BGPE Frontiers in Econometrics 2011
A. Colin Cameron, U.C.-Davis

Answers to Final Exam

1.(a) We have $\widehat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\mathbf{X}\boldsymbol{\beta}+\mathbf{u}) = \boldsymbol{\beta} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{u}$. So

$$\begin{aligned}\widehat{\boldsymbol{\beta}} &= \boldsymbol{\beta} + \left(N^{-1}\mathbf{X}'\mathbf{X}\right)^{-1} N^{-1}\mathbf{X}'\mathbf{u} \\ &\xrightarrow{p} \boldsymbol{\beta} + \left(\text{plim}\, N^{-1}\mathbf{X}'\mathbf{X}\right)^{-1} \text{plim}\, N^{-1}\mathbf{X}'\mathbf{u} \\ &\xrightarrow{p} \boldsymbol{\beta} + \mathbf{A} \times \mathbf{0} \xrightarrow{p} \boldsymbol{\beta}.\end{aligned}$$

(b) We have

$$\begin{aligned}\sqrt{N}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}) &= \left(N^{-1}\mathbf{X}'\mathbf{X}\right)^{-1} N^{-1/2}\mathbf{X}'\mathbf{u} \\ &\xrightarrow{d} \boldsymbol{\beta} + \left(\text{plim}\, N^{-1}\mathbf{X}'\mathbf{X}\right)^{-1} \times \mathcal{N}[\mathbf{0},\, \mathbf{B}] \\ &\xrightarrow{d} \boldsymbol{\beta}\ + \mathbf{A}^{-1} \times \mathcal{N}[\mathbf{0},\, \mathbf{B}] \\ &\xrightarrow{d} \mathcal{N}[\mathbf{0},\, \mathbf{A}^{-1}\mathbf{B}\mathbf{A}^{-1}]\end{aligned}$$

(c) Use $\widehat{\mathbf{A}} = N^{-1}\mathbf{X}'\mathbf{X}$ and since $\mathbf{B} = \lim \text{E}[N^{-1}\mathbf{X}'\mathbf{u}\mathbf{u}'\mathbf{X}] = \lim \text{E}[N^{-1}\sum_{i=1}^N u_i^2 \mathbf{x}_i\mathbf{x}_i']$ use $\widehat{\mathbf{B}} = N^{-1}\sum_{i=1}^N \widehat{u}_i^2 \mathbf{x}_i\mathbf{x}_i'$. Then $\widehat{\mathbf{A}}^{-1}\widehat{\mathbf{B}}\widehat{\mathbf{A}}^{-1} = N(\mathbf{X}'\mathbf{X})^{-1}\sum_{i=1}^N \widehat{u}_i^2\mathbf{x}_i\mathbf{x}_i'(\mathbf{X}'\mathbf{X})^{-1}$.

(d) Now

$$\begin{aligned}\widehat{\boldsymbol{\beta}} &= \boldsymbol{\beta} + \left(N^{-1}\mathbf{X}'\mathbf{X}\right)^{-1} N^{-1}\mathbf{X}'\mathbf{u} \\ &= \boldsymbol{\beta} + \left(N^{-1}\mathbf{X}'\mathbf{X}\right)^{-1} (N^{-1}\mathbf{X}'\mathbf{X}\boldsymbol{\delta} + N^{-1}\mathbf{X}'\mathbf{v}) \\ &\xrightarrow{p} \boldsymbol{\beta} + \left(\text{plim}\, N^{-1}\mathbf{X}'\mathbf{X}\right)^{-1} (\text{plim}\, N^{-1}\mathbf{X}'\mathbf{X}\boldsymbol{\delta} + \text{plim}\, N^{-1}\mathbf{X}'\mathbf{v} \\ &\xrightarrow{p} \boldsymbol{\beta} + \mathbf{A} \times (\mathbf{0} + \mathbf{A}^{-1}\boldsymbol{\delta}) \xrightarrow{p} \boldsymbol{\beta} + \boldsymbol{\delta}.\end{aligned}$$

(e) Here $\mathbf{R} = \begin{bmatrix} 0 & 1 & -1 \end{bmatrix}$ and $r = 0$. $\mathbf{R}\widehat{\boldsymbol{\beta}} - r = \begin{bmatrix} 0 & 1 & -1 \end{bmatrix}\begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix} = 3 - 5 = -2.$

$$\mathbf{R}\widehat{\text{V}}[\widehat{\boldsymbol{\beta}}]\mathbf{R}' = \begin{bmatrix} 0 & 1 & -1 \end{bmatrix}\begin{bmatrix} 2 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 3 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 & 2 & -2 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} = 4.$$

W$= (\mathbf{R}\widehat{\boldsymbol{\beta}} - r)'(\mathbf{R}\widehat{\text{V}}[\widehat{\boldsymbol{\beta}}]\mathbf{R}')^{-1}(\mathbf{R}\widehat{\boldsymbol{\beta}} - r) = (-2)(4)^{-1}(-2) = 1 < \chi^2_{.05}(1) = 3.84.$
Do not reject H_0.
(Could instead use $\widehat{\text{V}}[\widehat{\boldsymbol{\beta}}_2 - \widehat{\boldsymbol{\beta}}_3] = \widehat{\text{V}}[\widehat{\boldsymbol{\beta}}_2] - 2\widehat{\text{Cov}}[\widehat{\boldsymbol{\beta}}_2, \widehat{\boldsymbol{\beta}}_2] + \widehat{\text{V}}[\widehat{\boldsymbol{\beta}}_3] = 3 - 2 \times 1 + 3 = 4.$)

(f) No. We would consider $\sqrt{N}\sum_{i=1}^N x_iu_i$. Lindeberg-Levy CLT requires x_iu_i to be i.i.d. which is not the case if x_i is fixed. Instead x_iu_i will have mean 0 and variance $\sigma^2x_i^2$ so is not i.i.d.

2.(a) OUTPUT1 from OLS is preferred to OUTPUT2 (2SLS) if there is no endogeneity i.e. if $\text{plim}\, N^{-1}\sum_{i=1}^N \mathbf{x}_iu_i = \mathbf{0}$.

(b) Perform a Hausman test (preferably robust form) that compares the OLS coefficients (OUTPUT1) to the 2SLS coefficients (OUTPUT2) with 2SLS preferred if a big difference.

(c) There is a big efficiency loss in 2SLS compared to OLS due to the instruments `firmsz` and `multlc` being weakly correlated with `hi_empunion` as clear from command `correlate`.

(d) Greater efficiency. "Optimal" GMM (OUTPUT3) is fully efficient with heteroskedastic errors whereas 2SLS (OUTPUT2) presumes homoskedastic errors.

(e) 2SLS: $\widehat{\boldsymbol{\beta}}_{2SLS} = (\mathbf{X}'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{X})^{-1} \times \mathbf{X}'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{y}$.
Now consider the second set of output on the final page, which comes from model $y_{it} = \alpha_i + \mathbf{x}_{it}'\boldsymbol{\beta} + u_{it}$ where y_{it} is `wage` (the level of hourly wage) and $\mathbf{x}_{it}$ is `wks` (annual weeks worked) and `ed` (years of schooling).

(f) RE (OUTPUT5) preferred to OLS (OUTPUT4) if equicorrelated homoskedastic errors (implication of RE model) is a better model for the error correlation than independent homoskedastic.

(g) FE (OUTPUT6) is preferred to RE (OUTPUT5) if the individual effect α_i is correlated with the regressors $\mathbf{x}_{it}$.

(h) The purpose of option `vce(robust)` is to allow for any error correlation or heteroskedasticity even after inclusion of the fixed effect α_i.

3.(a) Here $Q_N(\boldsymbol{\beta}) = \mathcal{L}_N(\boldsymbol{\beta})$ where

$$\begin{aligned} \mathcal{L}_N(\boldsymbol{\beta}) &= \tfrac{1}{N}\textstyle\sum_{i=1}^N \{-\theta_i/y_i + 2\ln\theta_i - 3\ln y_i - \ln 2\} \\ &= \textstyle\sum_{i=1}^N \{-\exp(\mathbf{x}_i'\boldsymbol{\beta})/y_i + 2\mathbf{x}_i'\boldsymbol{\beta} - 3\ln y_i - \ln 2\}. \end{aligned}$$

(b) Then

$$\begin{aligned} \frac{\partial \mathcal{L}_N(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} &= \textstyle\sum_{i=1}^N \left(-(\exp(\mathbf{x}_i'\boldsymbol{\beta})/y_i)\mathbf{x}_i + 2\mathbf{x}_i\right) \\ &= \textstyle\sum_{i=1}^N \left(\frac{-\exp(\mathbf{x}_i'\boldsymbol{\beta})}{y_i} + 2\right)\mathbf{x}_i \end{aligned}$$

(c) Here can use standard ML results that $\widehat{\boldsymbol{\beta}} \stackrel{a}{\sim} \mathcal{N}[\boldsymbol{\beta}, -(\mathrm{E}[\frac{\partial^2 \mathcal{L}_N(\boldsymbol{\beta})}{\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}'}\Big|_{\boldsymbol{\beta}_0}])^{-1}]$.

$$\begin{aligned} \frac{\partial^2 \mathcal{L}_N(\boldsymbol{\beta})}{\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}'} &= \frac{\partial}{\partial\boldsymbol{\beta}}\textstyle\sum_{i=1}^N \left(\frac{-\exp(\mathbf{x}_i'\boldsymbol{\beta})}{y_i} + 2\right)\mathbf{x}_i = \sum_{i=1}^N \frac{-\exp(\mathbf{x}_i'\boldsymbol{\beta})}{y_i}\mathbf{x}_i\mathbf{x}_i' \\ \mathrm{E}\left[\frac{\partial^2 \mathcal{L}_N(\boldsymbol{\beta})}{\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}'}\Big|_{\boldsymbol{\beta}_0}\right] &= \textstyle\sum_{i=1}^N \mathrm{E}\left[\frac{-\exp(\mathbf{x}_i'\boldsymbol{\beta}_0)}{y_i}\mathbf{x}_i\mathbf{x}_i'\right] = \sum_{i=1}^N -2\mathbf{x}_i\mathbf{x}_i', \text{ since } \mathrm{E}[1/y_i] = 2/\exp(\mathbf{x}_i'\boldsymbol{\beta}_0) \end{aligned}$$

So $\mathrm{V}[\widehat{\boldsymbol{\beta}}] = (\sum_{i=1}^N 2\mathbf{x}_i\mathbf{x}_i')^{-1}$.

(d) The estimator is consistent if $\mathrm{E}[\frac{-\exp(\mathbf{x}_i'\boldsymbol{\beta})}{y_i} + 2|\mathbf{x}_i] = 0$ which is the case given $\mathrm{E}[1/y_i|\mathbf{x}_i] = 2/\exp(\mathbf{x}_i'\boldsymbol{\beta}_0)$. NOTE: $\mathrm{E}[1/y] \neq 1/\mathrm{E}[y]$!!

(e) Newton Raphson

$$(\widehat{\boldsymbol{\beta}}_{s+1} - \widehat{\boldsymbol{\beta}}_s) = -\mathbf{H}_s{}^{-1}\mathbf{g}_s = \left(\textstyle\sum_{i=1}^N \frac{-\exp(\mathbf{x}_i'\widehat{\boldsymbol{\beta}}}{y_i}\mathbf{x}_i\mathbf{x}_i'\right)^{-1} \sum_{i=1}^N \left(\frac{-\exp(\mathbf{x}_i'\widehat{\boldsymbol{\beta}})}{y_i} + 2\right)\mathbf{x}_i.$$

Or could use method of scoring

$$(\widehat{\boldsymbol{\beta}}_{s+1} - \widehat{\boldsymbol{\beta}}_s) = -\mathrm{E}[\mathbf{H}_s]^{-1}\mathbf{g}_s = \left(\textstyle\sum_{i=1}^N 2\mathbf{x}_i\mathbf{x}_i'\right)^{-1} \sum_{i=1}^N \left(\frac{-\exp(\mathbf{x}_i'\widehat{\boldsymbol{\beta}})}{y_i} + 2\right)\mathbf{x}_i.$$

(f) $\mathrm{AME}_j = \frac{1}{N}\sum_{i=1}^N \partial\exp(\mathbf{x}_i'\boldsymbol{\beta})/\partial x_{ij} = \frac{1}{N}\sum_{i=1}^N \exp(\mathbf{x}_i'\boldsymbol{\beta})/\partial x_{ij}$.

4.(a) No. Unbiased as the average value from simulations 1.022277 is close to the d.g.p. value of 1.0 (and the 95% simulation interval of (.997, 1.047) include 1.0.

(b) The average of the standard errors is 0.362 which differs somewhat from the simulation estimate of 0.40636. So within 10%. Judgement call as to whether this is a problem. Perhaps some bias (perhaps due to $N = 50$) but not bad.

(c) In theory it should be 0.05 given the simulation code. Perhaps expect a little more than 0.05 since from (b) the standard errors may be underestimated somewhat, leading to mild over-rejection.

(d) Just crank up the sample size

```
set obs 1000000
generate double x = rnormal(0,1)
generate mu = exp(-2 + 1*x)
generate double y = rpoisson(mu)
poisson y x
```

(e) Do a bootstrap.
400 times (for example) resample with replacement from the original sample and do Poisson regression in each resample. This gives 400 $\widehat{\boldsymbol{\beta}}'s$. Then use the standard deviation of these 400 $\widehat{\boldsymbol{\beta}}'s$.

(f) Use Stata command `ml` method `lf` as in the assignment.

On this exam: Median 24.5 out of 30 and range 17.5 to 28.
I would give A for 26+; A- for 22.5+; B+ for 17.5+