We have to estimate the set of parameters $\theta = (\beta', \gamma')'$, where β is $k_1 \times 1$ and γ is $k_2 \times 1$.The code exploits the fact that the model can be written as

$$\nu_{jt} = y_{jt}(\beta) - w_{jt}(\beta)\gamma,$$

where y_{jt} and w_{jt} (an scalar and a $1 \times k_2$ vector respectively) are "composite" variables which depend on the set of parameters β. The vector of parameters γ enters the equation linearly and "concentrating out" amounts to replacing them by its optimal expression given β.

To find the optimal expression notice that GMM minimizes

$$(\sum_j z_j' \nu_j(\theta))' A \sum_j z_j' \nu_j(\theta)$$

where z_j is the $L \times T_j$ matrix of instruments for firm j, ν_j is $T_j \times 1$, and A is the current weighting matrix. FOCs of the problem can be written as

$$\left[\sum_j \frac{\partial(z_j' \nu_j(\theta))}{\partial \beta}, \sum_j \frac{\partial(z_j' \nu_j(\theta))}{\partial \gamma} \right]' A \sum_j z_j' \nu_j(\theta) = 0$$

where the transponse matrix of derivatives is $(k_1+k_2)\times L$. This is hence a system of (k_1+k_2) equations. Using $\nu_j(\theta) = y_j(\beta) - w_j(\beta)\gamma$ and noticing that $\sum_j \frac{\partial(z_j'\nu_j(\theta))}{\partial\gamma} = -\sum_j z_j' w_j(\beta)$ the second block of k_2 equations gives

$$\gamma = \left[(\sum_j z_j' w_j(\beta))' A \sum_j z_j' w_j(\beta) \right]^{-1} (\sum_j z_j' w_j(\beta))' A \sum_j z_j' y_j(\beta).$$

The estimator of the asymptotic variance of the whole model can be written $Avar(\widehat{\theta}) = (\widehat{G}' A \widehat{G})^{-1} \widehat{G}' A (\sum_j z_j' \widehat{\nu}_j \widehat{\nu}_j' z_j) A \widehat{G} (\widehat{G}' A \widehat{G})^{-1}$, which becomes $Avar(\widehat{\theta}) = (\widehat{G}' (\sum_j z_j' \widehat{\nu}_j \widehat{\nu}_j' z_j)^{-1} \widehat{G})^{-1}$ when A is chosen optimally. Both variances depend of the estimator $\widehat{G} = \sum_j \frac{\partial(z_j'\nu_j(\widehat{\theta}))}{\partial\theta}$. Using the fact that $\frac{d(z_j'\nu_j)}{d\beta} = \frac{\partial(z_j'\nu_j)}{\partial\beta} - z_j' w_j \frac{\partial\gamma}{\partial\beta}$, we compute this estimator as

$$\widehat{G} = \left[\sum_j \frac{d(z_j'\nu_j)}{d\beta} + (\sum_j z_j' w_j) \frac{\partial\gamma}{\partial\beta}, \quad -\sum_j z_j' w_j \right].$$