

difference is that the regressand has not been divided by an estimate of σ . As we will see below, the test statistic is no more difficult to calculate by running (6.17) than by running (3.49).

Limiting our attention to zero restrictions makes it possible for us to gain a little more insight into the connection between the GNR and LM tests. Using the FWL Theorem, we see that regression (6.17) will yield exactly the same estimates of \mathbf{b}_2 , namely $\tilde{\mathbf{b}}_2$, and exactly the same sum of squared residuals as the regression

$$\mathbf{y} - \tilde{\mathbf{x}} = \tilde{\mathbf{M}}_1 \tilde{\mathbf{X}}_2 \mathbf{b}_2 + \text{residuals}, \quad (6.18)$$

where $\tilde{\mathbf{M}}_1$ is the matrix that projects onto $\mathcal{S}^\perp(\tilde{\mathbf{X}}_1)$. The regressand here is not multiplied by $\tilde{\mathbf{M}}_1$ because the first-order conditions imply that $\mathbf{y} - \tilde{\mathbf{x}}$ already lies in $\mathcal{S}^\perp(\tilde{\mathbf{X}}_1)$, which in turn implies that $\tilde{\mathbf{M}}_1(\mathbf{y} - \tilde{\mathbf{x}}) = \mathbf{y} - \tilde{\mathbf{x}}$. The sum of squared residuals from regression (6.18) is

$$(\mathbf{y} - \tilde{\mathbf{x}})^\top (\mathbf{y} - \tilde{\mathbf{x}}) - (\mathbf{y} - \tilde{\mathbf{x}})^\top \tilde{\mathbf{X}}_2 (\tilde{\mathbf{X}}_2^\top \tilde{\mathbf{M}}_1 \tilde{\mathbf{X}}_2)^{-1} \tilde{\mathbf{X}}_2^\top (\mathbf{y} - \tilde{\mathbf{x}}).$$

Since $\mathbf{y} - \tilde{\mathbf{x}}$ lies in $\mathcal{S}^\perp(\tilde{\mathbf{X}}_1)$, it is orthogonal to $\tilde{\mathbf{X}}_1$. Thus, if we had not included $\tilde{\mathbf{X}}_2$ in the regression, the SSR would have been $(\mathbf{y} - \tilde{\mathbf{x}})^\top (\mathbf{y} - \tilde{\mathbf{x}})$. Hence the reduction in the SSR of regression (6.17) brought about by the inclusion of $\tilde{\mathbf{X}}_2$ is

$$(\mathbf{y} - \tilde{\mathbf{x}})^\top \tilde{\mathbf{X}}_2 (\tilde{\mathbf{X}}_2^\top \tilde{\mathbf{M}}_1 \tilde{\mathbf{X}}_2)^{-1} \tilde{\mathbf{X}}_2^\top (\mathbf{y} - \tilde{\mathbf{x}}). \quad (6.19)$$

This quantity is also the explained sum of squares (around zero) from regression (6.17), again because $\tilde{\mathbf{X}}_1$ has no explanatory power. We can now show directly that this quantity, divided by any consistent estimate of σ^2 , is asymptotically distributed as $\chi^2(r)$ under the null hypothesis. We already showed this in Section 5.7, but the argument that the number of degrees of freedom is r was an indirect one.

First, observe that

$$n^{-1/2}(\mathbf{y} - \tilde{\mathbf{x}})^\top \tilde{\mathbf{X}}_2 \stackrel{a}{=} n^{-1/2} \mathbf{u}^\top \mathbf{M}_1 \mathbf{X}_2 \equiv \boldsymbol{\nu}^\top,$$

where $\mathbf{M}_1 \equiv \mathbf{M}_1(\boldsymbol{\beta}_0)$ and $\mathbf{X}_2 \equiv \mathbf{X}_2(\boldsymbol{\beta}_0)$. The asymptotic equality here follows from the fact that $\tilde{\mathbf{u}} \stackrel{a}{=} \mathbf{M}_1 \mathbf{u}$, which is the result (6.09) for the case in which the model is estimated subject to the restrictions that $\boldsymbol{\beta}_2 = \mathbf{0}$. The covariance matrix of the $r \times 1$ random vector $\boldsymbol{\nu}$ is

$$\begin{aligned} E(\boldsymbol{\nu} \boldsymbol{\nu}^\top) &= E(n^{-1} \mathbf{X}_2^\top \mathbf{M}_1 \mathbf{u} \mathbf{u}^\top \mathbf{M}_1 \mathbf{X}_2) = n^{-1} \mathbf{X}_2^\top \mathbf{M}_1 (\sigma_0^2 \mathbf{I}) \mathbf{M}_1 \mathbf{X}_2 \\ &= n^{-1} \sigma_0^2 (\mathbf{X}_2^\top \mathbf{M}_1 \mathbf{X}_2) \equiv \sigma_0^2 \mathbf{V}. \end{aligned}$$

The consistency of $\tilde{\boldsymbol{\beta}}$ and the regularity conditions for Theorem 5.1 imply that

$$n^{-1} \tilde{\mathbf{X}}_2^\top \tilde{\mathbf{M}}_1 \tilde{\mathbf{X}}_2 \stackrel{a}{=} n^{-1} \mathbf{X}_2^\top \mathbf{M}_1 \mathbf{X}_2 = \mathbf{V}.$$