Econometrics
Problem Set 4.
Instructor: Alexei Onatski

1. Suppose that y_i, $i = 1, ..., n$ are i.i.d. random variables distributed as $N\left(\mu, \mu^4\right)$. That is, the variance of y_i equals the fourth power of its mean.

   (a) Write down two different moment conditions that can be used to estimate μ.

   (b) Describe a (not necessarily optimal) GMM estimator of μ that uses the two moment conditions given in a).

   (c) What is the optimal GMM weighting matrix? How would you estimate it?

   (d) Give an explicit expression for the asymptotic variance of $\sqrt{n}\left(\hat{\mu}_{GMM} - \mu\right)$ in terms μ. Here $\hat{\mu}_{GMM}$ is an optimal GMM that is based on the weighting matrix from (c). Compare this variance to the asymptotic variances of $\sqrt{n}\left(\hat{\mu}_{OLS} - \mu\right)$ and $\sqrt{n}\left(\sqrt{\hat{\sigma}_{OLS}} - \mu\right)$, where $\hat{\mu}_{OLS} = \frac{1}{n}\sum_{i=1}^{n} y_i$ and $\hat{\sigma}^2_{OLS} = \frac{1}{n-1}\sum_{i=1}^{n}\left(y_i - \hat{\mu}_{OLS}\right)^2$.

   (e) Find the asymptotic variance of $\sqrt{n}\left(\hat{\mu}_{ML} - \mu\right)$, and compare it to that of $\sqrt{n}\left(\hat{\mu}_{GMM} - \mu\right)$.

2. The file ccapm.dta is the STATA file that contains 238 observations on the real USA consumption ratio c_{t+1}/c_t (cratio), the real gross return on Treasury bills R_{t+1} (rrate), and the real value weighted returns e_{t+1} (e). This is the adjusted Hansen and Singleton (1982) data set, used in their paper "Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models" (Econometrica 50, 1269-1286). Consider the first order conditions of the Consumption Capital Asset Pricing Model

$$E_t\left[\beta\left(c_{t+1}/c_t\right)^{-\gamma} R_{t+1} - 1\right] = 0.$$

   The parameters are the discount factor, β, and the relative risk aversion coefficient γ.

   (a) Estimate parameters β and γ by GMM using moment conditions

$$\begin{aligned}
E\left[\beta\left(c_{t+1}/c_t\right)^{-\gamma} R_{t+1} - 1\right] &= 0, \\
E\left[\left(\beta\left(c_{t+1}/c_t\right)^{-\gamma} R_{t+1} - 1\right) c_t/c_{t-1}\right] &= 0, \\
E\left[\left(\beta\left(c_{t+1}/c_t\right)^{-\gamma} R_{t+1} - 1\right) R_t\right] &= 0, \\
E\left[\left(\beta\left(c_{t+1}/c_t\right)^{-\gamma} R_{t+1} - 1\right) e_t\right] &= 0.
\end{aligned}$$

   Use command gmm in STATA to do the estimation. Specify starting values $\beta = 1$ and $\gamma = 1$.

(b) The default STATA setting is the optimal two-step GMM with heteroskedasticity-robust estimation of the optimal weighting matrix. Use option wmatrix(hac ba 5) to change the default estimation of the optimal weight to the Newey-West-type HAC estimation with truncation lag equal to 5 (first, you will need to tell stata that time=[_n]). Compare your results with those obtained in (a).

(c) Using the setting in (b), test the hypothesis that $\beta = 0.98$ at 95% significance level.

3. Consider the so-called Correlated Random Effects model

$$y_{it} = \alpha + x_{it}\beta + \bar{x}_i\gamma + \eta_i + \varepsilon_{it},$$

where for simplicity x_{it} is a scalar explanatory variable, and $\bar{x}_i = \frac{1}{T}\sum_{t=1}^{T} x_{it}$. The Correlated Random Effects GLS estimator $\hat{\beta}_{CRE}$ is the OLS estimator of β in the quasi-demeaned regression

$$\tilde{y}_{it} = \delta + \tilde{x}_{it}\beta + \bar{x}_i\rho + u_{it}$$

where

$$\tilde{y}_{it} = y_{it} - \theta\bar{y}_i,\ \ \tilde{x}_{it} = x_{it} - \theta\bar{x}_i,$$

$$\delta = \alpha\left(1-\theta\right),\ \ \rho = \gamma\left(1-\theta\right),$$

and

$$\theta = 1 - \left(\sigma_\varepsilon^2 / \left(\sigma_\varepsilon^2 + T\sigma_\eta^2\right)\right)^{1/2}.$$

(a) Argue that $\hat{\beta}_{CRE}$ is the same as the OLS estimator of β in the model

$$\tilde{y}_{it} = \delta + \left(x_{it} - \bar{x}_i\right)\beta + \bar{x}_i\lambda + u_{it}$$

where

$$\lambda = \rho + \beta\left(1-\theta\right)$$

(b) Show that the residuals from the regression of $x_{it} - \bar{x}_i$ on constant and $\bar{x}_i$ is just $x_{it} - \bar{x}_i$ itself.

(c) Show that

$$\sum_{i,t} \tilde{y}_{it}\left(x_{it} - \bar{x}_i\right) = \sum_{i,t}\left(y_{it} - \bar{y}_i\right)\left(x_{it} - \bar{x}_i\right)$$

(d) Using the Frisch-Waugh theorem together with the results from b) and c) demonstrate that $\hat{\beta}_{CRE}$ is numerically the same as $\hat{\beta}_{FE}$.

4. The file MURDER.DTA contains US state-level data on murder rates ($mrdrte_{it}$) and the total number of executions ($exec_{it}$) for the past three years. The data are available for 1987, 1990, and 1993.

   (a) Run regression of $mrdrte_{it}$ on the constant, $exec_{it}$, $unem_{it}$ (the annual unemployment rate), and two dummies: the dummy for the 1990 (equal to 1 for this year), and the dummy for 1993. Report your results.

   (b) Using xtset command, declare that the data are panel with "panel variable" id (this variable identifies states), and "time variable" $year$. Run FE regression, using xtreg command (with option fe). Report the results, compare with (a). For further use, save the coefficient vector and the variance-covariance matrix of this vector in matrices betafe and Vfe (use commands "mat betafe=get(_b)" and "mat Vfe=get(VCE)").

   (c) Now run regression of $mrdrte_{it}$ on the dummies for different state (use command xi: reg mrdrte i.id), get the residuals, say $rmrdrte$. Do the same regression using as dependent variables $exec$, $unem$, $d90$, and $d93$, each time getting the residuals. Regress $rmrdrte$ on $rexec$, $runem$, $rd90$, and $rd93$. Do not include constant. Compare your results with those you obtain in (b). Explain similarities and dis-similarities. How are the standard errors in (b) and (c) related?

   (d) Run RE regression using xtreg command (without any options). Using the obtained coefficients and their variance-covariance matrix and the values of the fixed effects coefficients and their variance-covariance matrix stored in (b), compute the value of the Hausman statistic for testing the RE specification (use only the difference between the four coefficients on non-constant regressors in your computation). What do you conclude?

5. Consider the following probit regression model

$$\Pr\left(Deny_i = 1 | P/Iratio_i, Black_i\right) = \Phi(\beta_1 + \beta_2 * P/Iratio_i + \beta_3 * Black_i),$$

   where $Deny_i = 1$ if person i who applied for a mortgage has been denied the mortgage, $P/Iratio_i$ is the applicant's anticipated total monthly loan payments to his or her monthly income, and $Black_i = 1$ if person i is black. You are given the following results

$$\hat{\beta}_{ML} = \begin{pmatrix} -2.2587 \\ 2.7416 \\ 0.7082 \end{pmatrix}, \widehat{Var}\left(\hat{\beta}_{ML}\right) = \begin{pmatrix} 0.0169 & -0.0445 & -0.0010 \\ -0.0445 & 0.1293 & -0.0018 \\ -0.0010 & -0.0018 & 0.0070 \end{pmatrix}$$

(a) Estimate the marginal effect of a change in $P/Iratio$ on the probability of mortgage denial for a black person with $P/Iratio = 0.3$.

(b) Compute the standard error of your estimate in (a) using the Delta method.

(c) Estimate the effect of the change from $Black = 1$ to $Black = 0$ on the probability of mortgage denial given that $P/Iratio = 0.3$. (Note that the change of the value of Black is not marginal)

(d) Compute the standard error of your estimate in (c) using the Delta method.