and that, as we have seen, is both Hansen's statistic and the LM statistic in these circumstances.

Finally, we consider $C(\alpha)$ tests. Let $\acute{\boldsymbol{\theta}}$ be a parameter vector satisfying the restrictions $\boldsymbol{r}(\acute{\boldsymbol{\theta}}) = \mathbf{0}$. Then the test statistic can be formed as though it were the difference of two LM statistics, one for the restricted and one for the unrestricted model, both evaluated at $\acute{\boldsymbol{\theta}}$. Suppose, for simplicity, that the parameter vector $\boldsymbol{\theta}$ can be partitioned as $[\boldsymbol{\theta}_1 \vdots \boldsymbol{\theta}_2]$ and that the restrictions can be written as $\boldsymbol{\theta}_2 = 0$. The first term of the $C(\alpha)$ statistic has the form (17.72) but is evaluated at $\acute{\boldsymbol{\theta}}$ rather than the genuine constrained estimator $\tilde{\boldsymbol{\theta}}$. The second term should take the form of an LM statistic appropriate to the constrained model, for which only $\boldsymbol{\theta}_1$ may vary. This corresponds to replacing the matrix $\tilde{\boldsymbol{D}}$ in (17.72) by $\acute{\boldsymbol{D}}_1$, where the partition of $\boldsymbol{D}$ as $[\boldsymbol{D}_1 \;\; \boldsymbol{D}_2]$ corresponds to the partition of $\boldsymbol{\theta}$. The $C(\alpha)$ test statistic is therefore

$$\begin{aligned} C(\alpha) = {} & \frac{1}{n}\boldsymbol{\iota}^\top \acute{\boldsymbol{F}} \hat{\boldsymbol{\Phi}}^{-1} \acute{\boldsymbol{D}} \big(\acute{\boldsymbol{D}}^\top \hat{\boldsymbol{\Phi}}^{-1} \acute{\boldsymbol{D}}\big)^{-1} \acute{\boldsymbol{D}}^\top \hat{\boldsymbol{\Phi}}^{-1} \acute{\boldsymbol{F}}^\top \boldsymbol{\iota} \\ & - \frac{1}{n}\boldsymbol{\iota}^\top \acute{\boldsymbol{F}} \hat{\boldsymbol{\Phi}}^{-1} \acute{\boldsymbol{D}}_1 \big(\acute{\boldsymbol{D}}_1^\top \hat{\boldsymbol{\Phi}}^{-1} \acute{\boldsymbol{D}}_1\big)^{-1} \acute{\boldsymbol{D}}_1^\top \hat{\boldsymbol{\Phi}}^{-1} \acute{\boldsymbol{F}}^\top \boldsymbol{\iota}. \end{aligned} \tag{17.75}$$

Here, as before, $\hat{\boldsymbol{\Phi}}$ is a suitable estimate of $\boldsymbol{\Phi}$. To show that (17.75) is asymptotically equivalent to the true LM statistic, it is enough to modify the details of the proof of the corresponding asymptotic equivalence in Section 13.7.

In the general case in which the restrictions are expressed as $\boldsymbol{r}(\boldsymbol{\theta}) = \mathbf{0}$, another form of the $C(\alpha)$ test may be more convenient, since forming a matrix to correspond to $\boldsymbol{D}_1$ may not be simple. This other form is

$$\boldsymbol{\iota}^\top \acute{\boldsymbol{F}} \hat{\boldsymbol{\Phi}}^{-1} \acute{\boldsymbol{D}} \big(\acute{\boldsymbol{D}}^\top \hat{\boldsymbol{\Phi}}^{-1} \acute{\boldsymbol{D}}\big)^{-1} \acute{\boldsymbol{R}}^\top \Big(\acute{\boldsymbol{R}} \big(\acute{\boldsymbol{D}}^\top \hat{\boldsymbol{\Phi}}^{-1} \acute{\boldsymbol{D}}\big)^{-1} \acute{\boldsymbol{R}}^\top\Big)^{-1} \acute{\boldsymbol{R}} \big(\acute{\boldsymbol{D}}^\top \hat{\boldsymbol{\Phi}}^{-1} \acute{\boldsymbol{D}}\big)^{-1} \acute{\boldsymbol{D}}^\top \hat{\boldsymbol{\Phi}}^{-1} \acute{\boldsymbol{F}}^\top \boldsymbol{\iota}.$$

For this statistic to be useful, the difficulty of computing the actual constrained estimate $\tilde{\boldsymbol{\theta}}$ must outweigh the complication of the above formula. The formula itself can be established, at the cost of some tedious algebra, by adapting the methods of Section 8.9. We leave the details to the interested reader.

The treatment we have given of LM, LR, and Wald tests has largely followed that of Newey and West (1987b). This article may be consulted for more details of regularity conditions sufficient for the results merely asserted here to hold. Another paper on testing models estimated by GMM is Newey (1985b). Nonnested hypothesis tests for models estimated by GMM are discussed by Smith (1992). These papers do not deal with $C(\alpha)$ tests, however.

An interesting question is whether the conditional moment tests discussed in the last chapter in the context of models estimated by maximum likelihood have any counterpart for models estimated by GMM. For simplicity, suppose that there is a single conditional moment of which the expectation is zero if the model is correctly specified. If the corresponding empirical moment is used as an overidentifying restriction, then it can be tested in the same way