the residual $\hat{u}_t$. But this expansion is still unnecessarily complicated, because we have

$$\boldsymbol{X}_t^* = \boldsymbol{X}_{0t} + (\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0)^\top \boldsymbol{A}_t^* = \boldsymbol{X}_{0t} + O(n^{-1/2})$$

by Taylor's Theorem and the fact that $\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0 = O(n^{-1/2})$; recall that $\boldsymbol{A}_t$ is the Hessian of the regression function $x_t(\boldsymbol{\beta})$. Thus (5.56) can be written more simply as

$$\hat{u}_t = u_t - n^{-1/2}\boldsymbol{X}_{0t}\big(n^{-1}\boldsymbol{X}_0^\top\boldsymbol{X}_0\big)^{-1} n^{-1/2}\boldsymbol{X}_0^\top\boldsymbol{u} + o(n^{-1/2}).$$

Since this is true for all t, we have the vector equation

$$\hat{\boldsymbol{u}} = \boldsymbol{u} - \boldsymbol{X}_0\big(\boldsymbol{X}_0^\top\boldsymbol{X}_0\big)^{-1}\boldsymbol{X}_0^\top\boldsymbol{u} + o(n^{-1/2}),$$

where the small-order symbol is now to be interpreted as an n–vector, each component of which is $o(n^{-1/2})$. This equation can be rewritten in terms of the projection $\boldsymbol{P}_0 \equiv \boldsymbol{X}_0(\boldsymbol{X}_0^\top\boldsymbol{X}_0)^{-1}\boldsymbol{X}_0^\top$ and its complementary projection $\boldsymbol{M}_0 \equiv \mathbf{I} - \boldsymbol{P}_0$:

$$\hat{\boldsymbol{u}} = \boldsymbol{u} - \boldsymbol{P}_0\boldsymbol{u} + o(n^{-1/2}) = \boldsymbol{M}_0\boldsymbol{u} + o(n^{-1/2}). \tag{5.57}$$

This is the asymptotic equivalent of the exact result that, for linear models, the OLS residuals are the orthogonal projection of the disturbances off the regressors. Recall that if one runs the regression $\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{u}$, and the DGP is indeed a special case of this model, then we have exactly that

$$\hat{\boldsymbol{u}} = \boldsymbol{M}_X\boldsymbol{u}. \tag{5.58}$$

The result (5.57) reduces to this when the model is linear. The projection matrix $\boldsymbol{M}_0$ is now equal to $\boldsymbol{M}_X$, and the $o(n^{-1/2})$ term, which was due only to the nonlinearity of $\boldsymbol{x}(\boldsymbol{\beta})$, no longer appears.

Now let us substitute the right-most expression of (5.57) into (5.53). The latter becomes

$$n^{-1/2}\boldsymbol{a}^\top\hat{\boldsymbol{u}} = n^{-1/2}\boldsymbol{a}^\top\boldsymbol{M}_0\boldsymbol{u} + n^{-1/2}\sum_{t=1}^{n} o(n^{-1/2}). \tag{5.59}$$

The first term on the right-hand side here is clearly $O(1)$, while the second is $o(1)$. Thus, in contrast to what happened when we simply replaced $\hat{u}_t$ by u_t, we can ignore the second term on the right-hand side of (5.59). So the result (5.57) provides what we need if we are to undertake asymptotic analysis of expressions like (5.53).

We should pause for a moment here in order to make clear the relation between the asymptotic result (5.57), the exact linear result (5.58), and two other results. These other results are (1.03), which states that the OLS residuals are orthogonal to the regressors, and (2.05), which we may express