IIA-3 Econometrics: Supervision 3

Emre Usenmez

2024-11-17

SUPPLEMENTARY QUESTIONS

QUESTION 1

In a study of the Cobb-Douglass production function, a researcher suspects that the parameters are subject to change over time. Data on output, Y, labor input, X_1, and capital stock, X_2, are available for years 1929 to 1967. T represents the time trend. The results obtained are as follows (t-values in parantheses):

$$
\begin{array}{rrlllll}
\text{Full Sample:} & \widehat{logY} = & -3.02 & +1.34\ logX_1 & +0.29\ logX_2 & +0.0052\ T \\
 & & (-6.65) & (14.68) & (4.89) & (2.34) \\
 & & & R^2 = 0.99535 & & \hat{\sigma} = 0.03274 \\
\text{1929-48:} & \widehat{logY} = & -3.22 & +1.36\ logX_1 & +0.32\ logX_2 & +0.0051\ T \\
 & & (-4.63) & (4.95) & (1.36) & (1.40) \\
 & & & R^2 = 0.97853 & & \hat{\sigma} = 0.04449 \\
\text{1949-67:} & \widehat{logY} = & -1.56 & +1.02\ logX_1 & +0.33\ logX_2 & +0.0095\ T \\
 & & (-2.21) & (7.58) & (2.33) & (1.85) \\
 & & & R^2 = 0.99565 & & \hat{\sigma} = 0.0135
\end{array}
$$

a) Conduct a test of the hypothesis that the four regression coefficients are jointly the same in both sub-periods, against the alternative that they differ.

Answer: This question is effectively testing if there is a structural break from the start of 1949 which may be caused by different intercept, different slope coefficient, or both. Suppose the coefficients of the full sample regression are βs, and coefficients of the 1929-48 regression are ψs, and coefficients of the 1949-67 are γs. Our hypothesis is therefore:

$$\mathbb{H}_0 : (\psi_0 = \gamma_0) \cap (\psi_1 = \gamma_1) \cap (\psi_2 = \gamma_2) \cap (\psi_3 = \gamma_3)$$

We can test this in two ways. First is to reparameterize the model and then run an F-test, and the second is to run a Chow test.

First Approach:

We can create a new coefficient $\delta = \psi - \gamma$ whereby our model becomes: $\widehat{logY} = \delta_0 + \delta_1\ logX_1 + \delta_2\ logX_2 + \delta_3 T + \varepsilon$ for which the hypothesis becomes:

$$\mathbb{H}_0 : (\delta_0 = 0) \cap (\delta_1 = 0) \cap (\delta_2 = 0) \cap (\delta_3 = 0)$$

We would then use the F-test for this join hypothesis.

Second Approach:

An alternative appraoch is to use *Chow Test*.[1] This test assumes that:

- The error terms in the subperiod regressions are normally distributed with the same, i.e. homoskedastic, variance σ^2. That is, $u_{1929-48t} \sim N(0, \sigma^2)$ and $u_{1949-67t} \sim N(0, \sigma^2)$.
- The two error terms $u_{1929-48t}$ and $u_{1949-67t}$ are independently distributed.

The Chow test is an F-ratio which means we will need the RSS of both unrestricted and restricted models. Here, *the full sample model is the restricted model* since that is the model we have by imposing the restrictions that all $\psi_j = \gamma_j$ for $j = 0, \ldots, 3$. The RSS of the unrestricted model, on the other hand - and this is a key insight - is the combination of the two sub-sample RSSs.

Why RSS and not R^2 form of F-Test

Note that the Chow test uses RSS and that there is no simple R^2 form of the F-test if separate regressions have been estimated for each group. This is because the TSSs are not the same as $\bar{Y}$ is not the same in both samples.

The steps to carry out a Chow test is as follows:

1. Obtain the restricted model's residual sum of squares, RSS_R, by estimating the regression for the full sample model with $(n - k - 1)$ degrees of freedom, where $n = n_1 + n_2$ with n_1 being the sample size of the first sub-sample, and n_2 being he sample size of the second sub-sample, and where k is the number of regressors in that model.
2. Estimate the first sub-sample model to obtain its residual sum of squares, RSS_1 with $n_1 - k - 1$ degrees of freedom.
3. Do the same for the second sub-sample model to obtain RSS_2 with $n_2 - k - 1$ degrees of freedom.
4. Add the two RSSs to compute the unrestricted model's residual sum of squares: $RSS_{UR} = RSS_1 + RSS_2$.
5. Compute the F-ratio:

$$F = \frac{\dfrac{RSS_R - RSS_{UR}}{k+1}}{\dfrac{RSS_{UR}}{n - 2(k+1)}} = \frac{\dfrac{\text{improvement in fit}}{\text{extra degrees of freedom used up}}}{\dfrac{\text{residual sum of squares remaining}}{\text{degrees of freedom remaining}}}$$

   $\hookrightarrow$ Because we are splitting the sample into two, we are estimating k regressors for each sample plus their intercepts, so we are using up $(k + 1)$ degrees of freedom twice.
6. Compare the F-ratio to the critical F value with $\big((k + 1), n - 2(k + 1)\big)$ degrees of freedom and fail to reject the null hypothesis of *parameter stability*, i.e. no structural change, if F-ratio does not exceed the critical value at the chosen significance level.

Accordingly, we first need to calculate the RSSs. For that, recall that $RSS = \hat{\sigma}^2(n - k - 1)$ where $n = 39, n_1 = 20, n_2 = 19$ because the dates are inclusive. Therefore:

[1] Chow, C Gregory (1960) *Tests of Equality Between Sets of Coefficients in Two Linear Regressions*, Econometrica, 28(3) 591:605

$$
\begin{aligned}
RSS_R &= 0.03274^2 \times 35 && = 0.03751677 \\
RSS_U R &= RSS_1 + RSS_2 \\
&= 0.04449^2 \times 16 + 0.0135^2 \times 15 \\
&= 0.03440351
\end{aligned}
$$

With these we can now calculate our F-ratio:

$$
F = \frac{\frac{0.03751677 - 0.04304995}{4}}{\frac{RSS_{UR}}{n - 2(k+1)}} = \frac{\frac{0.03751677 - 0.03440351}{4}}{\frac{0.03440351}{31}} = 0.7013157
$$

The F-statistic for $\alpha = 0.05$ is 2.678667 and for $\alpha = 0.01$ is 3.992811, and thus we fail to reject the null hypothesis of parameter stability at either of the ψ values.

```r
qf(p=c(0.05, 0.01), df1=4, df2=31, lower.tail = FALSE)
```

```
## [1] 2.678667 3.992811
```

Why RSS equals $\hat{\sigma}^2(n-k-1)$?

Consider how we estimate the error variance, σ^2. First notice that $\sigma^2 = \mathbb{E}(u^2)$, so an unbiased estimator of σ^2 is $\frac{1}{n}\sum_{i=1}^{n} u_i^2$. However, since we do not observe the errors u_i this is not a true estimator. What we have, though, is the estimates of the errors u_i which are the OLS residuals $\hat{u}_i$. If we replace the errors with the OLS residuals then we have

$$
\sigma^2 = \frac{\sum_{i=1}^{n} u_i^2}{n} = \frac{RSS}{n}
$$

which is a true estimator because it gives a computable rule for any sample of data on Xs and Y. However, this is biased because it does not account for the restrictions that must be satisfied by the OLS residuals. These restrictions are given by the two OLS first order conditions:

$$
\sum_{i=1}^{n} \hat{u}_i = 0, \qquad \sum_{i=1}^{n} X_i \hat{u}_i = 0
$$

for a simple regression with one regressor. In a way, if we know $n-k-1$ residuals, we can always get the other remaining residuals by using the restrictions implied by the first order conditions. Therefore there are only $n-k-1$ degrees of freedom in the OLS residuals, as opposed to n degrees of freedom in the errors.
The unbiased estimator of the error variance is therefore:

$$
\hat{\sigma}^2 = \frac{\sum_{i=1}^{n} \hat{u}_i^2}{n-k-1} = \frac{RSS}{n-k-1}
$$

where in simple regression with one regressor, $k = 1$.

(b) It is believed that irrespective of whether the form of the relationship has changed over the two periods (i.e. whether the coefficients of the equation have changed), there has still been a structural break. Test this hypothesis and use this result to comment on the assumptions made in part (a). What limitations are there to these methods for testing for stability over the whole period.

Answer: What this question is effectively asking is that whether the error variances of the two subperiod regressions are the same. Recall from part (a) one of the two key assumptions of the Chow test is that the variances of the errors are homoskedastic. Therefore the hypothesis is:

$$\mathbb{H}_0 : \sigma_\psi = \sigma_\gamma \qquad \text{or} \qquad \mathbb{H}_0 : \frac{\sigma_\psi}{\sigma_\gamma} = 1 \mathbb{H}_1 : \sigma_\psi \neq \sigma_\gamma \qquad textor \quad \mathbb{H}_0 : \frac{\sigma_\psi}{\sigma_\gamma} \neq 1$$

If the variances of the two subpopulations are the same, i.e. $\sigma_\psi = \sigma_\gamma$, as assumed by the Chow test, then the ratios of the ratios of estimated error variance to population error variance has an F distribution with $(n_1 - k - 1)$ and $(n_2 - k - 1)$ degrees of freedom in the numerator and denominator, respectively. That is,

$$\frac{\frac{\hat{\sigma}^2_\psi}{\sigma^2_\psi}}{\frac{\hat{\sigma}^2_\gamma}{\sigma^2_\gamma}} \sim F_{(n_1-k-1),(n_2-k-1)}.$$

Notice that if $\sigma_\psi = \sigma_\gamma$ then this ratio and thus the F-test becomes:

$$F = \frac{\sigma^2_\psi}{\sigma^2_\gamma}$$

where by convention the larger estimated variance is in the numerator.

Accordingly, our F-statistic is:

$$F = \frac{0.04449^2}{0.0135^2} = 10.86069.$$

Since this is a two-tailed test, but we are putting the higher variance in the numerator, then we can treat it as one-sided test with alternative hypothesis is greater than 1, The critical values for $\alpha = 0.05$ with $(16, 15)$ degrees of freedom is 2.384875, and for $\alpha = 0.01$ it is 3.485246.

Thus at both α levels we can reject the null hypothesis and conclude that the subperiod variances are not the same at $\alpha = 0.01$. This means, the assumption of Chow test does not hold and we shouldn't use the Chow test, at least not in this form. There are modifications to Chow test that can be utilized but that is beyond this class.

Another point regarding the Chow test to bear in mind is that it is sensitive to the choice of the time at which we divide the subperiods. The F values would be different if the cut-off point was 1947 or 1949.

Finally, the Chow test will tell us only if the two regressions are different but not whether the difference is due to the intercepts, the slopes, or both. We can use dummy variables for that, though.

```r
qf(c(0.05, 0.01), df1 = 16, df2 = 15, lower.tail = FALSE)
```

```
## [1] 2.384875 3.485246
```

QUESTION 2

The following demand for money function was estimated from 60 observations, for which $\sum(M-\bar{M})^2 = 45600$:

$$\hat{M}_t = 284 + 0.56Y_t - 0.43M_{t-1} \qquad R^2 = 0.841$$

When a further 8 observations became available, the equation was re-estimated. The pooled data had $\sum(M-\bar{M})^2 = 50100$, and the re-estimated equation $R^2 = 0.818$. Carry out a Chow test for predictive failure. What do you conclude from your results? Explain carefully the role of the dummy variable in this test.

Answer: The approach to the Chow test is similar to Question 1 whereby the RSS for restricted model is the pooled data. Hwoever, notice we are not running a two separate regressions on the subsamples. So we follow the first approach from Question 1(a) by taking the differences of the model:

$$\mathbb{H}_0 : (\delta_0 = 0) \cap (\delta_1 = 0) \cap (\delta_2 = 0)$$

where δs are the coefficients of model D_t which is the difference between the initial model and the model after the new observations.

Accordingly, we first need to derive the respective RSSs using the identity $R^2 = 1 - \frac{RSS}{TSS}$ or $RSS = (1-R^2)TSS$ where $TSS = \sum(M-\bar{M})^2$:

$$RSS_R = (1 - 0.841) \times 45,600 = 7,250.4$$
$$RSS_{UR} = (1 - 0.818) \times 50,100 = 9,118.2$$

We can then calculate the Chow test whereby

$$F = frac\frac{9118.2 - 7250.4}{8}\frac{7250.4}{57} = 1.835495$$

The F critical values for $\alpha = 0.05$ and for $\alpha = 0.01$ with $(8, 57)$ degrees of freedom are 2.105599 and 2.840694 respectively. Accordingly we cannot reject the null hypothesis and conclude that there is no predictive failure.

```r
qf(c(0.05, 0.01), df1 = 8, df2 = 57, lower.tail = FALSE)
```

```
## [1] 2.105599 2.840694
```

— —

QUESTION 3

(a) Use the dataset sup3.xls to import the following variables for the period 1955-1990:

W = Wages and Salaries (£ million), CSO code: CFAJ_AU

P = Implied Deflator for Consumers Expenditure, CSO code: GIEF_AU

E = Employees in Employment, CSO code: BCAD_AU

WF = Workforce, CSO code: DYDB_AU

U = Unemployed, CSO code: BCAB_AU

Load the libraries:

```
libraries <- c("haven",       # to import/export SPSS, STATA, SAS files
               "readxl",      # to import/export Excel files
               "tidyverse",   # for tidy data
               "ggplot2",     # for visualization
               "gridExtra",   # to plot graphs in grids
                "rstatix")     # converts stats functions to a tidyverse-friendly format, and can use `

# lapply(libraries, library, character.only=TRUE) will load the libraries
```

Load the data:

```
salaries_df <- read_excel("../Data/sup3.xls", sheet = 1)
```

Briefly examine the data frame

```
# Yuo can use any of the following to examine data frame (df):
# `dim()`: for its dimensions, by row and column
# `str()`: for its structure
# `summary()`: for summary statistics on its columns
# `colnames()`: for the name of each column
# `head()`: for the first 6 rows of the data frame
# `tail()`: for the last 6 rows of the data frame
# `View()`: for a spreadsheet-like display of the entire data frame

summary(salaries_df)
```

```
##       Year         CFAJ_AU          GIEF_AU          BCAD_AU
##  Min.   :1955   Min.   : 10210   Min.   : 12.20   Min.   :21067
##  1st Qu.:1964   1st Qu.: 17430   1st Qu.: 15.55   1st Qu.:21890
##  Median :1972   Median : 35890   Median : 24.75   Median :22466
##  Mean   :1972   Mean   : 76269   Mean   : 47.06   Mean   :22285
##  3rd Qu.:1981   3rd Qu.:127400   3rd Qu.: 81.03   3rd Qu.:22740
##  Max.   :1990   Max.   :271394   Max.   :127.90   Max.   :23257
##     DYDB_AU         BCAB_AU
##  Min.   :24180   Min.   : 210.0
##  1st Qu.:25211   1st Qu.: 406.0
##  Median :25634   Median : 626.5
##  Mean   :26005   Mean   :1175.9
##  3rd Qu.:26692   3rd Qu.:1602.8
##  Max.   :28437   Max.   :3229.0
```

We can change the column names to something more memorable:

```r
colnames(salaries_df) <- c("Year", "W", "P", "E", "WF", "U")
```

(b) Use the data set to estimate the following wage equation by the OLS method:

$$\Delta lnW_t = \beta_0 + \beta_1 lnP + \beta_2 lnP_{t-1} + \beta_3 lnE_t + \beta_4 lnE_{t-1} + \beta_5 lnUR_t + \beta_6 lnUR_{t-1} + \varepsilon_t$$

\textbf{where β_i are constants, Δ denotes the first difference operator and $UR_t = \frac{U_i}{WF_i}$.

Verify that your estimate of β_1 is .91170. Also show that the above is a generalized version of the following hypothesis:

$$\Delta ln\left(\frac{W}{EP}\right)_t = \beta_0 + \varepsilon.$$

How would you explain the inclusion of the UR terms?

Answer:

```r
#Transform the data to get it ready for the regression
salaries_new_df<- salaries_df %>%
  mutate(lnW = log(W),
         DlnW = lnW - lag(lnW,1),
         lnP = log(P),
         lag_lnP = lag(lnP,1),
         lnE = log(E),
         lag_lnE = lag(lnE,1),
         lnUR = log(U/WF),
         lag_lnUR = lag(lnUR,1))
# Run the regression
lm1 <- lm(DlnW ~ lnP + lag_lnP + lnE + lag_lnE + lnUR + lag_lnUR, data = salaries_new_df)
summary(lm1)
```

```
##
## Call:
## lm(formula = DlnW ~ lnP + lag_lnP + lnE + lag_lnE + lnUR + lag_lnUR,
##     data = salaries_new_df)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.052235 -0.010767  0.001441  0.007637  0.046109
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.87957    1.96811   0.447   0.6584
## lnP          0.91170    0.10094   9.032 8.67e-10 ***
## lag_lnP     -0.89866    0.10072  -8.923 1.12e-09 ***
## lnE          0.62385    0.38833   1.607   0.1194
```

```
## lag_lnE     -0.71830    0.35308  -2.034   0.0515 .
## lnUR        -0.03547    0.02731  -1.299   0.2046
## lag_lnUR     0.01994    0.02391   0.834   0.4114
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.02058 on 28 degrees of freedom
##   (1 observation deleted due to missingness)
## Multiple R-squared:  0.8324, Adjusted R-squared:  0.7965
## F-statistic: 23.18 on 6 and 28 DF,  p-value: 1.175e-09
```

We can therefore see that only lnP and lag_lnP coefficients are significant. However, F-stat is pretty high which suggests presence of multicollinearity. We can look at the variance inflation factor, VIF:

```r
car::vif(lm1)
```

```
##        lnP    lag_lnP        lnE    lag_lnE       lnUR   lag_lnUR
## 560.948084 539.661549   8.784345   7.269653  40.557237  32.997711
```

where we see each of the VIF value is greater than 5 suggesting serious collinearity.

(c) Append the variable IP (strength of Income Policy Index) in Table 1 to your data set (use the generate command in Stata to generate a new variable, or copy and paste the data from the question sheet) and run the above regression including IP as an additional regressor and give both a statistical and an economic interpretation of your results.

Answer:

```r
#Append IP with data from the question

salaries_new_df <- salaries_new_df %>%
  mutate(IP = 0)
salaries_new_df$IP[salaries_new_df$Year == 1962] <- 1.0
salaries_new_df$IP[salaries_new_df$Year == 1965] <- 1.0
salaries_new_df$IP[salaries_new_df$Year == 1966] <- 1.5
salaries_new_df$IP[salaries_new_df$Year == 1967] <- 1.0
salaries_new_df$IP[salaries_new_df$Year == 1968] <- 1.5
salaries_new_df$IP[salaries_new_df$Year == 1969] <- 1.75
salaries_new_df$IP[salaries_new_df$Year == 1975] <- 1.0
salaries_new_df$IP[salaries_new_df$Year == 1976] <- 3.0
salaries_new_df$IP[salaries_new_df$Year == 1977] <- 4.5
salaries_new_df$IP[salaries_new_df$Year == 1978] <- 1.0
salaries_new_df$IP[salaries_new_df$Year == 1979] <- 1.0

#Run the regression
lm2 <- lm(DlnW ~ lnP + lag_lnP + lnE + lag_lnE + lnUR + lag_lnUR + IP, data = salaries_new_df)
summary(lm2)
```

```
##
## Call:
## lm(formula = DlnW ~ lnP + lag_lnP + lnE + lag_lnE + lnUR + lag_lnUR +
##     IP, data = salaries_new_df)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.024935 -0.009426 -0.001330  0.006936  0.037838
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -2.008500   1.700503  -1.181  0.24785
## lnP          0.951886   0.080394  11.840 3.37e-12 ***
## lag_lnP     -0.961580   0.081034 -11.866 3.20e-12 ***
## lnE          1.098018   0.327043   3.357  0.00235 **
## lag_lnE     -0.887196   0.282076  -3.145  0.00401 **
## lnUR        -0.003416   0.022894  -0.149  0.88250
## lag_lnUR     0.012673   0.018987   0.667  0.51016
## IP          -0.014722   0.003492  -4.216  0.00025 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.01627 on 27 degrees of freedom
##   (1 observation deleted due to missingness)
## Multiple R-squared:  0.899,  Adjusted R-squared:  0.8728
## F-statistic: 34.31 on 7 and 27 DF,  p-value: 7.754e-12
```

With the introduction of IP, we now see that all are statistically significant except for UR and its lag, as well as the intercept. Thus we can see that by adding a required variable, IP, we reduce the standard error for the whole equation, which in turn, increases the t-values of the other coefficients to rise; except for UR which, as we will see, is because UR has no effect.

R^2 is not helpful here since it will always increase when a new variable is added. Adjusted R^2 also increases from 0.7965 to 0.8728. However, a more useful indicated here is the F-statistic which tests that all slope coefficients are zero. If this is higher, then the equation is an improvement. With the introduction of IP, we see an increase from 23.18 to 34.31.

We can also check the significance of IP by looking at its t-value which is -4.216. Since the p-value is 0.00025 we know it is statistically significant and we don't need to calculate the t-statistic. As long as p-value is less than 0.05 it is significant at the 5 level.

(d) Test the following set of restrictions jointly by re-specifying your preferred equation:

$$\beta_1 + \beta_2 = 0; \quad \beta_3 + \beta_4 = 0; \quad \beta_5 + \beta_6 = 0$$

Verify your results using the Wald Test (the 'test' command in Stata). Interpret your results.

Answer: We can do this in two different ways. First is to use the 'test' command in Stata, or 'linearHypothesis' command from the 'car' package in R. The second way is to impose the constraint on the equation and then compare the constrained and unconstrained equations, testing the restriction on the constrained equation.

```
# Approach 1:

car::linearHypothesis(lm2, c("lnP=-lag_lnP", "lnE=-lag_lnE", "lnUR=-lag_lnUR"))
```

```
## Linear hypothesis test
##
## Hypothesis:
## lnP  + lag_lnP = 0
## lnE  + lag_lnE = 0
## lnUR  + lag_lnUR = 0
##
## Model 1: restricted model
## Model 2: DlnW ~ lnP + lag_lnP + lnE + lag_lnE + lnUR + lag_lnUR + IP
##
##   Res.Df       RSS Df  Sum of Sq      F Pr(>F)
## 1     30 0.0077553
## 2     27 0.0071515  3 0.00060376 0.7598 0.5265
```

```
#Approach 2

#Modify the data frame for constrained version:
salaries_new_diff_df <- salaries_new_df %>%
  mutate(DlnP = lnP - lag_lnP,
         DlnE = lnE - lag_lnE,
         DlnUR = lnUR - lag_lnUR)

#Estimate the constrained model:
lm2_constrained <- lm(DlnW ~ DlnP + DlnE + DlnUR + IP, data = salaries_new_diff_df)
```

We will then need to calculate the F-statistic:

$$F = \frac{\frac{RSS_R - RSS_{UR}}{df_1}}{\frac{RSS_{UR}}{df_2}}$$

which means we need to obtain the RSSs from both regressions which can be done either by `deviance` function in R or by calculating the RSS manually:

```
#RSS unrestricted
sum(resid(lm2)^2)
```

```
## [1] 0.00715151
```

```
RSS_unr <- deviance(lm2)

#RSS restricted
sum(resid(lm2_constrained)^2)
```

```
## [1] 0.007755266
```

```r
RSS_res <- deviance(lm2_constrained)

#F-statistic:
Fstat_Q3d <- ((RSS_res - RSS_unr)/3)/(RSS_unr/27)

#F Critical value:
Crit_Value_Q3d <- qf(0.05,df1=3,df2=27,lower.tail = FALSE)

as.table(c("RSS_unr"=RSS_unr, "RSS_res"=RSS_res, "F-Stat"=Fstat_Q3d, "Critical_Value"=Crit_Value_Q3d))
```

```
##        RSS_unr        RSS_res         F-Stat Critical_Value
##    0.007151510    0.007755266    0.759812374    2.960351318
```

In both approaches we obtain F-statistic of 0.7598 which is much lower than the critical value of 2.96, thus we cannot reject the null hypothesis. That is, the constraints seem to be correct.

(e) Test the further restrictions that $\beta_5 = \beta_6 = 0$. (N.B. your unrestricted equation is always your currently preferred equation, so in this case you only need to test $\beta_5 = 0$ explicitly.)