C Extension to randomly missing transfers

In some applications, data will come from surveys where typically non response to questions about earnings are frequently encountered. Our proposed estimation strategy extends to the case where, for some random matches, transfers are missing. The log-likelihood expression presented in theorem 1 offers a very intuitive way of understanding how missing transfers for some random observations will impact the estimation. To formalize ideas, let p be the probability that for any arbitrary match the transfer is missing. The sample is still representative of the population of matches, but a random part of the sample consists of matches with observed transfers , i.e. $(X_i, Y_i, W_i)_{i=1}^{n^o}$, and the other part of matches with missing transfers, i.e. $(X_i, Y_i, .)_{i=n^o+1}^{n}$ where n^o is the number of matches with observed transfers and n is as before the size of our sample of matches (we have re-ordered the observations such that those matches with observed transfers are indexed first). The log-likelihood in this situation is therefore

$$\log \hat{L}(\theta) = \log \hat{L}_1(\theta) + \log \hat{L}_2(\theta) + n^o \log p + (n - n^o) \log(1-p)$$

where $\log \hat{L}_1(\theta)$ is given as in equation (3.11) and $\log \hat{L}_2(\theta)$ reads now as

$$\log \hat{L}_2(\theta) = -\sum_{i=1}^{n^o} \frac{(W_i - w_i(\theta))^2}{2s^2} - \frac{n^o}{2} \log s^2 \tag{C.1}$$

thus $p = n^0/n$. As n^o tends to 0, and hence p tends to 0, the log-likelihood function tends to $\log \hat{L}_1(\theta)$. In contrast, when n^o tends to n, and hence p tends to 1, the expression of $\log \hat{L}_2(\theta)$ in equation (C.1) tends to that of $\log \hat{L}_2(\theta)$ in equation (3.12) such that the log-likelihood function tends to equation (3.10).

D Concentrated Likelihood

In most applications, the parameters of primary interest are those governing workers' deterministic values of amenities and firms' deterministic values of productivity, i.e. A and Γ respectively. The remaining parameters $(\sigma_1, \sigma_2, t, s^2)$ are auxiliary, and the focus of attention is the *concentrated log-likelihood*, which is given by

$$\log l\,(A, \Gamma) := \max_{\sigma_1,\sigma_2,t,s^2} \log \hat{L}\,(\theta) = \log \hat{L}_1\,(\Phi) + \max_{\sigma_1,\sigma_2,t,s^2} \log \hat{L}_2\left(A, \Gamma, \sigma_1, \sigma_2, t, s^2\right).$$

where as usual, $\Phi = A + \Gamma$. Denoting $\sigma_1^*, \sigma_2^*, t^*$ and s^{*2} the optimal value of the corresponding parameters given A and Γ, one gets

$$(\sigma_1^*, \sigma_2^*, t^*) = \arg\min_{\sigma_1,\sigma_2,t} \sum_{i=1}^{n} \left(W_i - w_i\,(\theta)\right)^2, \tag{D.1}$$

which is the solution to a Nonlinear Least Squares problem which is readily implemented in standard statistical packages, and $s^{*2} = n^{-1} \sum_{i=1}^{n} \left(W_i - w_i\,(\theta^*)\right)^2$. The partial derivative of the concentrated log-likelihood with respect to A_k is given by

$$\frac{\partial \log l\,(A, \Gamma)}{\partial A_k} = \frac{\partial \log \hat{L}_1\,(\Phi)}{\partial \Phi_k} + \frac{\partial \log \hat{L}_2\,(A, \Gamma, \sigma_1^*, \sigma_2^*, t^*, s^{*2})}{\partial A_k}$$

and a similar expression holds for $\partial \log l / \partial \Gamma_k$. These formulas are derived in the following proof.

Proof. Recall $\theta = (A, \Gamma, \sigma_1, \sigma_2, t, s^2)$. The maximum likelihood problem can be written as

$$\max_{\theta} \log \hat{L}\,(\theta) = \max_{A,\Gamma} \log l\,(A, \Gamma)$$

where $\log l\,(A, \Gamma) = \max_{\sigma_1,\sigma_2,t,s^2} \log \hat{L}\,(\theta)$ is the concentrated log-likelihood which can be rewritten as

$$\log l\,(A, \Gamma) = \log \hat{L}_1\,(\theta) + \max_{\sigma_1,\sigma_2,t,s^2} \log \hat{L}_2\,(\theta)\,. \tag{D.2}$$

where

$$\max_{\sigma_1,\sigma_2,t,s^2} \log \hat{L}_2(\theta) = -\min_{s^2} \left(\frac{n}{2} \log s^2 + \frac{1}{2s^2} \min_{\sigma_1,\sigma_2,t} \sum_{i=1}^{n} (W_i - w_i(\theta))^2 \right) \tag{D.3}$$

The second minimization in equation (D.3) is an Ordinary Least Squares problem whose solution given A, Γ, denoted $(\sigma_1^*, \sigma_2^*, t^*)$, is the vector of coefficients of the OLS regression of W on $(\gamma - b, a - \alpha, 1)$. The value of s^2, denoted s^{*2}, is given by

$$s^{*2} = \frac{\sum_{i=1}^{n} (W_i - w_i(\theta^*))^2}{n}.$$

The envelope theorem yields an expression for the gradient of the concentrated log-likelihood with respect to the concentrated parameters A and Γ, that is

$$\nabla_{A,\Gamma} \log l(A, \Gamma) = \nabla_{A,\Gamma} \log \hat{L}_1(\theta^*) + \nabla_{A,\Gamma} \log \hat{L}_2(\theta^*).$$

The elements of the first part of the gradient are given in theorem 2 part (i) whereas parts (ii), (iv) and (v) of theorem 2 provide the building blocks for the elements of the second part of the gradient. ∎