Economics 481
Mock Exam

Oliver Linton

May 2, 2000

Solutions

1. Bookwork. ■

2. Let $m_1(z, \theta) = (y - \theta)$ and $m_2(z, \theta) = \{(y - \theta)^2 - \theta^2\}$. The sample equivalent is

$$G_n(\theta) = \frac{1}{n} \sum_{i=1}^{n} \begin{pmatrix} (y_i - \theta) \\ \{(y_i - \theta)^2 - \theta^2\} \end{pmatrix}$$

The GMM estimator then minimizes a quadratic form in $G_n(\theta)$, i.e., minimizes

$$Q_n(\theta) = G_n(\theta)' W G_n(\theta)$$

with respect to θ for some 2×2 matrix W. The Maximum Likelihood Estimator maximizes the log likelihood function

$$\ell_n(\theta) = -\frac{1}{2} \sum_{i=1}^{n} \log \theta^2 - \frac{1}{2} \sum_{i=1}^{n} \frac{(y_i - \theta)^2}{\theta^2}.$$

This estimator has first order condition

$$\frac{\partial \ell_n(\theta)}{\partial \theta} = \frac{1}{2} \sum_{i=1}^{n} \left[\frac{(y_i - \theta)^2}{\theta^2} - 1 \right] \frac{2}{\theta} + \sum_{i=1}^{n} \frac{(y_i - \theta)}{\theta^2}.$$

Both of these estimation methods are nonlinear in the parameters. However, the OLS estimator is consistent in this case, and it is very simple to compute. This is what motivates doing a two-step approach here. First, let

$$\widehat{\theta}_{OLS} = \frac{1}{n} \sum_{i=1}^{n} y_i = \overline{y}.$$

To motivate the two-step method we linearize the score function about $\widehat{\theta}_{OLS}$, i.e.,

$$\frac{\partial \ell_n(\theta)}{\partial \theta} \simeq \frac{\partial \ell_n(\widehat{\theta}_{OLS})}{\partial \theta} + \frac{\partial^2 \ell_n(\widehat{\theta}_{OLS})}{\partial \theta^2} (\theta - \widehat{\theta}_{OLS}),$$

and then let

$$G_n(\theta; \widehat{\theta}_{OLS}) = \frac{\partial \ell_n(\widehat{\theta}_{OLS})}{\partial \theta} + \frac{\partial^2 \ell_n(\widehat{\theta}_{OLS})}{\partial \theta^2}(\theta - \widehat{\theta}_{OLS}).$$

This has zero

$$\widehat{\theta}_{2-step} = \widehat{\theta}_{OLS} - \left[\frac{\partial^2 \ell_n(\widehat{\theta}_{OLS})}{\partial \theta^2}\right]^{-1} \frac{\partial \ell_n(\widehat{\theta}_{OLS})}{\partial \theta}.$$

This estimator will be asymptotically efficient because

$$\begin{aligned} \frac{\partial \ell_n(\widehat{\theta}_{OLS})}{\partial \theta} &\simeq \frac{\partial \ell_n(\theta_0)}{\partial \theta} + \frac{\partial^2 \ell_n(\theta_0)}{\partial \theta^2}(\widehat{\theta}_{OLS} - \theta_0) \\ \frac{\partial^2 \ell_n(\widehat{\theta}_{OLS})}{\partial \theta^2} &\simeq \frac{\partial^2 \ell_n(\theta_0)}{\partial \theta^2} \end{aligned}$$

so that

$$\begin{aligned} \sqrt{n}(\widehat{\theta}_{2-step} - \theta_0) &\simeq \sqrt{n}(\widehat{\theta}_{OLS} - \theta_0) - \left[\frac{\partial^2 \ell_n(\theta_0)}{\partial \theta^2}\right]^{-1} \sqrt{n}\left[\frac{\partial \ell_n(\theta_0)}{\partial \theta} + \frac{\partial^2 \ell_n(\theta_0)}{\partial \theta^2}(\widehat{\theta}_{OLS} - \theta_0)\right] \\ &\simeq -\left[\frac{\partial^2 \ell_n(\theta_0)}{\partial \theta^2}\right]^{-1} \sqrt{n}\frac{\partial \ell_n(\theta_0)}{\partial \theta}, \end{aligned}$$

which is asympotically normal with mean zero and variance given by the inverse information.

■

3. First, notice that h need not be a density function. However, we can write

$$m(z, \alpha) = \int_A \frac{h(x; z, \alpha)}{p(x)} p(x) dx$$

for any density function p with support A. We then compute

$$\widehat{m}(z, \alpha) = \frac{1}{R}\sum_{r=1}^{R} \frac{h(X_r; z, \alpha)}{p(X_r)},$$

where $X_1, \ldots, X_R$ are i.i.d. draws from p. By construction

$$E\widehat{m}(z, \alpha) = m(z, \alpha)$$

for all z, α. Actually, could choose p to be uniform on A in which case we dont need to divide out.

The conditional moment condition implies that

$$E\left[(y_i - m(z_i, \alpha_0) - \beta_0' w_i)\, g(z_i, w_i)\right] = 0$$

for any function g. Therefore let

$$\widehat{G}_n(\theta) = \frac{1}{n}\sum_{i=1}^{n} (y_i - \widehat{m}(z_i, \alpha) - \beta' w_i)\, g(z_i, w_i),$$

where $\widehat{m}(z,\alpha)$ was defined above. Since $\widehat{m}(z,\alpha)$ is unbiased, we have $E\widehat{G}_n(\theta) = G_n(\theta)$ for all θ, so that consistent estimation can be based on this first order condition if g has at least as many components as θ. The nonlinear least squares estimator works from the population condition where

$$g(z,w) = \begin{pmatrix} \frac{\partial m(z,\alpha_0)}{\partial \alpha} \\ w \end{pmatrix}$$

and so in this case we have

$$\widehat{G}_n(\theta) = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - \widehat{m}(z_i,\alpha) - \beta' w_i\right)\widehat{g}(z_i,w_i;\alpha),$$

where

$$\widehat{g}(z,w;\alpha) = \begin{pmatrix} \frac{1}{R}\sum_{r=1}^{R}\frac{\partial h(X_r,z,\alpha)}{\partial \alpha}\frac{1}{p(X_r)} \\ w \end{pmatrix}.$$

Now, $E\widehat{G}_n(\theta) \neq G_n(\theta)$ unless we make $\widehat{m}$ and $\widehat{g}$ independent - by taking a different sample in each. Therefore we do this.

The asymptotic distribution depends on

$$\sqrt{n}\widehat{G}_n(\theta_0) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(\varepsilon_i - \left(\widehat{m}(z_i,\alpha_0) - m(z_i,\alpha_0)\right)\right) g(z_i,w_i;\alpha_0) + o_p(1)$$

because

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(\varepsilon_i - \left(\widehat{m}(z_i,\alpha_0) - m(z_i,\alpha_0)\right)\right)\begin{pmatrix} \frac{1}{R}\sum_{r=1}^{R}\frac{\partial h(X_r',z_i,\alpha_0)}{\partial \alpha}\frac{1}{p(X_r)} - g(z_i,w_i;\alpha_0) \\ 0 \end{pmatrix} = o_p(1)$$

We have

$$\sqrt{n}\widehat{G}_n(\theta_0) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(\varepsilon_i - \left(\widehat{m}(z_i,\alpha_0) - m(z_i,\alpha_0)\right)\right) g(z_i,w_i;\alpha_0) \Longrightarrow N(0, V + \frac{W}{R}),$$

where

$$\begin{aligned} V &= E\left[\varepsilon_i^2 \begin{pmatrix} \frac{\partial m(z_i,\alpha_0)}{\partial \alpha}\frac{\partial m(z_i,\alpha_0)}{\partial \alpha^T} & \frac{\partial m(z_i,\alpha_0)}{\partial \alpha} w_i^T \\ w_i \frac{\partial m(z_i,\alpha_0)}{\partial \alpha^T} & w_i w_i^T \end{pmatrix}\right] \\ W &= RE\left(\mathrm{var}[\widehat{m}(z_i,\alpha_0)|z_i] \begin{pmatrix} \frac{\partial m(z_i,\alpha_0)}{\partial \alpha}\frac{\partial m(z_i,\alpha_0)}{\partial \alpha^T} & \frac{\partial m(z_i,\alpha_0)}{\partial \alpha} w_i^T \\ w_i \frac{\partial m(z_i,\alpha_0)}{\partial \alpha^T} & w_i w_i^T \end{pmatrix}\right), \end{aligned}$$

where

$$R\mathrm{var}[\widehat{m}(z_i,\alpha_0)|z_i] = \int \frac{h(x;z_i,\alpha_0)^2}{p(x)}dx - \left(\int h(x;z_i,\alpha_0)dx\right)^2.$$

Also,

$$\Gamma = \frac{\partial}{\partial \theta} EG_n(\theta_0) = \begin{pmatrix} \frac{\partial m(z_i,\alpha_0)}{\partial \alpha}\frac{\partial m(z_i,\alpha_0)}{\partial \alpha^T} & \frac{\partial m(z_i,\alpha_0)}{\partial \alpha} w_i^T \\ w_i \frac{\partial m(z_i,\alpha_0)}{\partial \alpha^T} & w_i w_i^T \end{pmatrix}.$$

Therefore,

$$\sqrt{n}(\widehat{\theta}-\theta_0) \Longrightarrow N(0, \Gamma^{-1}(V+\frac{W}{R})\Gamma^{-1\prime}).$$

■

4. The efficient GMM estimator of θ minimizes the quadratic form

$$G_n(\theta)' W_{opt} G_n(\theta),$$

where

$$G_n(\theta) = \frac{1}{n}\sum_{i=1}^{n} m(Z_i, \theta)$$

and

$$W_{opt} = \left(E\left[m(Z_i,\theta_0)m(Z_i,\theta_0)'\right]\right)^{-1}.$$

This estimator has asymptotic variance

$$(\Gamma' W_{opt}\Gamma)^{-1} = \left(\sum_{j=1}^{J} \Gamma_j^2 E\left[m_j(Z_i,\theta_0)^2\right]\right)^{-1}$$

The estimator $\widehat{\theta}_j$ that solve the first order condition

$$G_{nj}(\widehat{\theta}_j) = \frac{1}{n}\sum_{i=1}^{n} m_j(Z_i, \widehat{\theta}_j) = 0$$

has asymptotic variance

$$\frac{\text{var}[\sqrt{n}G_{nj}(\theta_0)]}{\Gamma_j^2},$$

where

$$\text{var}[\sqrt{n}G_{nj}(\theta_0)] = E[m_j(Z_i,\theta_0)^2]$$

$$\Gamma_j = E\left[\frac{\partial m_j(Z_i,\theta_0)}{\partial\theta}\right].$$

The estimators have asymptotic covariance

$$E\left[\frac{\sqrt{n}G_{nj}(\theta_0)}{\Gamma_j}\frac{\sqrt{n}G_{nk}(\theta_0)}{\Gamma_k}\right] = \frac{E\left[m_j(Z_i,\theta_0)m_k(Z_i,\theta_0)\right]}{\Gamma_j\Gamma_k} = 0.$$

The weighted estimator has asymptotic variance

$$\text{var}\widehat{\theta}(w) = \text{var}\left[\sum_{j=1}^{q} w_j\widehat{\theta}_j\right] = \text{var}\left[\sum_{j=1}^{q} w_j\widehat{\theta}_j\right] = \sum_{j=1}^{q} w_j^2\frac{E[m_j(Z_i,\theta_0)^2]}{\Gamma_j^2}.$$

Let

$$w_j = \frac{\Gamma_j^2 / E[m_j(Z_i, \theta_0)^2]}{\sum_{j=1}^q \Gamma_j^2 / E[m_j(Z_i, \theta_0)^2]}.$$

Then these weights sum to one and

$$\text{var}\widehat{\theta}(w) = \frac{1}{\sum_{j=1}^J \Gamma_j^2 E\left[m_j(Z_i, \theta_0)^2\right]}.$$

■

5. The NLLSE solves the equation $G_n(\widehat{\beta}) = 0$ where

$$G_n(\beta) = \sum_{i=1}^n \left(y_i - \frac{1}{1 + \beta i^\alpha}\right) \frac{i^\alpha}{(1 + \beta i^\alpha)^2}.$$

We deal with the three cases separately.

if $\alpha = 0$

$$G_n(\beta) = \sum_{i=1}^n \left(y_i - \frac{1}{1 + \beta}\right) \frac{1}{(1 + \beta)^2}.$$

In this case, we let $\theta = 1/(1+\beta)$ in which case there is an exact solution to the first order condition for θ, i.e., $\widehat{\theta} = \overline{y}$. Substituting back we find $\beta = (1 - \theta)/\theta$. Therefore,

$$\widehat{\beta} = \frac{1 - \overline{y}}{\overline{y}}.$$

Clearly, $\widehat{\beta}$ is consistent unless $\theta = 0$ or $\theta = \infty$ [which corresponds to $\beta = \infty$ or $\beta = -1$]. The asymptotic distribution follows from the delta method. Thus

$$\sqrt{n}(\widehat{\beta} - \beta) = \sqrt{n}\left(\frac{1 - \overline{y}}{\overline{y}} - \frac{1 - \theta}{\theta}\right) \simeq \frac{-1}{\theta^2}\sqrt{n}(\overline{y} - \theta) \Longrightarrow N(0, \frac{\sigma^2}{\theta^4}) = N(0, \sigma^2(1 + \beta)^4).$$

The case $\alpha > 0$. We have

$$\frac{1}{1 + \beta i^\alpha} \simeq \frac{1}{\beta}\frac{1}{i^\alpha},$$

and so

$$G_n(\beta) \simeq \sum_{i=1}^n \left(y_i - \frac{1}{\beta i^\alpha}\right) \frac{1}{\beta^2 i^\alpha}.$$

Therefore,

$$G_n(\beta) = \sum_{i=1}^n \varepsilon_i \frac{1}{\beta^2 i^\alpha} + \sum_{i=1}^n \left(\frac{1}{\beta_0 i^\alpha} - \frac{1}{\beta i^\alpha}\right) \frac{1}{\beta^2 i^\alpha}.$$

The first term is

$$\sum_{i=1}^n \varepsilon_i \frac{1}{\beta^2 i^\alpha} = O_p(\sqrt{\sum_{i=1}^n i^{-2\alpha}}),$$

while the second term is approximately

$$\left(\frac{\beta - \beta_0}{\beta_0 \beta^3}\right) \sum_{i=1}^n \frac{1}{i^{2\alpha}} = O(\sum_{i=1}^n i^{-2\alpha}).$$

Therefore, if we re-normalize

$$G_n(\beta) = \frac{\sum_{i=1}^n \left(y_i - \frac{1}{1+\beta i^\alpha}\right) \frac{i^\alpha}{(1+\beta i^\alpha)^2}}{\sum_{i=1}^n \frac{1}{i^{2\alpha}}},$$

we find that provided $\sum_{i=1}^n i^{-2\alpha} \to \infty$

$$G_n(\beta) \to_p \frac{\beta - \beta_0}{\beta_0 \beta^3} \equiv G(\beta),$$

which has a unique zero at $\beta = \beta_0$. Therefore a sufficient condition for consistency is that $\sum_{i=1}^n i^{-2\alpha} \to \infty$, which is equivalent to the requirement that $\alpha \leq 1/2$. The asymptotic distribution is now determined by $\Gamma = G'(\beta_0) = 1/\beta_0^4$ and $\mathrm{var}[\sqrt{\sum_{i=1}^n i^{-2\alpha}} G_n(\beta_0)]$, specifically,

$$\sqrt{\sum_{i=1}^n i^{-2\alpha}} (\widehat{\beta} - \beta_0) \Longrightarrow N(0, \frac{\sigma^2}{\beta_0^4}).$$

The case $\alpha < 0$. Now we have

$$\frac{1}{1+\beta i^\alpha} \simeq 1 - \beta i^\alpha$$

and so

$$G_n(\beta) \simeq \sum_{i=1}^n (y_i - 1 + \beta i^\alpha)\, i^\alpha (1 - 2\beta i^\alpha) \simeq \sum_{i=1}^n \varepsilon_i i^\alpha + \sum_{i=1}^n (\beta_0 - \beta) i^{2\alpha}.$$

The first term is

$$\sum_{i=1}^n \varepsilon_i i^\alpha = O_p(\sqrt{\sum_{i=1}^n i^{2\alpha}}).$$

Therefore we renormalize by $\sum_{i=1}^n i^{2\alpha}$ and define

$$G_n(\beta) = \frac{\sum_{i=1}^n \left(y_i - \frac{1}{1+\beta i^\alpha}\right) \frac{i^\alpha}{(1+\beta i^\alpha)^2}}{\sum_{i=1}^n i^{2\alpha}}.$$

Provided $\sum_{i=1}^n i^{2\alpha} \to \infty$ we have

$$G_n(\beta) \to_p \beta_0 - \beta \equiv G(\beta),$$

which is uniquely zeroed at $\beta = \beta_0$. Provided $\alpha \geq -1/2$ we have consistency. As for asymptotic normality, we have

$$\sqrt{\sum_{i=1}^n i^{2\alpha}} (\widehat{\beta} - \beta_0) \Longrightarrow N(0, \sigma^2).$$

■

6. (a) We have

$$\begin{aligned} E\{y_i \in [\mu_0 - \sigma, \mu_0 + \sigma]\} &= \Pr(y_i \in [\mu_0 - \sigma, \mu_0 + \sigma]) \\ &= F_y(\mu_0 + \sigma) - F_y(\mu_0 - \sigma). \end{aligned}$$

Therefore

$$EG_n(\sigma) = F_y(\mu_0 + \sigma) - F_y(\mu_0 + \sigma_0) - (F_y(\mu_0 - \sigma) - F_y(\mu_0 - \sigma_0)).$$

(b) We can show that

$$G_n(\sigma) \to_p G(\sigma) = EG_n(\sigma)$$

uniformly in σ [the criterion is bounded and continuous with prob one- use the theorem in the notes]. Clearly, $G(\sigma_0) = 0$. The question is, does there exist another σ^* for which $G(\sigma^*) = 0$? Provided F is strictly monotonic in a neighborhood of σ_0, the answer is no. If $\sigma^* < \sigma_0$, then $F_y(\mu_0+\sigma_0) > F_y(\mu_0+\sigma^*)$ and $F_y(\mu_0-\sigma^*) > F_y(\mu_0-\sigma_0)$ so that $EG_n(\sigma^*) > 0$, while if $\sigma^* > \sigma_0$, then $F_y(\mu_0+\sigma_0) < F_y(\mu_0+\sigma^*)$ and $F_y(\mu_0 - \sigma^*) < F_y(\mu_0 - \sigma_0)$ so that $EG_n(\sigma^*) < 0$.

(c) The asymptotic distribution requires we compute

$$\Gamma = \frac{\partial}{\partial\sigma} G(\sigma_0) = f_y(\mu_0 + \sigma_0) + f_y(\mu_0 - \sigma_0)$$

$$\begin{aligned} V &= \text{var}[\sqrt{n} G_n(\sigma_0)] = \text{var}[\frac{1}{\sqrt{n}} \sum_{i=1}^n (\{y_i \in [\mu_0 - \sigma_0, \mu_0 + \sigma_0]\} - 0.5)] \\ &= \Pr(y_i \in [\mu_0 - \sigma_0, \mu_0 + \sigma_0]) - \Pr(y_i \in [\mu_0 - \sigma_0, \mu_0 + \sigma_0])^2 \\ &= \frac{1}{4}. \end{aligned}$$

Therefore,

$$\sqrt{n}(\widehat{\beta} - \beta_0) \Longrightarrow N(0, \frac{1}{4(f_y(\mu_0 + \sigma_0) + f_y(\mu_0 - \sigma_0))^2}).$$

(d) The first order condition we choose is

$$G_n(\theta) = \frac{1}{n} \sum_{i=1}^n \begin{pmatrix} x_i \text{sign}(y_i - \beta x_i) \\ \{y_i \in [\beta x_i - \sigma, \beta x_i + \sigma]\} - 0.5 \end{pmatrix},$$

where $\theta = (\beta, \sigma)$.