Tutorial 12

keywords: exam revision

estimated reading time: 36 minutes

Quang Bui

October 16, 2018

Introduction

The *Quotes Information for Asset Risk - September 2018* contains raw data on all quotes created during the month of September for LeasePlan NZ's existing and prospective clients. The status of these quotes is as of 12th of October 2018. Although each quote was created on different days in September, we will assume that this data set was captured in a single time period i.e. a cross-sectional data set. Information about each quote is displayed in the following variables:

Question 2

This question is based on a question from S1, 2018 final emxa: A researcher wants to test the Efficient Market Hypothesis (EMH) using weekly percentage returns, denoted by r_t, on the New York Stock Exchange composite index. In its strict form the EMH states that information observable to the market prior to week t should not help to predict the return during week t. If we use only past information on r, the EMH stated as

$$E(r_t|r_{t-1}, r_{t-2}, \dots) = E(r_t)$$

One simple way to test that the EMH holds is to specify the following alternative AR(1) model to describe r_t:

$$r_t = \beta_0 + \beta_1 r_{t-1} + u_t$$

where $E(u_t|r_{t-1}, r_{t-2}, \dots) = 0$ and $Var(u_t|r_{t-1}, r_{t-2}, \dots) = \sigma^2$.

Using data from the first week of January 2004 to the third week of April 2018 gives the following estimation:

$$\hat{r}_t = \underset{(0.096)}{0.086} - \underset{(0.038)}{0.059} r_{t-1}$$
$$n = 689, \quad R^2 = 0.0035, \quad \bar{R}^2 = 0.0020$$

(a)

i. How would you formulate the null hypothesis that the EMH holds based on (4)? Briefly explain your intuition behind your choice of H_0.

The EMH states that based returns will not help to explain returns at time t $\therefore$ the null that EMH holds is given by,

$$H_0 : \beta_1 = 0$$

i.e. weekly percentage returns at time $t-1$ should not help to explain results at time t.

ii. Given the OLS regression results in (5) do you reject or not reject the H_0 from i.? Briefly explain.

$$H_0 : \beta_1 = 0 \quad (EMH\ does\ hold)$$
$$H_1 : \beta_1 \neq 0 \quad (EMH\ does\ not\ hold)$$

We are performing a two-sided t-test.

$$t = \frac{\hat{\beta}_1 - \beta_1}{se(\hat{\beta}_1)} = \frac{\hat{\beta}_1}{se(\hat{\beta}_1)} \sim t_{n-k-1} \quad under\ H_0$$

$$n - k - 1 = 689 - 1 - 1 = 687$$

$$t_{calc} = \frac{-0.059}{0.038} = 1.553$$

We perform this two-sided t-test at the 5% level of significance $\alpha = 0.05$. The critical values can be found from the Stats Table or EViews,

$$t_{crit} \approx 1.96$$
$$-t_{crit} \approx -1.96$$

For this two-sided t-test, we reject H_0 if,

$$t_{calc} > t_{crit}$$

$$OR$$

$$t_{calc} < -t_{crit}$$

Since $t_{calc} = 1.553 < t_{crit} = 1.96$, we fail to reject the null and conclude that there is insufficient evidence from out sample to suggest that the EMH does not hold.

(b) The alternative AR(1) model does not preclude that potentially there could be dependence between returns that are more than one week apart.

i. If the AR(1) model were the correct specification for describing returns, what type of process would you expect u_t to follow? Provide the properties of this process.

If the AR(1) model were the correct specification for describing returns, we would expect the errors in the model to be serially uncorrelated,

$$cov(u_t, u_{t-j}) = 0 \quad for\ all\ t\ and\ j \neq 0.$$

Specifically, we would expect the errors to follow a *white noise* process.

If u_t follows a white noise process,

$$u_t \sim WN(0, \sigma^2)$$

then u_t would have the following white noise properties:

- $E(u_t) = 0 < \infty$ for all t, i.e. the error term has zero mean
- $Var(u_t) = \sigma^2 < \infty$ for all t, i.e. the error term has constant variance
- $cov(u_t, u_{t-j}) = 0$ for all t and $j \neq 0$, i.e. there is no serial correlation in the error term

ii. If the researcher suspects that returns 3 and 4 weeks apart individually add power to the prediction of r_t, what problem do you think that he would be worried about with regard to the behaviour of u_t? Briefly explain.

If r_{t-3} and r_{t-4} helps to predict r_t but are not included as regressors in the model of r_t, then the error term in model of y_t is not longer white noise but exhibits some time dependence,

$$cov(u_t, u_{t-j}) \neq 0 \quad j \neq 0.$$

The research would be worried about serial correlation in the error term as this would affect the validity of the OLS standard errors and by consequence inference based on this OLS regression.

(c) The researcher is also interested in the behaviour of the squared residuals from

$$\hat{r}_t = \underset{(0.096)}{0.086} - \underset{(0.038)}{0.059} r_{t-1}$$

because he is concerned that the variance given past information might not be constant. For this purpose he runs a regression of $\hat{u}_t^2$ of r_{t-1} and obtains the following results:

$$\hat{u}_t^2 = \underset{(0.43)}{4.66} - \underset{(0.201)}{1.104} r_{t-1} + \hat{v}_t$$
$$n = 689, \quad R^2 = 0.042.$$

i. Which problem is the researcher worried about in this case? Define the problem and set up a formal test that makes use of the goodness-of-fit of

$$\hat{u}_t^2 = \underset{(0.43)}{4.66} - \underset{(0.201)}{1.104} r_{t-1} + \hat{v}_t.$$

Clearly state the steps involved in the implementation of this test, the null and alternative hypothesis of the test, the statistic(c) of interest and corresponding distribution(s).

The researcher is concerned that the variance of the error term give past information might not be constant i.e. he is concern that the error term is heteroskedastic.

The regression of $\hat{u}_t^2$ of r_{t-1}:

$$\hat{u}_t^2 = \underset{(0.43)}{4.66} - \underset{(0.201)}{1.104} r_{t-1} + \hat{v}_t$$
$$n = 689, \quad R^2 = 0.042.$$

suggests that the researcher is implementing a Breusch-Pagan test for heteroskedasticity in the error term, where the null hypothesis is that the error term is homoskedastic

$$H_0 : E(u_t^2 | r_{t-1}) = \sigma^2$$

and the alternative hypothesis it is heteroskedastic

$$H_1 : E(u_t^2|r_{t-1}) = \alpha_0 + \alpha_1 r_{t-1} \neq \sigma^2.$$

To perform the BP test

- Estimate the regression model of r_t to obtain the residuals $\hat{u}_t$
- Estimate the auxiliary regression model

  $$\hat{u}_t = \alpha_0 + \alpha_1 r_{t-1} + v_t$$

  to obtain the R^2_{aux}.
- Calculate the test statistics

  $$BG_{calc} = n^* R^2_{aux} = 689^* 0.042 = 28.94$$

- Compared BP_{calc} with the critical value from a chi-squared distribution with $q = 1$ degrees of freedom

  $$\alpha = 0.05$$
  $$BG_{crit} = \chi^2_{0.95, q=1} = 3.84$$

  ($q = 1$ because our BP auxiliary regression contains 1 independent variable).
- Reject H_0 if $BG_{calc} > critical\ value$

ii. What advice would you give this researcher based on your analysis of (c) i.? Briefly explain your answer.

Since $BP_{calc} = 28.94 > cv = 3.84$, we reject the null can conclude that there is sufficient evidence to suggest that the error term is heteroskedastic.

Advice:

- Use heteroskedastic robust standard errors instead of the usual OLS standard errors
- Log transformation of the dependent variable is not valid because returns can be negative
- WLS is tricky because the variance of the error (which is also the variance of the dependent variable i.e. weekly returns) is not proportional to a single variable. It cannot be proportional to r_{t-1},

  $$\cancel{var(\boldsymbol{u}|\boldsymbol{X}) = \sigma^2 \times r_{t-1}}$$

  because $var(\boldsymbol{u}|\boldsymbol{X})$ must be positive and this will not hold since $\sigma^2 > 0$ and r_{t-1} can be negative.

So, among the remedies that we have learnt in this unit, robust standard errors is the only one feasible.

Background

Weighted Least Squares Estimator (Tutorial 9 notes)

It is helpful to consider the WLS estimator as a 2-step estimator:

- At step 1, apply some weighting/transformation to the original model to obtain the weighted model.
- At step 2, estimate the weighted model by OLS.

If the variance of the error has the following known functional form,

$$Var(u_i|x_{i1}, x_{i2}, ...) = \sigma^2 \times h_i$$

then weighing the original model by,

$$w_i = \frac{1}{\sqrt{h_i}}$$

produces the following weighted model,

$$w_i y_i = \beta_0 w_i + \beta_1 w_i x_{i1} + \beta_2 w_i x_{i2} + \cdots + w_i u_i$$

with a constant error variance (homoskedastic error),

$$\begin{aligned} Var(w_i u_i|x_{i1}, x_{i2}, \dots) &= w_i^2 Var(u_i|x_{i1}, x_{i2}, \dots) \\ &= \frac{1}{h_i} Var(u_i|x_{i1}, x_{i2}, \dots) \\ &= \frac{1}{h_i}\sigma^2 h_i \\ &= \sigma^2 \end{aligned}$$

Question 3

(a) A researcher who wished to study the behaviour of a stationary time series $\{y_t\}$ estimated both an AR(2) model,

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + u_t$$

and an ADL(2,1) model,

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \beta_1 x_{t-1} + u_t.$$

The researcher obtained the results reported in Table 1 below (standard errors are reported below the estimated coefficients). Based on the information in Table 1,

Table 1		
	AR(2)	ADL(2,1)
$\widehat{c}$	1.28 (0.53)	1.30 (0.44)
$\widehat{\phi}_1$	-0.31 (0.09)	-0.42 (0.08)
$\widehat{\phi}_2$	-0.39 (0.08)	-0.37 (0.08)
$\widehat{\beta}_1$	–	-2.64 (0.46)
$\overline{R}^2$	0.55	0.71
SSR	475	462
AIC	1.08	1.04
BIC	1.09	1.11
n	200	200

which model do you prefer? Briefly explain.

Since both models have the same dependent variable y, we can use the $\bar{R}^2$, AIC, and BIC to choose a preferred model for y. We prefer the model with:

- A higher $\bar{R}^2$
- A lower AIC
- A lower BIC

Since the ADL(2,1) model has the higher $\bar{R}^2$ and lower AIC,

$$\bar{R}^2_{ARDL(2,2)} = 0.71 > \bar{R}^2_{AR(2)} = 0.55$$
$$AIC_{ARDL(2,2)} = 1.04 < AIC_{AR(2)} = 1.08$$

we prefer the ADL(2,1). We are further in favour of ADL(2,1) model of y because $\hat{\beta}_1$ has a test statistics equal to

$$\frac{-2.64}{0.46} = -5.74$$

which makes x_{t-1} statistically significant in explain y_t (at any reasonable level of significance).

However, it is not incorrect to prefer the AR(2) model because it is more parsimonious and has a lower BIC than the ADL(2,1) model.

There are no unique correct answer, but there are wrong answers.

Explain why it is incorrect to say:

My preferred model is ADL(2,1) because it has the smaller Sum of Squared Residuals (SSR)

Because the ADL(2,1) model has the same regressors as the AR(2) model plus 1 additional regressor so its SSR is necessarily smaller.

(b) When the researcher estimated the model

$$y_t = c + \beta_0 D_t + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \beta_1 x_{t-1} + \beta_2 (D_t x_{t-1}) + u_t$$

by OLS where,

$$D_t = \begin{cases} 1 & for\ t = 1, 2, \ldots, 100 \\ 0 & for\ t = 101, 102, \ldots, 200 \end{cases}$$

he obtain the results reported below (standard errors are reported in parenthesis):

$$\hat{y}_t = \underset{(0.44)}{1.32} + \underset{(0.04)}{0.30} D_t - \underset{(0.07)}{0.39} y_{t-1} - \underset{(0.06)}{0.31} y_{t-2} + \underset{(0.41)}{2.15} x_{t-1} + \underset{(0.07)}{0.15} (D_t x_{t-1})$$

$$SSR = 450 \quad \bar{R}^2 = 0.69 \quad n = 200$$

i. What is

$$\hat{E}(y_t | y_{t-1}, y_{t-2}, x_{t-1})$$

for different values of t. Briefly explain.

Since

$$\begin{aligned} y_t &= c + \beta_0 D_t + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \beta_1 x_{t-1} + \beta_2 (D_t x_{t-1}) + u_t \\ &= E(y_t | y_{t-1}, y_{t-2}, x_{t-1}, D_t) + u_t \end{aligned}$$

that is,

$$E(y_t|y_{t-1}, y_{t-2}, x_{t-1}, D_t) = c + \beta_0 D_t + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \beta_1 x_{t-1} + \beta_2(D_t x_{t-1})$$

so the estimated conditional expectation of y_t is given by,

$$\begin{aligned}\hat{E}(y_t|y_{t-1}, y_{t-2}, x_{t-1}, D_t) &= \hat{c} + \hat{\beta}_0 D_t + \hat{\phi}_1 y_{t-1} + \hat{\phi}_2 y_{t-2} + \hat{\beta}_1 x_{t-1} + \hat{\beta}_2(D_t x_{t-1}) \\ &= 1.32 + 0.30 D_t - 0.39 y_{t-1} - 0.31 y_{t-2} + 2.15 x_{t-1} + 0.15(D_t x_{t-1})\end{aligned}$$

It follows that

$$\hat{E}(y_t|y_{t-1}, y_{t-2}, x_{t-1})$$

for different values of t i.e.

- $D_t = 1$ for $t = 1, 2, \ldots, 100$
- $D_t = 0$ for $t = 101, 102, \ldots, 200$

is given by,

$$\hat{E}(y_t|y_{t-1}, y_{t-2}, x_{t-1}) = \begin{cases} 1.62 - 0.39 y_{t-1} - 0.31 y_{t-2} + 2.30 x_{t-1} & for\ t = 1, 2, \ldots, 100 \\ 1.32 - 0.39 y_{t-1} - 0.31 y_{t-2} + 2.15 x_{t-1} & for\ t = 101, 102, \ldots, 200 \end{cases}$$

ii. What is the estimated immediate and long run effect of a one unit increase in x on y before and after time $t = 100$. Briefly explain.

Since x_t is not included in either equations, x does not have an immediate effect on y before or after time 100.

The estimated long run effect on y of a unit increase in x

$$= \frac{sum\ of\ estimated\ coefficients\ x_t\ and\ its\ lags}{1 - sum\ of\ estimated\ coefficients\ of\ lags\ of\ y_t}$$

$\therefore$ the estimated long run effect on y of a unit increase in x before and after time 100

$$\frac{2.30}{1 - (-0.39 - 0.31)} = 1.35 \quad for\ t = 1, 2, \ldots, 100$$
$$\frac{2.15}{1 - (-0.39 - 0.31)} = 1.26 \quad for\ t = 101, 102, \ldots, 200$$

iii. Test the null **hypothesis** that there is no structural break in either the intercept or the coefficient attached to x_{t-1} in the ADL(2,1) model. State the null and alternative hypothesis, the form of the asymptotic distribution of the test statistic under the null, the sample value and critical value of the test statistic and your conclusion.

Structural break - change in the coefficient after some time period

$$y_t = c + \beta_0 D_t + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \beta_1 x_{t-1} + \beta_2 (D_t x_{t-1}) + u_t$$

$$H_0 : \beta_0 = \beta_2 = 0$$
$$H_1 : \beta_0 \ and/or \ \beta_2 \neq 0$$
$$F = \frac{(SSR_r - SSR_{ur})/2}{SSR_{ur}/(200 - 5 - 1)} \sim F_{2,200-5-1} \quad under \ H_0$$

The SSR_{ur} is given in (b) and SSR_r is given in the ADL(2,1) model in (a) Table 1,

$$SSR_{ur} = 450 \quad SSR_r = 462$$

$$F_{calc} = \frac{(462 - 450)/2}{450/194} = 2.59$$

$$F_{crit} \approx 3.07$$

Since $F_{calc} = 2.59 < F_{crit} \approx 3.07$ we fail to reject the null and conclude that there is insufficient evidence from our sample of a structural break in either the intercept or the coefficient of x_{t-1}. (Correct to say that sample size used to estimate the model is 198 $\therefore$ $F_{2,198-5-1}$ because 2 observations were loss by including a variable lagged 2 periods as a regressor.)

(c) Carefully describe the steps involved in performing a **Breusch-Godfrey test** for autocorrelation up to order 2 in the error term in the ADL(2,1) model

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \beta_1 x_{t-1} + u_t.$$

Make sure you specify the form and the distribution of the test statistic.

We are testing for serial correlation in u_t up to order 2, so we specify a model of $dlgdp$ with errors that follow an AR(2) process.

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \beta_1 x_{t-1} + u_t$$
$$u_t = \rho_1 u_{t-1} + \rho_2 u_{t-2} + e_t$$

$$e_t \sim WN(0, \sigma^2)$$

$$(e_t \sim i.i.d(0, \sigma^2) \ also \ okay)$$

$$H_0 : \rho_1 = \rho_2 = 0$$
$$H_1 : at \ least \ one \ of \ the \ above \ \rho \neq 0$$

To perform the Breusch-Godfrey test for serial correlation in the errors of model y_t up to order 2,

- Estimate the model

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \beta_1 x_{t-1} + u_t$$

- Save the residuals from the above estimated model
- Then estimate the following *auxiliary regression* ...

$$\hat{u}_t = \alpha_0 + \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \alpha_3 x_{t-1} + \rho_1 \hat{u}_{t-1} + \rho_2 \hat{u}_{t-2} + v_t$$

- The test statistic is

$$BG = n_{aux} \times R^2_{aux} \overset{asy}{\sim} \chi^2_2 \quad under\ H_0$$

where n_{aux} is the number of observations in the auxiliary regression

- Calculate the test statistic

$$BG_{calc}$$

and we reject the null of

$$BG_{calc} > BG_{crit}$$

where

$$BG_{crit} = \chi^2_{0.95,2} = 5.99$$

. If we reject the null then there is evidence of serial correlation in the errors.