Advanced Econometrics 1

University of Amsterdam 2017-2018
Computer Lab Exercise Week 7
GMM with heteroskedasticity

We consider 2SLS and GMM estimation in the model

$$\begin{aligned} y_i &= x_i\beta + u_i, \\ x_i &= \mathbf{z}_i'\boldsymbol{\pi} + v_i, \end{aligned}$$

with a single regressor x_i and a 3×1 instrument vector $\mathbf{z}_i$, similar to the assignment of Week 3. The difference is that we now allow heteroskedasticity of the form

$$\sigma_i^2 = \mathrm{E}[u_i^2|\mathbf{z}_i] = \exp(2 \cdot \mathbf{z}_i'\boldsymbol{\gamma}).$$

We fix the sample size N, the degree of endogeneity $\rho = \mathrm{Corr}[u_i, v_i] = 0.5$ and instrument strength $\boldsymbol{\pi} = (1, 0, 0)'$, but vary the heteroskedasticity parameter vector $\boldsymbol{\gamma} = (\gamma_1, 0, 0)'$ and study its effect on the bias and Root Mean Squared Error (RMSE) of 2SLS and two implementations of GMM.

1. Open the Matlab program `CompLabEx4.m`. The program initially only considers the 2SLS estimator, and the parameter γ_1 has been set to 0, so that $\sigma_i^2 = 1$ for all i (i.e., we first consider homoskedasticity).
   (a) Run the program a few times with $\gamma_1 = 0$. Why do you get (slightly) different answers all the time? How can you reduce this difference between runs?
   (b) for various values of γ_1, ranging from 0 to 1, and discuss the effect of the strenght of the heteroskedasticity on the bias and RMSE of $\widehat{\beta}_{\mathrm{2SLS}}$.

2. The two-step GMM estimator minimizes

$$Q_N(\beta) = \left(\frac{1}{N}\mathbf{Z}'(\mathbf{y} - \mathbf{x}\beta)\right)' \widehat{\mathbf{S}}^{-1} \left(\frac{1}{N}\mathbf{Z}'(\mathbf{y} - \mathbf{x}\beta)\right),$$

   where $\widehat{\mathbf{S}} = N^{-1}\sum_{i=1}^N \widehat{u}_i^2\mathbf{z}_i\mathbf{z}_i'$, and $\widehat{u}_i = y_i - x_i\widehat{\beta}_{\mathrm{2SLS}}$.
   - Give an explicit expression for the solution $\widehat{\beta}_{\mathrm{GMM}}$ of this minimization problem
   - Extend the Matlab program to calculate this estimator for each replication and to compare its bias and RMSE with that of the 2SLS estimator. (you can use the row-by-row procedure rbr(u,Z) to construct a vector eZ with elements $\widehat{u}_i\mathbf{z}_i$ to calculate $\widehat{\mathbf{S}}$.
   - Theoretically, the GMM estimator should be asymptotically more efficient than the 2SLS estimator if $\gamma_1 \neq 0$. Do you find this for all values of γ_1? Does the degree of heteroskedasticity affect the (relative) bias of the two estimators?

PTO

3. The conditional moment condition $\mathrm{E}[u_i|\mathbf{z}_i] = 0$, together with the fact that $\sigma_i^2 = \exp(\mathbf{z}_i'\boldsymbol{\gamma})$ is a function of the instruments, implies the unconditional moment condition

$$\mathbf{h}_i(\beta) = \mathrm{E}\left[\frac{\mathbf{z}_i(y_i - x_i\beta)}{\sigma_i^2}\right] = \mathbf{0},$$

because $\mathrm{E}[\mathbf{z}_i u_i/\sigma_i^2] = \mathrm{E}[\mathrm{E}(\mathbf{z}_i u_i/\sigma_i^2|\mathbf{z}_i)] = \mathrm{E}[\mathbf{z}_i/\sigma_i^2 \mathrm{E}(u_i|\mathbf{z}_i)] = \mathbf{0}$. We may use this moment condition instead of $\mathrm{E}[\mathbf{z}_i(y_i - x_i\beta)] = \mathbf{0}$ to estimate β by GMM, with optimal weighting matrix

$$\mathbf{S}_0 = \mathrm{E}[\mathbf{h}_i(\beta_0)\mathbf{h}_i(\beta_0)'] = \mathrm{E}\left[\frac{\mathbf{z}_i\mathbf{z}_i'}{\sigma_i^2}\right].$$

The discussion around equation (6.45), Section 6.4.2 of Cameron & Trivedi suggests that this leads to an even more efficient estimator. A feasible, two-step version is obtained as follows:

(a) estimate $\boldsymbol{\gamma}$ by the least-squares estimator in the linear regression $\ln \widehat{u}_i^2 = \mathbf{z}_i'\boldsymbol{\gamma} + e_i$, and use this to construct $\widehat{\sigma}_i^2 = \exp(\mathbf{z}_i'\widehat{\boldsymbol{\gamma}})$;

(b) estimate β by minimizing

$$Q_N(\beta) = \left(\frac{1}{N}\sum_{i=1}^N \frac{\mathbf{z}_i(y_i - x_i\beta)}{\widehat{\sigma}_i^2}\right)' \left(\frac{1}{N}\sum_{i=1}^N \frac{\mathbf{z}_i\mathbf{z}_i'}{\widehat{\sigma}_i^2}\right)^{-1} \left(\frac{1}{N}\sum_{i=1}^N \frac{\mathbf{z}_i(y_i - x_i\beta)}{\widehat{\sigma}_i^2}\right),$$

The solution can again be expressed in closed form. Add this estimator to you Matlab program, and again compare its bias and RMSE to the other two estimators for various values of γ_1.