Solution to Exercise 15.20

⋆15.20 Consider the nonnested linear regression models (15.37) and suppose that the data are generated by a DGP in H_1 with parameters $\boldsymbol{\beta}$ and σ_1^2 and with IID normal errors. Calculate the statistics T_1 and T_1', which were defined in equations (15.52) and (15.53), respectively, and show that they are both asymptotically equal to

$$T_1^{\rm OLS} = n^{1/2}\log\biggl(\frac{\hat\sigma_1^2+\hat\sigma_a^2}{\hat\sigma_2^2}\biggr), \tag{15.90}$$

where $\hat\sigma_i^2$, for $i = 1, 2$, is the ML estimate of the error variance from estimating model H_i, and $\hat\sigma_a^2 \equiv n^{-1}\|\boldsymbol{M_Z P_X y}\|^2$.

Show that the statistic $T_1^{\rm OLS}$ is asymptotically proportional to the J statistic for testing H_1 and also, therefore, to the variance encompassing test statistic of Exercise 15.18. Why is it not surprising that the Cox test, which can be interpreted as an encompassing test based on the maximized loglikelihood, should be asymptotically equivalent to the variance encompassing test?

Show that the asymptotic variance of the statistic (15.90) is

$$\frac{4\sigma_1^2}{(\sigma_1^2+\sigma_a^2)^2}\operatorname*{plim}_{n\to\infty}\frac{1}{n}\|\boldsymbol{M_X P_Z X\beta}\|^2,$$

where $\sigma_a^2 \equiv \operatorname{plim}\hat\sigma_a^2$. Use this result to write down a Cox statistic that is asymptotically distributed as $\mathrm{N}(0,1)$.

The loglikelihood function for model H_1 is

$$-\frac{n}{2}\log 2\pi - \frac{n}{2}\log\sigma_1^2 - \frac{1}{2\sigma_1^2}(\boldsymbol{y}-\boldsymbol{X\beta})^\top(\boldsymbol{y}-\boldsymbol{X\beta}).$$

When this is evaluated at the ML estimates $\hat{\boldsymbol{\beta}}$ and $\hat\sigma_1^2$, it becomes, according to equation (10.12),

$$\begin{aligned}\hat\ell_1 &\equiv -\frac{n}{2}(1+\log 2\pi-\log n) - \frac{n}{2}\log \mathrm{SSR}(\hat{\boldsymbol{\beta}})\\ &= -\frac{n}{2}(1+\log 2\pi) - \frac{n}{2}\log\hat\sigma_1^2.\end{aligned} \tag{S15.38}$$

Notice that, for any deterministic σ_1^2,

$$\log\hat\sigma_1^2 = \log\sigma_1^2 + \log\biggl(1+\frac{\hat\sigma_1^2-\sigma_1^2}{\sigma_1^2}\biggr).$$

If σ_1^2 is the true error variance, then, since $\hat\sigma_1^2$ is root-n consistent, we may Taylor expand the logarithm to obtain

$$\log\biggl(1+\frac{\hat\sigma_1^2-\sigma_1^2}{\sigma_1^2}\biggr) = \frac{\hat\sigma_1^2-\sigma_1^2}{\sigma_1^2} + O_p(n^{-1}).$$

On taking expectations, we see that

$$\mathrm{E}(\log\hat\sigma_1^2) = \log\sigma_1^2 + \frac{1}{\sigma_1^2}\mathrm{E}(\hat\sigma_1^2 - \sigma_1^2) + O(n^{-1}). \tag{S15.39}$$

In fact, the second term on the right-hand side above is also just of order n^{-1}. To see why, recall that $\mathrm{E}(\hat\sigma_1^2) = n^{-1}\mathrm{E}(\boldsymbol{u}^\top\boldsymbol{M_X u}) = \sigma_1^2(1 - k_1/n)$. It follows that $\mathrm{E}(\hat\sigma_1^2 - \sigma_1^2) = -k_1\sigma_1^2/n = O(n^{-1})$.

The results (S15.38) and (S15.39) imply that, under any DGP in H_1 with error variance σ_1^2,

$$2n^{-1/2}\big(\hat\ell_1 - \mathrm{E}(\hat\ell_1)\big) = -n^{1/2}(\log\hat\sigma_1^2 - \log\sigma_1^2) + O_p(n^{-1/2}).$$

For the computation of the Cox statistic T_1, the above expression is to be evaluated at $\sigma_1^2 = \hat\sigma_1^2$. When we do that, the first term on the right-hand side above vanishes. Thus the whole expression is of order $n^{-1/2}$, and so it can be ignored asymptotically. What is left of the statistic T_1 is thus nothing other than the statistic T_1'.

For the model H_2, the maximized loglikelihood function is

$$\hat\ell_2 \equiv -\tfrac{n}{2}(\log 2\pi + 1) - \tfrac{n}{2}\log\hat\sigma_2^2.$$

If the true parameters are $\boldsymbol\beta$ and σ_1^2, we have, from a result which we obtained in the course of answering Exercise 15.18, that

$$\mathrm{E}(\hat\sigma_2^2) = \frac{n-k_2}{n}\,\sigma_1^2 + \tfrac{1}{n}\boldsymbol\beta^\top\boldsymbol X^\top\boldsymbol{M_Z X}\boldsymbol\beta.$$

Note that, by the definition of σ_a^2, $\operatorname{plim} n^{-1}\boldsymbol\beta^\top\boldsymbol X^\top\boldsymbol{M_Z X}\boldsymbol\beta = \sigma_a^2$. It is easy to verify that $\hat\sigma_2^2 - \mathrm{E}(\hat\sigma_2^2) = O_p(n^{-1/2})$, and from this it follows that

$$\hat\sigma_2^2 = \sigma_1^2 + \sigma_a^2 + O_p(n^{-1/2}).$$

An argument just like the one above, based on a Taylor expansion of the logarithm, then shows that

$$\mathrm{E}(\log\hat\sigma_2^2) = \log(\sigma_1^2 + \sigma_a^2) + O(n^{-1}).$$

Thus, under the DGP in H_1 with parameters $\boldsymbol\beta$ and σ_1^2, we have

$$2n^{-1/2}\big(\hat\ell_2 - \mathrm{E}(\hat\ell_2)\big) = -n^{1/2}\log\bigg(\frac{\hat\sigma_2^2}{\sigma_1^2 + \sigma_a^2}\bigg) + O_p(n^{-1/2}).$$

When the right-hand side of this equation is evaluated at $\boldsymbol\beta = \hat{\boldsymbol\beta}$ and $\sigma_1^2 = \hat\sigma_1^2$, we obtain

$$2n^{-1/2}\big(\hat\ell_2 - \mathrm{E}(\hat\ell_2)\big) = -n^{1/2}\log\bigg(\frac{\hat\sigma_2^2}{\hat\sigma_1^2 + \hat\sigma_a^2}\bigg) + O_p(n^{-1/2}), \tag{S15.40}$$

where $\hat\sigma_a^2$ is $n^{-1}\hat{\boldsymbol\beta}^\top\boldsymbol X^\top\boldsymbol M_{\boldsymbol Z}\boldsymbol X\hat{\boldsymbol\beta}$, which is equal to $n^{-1}\|\boldsymbol M_{\boldsymbol Z}\boldsymbol P_{\boldsymbol X}\boldsymbol y\|^2$, as in the statement of the exercise.

As we saw above, the Cox test can be based on the statistic T_1' as well as on T_1. Using (S15.40), we see that, to leading order asymptotically,

$$T_1' = n^{1/2}\log\left(\frac{\hat\sigma_1^2+\hat\sigma_a^2}{\hat\sigma_2^2}\right),$$

which is the statistic $T_1^{\rm OLS}$ defined in equation (15.90).

The next task is to show that $T_1^{\rm OLS}$ is asymptotically proportional to the J statistic. If we define $\tilde\sigma_2^2$ as $\hat\sigma_1^2+\hat\sigma_a^2$, as we did in Exercise 15.18, then, by yet another Taylor expansion of the logarithm, we see that

$$\log\left(\frac{\tilde\sigma_2^2}{\hat\sigma_2^2}\right) = \log\left(1+\frac{\tilde\sigma_2^2-\hat\sigma_2^2}{\hat\sigma_2^2}\right) \overset{a}{=} \frac{\tilde\sigma_2^2-\hat\sigma_2^2}{\hat\sigma_2^2}. \qquad \text{(S15.41)}$$

In Exercise 15.18, we showed that, under the null hypothesis,

$$n^{1/2}(\tilde\sigma_2^2-\hat\sigma_2^2) \overset{a}{=} 2n^{-1/2}\boldsymbol u^\top\boldsymbol M_{\boldsymbol X}\boldsymbol P_{\boldsymbol Z}\boldsymbol X\boldsymbol\beta. \qquad \text{(S15.42)}$$

Thus the numerator of the rightmost expression in (S15.41) is proportional to the expression $n^{-1/2}\boldsymbol u^\top\boldsymbol M_{\boldsymbol X}\boldsymbol P_{\boldsymbol Z}\boldsymbol X\boldsymbol\beta$, to which the numerator of the J statistic is proportional asymptotically; recall equation (15.41). The denominator of the rightmost expression in (S15.41) tends to a nonstochastic plim as $n\to\infty$. Therefore, the Cox test must be asymptotically equivalent to both the J test and the encompassing test of Exercise 15.18. Note that, if we had defined the Cox statistic in the less intuitive way that Cox did originally, the Cox and J statistics would still have been asymptotically equivalent, but they would have had opposite signs.

It was noted just after the definition (15.53) of the statistic T_1' that the test based on T_1' can be interpreted as an encompassing test for the maximized loglikelihood function. Expression (S15.38) expresses the latter as a deterministic function of the estimate $\hat\sigma_1^2$ of the error variance. It is thus not at all surprising to find that the Cox test is asymptotically equivalent to the variance encompassing test.

To find the asymptotic variance of $T_1^{\rm OLS}$, we use the results (S15.41) and (S15.42). Along with the the fact that $\hat\sigma_2^2\to\sigma_1^2+\sigma_a^2$, these imply that

$$n^{1/2}\log\left(\frac{\hat\sigma_1^2+\hat\sigma_a^2}{\hat\sigma_2^2}\right) \overset{a}{=} \frac{2n^{-1/2}\boldsymbol u^\top\boldsymbol M_{\boldsymbol X}\boldsymbol P_{\boldsymbol Z}\boldsymbol X\boldsymbol\beta}{\sigma_1^2+\sigma_a^2}.$$

Therefore, the asymptotic variance must be $4/(\sigma_1^2+\sigma_a^2)^2$ times the asymptotic variance of $n^{-1/2}\boldsymbol u^\top\boldsymbol M_{\boldsymbol X}\boldsymbol P_{\boldsymbol Z}\boldsymbol X\boldsymbol\beta$, which is

$$\sigma_1^2 \operatorname*{plim}_{n\to\infty}\frac{1}{n}\boldsymbol\beta^\top\boldsymbol X^\top\boldsymbol P_{\boldsymbol Z}\boldsymbol M_{\boldsymbol X}\boldsymbol P_{\boldsymbol Z}\boldsymbol X\boldsymbol\beta = \sigma_1^2\operatorname*{plim}_{n\to\infty}\frac{1}{n}\big\|\boldsymbol M_{\boldsymbol X}\boldsymbol P_{\boldsymbol Z}\boldsymbol X\boldsymbol\beta\big\|^2.$$

Thus we conclude that the asymptotic variance of $T_1^{\rm OLS}$ is

$$\frac{4\sigma_1^2}{(\sigma_1^2+\sigma_a^2)^2}\operatorname*{plim}_{n\to\infty}\frac{1}{n}\big\|\boldsymbol{M_X P_Z X\beta}\big\|^2, \tag{S15.43}$$

as stated in the exercise.

To obtain a test statistic that is asymptotically distributed as $\mathrm{N}(0,1)$, we need to estimate the quantities that appear in expression (S15.43). We can use $\hat\sigma_1^2$ and $\hat\sigma_a^2$ to estimate σ_1^2 and σ_a^2, and we can use

$$\frac{1}{n}\big\|\boldsymbol{M_X P_Z X\hat\beta}\big\|^2=\frac{1}{n}\big\|\boldsymbol{M_X P_Z P_X y}\big\|^2$$

to estimate the plim. Then $T_1^{\rm OLS}$ divided by the square root of the estimated variance is asymptotically distributed as $\mathrm{N}(0,1)$ under the null hypothesis.