Econ 203c, Problem Set 6, Spring 2003

Question 1:

1)-4) Matlab code is attached.

The one part that is not straight forward is the computation of the restricted estimates.

The restricted estimate is defined as
$\widetilde{\beta}_{gmm} = \arg\min m_n(\beta)' V_n^{-1} m_n(\beta)$
such that $\beta_2\beta_3 + \beta_4 = 0$,
where $m_n(\beta) = 1/n \sum z_i(y_i - x_i'\beta)$ and V_n^{-1} is the optimal weight matrix defined previously.

The Lagrangian for this problem is
$m_n(\beta)' V_n^{-1} m_n(\beta) + \lambda(\beta_2\beta_3 + \beta_4)$.

The first order conditions are
(1) $\frac{\partial m_n(\beta)}{\partial \beta_1} V_n^{-1} m_n(\beta_1) = 0$
(2) $\frac{\partial m_n(\beta)}{\partial \beta_2} V_n^{-1} m_n(\beta_2) + \lambda\beta_3 = 0$
(3) $\frac{\partial m_n(\beta)}{\partial \beta_3} V_n^{-1} m_n(\beta_3) + \lambda\beta_2 = 0$

(4) $\frac{\partial m_n(\beta)}{\partial \beta_4} V_n^{-1} m_n(\beta_4) + \lambda = 0$
(5) $\beta_2\beta_3 + \beta_4 = 0$,

where $\frac{\partial m_n(\beta)}{\partial \beta_k} = (x_k' z)$, $m_n(\beta_k) = (z'y) - (z'x_k)\beta_k$, and $x_k = [x_{k1}, \ldots, x_{kn}]'$ (nx1).

Dividing (2) by (3)

$$\Rightarrow \frac{\frac{\partial m_n(\beta)}{\partial \beta_2} m_n(\beta_2) V_n^{-1}}{\frac{\partial m_n(\beta)}{\partial \beta_3} m_n(\beta_3) V_n^{-1}} = 0$$

$$\Rightarrow \frac{(x_2' z) V_n^{-1}[(z'y) - (z'x_2)\beta_2]}{(x_3' z) V_n^{-1}[(z'y) - (z'x_3)\beta_3]} - \frac{\beta_3}{\beta_2} = 0 \qquad (6)$$

Dividing (3) by (4)

$$\Rightarrow \frac{\frac{\partial m_n(\beta)}{\partial \beta_3} m_n(\beta_3) V_n^{-1}}{\frac{\partial m_n(\beta)}{\partial \beta_4} m_n(\beta_4) V_n^{-1}} + \beta_2 =$$

$$\Rightarrow \frac{(x_3' z) V_n^{-1}[(z'y) - (z'x_3)\beta_3]}{(x_4' z) V_n^{-1}[(z'y) - (z'x_4)\beta_4]} + \beta_2 = 0 \qquad (7)$$

(6), (7), (1), (5) is a non-linear system of equations with unkown parameters.

There is a unique solution to this, but I cannot solve for it analytically.
In Matlab, I will use fmins to solve it.

My Matlab code is attached. I won't print out my histograms. But since the null hypothesis is true, most of the test statistics you compute should be less than 3.84 which is the critical value at 95 percent confidence level.

Type 1 Error is the probability of rejecting the null when it is true. At a confidence level of 95 percent, this would be the percentage of test statistics that exceed 3.84. Because this is simulation involving random numbers, you may have different answers than mine, but the type 1 error should be very low.

My results (means over 500 draws):
1) Mean of unrestricted betas
0.9936
0.5008
-0.5014
0.2519
2) Mean of unrestricted covariance matrices
17.5389 -1.8130 -1.5628 -2.1980
-1.8130 0.8987 0.0425 -0.0360

-1.5628 0.0425 0.7442 -0.0026
-2.1980 -0.0360 -0.0026 0.7445

3) mean LM: 0.8863
mean wald: 1.4161
mean LR: 0.8050

Question 2: MLE is GMM

Definition of GMM:
Given a set of moment condtions, $E[\Psi(x,\theta_0)] = 0$, we can define the GMM estimator of the true population parameter, θ_0, as
$\widehat{\theta}_{gmm} = \arg\min m_n(\theta)' V_n^{-1} m_n(\theta)$,
where $m_n(\theta) = 1/n \sum \Psi(x_i,\theta)$ and V_n^{-1} is some weight matrix.

Definition of MLE:
Given a distribution of $f(x,\theta_0)$, we can define the ML estimator of the true population parameter, θ_0, as
$\widehat{\theta}_{ml} = \arg\max \sum f(x_i,\theta)$.

Using the logarithm transformation, the FOC for the MLE problem is
$$\sum \frac{\partial \ln f(x_i,\theta)}{\partial\theta} = 0.$$

The GMM FOC is
$$\frac{\partial m_n(\theta)}{\partial\theta} V_n^{-1} m_n(\theta) = 0.$$
In the previous problem set, I showed that if GMM is just identified, this FOC is equivalent to
$$\sum \Psi(x_i,\theta) = 0.$$

Since MLE is always a just identified system of equations, if we let
$$\Psi(x_i,\theta) = \frac{\partial \ln f(x_i,\theta)}{\partial\theta},$$
we can solve for the MLE using GMM.
The equivalent population moment condition for GMM is $E[\frac{\partial \ln f(x,\theta_0)}{\partial\theta}] = 0$.

Question 3:

Model: $y_i = g(x_i;\theta_0) + u_i$ and $E[u_i|x_i] = 0$

1) The population parameter is
$\theta_0 = \arg\min_{\theta\in\Theta} E[(y - g(x;\theta))^2]$

FOC
$$\frac{\partial E[(y - g(x;\theta_0))^2]}{\partial\theta} = 0$$
$$\Rightarrow 2E[(y - g(x;\theta_0))\frac{\partial g(x;\theta_0)}{\partial\theta}] = 0$$
By LIE,
$$\Rightarrow E_x\{E(y|x) - g(x;\theta_0)]\frac{\partial g(x;\theta_0)}{\partial\theta}\} = 0.$$
The assumption $E[u|x] = 0$ implies
$E[y|x] = g(x;\theta_0)$.

Substituting,
$$E_x\{g(x;\theta_0) - g(x;\theta_0)]\frac{\partial g(x;\theta_0)}{\partial\theta}\} = 0.$$

This FOC is zero only if evaluated at the population parameter, θ_0.

2) θ_0 is unique if the objective function is continuous and Θ is compact (Weierstrass theorem).

Compactness of Θ is assumed.

Continuity of $E[(y - g(x;\theta))^2]$ could be checked with some assumptions about the distribution of y and x and knowledge of $g(x;\theta)$.

3) Sample analog of population moments
$$\varphi_1(y,x,\theta) = 1/n\sum(y_i - g(x_i;\theta)) = 0$$

4) Additional K moment conditions
$$\varphi_2(y,x,\theta) = 1/n\sum z_i(y_i - g(x_i;\theta)) = 0,$$

where $z_i = x_i^2$.

(Of course you may need to alter this if the stacked moments are not full rank.)
$E[\varphi_2(y,x,\theta)] = 1/n \sum E[z_i(y_i - g(x_i;\theta))]$

$E[(y_i - g(x_i;\theta)^2)] = E_x\{z_i[E(y_i|x_i) - g(x_i;\theta)]\}$
$= 0$ since $E(y_i|x_i) = g(x_i;\theta)$
Thus, $E[\varphi_2(y,x,\theta)] = 0.$

5) Optimal GMM estimator based on $\varphi_1(y,x,\theta)$ and $\varphi_2(y,x,\theta)$

First, compute
$\widehat{\theta}_{init} = \arg\min m_n(y,x,\theta)' m_n(y,x,\theta),$
where $m_n(\theta) = 1/n \sum \begin{bmatrix} (y_i - g(x_i;\theta)) \\ z_i(y_i - g(x_i;\theta)) \end{bmatrix} = 1/n \sum \varphi_3(y_i,x_i,\theta).$

Next, compute
$V_n = 1/n \sum [\varphi_3(y_i,x_i,\widehat{\theta}_{init})\varphi_3(y_i,x_i,\widehat{\theta}_{init})']$

Finally, compute
$\widehat{\theta}_{gmm} = \arg\min m_n(y,x,\theta)' V_n^{-1} m_n(y,x,\theta)$

6) Asymptotic distribution for $\widehat{\theta}_{gmm}$

$\sqrt{n}\,(\widehat{\theta}_{gmm} - \theta_0) \xrightarrow{d} N(0,\Lambda(\theta_0)),$
where $\Lambda(\theta_0) = [A(\theta_0)V^{-1}A(\theta_0)']^{-1}$, $\dfrac{\partial m_n(\widehat{\theta}_{gmm})}{\partial\theta} \xrightarrow{p} A(\theta_0)$, and $V_n^{-1} \xrightarrow{p} V^{-1}$.
7) Test statistics

$H_0 : \theta_1^2 + \cdots + \theta_K^2 = 1$
$H_1 : \theta_1^2 + \cdots + \theta_K^2 \neq 1$

Wald statistic

$$W_n = \frac{[\widehat{\theta}^2_{gmm,1} + \cdots + \widehat{\theta}^2_{gmm,K} - 1]^2}{cov(\widehat{\theta}^2_{gmm,1} + \cdots + \widehat{\theta}^2_{gmm,K})},$$
where $cov(\widehat{\theta}^2_{gmm,1} + \cdots + \widehat{\theta}^2_{gmm,K}) = \Delta\widehat{\Omega}\Delta'$, $\Delta = [2\widehat{\theta}_{gmm,1},\ldots,2\widehat{\theta}_{gmm,K}]'$, and

$$\widehat{\Omega} = \frac{\widehat{\Lambda}(\theta_0)}{n} = 1/n[\frac{\partial m_n(\widehat{\theta}_{gmm})}{\partial \theta} V_n^{-1} \frac{\partial m_n(\widehat{\theta}_{gmm})}{\partial \theta'}]^{-1}.$$

You could re-write this as

$W_n = n\frac{[\widehat{\theta}^2_{gmm,1} + \bullet\bullet\bullet + \widehat{\theta}^2_{gmm,K} - 1]^2}{cov(\widehat{\theta}^2_{gmm,1} + \bullet\bullet\bullet + \widehat{\theta}^2_{gmm,K})}$, where $\widehat{\Omega} = \widehat{\Lambda}(\theta_0)$ (not $\frac{\widehat{\Lambda}(\theta_0)}{n}$).

Lagrange Multiplier statistic

First, need to calculate optimal weighting matrix under null hypothesis: W_r.

$\widehat{\theta}_{init,r} = \arg\min m_n(y,x,\theta)' m_n(y,x,\theta)$ s.t. $\theta_1^2 + \bullet\bullet\bullet + \theta_K^2 = 1.$

$W_r = \sum[\varphi_3(y_i,x_i,\widehat{\theta}_{init,r})\varphi_3(y_i,x_i,\widehat{\theta}_{init,r})']$

Then, calculate optimal GMM under the null: $\widetilde{\theta}$.

$\widetilde{\theta} = \arg\min m_n(y,x,\theta)' W_r^{-1} m_n(y,x,\theta)$ s.t. $\theta_1^2 + \bullet\bullet\bullet + \theta_K^2 = 1.$

Estimator for asymptotic covariance is $V_r = \sum \varphi_3(y_i,x_i,\widetilde{\theta})\varphi_3(y_i,x_i,\widetilde{\theta})'$

$$LM_n = n m_n(\widetilde{\theta})' V_r^{-1} \frac{\partial m_n(\widetilde{\theta})}{\partial \theta} [\frac{\partial m_n(\widetilde{\theta})}{\partial \theta} W_r^{-1} \frac{\partial m_n(\widetilde{\theta})}{\partial \theta'}]^{-1} \frac{\partial m_n(\widetilde{\theta})}{\partial \theta'} V_r^{-1} m_n(\widetilde{\theta}).$$

Likelihood Ratio statistic

$LR_n = n(m_n(y,x,\widetilde{\theta}' W_r^{-1} m_n(y,x,\widetilde{\theta}) - m_n(y,x,\widehat{\theta}_{gmm})' V_n^{-1} m_n(y,x,\widehat{\theta}_{gmm})).$