*, 2007).

Multiple Structural Breaks

- Hence, The Perron-Vogelsang and Clemente-Montanes-Reyes unit root tests are more preferable.

- Both of these tests offer two models:
- (1) an additive outliers (AO) model, which captures a sudden change in the mean of a series; and

- (2) an innovational outliers (IO) model, which allows for a gradual shift in the mean of the series.

Multiple Structural Breaks

- According to Baum (2004), if the estimates of the Perron-Vogelsang and Clemente-Montanes-Reyes unit root tests provide evidence of significant additive or innovational outliers in the time series, the results derived from ADF and PP tests are doubtful, as this is evidence that the model excluding structural breaks is misspecified.

- Therefore, in applying unit root tests in time series that exhibit structural breaks, only the results from the Clemente-Montanes-Reyes unit root tests should be considered if the two structural breaks indicated by the respective tests are statistically significant (at the 5% level as used by STATA).

Multiple Structural Breaks

- On the other hand, if the results of the Perron-Vogelsang and Clemente-Montanes-Reyes unit root tests show no evidence of two significant breaks in the series, the results from the Perron–Vogelsang unit root tests are considered.

- If these tests show no evidence of a structural break, the ADF and PP tests can be considered.