introduced in (9.52), as follows:

$$\boldsymbol{X}_i(\boldsymbol{\beta}) = \sum_{j=1}^{m} \boldsymbol{Z}_j(\boldsymbol{\beta})\psi_{ji}.$$

Then the stacked GNR is

$$\begin{bmatrix} (\boldsymbol{Y} - \boldsymbol{\xi}(\boldsymbol{\beta}))\boldsymbol{\psi}_1 \\ \vdots \\ (\boldsymbol{Y} - \boldsymbol{\xi}(\boldsymbol{\beta}))\boldsymbol{\psi}_m \end{bmatrix} = \begin{bmatrix} \boldsymbol{X}_1(\boldsymbol{\beta}) \\ \vdots \\ \boldsymbol{X}_m(\boldsymbol{\beta}) \end{bmatrix} \boldsymbol{b} + \text{residuals}. \tag{9.58}$$

The OLS estimates from the GNR (9.58) will be defined by the first-order conditions

$$\Bigg(\sum_{i=1}^{m} \boldsymbol{X}_i^{\top}(\boldsymbol{\beta})\boldsymbol{X}_i(\boldsymbol{\beta})\Bigg)\ddot{\boldsymbol{b}} = \sum_{i=1}^{m} \boldsymbol{X}_i^{\top}(\boldsymbol{\beta})\big(\boldsymbol{Y} - \boldsymbol{\xi}(\boldsymbol{\beta})\big)\boldsymbol{\psi}_i. \tag{9.59}$$

Some manipulation of (9.59) based on the definition of the $\boldsymbol{X}_i$'s and of $\boldsymbol{\psi}$ shows that this is equivalent to

$$\sum_{i=1}^{m}\sum_{j=1}^{m} \sigma^{ij}\boldsymbol{Z}_i^{\top}(\boldsymbol{\beta})\big(\boldsymbol{y}_j - \boldsymbol{x}_j(\boldsymbol{\beta}) - \boldsymbol{Z}_j(\boldsymbol{\beta})\boldsymbol{b}\big) = \boldsymbol{0}. \tag{9.60}$$

Thus we see that regression (9.58) has all the properties we have come to expect from the Gauss-Newton regression. If we evaluate it at $\boldsymbol{\beta} = \tilde{\boldsymbol{\beta}}$, the regression will have no explanatory power at all, because (9.60) is satisfied with $\boldsymbol{b} = \boldsymbol{0}$ by the first-order conditions (9.53). The estimated covariance matrix from regression (9.58) with $\boldsymbol{\beta} = \tilde{\boldsymbol{\beta}}$ will be

$$\tilde{s}^2\Bigg(\sum_{i=1}^{m}\sum_{j=1}^{m} \sigma^{ij}\tilde{\boldsymbol{Z}}_i^{\top}\tilde{\boldsymbol{Z}}_j\Bigg)^{-1}, \tag{9.61}$$

where $\tilde{s}^2$ is the estimate of the variance that the regression package will generate, which will evidently tend to 1 asymptotically if $\boldsymbol{\Sigma}$ is in fact the contemporaneous covariance matrix of $\boldsymbol{U}_t$. If (9.61) is rewritten as a sum of contributions from the successive observations, the result is

$$\tilde{s}^2\Bigg(\sum_{t=1}^{n} \tilde{\boldsymbol{\Xi}}_t\boldsymbol{\Sigma}^{-1}\tilde{\boldsymbol{\Xi}}_t^{\top}\Bigg)^{-1},$$

from which it is clear that (9.61) is indeed the proper GNLS covariance matrix estimator.