Session 6
Quantitative Analysis of Financial Markets
Diagnostic Tests

Christopher Ting

http://www.mysmu.edu/faculty/christophert/

Christopher Ting

: christophert@smu.edu.sg

: 6828 0364

: LKCSB 5036

March 17, 2018

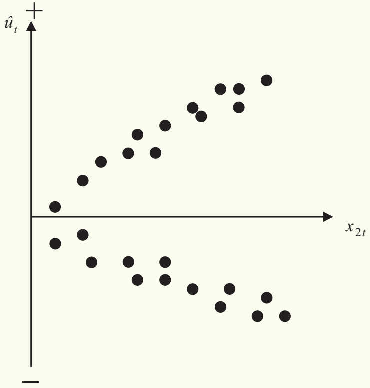
Broad Lesson Plan

Learning Objectives

- Describe the steps involved in testing regression residuals for heteroscedasticity.
- Explain the impact of heteroscedasticity on the ordinary least squares (OLS) parameter and standard error estimation.
- Test for whether the functional form of the model employed is appropriate.
- Describe tests to investigate whether the model parameters are stable.

Homoscedasticity versus Heteroscedasticity

- We have so far assumed that the variance σ_u^2 of the errors is constant.
- But most likely, the errors do not have a constant variance.



The Goldfeld-Quandt (GQ) Test

- Split the total sample of length T into two sub-samples of length T_1 and T_2. The regression model is estimated on each sub-sample and the two residual variances are calculated.
- The null hypothesis is that the variances of the disturbances are equal, $H_0 : \sigma_1^2 = \sigma_2^2$
- The test statistic, denoted GQ, is simply the ratio of the two residual variances where the larger of the two variances must be placed in the numerator.

$$\text{GQ} = \frac{s_1^2}{s_2^2} \sim F\big(T_1 - K, T_2 - K\big)$$

- A problem with the test is that the choice of where to split the sample is usually arbitrary and may crucially affect the outcome of the test.

White's Test

- White's general test for heteroscedasticity is one of the best approaches because it makes few assumptions about the form of the heteroscedasticity.
- Suppose the regression we have carried out is as follows

$$y_t = \beta_1 + \beta_2 x_{2,t} + \beta_3 x_{3,t} + u_t$$

- We want to test $\mathbb{V}(u_t) = \sigma^2$. We estimate the model, obtaining the residuals, $\hat{u}_t$.
- Then run the auxiliary regression

$$\hat{u}_t^2 = \alpha_1 + \alpha_2 x_{2,t} + \alpha_3 x_{3,t} + \alpha_4 x_{2,t}^2 + \alpha_5 x_{3,t}^2 + \alpha_6 x_{2,t} x_{3,t} + v_t.$$

White's Test and χ^2 Statistic

- Obtain R^2 from the auxiliary regression and multiply it by the number of observations, T.

- White shows that

$$T \times R^2 \sim \chi^2(m).$$

  where m is the number of regressors in the auxiliary regression excluding the constant term.

- If the χ^2 test statistic is greater than the corresponding value from the statistical table then reject the null hypothesis that the disturbances are homoscedastic.

Consequences of Heteroscedasticity

- OLS estimation still gives unbiased coefficient estimates, but they are no longer BLUE.

- The implication is that if we still use OLS in the presence of heteroscedasticity, our standard errors could be inappropriate and hence any inferences we make could be misleading.

- Whether the standard errors calculated using the usual formulas are too big or too small will depend upon the form of the heteroscedasticity.

How to Deal with Heteroscedasticity?

- *If* the form (i.e. the cause) of the heteroscedasticity is known, then we can use the generalised least squares method.
- Suppose the error variance is related to another variable z_t by

$$\mathbb{V}(u_t) = \sigma^2 z_t^2$$

- To remove the heteroscedasticity, divide the regression equation by z_t

$$\frac{y_t}{z_t} = \beta_1 \frac{1}{z_t} + \beta_2 \frac{x_{2,t}}{z_t}, + \beta_3 \frac{x_{3,t}}{z_t} + v_t$$

where $v_t = \dfrac{u_t}{z_t}$ is an error term.

- Hence, $\mathbb{V}(u_t) = \sigma^2 z_t^2$, $\mathbb{V}(v_t) = \mathbb{V}\left(\dfrac{u_t}{z_t}\right) = \dfrac{\mathbb{V}(u_t)}{z_t^2} = \dfrac{\sigma^2 z_t^2}{z_t^2} = \sigma^2 \implies$ homoscedastic!

Other Approaches

- Transform the variables into logs or reducing by some other measure of “size”.
- Use White’s heteroscedasticity consistent standard error estimates.
- The effect of using White’s correction is that in general the standard errors for the slope coefficients are increased relative to the usual OLS standard errors.
- The goal is to be “conservative” in hypothesis testing, so that we would need more evidence against the null hypothesis before we could reject it.

Multicollinearity

- Multicollinearity occurs when the explanatory variables are very highly correlated with each other.

- Perfect multicollinearity

  Cannot estimate all the coefficients

  - e.g. suppose $x_3 = 2x_2$
    and the model is $y_t = \beta_1 + \beta_2 x_{2t} + \beta x_{3t} + \beta_4 x_{4t} + u_t$

- Problems if near multicollinearity is present
  - R^2 will be high but the individual coefficients will have high standard errors.
  - The regression becomes very sensitive to small changes in the specification.
  - Thus confidence intervals for the parameters will be very wide, and significance tests might therefore give inappropriate conclusions.

Measuring Multicollinearity

- The easiest way to measure the extent of multicollinearity is simply to look at the matrix of correlations between the individual variables. e.g.

corr	x_2	x_3	x_4
x_2	–	0.2	0.8
x_3	0.2	–	0.3
x_4	0.8	0.3	–

- But another problem: if 3 or more variables are linear
  - e.g. $x_{2t} + x_{3t} = x_{4t}$,

  then this method will not work.

- Note that high correlation between y and one of the x's is not muticollinearity.

Solutions to the Problem of Multicollinearity

- “Traditional” approaches, such as ridge regression or principal component, usually bring more problems than they solve.

- Some econometricians argue that if the model is otherwise OK, just ignore it.

- The easiest ways to “cure” the problems are
  - drop one of the collinear variables
  - transform the highly correlated variables into a ratio
  - go out and collect more data e.g.
    - a longer run of data
    - switch to a higher frequency

Wrong Functional Form

- We have previously assumed that the appropriate functional form is linear, which may not always be true.
- Ramsey's RESET test is a general test for mis-specification of functional form.
- Essentially the method works by adding higher order terms of the fitted values (e.g. $\hat{y}_t^2, \hat{y}_t^3$, etc.) into an auxiliary regression:
  Regress $\hat{u}_t$ on powers of the fitted values:

$$\hat{u}_t = \beta_0 + \beta_1 \hat{y}_t^2 + \beta_2 \hat{y}_t^3 + \cdots + \beta_{p-1} \hat{y}_t^p + v_t$$

  Obtain R^2 from this regression.
- The test statistic is given by $TR^2 \sim \chi^2_{p-1}$.
- If the value of the test statistic is greater than a χ^2 critical value, then reject the null hypothesis that the functional form is correct.

Dealing with Wrong Functional Form

- The RESET test gives us no guide as to what a better specification might be.
- One possible cause of rejection of the test is if the true model is

$$y_t = \beta_1 + \beta_2 x_{2,t} + \beta_3 x_{2,t}^2 + \beta_4 x_{2,t}^3 + u_t$$

  In this case the remedy is obvious.
- Another possibility is to transform the data into logarithms, which will linearise many previously multiplicative models into additive ones:

$$y_t = A x_t^{\beta} e^{u_t} \Longleftrightarrow \ln(y_t) = \alpha + \beta \ln(x_t) + u_t$$

Omission of an Important Variable

- The estimated coefficients on all the other variables will be biased and inconsistent unless the excluded variable is uncorrelated with all the included variables.
- Even if the condition of uncorrelatedness is satisfied, the estimate of the coefficient on the constant term will be biased.
- The standard errors will also be biased.

Inclusion of an Irrelevant Variable

- Also called the error of commission.
- Coefficient estimates will still be consistent and unbiased
- But the variance of the estimators will be inefficient.

Measurement Errors

- If there is measurement error in one or more of the explanatory variables, the assumption that the explanatory variables are non-stochastic is violated.
- Known also as the errors-in-variables problem
- Measurement errors can occur in a variety of circumstances, e.g.
  - Macroeconomic variables are almost always estimated quantities (GDP, inflation, and so on), as is most information contained in company accounts.
  - Sometimes we cannot observe or obtain data on a variable we require and so we need to use a proxy variable – for instance, many models include expected quantities (e.g., expected inflation) but we cannot typically measure expectations.

Measurement Error in the Explanatory Variable(s)

- Suppose we estimate a model $y_t = \beta_1 + \beta_2 x_t + u_t$, where u_t is a disturbance term.
- Suppose further that x_t is measured with error so that instead of observing its true value, we observe a noisy version, $\tilde{x}$, that comprises the actual x_t plus some additional noise, v_t, which is independent of x_t and u_t:

$$\tilde{x}_t = x_t + v_t$$

- Taking the first equation and substituting in for x_t from the second:

$$y_t = \beta_1 + \beta_2(\tilde{x}_t - v_t) + u_t$$

- We can rewrite this equation by separately expressing the composite error term, $(u_t - \beta_2 v_t)$

$$y_t = \beta_1 + \beta_2 \tilde{x}_t + (u_t - \beta_2 v_t)$$

Consequence of Errors-in-Variables Problem

- It should be clear that $\tilde{x}_t$ and the composite error term $(u_t - \beta_2 v_t)$, are correlated since both depend on v_t.
- Thus the requirement that the explanatory variables are non-stochastic does not hold, causing the parameters to be estimated inconsistently
- The size of the bias in the estimates will be a function of the variance of the noise in x_t as a proportion of the overall disturbance variance
- If β_2 is positive, the bias will be negative but if β_2 is negative, the bias will be positive.
- So the parameter estimate will always be biased towards zero as a result of the measurement noise.

Measurement Error and Tests of the CAPM

- The standard approach to testing the CAPM pioneered by Fama and MacBeth (1973) comprises two stages.
- Since the betas are estimated at the first stage rather than being directly observable, they will surely contain measurement error.
- The effect of this has sometimes been termed attenuation bias.
- Tests of the CAPM showed that the relationship between beta and returns was smaller than expected, and this is precisely what would happen as a result of measurement error.
- Various approaches to solving this issue have been proposed, the most common of which is to use portfolio betas in place of individual betas.
- An alternative approach (Shanken,1992) is to modify the standard errors in the second stage regression to adjust directly for the measurement errors.

Measurement Error in the Explained Variable

- Measurement error in the explained variable is much less serious than in the explanatory variable(s).
- When the explained variable is measured with error, the disturbance term will in effect be a composite of the usual disturbance term and another source of noise from the measurement error.
- Then the parameter estimates will still be consistent and unbiased and the usual formulas for calculating standard errors will still be appropriate.
- The only consequence is that the additional noise means the standard errors will be enlarged relative to the situation, where there was no measurement error in y.

Parameter Stability Tests

- So far, we have estimated regressions such as

$$y_t = \beta_1 + \beta_2 x_{2,t} + \beta_3 x_{3,t} + u_t$$

- We have implicitly assumed that the parameters (β_1, β_2, and β_3) are constant for the entire sample period.
- We can test this implicit assumption using parameter stability tests. The idea is essentially to **split the data into sub-periods** and then to estimate up to three models, for each of the sub-parts and for all the data and then to "compare" the RSS of the models.
- There are two types of test we can look at:
  - Chow test (analysis of variance test)
  - Predictive failure tests

The Chow Test

- Split the data into two sub-periods. Estimate the regression over the whole period and for the two sub-periods separately, 3 regressions altogether.
- The restricted regression is now the regression for the whole period while the "unrestricted regression" comes in two parts.
- The statistic is

$$\text{test statistic} = \frac{\text{RSS} - (\text{RSS}_1 + \text{RSS}_2)}{\text{RSS}_1 + \text{RSS}_2} \times \frac{T - 2K}{K} \sim F(K, T - 2K)$$

  where

  RSS = RSS for whole sample
  RSS_1 = RSS for sub-sample 1
  RSS_2 = RSS for sub-sample 2

- If the value of the test statistic is greater than the critical value from the F-distribution, then reject the null hypothesis that the parameters are stable over time.

A Chow Test Example

- Consider the following regression for the CAPM β for the returns on Glaxo.
- We estimate beta using monthly data from 1981–1992.
- Sub-Period 1: 1981M1–1987M10

$$\hat{r}_{gt} = 0.24 + 1.2 r_{mt} \quad T = 82 \quad \text{RSS}_1 = 0.03555$$

- Sub-Period 2: 1987M11–1992M12

$$\hat{r}_{gt} = 0.68 + 1.53 r_{mt} \quad T = 62 \quad \text{RSS}_2 = 0.00336$$

- Whole Period: 1981M1–1992M12

$$\hat{r}_{gt} = 0.39 + 1.37 r_{mt} \quad T = 144 \quad \text{RSS} = 0.0434$$

A Chow Test Example - Results

- The null hypothesis is

$$H_0: \ \alpha_1 = \alpha_2 \ \text{ and } \ \beta_1 = \beta_2$$

- The unrestricted model is the model where this restriction is not imposed

$$\begin{aligned} \text{test statistic} &= \frac{0.0434 - (0.0355 + 0.00336)}{0.0355 + 0.00336} \times \frac{144 - 4}{2} \\ &= 7.698 \end{aligned}$$

- Compare with 5% F(2,140) = 3.06
- We reject H_0 at the 5% level, i.e., the restriction that the coefficients are the same in the two periods is rejected.

The Predictive Failure Test

- Problem with the Chow test is that we need to have enough data to do the regression on both sub-samples, i.e. $T_1 \gg k$, $T_2 \gg k$.
- An alternative formulation is the predictive failure test.
- What we do with the predictive failure test is estimate the regression over a "long" sub-period (i.e. most of the data) and then we predict values for the other period and compare the two.

Predictive Failure Test Statistic

- Run the regression for the whole period (the restricted regression) and obtain the RSS.

- Run the regression for the "large" sub-period and obtain the RSS (called RSS_1). Note we call the number of observations T_1 (even though it may come second).

$$\text{test statistic} = \frac{\text{RSS} - \text{RSS}_1}{\text{RSS}_1} \times \frac{T_1 - K}{T_2},$$

where $T_2 =$ number of observations that the model is attempting to 'predict'.

- The test statistic is distributed as $F(T_2, T_1 - K)$.

Backwards vs Forwards Predictive Failure Tests

- There are 2 types of predictive failure tests:
  - Forward predictive failure tests, where we keep the last few observations for forecast testing
    - Example: We have observations for 1970Q1–1994Q4.
    - So estimate the model over 1970Q1–1993Q4 and forecast 1994Q1–1994Q4.
  - Backward predictive failure tests, where we attempt to “back-cast” the first few observations
    - Example: We have data for 1970Q1–1994Q4
    - We estimate the model over 1971Q1–1994Q4 and backcast 1970Q1–1970Q4.

Predictive Failure Tests – An Example

- We have the CAPM β for Glaxo.
- 1981M1–1992M12 (whole sample)

$$\hat{r}_{gt} = 0.39 + 1.37 r_{Mt} \qquad T = 144 \qquad \text{RSS} = 0.0434$$

- 1981M1–1990M12 (long sub-sample)

$$\hat{r}_{gt} = 0.32 + 1.31 r_{Mt} \qquad T = 120 \qquad \text{RSS}_1 = 0.0420$$

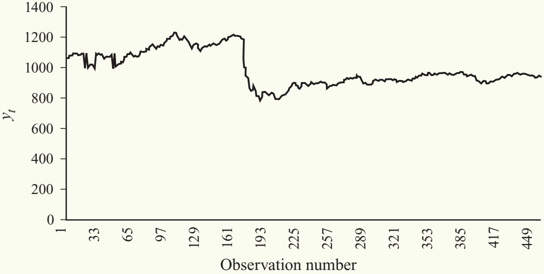
- The test statistic is given by

$$\text{test statistic} = \frac{0.0434 - 0.0420}{0.0420} \times \frac{120 - 2}{24} = 0.164$$

- Compare the test statistic with an $F(24{,}118) = 1.66$ at the 5% level.
- We cannot reject the null hypothesis that the model can adequately predict the last few observations.

How do we decide the sub-parts to use?

- Plot the dependent variable over time and split the data accordingly to any obvious structural changes in the series, e.g.



- Split the data according to any known important historical events (e.g. stock market crash, new government elected)
- Use all but the last few observations and do a predictive failure test on those.

Takeaways

- Many things could go wrong in modeling.
- The most serious ones cause the parameter estimates to be biased and they include
  - Wrong function form
  - Omission of important variables
  - Parameter instability
  - Multicolinearity
  - Measurement errors
- Less serious ones make the variance of the estimate inaccurate and they include
  - Serial correlation
  - Heteroskedasticity
  - Departure from normality
  - Commission error of irrelevant variables
- Nevertheless, multiple linear regression is robust and is the work horse in modeling.