Problem Set 9 – Solutions

ECON 837

Brian Krauth (adapted from problems by Simon Woodcock), Spring 2010

1. Rewrite the first equation as

$$\begin{aligned}\mathbf{y}_1 &= \gamma_{21}^*\mathbf{y}_2 + \beta_{21}^*\mathbf{x}_2 + \beta_{31}^*\mathbf{x}_3 + \varepsilon_1 \\ &= \mathbf{Z}\delta + \varepsilon_1\end{aligned}$$

where $\gamma_{21}^* = -\gamma_{21}, \beta_{21}^* = -\beta_{21}, \beta_{31}^* = -\beta_{31}, \mathbf{Z} = \begin{bmatrix} \mathbf{y}_2 & \mathbf{x}_2 & \mathbf{x}_3 \end{bmatrix}$,and $\delta = \begin{bmatrix} \gamma_{21}^* & \beta_{21}^* & \beta_{31}^* \end{bmatrix}'$. The 2SLS estimator is given by

$$\hat{\delta}_{2SLS} = \left[\mathbf{Z}'\mathbf{X}\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{Z}\right]^{-1}\mathbf{Z}'\mathbf{X}\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{y}_1$$

for $\mathbf{X} = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \mathbf{x}_3 & \mathbf{x}_4 \end{bmatrix}$. Now

$$\mathbf{Z}'\mathbf{X} = \begin{bmatrix} \mathbf{y}_2 & \mathbf{x}_2 & \mathbf{x}_3 \end{bmatrix}'\begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \mathbf{x}_3 & \mathbf{x}_4 \end{bmatrix} = \begin{bmatrix} \mathbf{y}_2'\mathbf{x}_1 & \mathbf{y}_2'\mathbf{x}_2 & \mathbf{y}_2'\mathbf{x}_3 & \mathbf{y}_2'\mathbf{x}_4 \\ \mathbf{x}_2'\mathbf{x}_1 & \mathbf{x}_2'\mathbf{x}_2 & \mathbf{x}_2'\mathbf{x}_3 & \mathbf{x}_2'\mathbf{x}_4 \\ \mathbf{x}_3'\mathbf{x}_1 & \mathbf{x}_3'\mathbf{x}_2 & \mathbf{x}_3'\mathbf{x}_3 & \mathbf{x}_3'\mathbf{x}_4 \end{bmatrix} = \begin{bmatrix} -0.5 & 1.5 & 0.5 & -1 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

so that

$$\begin{aligned}\hat{\delta}_{2SLS} &= \left[\begin{bmatrix} -0.5 & 1.5 & 0.5 & -1 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0.5 \end{bmatrix}^{-1}\begin{bmatrix} -0.5 & 0 & 0 \\ 1.5 & 2 & 0 \\ 0.5 & 0 & 1 \\ -1 & 0 & 0 \end{bmatrix}\right]^{-1} \\ &\quad\times\begin{bmatrix} -0.5 & 1.5 & 0.5 & -1 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0.5 \end{bmatrix}^{-1}\begin{bmatrix} 2 \\ 1 \\ -3 \\ -5 \end{bmatrix} \\ &= \begin{bmatrix} 4.64 \\ -2.98 \\ -5.32 \end{bmatrix}\end{aligned}$$

and therefore

$$\begin{bmatrix} \gamma_{21} \\ \beta_{21} \\ \beta_{31} \end{bmatrix} = \begin{bmatrix} -4.64 \\ 2.98 \\ 5.32 \end{bmatrix}.$$

As for the standard errors, let

$$\hat{\sigma}^2 = \frac{\left(\mathbf{y}_1 - \mathbf{Z}\hat{\delta}_{2SLS}\right)'\left(\mathbf{y}_1 - \mathbf{Z}\hat{\delta}_{2SLS}\right)}{n} = \frac{\mathbf{y}_1'\mathbf{y}_1 - 2\mathbf{y}_1'\mathbf{Z}\hat{\delta}_{2SLS} + \hat{\delta}_{2SLS}'\mathbf{Z}'\mathbf{Z}\hat{\delta}_{2SLS}}{100}.$$

The elements of this are

$$\begin{aligned}\mathbf{y}_1'\mathbf{y}_1 &= 80 \\ \mathbf{y}_1'\mathbf{Z}\hat{\delta}_{2SLS} &= \begin{bmatrix} \mathbf{y}_1'\mathbf{y}_2 & \mathbf{y}_1'\mathbf{x}_2 & \mathbf{y}_1'\mathbf{x}_3 \end{bmatrix}\begin{bmatrix} 4.64 \\ -2.98 \\ -5.32 \end{bmatrix} = \begin{bmatrix} -4 & 1 & -3 \end{bmatrix}\begin{bmatrix} 4.64 \\ -2.98 \\ -5.32 \end{bmatrix} = -5.58 \\ \hat{\delta}_{2SLS}'\mathbf{Z}'\mathbf{Z}\hat{\delta}_{2SLS} &= \hat{\delta}_{2SLS}'\begin{bmatrix} \mathbf{y}_2'\mathbf{y}_2 & \mathbf{y}_2'\mathbf{x}_2 & \mathbf{y}_2'\mathbf{x}_3 \\ \mathbf{y}_2'\mathbf{x}_2 & \mathbf{x}_2'\mathbf{x}_2 & \mathbf{x}_2'\mathbf{x}_3 \\ \mathbf{y}_2'\mathbf{x}_3 & \mathbf{x}_2'\mathbf{x}_3 & \mathbf{x}_3'\mathbf{x}_3 \end{bmatrix}\hat{\delta}_{2SLS} \\ &= \begin{bmatrix} 4.64 & -2.98 & -5.32 \end{bmatrix}\begin{bmatrix} 5 & 1.5 & 0.5 \\ 1.5 & 2 & 0 \\ 0.5 & 0 & 1 \end{bmatrix}\begin{bmatrix} 4.64 \\ -2.98 \\ -5.32 \end{bmatrix} = 87.55\end{aligned}$$

so that

$$\hat{\sigma}^2 = \frac{80 - 2\left(-5.58\right) + 87.55}{100} = 1.79.$$

Finally, we estimate the variance of the 2SLS estimator with

$$\widehat{Var}\left[\hat{\delta}_{2SLS}\right] = \hat{\sigma}^2\left[\mathbf{Z}'\mathbf{X}\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{Z}\right]^{-1} = \begin{bmatrix} 0.858 & -0.643 & -0.429 \\ -0.643 & 1.376 & 0.322 \\ -0.429 & 0.322 & 2.002 \end{bmatrix}.$$

2. (a) Under suitable regularity conditions, we know that the NLLS estimator of β is consistent and asymptotically normal (see Lecture 18). In fact, we know that

$$\sqrt{n}\left(\hat{\beta}-\beta^0\right) \overset{a}{\sim} N\left(\mathbf{0},\sigma^2\mathbf{Q}^{-1}\right)$$

or

$$\hat{\beta} \overset{a}{\sim} N\left(\beta^0,\frac{\sigma^2}{n}\mathbf{Q}^{-1}\right). \tag{1}$$

where β^0 is the vector of true parameter values, $\sigma^2 = Var\left[\varepsilon_i\right]$, n is the sample size, and

$$\text{plim}\left(\frac{\mathbf{G}\left(\beta^0\right)'\mathbf{G}\left(\beta^0\right)}{n}\right)=\mathbf{Q}$$

for gradient matrix $\mathbf{G}$. For the model in question, x_i is a scalar so the gradient is a vector:

$$\mathbf{g}\left(\beta\right)=\frac{\partial\exp\left(\mathbf{x}\beta\right)}{\partial\beta}=\begin{bmatrix} x_1\exp\left(x_1\beta\right)\\ x_2\exp\left(x_2\beta\right)\\ \vdots\\ x_n\exp\left(x_n\beta\right)\end{bmatrix}$$

and hence

$$\begin{aligned}\mathbf{g}\left(\beta\right)'\mathbf{g}\left(\beta\right) &= \begin{bmatrix} x_1\exp\left(x_1\beta\right) & x_2\exp\left(x_2\beta\right) & \cdots & x_n\exp\left(x_n\beta\right)\end{bmatrix}\begin{bmatrix} x_1\exp\left(x_1\beta\right)\\ x_2\exp\left(x_2\beta\right)\\ \vdots\\ x_n\exp\left(x_n\beta\right)\end{bmatrix}\\ &= \sum_{i=1}^{n} x_i^2\exp\left(2x_i\beta\right)\end{aligned}$$

and therefore

$$\sqrt{n}\left(\hat{\beta}-\beta\right)\overset{d}{\rightarrow}N\left(0,\sigma^2\left(\frac{\sum_{i=1}^{n}x_i^2\exp\left(2x_i\beta\right)}{n}\right)^{-1}\right)$$

or

$$\hat{\beta}\overset{a}{\sim}N\left(\beta,\frac{\sigma^2}{n}\left(\frac{1}{n}\sum_{i=1}^{n}x_i^2\exp\left(2x_i\beta\right)\right)^{-1}\right).$$

We can estimate the asymptotic variance with

$$Est.Asy.Var\left[\hat{\beta}\right]=\hat{\sigma}^2\left(\frac{1}{n}\sum_{i=1}^{n}x_i^2\exp\left(2x_i\hat{\beta}\right)\right)^{-1}$$

where

$$\hat{\sigma}^2=\frac{1}{n}\left(\mathbf{y}-\exp\left(\mathbf{x}\hat{\beta}\right)\right)'\left(\mathbf{y}-\exp\left(\mathbf{x}\hat{\beta}\right)\right).$$

(b) Since we've assumed a parametric joint distribution for the data, MLE is feasible here and we know it is "better" in the sense of asymptotic efficiency. The density of a single observation is

$$f\left(y_i|x_i,\beta\right)=\exp\left(-x_i\beta\right)\exp\left(-y_ie^{-x_i\beta}\right)$$

and hence the likelihood function:

$$L\left(\beta\right)=\prod_{i=1}^{n}\exp\left(-x_i\beta\right)\exp\left(-y_ie^{-x_i\beta}\right)$$

and log-likelihood

$$l\left(\beta\right)=\sum_{i=1}^{n}\left(-x_i\beta-y_i\exp\left(-x_i\beta\right)\right).$$

The score function is

$$\frac{\partial l(\beta)}{\partial \beta} = -\sum_{i=1}^{n} x_i + \sum_{i=1}^{n} y_i x_i \exp(-x_i\beta).$$

We can solve these numerically for the MLE $\hat{\beta}_{ML}$. The second derivative is

$$\frac{\partial^2 l(\beta)}{\partial \beta^2} = -\sum_{i=1}^{n} y_i x_i^2 \exp(-x_i\beta).$$

Therefore,

$$E\left[-\frac{\partial^2 l(\beta)}{\partial \beta^2}\right] = E\left[\sum_{i=1}^{n} y_i x_i^2 \exp(-x_i\beta)\right] = \sum_{i=1}^{n} E[y_i] x_i^2 \exp(-x_i\beta) = \sum_{i=1}^{n} \exp(x_i\beta) x_i^2 \exp(-x_i\beta) = \sum_{i=1}^{n} x_i^2$$

so that

$$\sqrt{n}\left(\hat{\beta}_{ML} - \beta\right) \xrightarrow{d} N\left(0, \sigma^2 \left(\frac{\sum_{i=1}^{n} x_i^2}{n}\right)^{-1}\right)$$

or

$$\hat{\beta}_{ML} \overset{a}{\sim} N\left(\beta, \frac{\sigma^2}{n}\left(\frac{1}{n}\sum_{i=1}^{n} x_i^2\right)^{-1}\right).$$

3. You can verify all of the following:

θ	(a) MLE θ	(b) OLS β
[1] $\bar{\mathbf{m}}(\theta)$	$\frac{\partial l(\theta)}{\partial \theta'}$	$\frac{1}{n}\mathbf{X}'(\mathbf{y} - \mathbf{X}\beta)$
[2] $\mathbf{G}(\theta)$	$\frac{\partial^2 l(\theta)}{\partial\theta\partial\theta'}$ or $\frac{\partial l(\theta)}{\partial \theta}\frac{\partial l(\theta)}{\partial \theta'}$	$-\frac{1}{n}\mathbf{X}'\mathbf{X}$
[3] $Var[\bar{\mathbf{m}}(\theta)]$	$-\frac{\partial^2 l(\theta)}{\partial\theta\partial\theta'}$	$\mathbf{X}'\mathbf{V}\mathbf{X}$
[4] $\hat{\theta}_{GMM}$	implicit solution to $\frac{\partial l(\theta)}{\partial \theta'} = \mathbf{0}$	$(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$
[5] $Var\left[\sqrt{n}\left(\hat{\theta} - \theta\right)\right]$	$E\left[-\frac{\partial^2 l(\theta)}{\partial\theta\partial\theta'}\right]^{-1}$	$(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}$

4.

(a) The log-likelihood takes the very simple form $l(\gamma) = \sum_i (\ln\gamma - t_i\gamma)$. The MLE solves the FOC $\sum_i \left(\gamma^{-1} - t_i\right) = 0$, so that the MLE is $\hat{\gamma} = n/\sum_i t_i = \bar{t}^{-1}$ where $\bar{t} = n^{-1}\sum_i t_i$ is the sample mean of t. Differentiating the log-likeliood again, $d^2 l(\gamma)/d\gamma^2 = -\sum_i \gamma^{-2} = -n\gamma^{-2}$. Hence the variance of the MLE is $Asy.Var(\hat{\gamma}) = -E\left[d^2 l(\gamma)/d\gamma^2\right]^{-1} = \gamma^2/n$.

(b) You can choose any number of moment conditions. You need at least one. You might, for example, use sample analogs of population moment conditions based on the mean and variance of t, e.g., $E[t] = \gamma^{-1}$ and $V[t] = \gamma^{-2}$. Or you might use the maximum likelihood FOC, which in this case is the same as the sample moment condition based on the population mean. In the latter case, MLE and GMM are equivalent.

(c) Define $d_i = 1$ if the observation is censored (i.e., the individual dropped out of the sample) and $d_i = 0$ otherwise. Since all we know is that $t > t_i$ for the censored observations, the log-likelihood is now

$$\begin{aligned} l(\gamma) &= \sum_i \left[(1 - d_i)(\ln\gamma - t_i\gamma) + d_i \ln \Pr(t > t_i)\right] \\ &= \sum_i \left[(1 - d_i)(\ln\gamma - t_i\gamma) + d_i \ln(1 - \Pr(t \le t_i))\right] \\ &= \sum_i \left[(1 - d_i)(\ln\gamma - t_i\gamma) + d_i \ln\left(1 - \int_0^{t_i} \gamma e^{-\gamma s} ds\right)\right] \\ &= \sum_i \left[(1 - d_i)(\ln\gamma - t_i\gamma) + d_i \ln e^{-\gamma t_i}\right] \\ &= \sum_i \left[(1 - d_i)(\ln\gamma - t_i\gamma) + d_i t_i \gamma\right]. \end{aligned}$$

Hence the MLE solves the FOC:

$$\sum_i \left[\left(\frac{1 - d_i}{\gamma}\right) - (1 - d_i)t_i + d_i t_i\right] = 0$$

and you can verify that the solution is

$$\hat{\gamma} = \frac{\sum (1 - d_i)}{\sum t_i} = \frac{\#\text{ complete spells}}{\sum t_i} \frac{n}{n} = \bar{t}^{-1} \frac{\#\text{ complete spells}}{n}$$

i.e., it is the MLE we found before, multiplied by the proportion of spells that are complete.