Mis-specification in Linear Regressions

Laurent Ferrara[1]

[1]U. Paris Nanterre

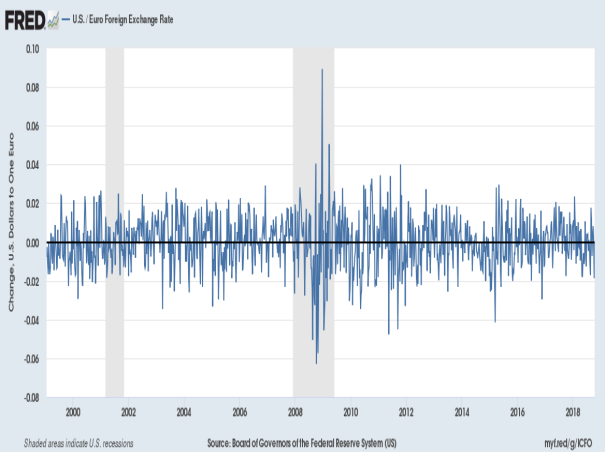
M1 U. Paris Nanterre, Feb. 2019

Plan

Heteroskedasticity

- Standard violation of hypothesis: σ^2 non-constant overtime
- Issue: errors are not observable
- A good proxy: residuals $\hat{\varepsilon}_t$.
- A first hint: Plot the residuals or squared-residuals and check the variability

USD/EUR daily exchange rate (log-returns, 1999-2018)



Heteroskedasticity: Goldfeld-Quandt test

- The null hypothesis is that errors are homoskedastic, ie:

$$H_0 : \sigma_i^2 = \sigma$$

vs

$$H_1 : \sigma_i^2 = cX_i^2$$

with $c > 0$, meaning that the variance increases with the explanatory variable X.

Heteroskedasticity: Goldfeld-Quandt test

Implementation:

1. The observations (y_i, x_i) for $i = 1, \ldots, T$ are ranked according to the values of x_i, the lowest values of x_i being in the first part of the sample.
2. The d central observations of the ordered sample are excluded (about 20% of the sample, ie: $d = 0.2T$) and the remaining observations are divided in 2 sub-samples of size $(T - d)/2$.
3. For each sub-sample, the Residuals Sum of Squares (RSS), such as $RSS = \sum \hat{\varepsilon}_i^2$, are computed and the following Golfeld-Quandt ratio is given by:

$$GQ = \frac{RSS_2}{RSS_1}$$

Heteroskedasticity: Goldfeld-Quandt test

- Assuming that all other hypotheses hold (including Normality of errors), then under the null H_0, the test statistics GQ is distributed as $F(p, p)$ where $p = [(T - d)/2] - k$
- Intuitively, if σ_i^2 is constant, then both RSS should be similar in size and GQ should be close to 1.

Heteroskedasticity: Breusch-Pagan-Godfrey test

- Test the null hypothesis:

$$H_0 : \sigma_i^2 = \sigma$$

vs

$$H_1 : \sigma_i^2 = \gamma + \delta Z$$

ie: we assume there exists an unknown relationship between the error's variance and one or a set of variables Z (possibly $Z = X$).

Heteroskedasticity: Breusch-Pagan-Godfrey test

Implementation:

1. Regress the squared OLS residuals on Z, ie: $\hat{\varepsilon}_t^2 = \gamma + \delta Z + \nu_i$
2. The BPG test statistics is given by :

$$BPG = TR^2$$

3. Under the null H_0, BPG is distributed as $\chi^2(q)$ where q is the dimension of Z.

Intuitively, under the null of homoskedasticity, R^2 and thus BPG should be very low.

Parameter instability

- In principle, the linear regression model satisfies the property that parameters are stable across time, ie $\beta_{it} = \beta_i$.
- In practice, there is often a break in this linear relationship.
- Assume that from $t = 1, \ldots, T_1$ $(t \in \mathbb{T}_1)$ we have:

$$Y_t = \beta_1 X_t + \varepsilon_{1t}$$

and from $t = T_1 + 1, \ldots, T$ $(t \in \mathbb{T}_2)$:

$$Y_t = \beta_2 X_t + \varepsilon_{2t}$$

with ε_{1t} and ε_{2t} being independent Gaussian WNs with respective variance σ_1^2 and σ_2^2

Parameter instability

- For forecasting purposes, assume that $\mathbb{T}_1$ is the estimation sample and $\mathbb{T}_2$ is the forecasting sample
- The optimal forecast is given by:

$$\hat{Y}_t = \beta_1 X_t$$

- The actual values are:

$$Y_t = \beta_2 X_t + \varepsilon_{2t}$$

- The forecast errors are:

$$e_t = (\beta_2 - \beta_1) X_t + \varepsilon_{2t}$$

Tests for parameter changes

- Let's consider the following sets of hypotheses:

$$H_0' : \beta_1 = \beta_2, \qquad H_0'' : \sigma_1^2 = \sigma_2^2,$$

vs

$$H_1' : \beta_1 \neq \beta_2, \qquad H_1'' : \sigma_1^2 \neq \sigma_2^2,$$

where $H_0 = H_0' \bigcup H_0''$

- Assume that the second subsample $\mathbb{T}_2$ is greater than the number of regressors k; so that all the models can be estimated for all subsamples leading to 3 RSS: RSS_T (Total), RSS_1 (for $\mathbb{T}_1$) and RSS_2 (for $\mathbb{T}_2$)

Tests for parameter changes

- The Chow test for the null $H_0' : \beta_1 = \beta_2 | H_0''$ is

$$CH1 = \frac{RSS_T - RRS_1 - RSS_2}{RSS_1 + RSS_2} \times \frac{T - 2k}{k}$$

and under H_0', $CH1 \sim F(k, T - 2k)$

- Intuitively, if parameters are stable (ie under H_0), then RSS_T should be close $RSS_1 + RSS_2$, then $CH1$ should be low. Otherwise $RSS_T > RSS_1 + RSS_2$

Tests for parameter changes

- The Chow test is built on the assumption that the variance is constant across the 2 subsamples, but this has to be checked
- The Chow test statistic for the null $H_0'': \sigma_1^2 = \sigma_2^2$ vs $H_0'': \sigma_1^2 < \sigma_2^2$

$$CH2 = \frac{RSS_2}{RSS_1} \frac{T_1 - k}{T_2 - k}$$

  and under H_0'', $CH2 \sim F(T_2 - k, T_1 - k)$

- As $CH1$ and $CH2$ are shown to be independent, it is convenient to first use CH_2, then CH_1.

Tests for parameter changes

- Note that the Chow test requires the specification of the date of the parameter change
- If the date is unknown, we can compute the test for all possible break dates and take the max of the resulting test statistics
- Unfortunately, the max-Chow test doesn't have a standard distribution, so the use of proper critical values must be tabulated (Andrews, 1993, or Andrews and Ploberger, 1994)
- Similar tests with multiple, known or unknown, break dates are available (Bai and Perron, 1998)

Recursive methods for detecting parameter changes

- Plotting the recursive OLS estimators of β, with confidence bands, is useful in detecting instability
- The general form of the recursive OLS estimator for $t = T_0, T_0 + 1, \ldots, T$, $(T_0 > k)$ is:

$$\hat{\beta}_t = (\bar{X}_t' \bar{X}_t)^{-1} \bar{X}_t' \bar{Y}_t$$

  where the matrices $\bar{X}_t'$ and vectors $\bar{Y}_t$ contain the first t observations.
- The corresponding variance estimator for $\hat{\beta}_t$ is

$$V(\hat{\beta}_t) = \hat{\sigma}_t^2 (\bar{X}_t' \bar{X}_t)^{-1}$$

Recursive methods for detecting parameter changes

- Formal tests for parameter instability can be computed from the 1-step ahead recursive residuals, corresponding to 1-step ahead forecast errors:

$$\tilde{\varepsilon}_t = (Y_t - \hat{\beta}_{t-1} X_t) = \varepsilon_t + (\beta - \hat{\beta}_{t-1}) X_t$$

with

$$\hat{\beta}_{t-1} = (\bar{X}'_{t-1} \bar{X}_{t-1})^{-1} \bar{X}'_{t-1} \bar{Y}_{t-1}$$

and

$$\tilde{\sigma}_t^2 = V(\tilde{\varepsilon}_t) = \sigma^2 \left(1 + X_t (\bar{X}'_{t-1} \bar{X}_{t-1})^{-1} X'_t\right)$$

- Standardized recursive residuals are for $t = k+1, \ldots, T$:

$$\tilde{\omega}_t = \tilde{\varepsilon}_t / \tilde{\sigma}_t$$

Recursive methods for detecting parameter changes

- The CUSUM statistics is given by:

$$CUSUM_t = \sum_{j=k+1}^{t} \frac{\tilde{\omega}_j}{\tilde{\sigma}_{\omega,T}}$$

with

$$\tilde{\sigma}^2_{\omega,T} = (T-k)^{-1} \sum_{t=k+1}^{T} (\tilde{\omega}_t - \bar{\omega})$$

- Under the null $H_0 : \beta_{k+1} = \ldots = \beta_T$, $CUSUM_t$ has zero mean and variance proportional to $t-k-1$

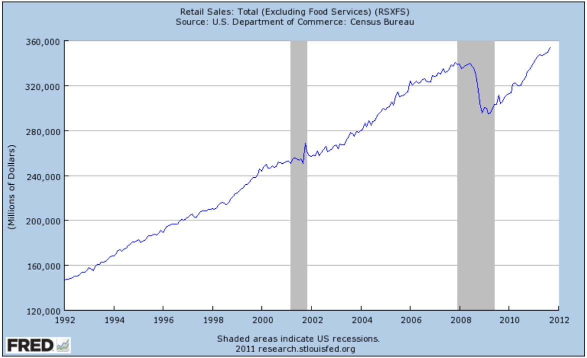
Recursive methods for detecting parameter changes

- The CUSUMQ statistic is given by:

$$CUSUMQ_t = \frac{\sum_{j=k+1}^{t} \tilde{\omega}_j^2}{\sum_{j=k+1}^{T} \tilde{\omega}_j^2}$$

- Under the null of parameter stability, $CUSUMQ_t$ has an asymptotic $\chi^2(t)$ distribution

Example: US sales during the Great Recession



Dummy variables

- An easy way to account for breaks in the linear relationship by using specific binary variables
- A permanent change (Shift):

$$D_t^{(t')} = \begin{cases} 0 & \text{if } t < t', \\ 1 & \text{if } t \geq t' \end{cases} \tag{1}$$

- A temporary change (Impulse):

$$P_t^{(t')} = \begin{cases} 0 & \text{if } t \neq t', \\ 1 & \text{if } t = t' \end{cases} \tag{2}$$

Dummy variables

- More generally

$$D_t = \begin{cases} 1 & \text{if} \qquad t \in \mathbb{T}, \\ 0 & \text{if} \qquad t \notin \mathbb{T}, \end{cases} \tag{3}$$

where $\mathbb{T}$ can represent many different periods over the sample, such as recessions, credit crunches, financial crises, banking crises ...

Dummy variables

- Example of a consumption equation where conso c_t depends on income inc_t:

$$c_t = \beta_1 + \beta_2 inc_t + \varepsilon_t$$

- Assume D_t is dummy variable that takes 1 during credit crunch periods and 0 otherwise
- D_t can play 3 various roles in the model:
  1. **M1:** Structural change in autonomous consumption
     $c_t = \beta_1 + \alpha D_t + \beta_2 inc_t + \varepsilon_t$
  2. **M2:** Structural change in the marginal propensity to consume
     $c_t = \beta_1 + \beta_2 inc_t + \gamma D_t inc_t + \varepsilon_t$
  3. **M3:** Structural change in both the autonomous consumption and the marginal propensity to consume
     $c_t = \beta_1 + \alpha D_t + \beta_2 inc_t + \gamma D_t inc_t + \varepsilon_t$

Dummy variables

- Model M3 can be rewritten as a time-varying parameter model:

$$c_t = \begin{cases} \beta_1 + \beta_2 inc_t + \varepsilon_t & \text{if} \quad t \notin \text{credit crunch}, \\ \gamma_1 + \gamma_2 inc_t + \varepsilon_t & \text{if} \quad t \in \text{credit crunch}, \end{cases} \tag{4}$$

where: $\gamma_1 = \beta_1 + \alpha$ and $\gamma_2 = \beta_2 + \gamma$

References

- Andrews D. (1993). Test for parameter instability and structural change with unknown *Econometrica* , 61, 821-856
- Andrews, D. and Ploberger, W. (1994). Optimal tests when a nuisance parameter is present only under the alternative *Econometrica* , 62, 1383-1414
- Bai, J. and Perron, P. (1998). Estimating and testing linear models with multiple structural changes *Econometrica* 66, 47-78
- Ghysels, E. and M. Marcellino (2018), *Applied Economic Forecasting using Time Series Methods*, Oxford University Press
- Hamilton, J. (1994). *Time Series Analysis*. Princeton University Press.