Answer key for the final exam in 2019

1a.) Choose the banks from which to purchase assets and the amounts purchased from each bank completely randomly.

b.) Government likely purchased more assets from banks that were in more trouble or had more troubled assets. These banks may have performed more poorly than other banks even if x_i had been zero. This would mean $E(x_i u_i) < 0$ and the estimate of β would be biased down.

c.) One cannot test this from OLS residuals because $\sum_{i=1}^n x_i \hat{u}_i = 0$ regardless of whether $E(x_i u_i) = 0$.

2a.) $\hat{\boldsymbol{\beta}} = \left(T^{-1}\sum_{t=1}^T \mathbf{x}_t \mathbf{z}_t'\right)^{-1}\left(T^{-1}\sum_{t=1}^T \mathbf{x}_t y_t\right) \xrightarrow{p} [E(\mathbf{x}_t\mathbf{z}_t')]^{-1}[E(\mathbf{x}_t y_t)]$

This requires only that $E(\mathbf{x}_t y_t)$ and $E(\mathbf{x}_t\mathbf{z}_t')$ exist and that $E(\mathbf{x}_t\mathbf{z}_t')$ has rank k.

b.) No, for purposes of forecasting we instead would want $\boldsymbol{\beta}^* = [E(\mathbf{z}_t\mathbf{z}_t')]^{-1}[E(\mathbf{z}_t y_t)]$.

3a.) $E(\mathbf{x}_t u_t) = \mathbf{0}$

b.)

$$\begin{bmatrix} E(\mathbf{z}_{t1}\mathbf{z}_{t1}') & E(\mathbf{z}_{t1}\mathbf{z}_{t2}') \\ E(\mathbf{x}_t\mathbf{z}_{t1}') & E(\mathbf{x}_t\mathbf{z}_{t2}') \end{bmatrix} \text{has rank } k_1 + k_2$$

c.) Regress ith element of $\mathbf{z}_{t2}$ on $(\mathbf{z}_{t1}, \mathbf{x}_t)$ and look at OLS F-test of hypothesis that coefficients on $\mathbf{x}_t$ are all zero. Want this F-test to be bigger than 10 to be convinced instruments are strong.

d.) Regress $y_t = \mathbf{z}_{t1}'\boldsymbol{\beta}_1 + \mathbf{x}_t'\boldsymbol{\gamma} + v_t$ and look at OLS F-test of null hypothesis that $\boldsymbol{\gamma} = \mathbf{0}$.

d.) Would have exact F-distribution if $u_t \sim$ i.i.d. $N(0,\sigma^2)$ with u_t independent of $(\mathbf{z}_{s1}', \mathbf{x}_s')'$ for all t, s.

4a.)

$$\mathbf{s}_t(\boldsymbol{\theta}) = \frac{\partial \ell_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} = \begin{bmatrix} \frac{(y_t - \mathbf{x}_t'\boldsymbol{\beta})\mathbf{x}_t}{\sigma^2} \\ -\frac{1}{2\sigma^2} + \frac{(y_t - \mathbf{x}_t'\boldsymbol{\beta})^2}{2\sigma^4} \end{bmatrix}$$

$$E[\mathbf{s}_t(\boldsymbol{\theta}_0)|\mathbf{s}_{t-1}(\boldsymbol{\theta}_0), \ldots, \mathbf{s}_1(\boldsymbol{\theta}_0)] = \begin{bmatrix} \frac{E(u_t\mathbf{x}_t|y_{t-1},\mathbf{x}_{t-1},\ldots,y_1,\mathbf{x}_1)}{\sigma_0^2} \\ \frac{E[(u_t^2-\sigma_0^2)|y_{t-1},\mathbf{x}_{t-1},\ldots,y_1,\mathbf{x}_1]}{2\sigma_0^4} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ 0 \end{bmatrix}.$$

b.) $\hat{\boldsymbol{\theta}}_{MLE}$ satisfies $\sum_{t=1}^T s_t(\hat{\boldsymbol{\theta}}_{MLE}) = \mathbf{0}$ or

$\sum_{t=1}^T (y_t - \mathbf{x}_t'\hat{\boldsymbol{\beta}}_{MLE})\mathbf{x}_t = \mathbf{0} \Rightarrow \hat{\boldsymbol{\beta}}_{MLE} = \left(\sum_{t=1}^T \mathbf{x}_t\mathbf{x}_t'\right)^{-1}\left(\sum_{t=1}^T \mathbf{x}_t y_t\right)$ and

$-T\hat{\sigma}_{MLE}^2 + \sum_{t=1}^T (y_t - \mathbf{x}_t'\hat{\boldsymbol{\beta}}_{MLE})^2 = 0 \Rightarrow \hat{\sigma}_{MLE}^2 = T^{-1}\sum_{t=1}^T \hat{u}_t^2$ for $\hat{u}_t = y_t - \mathbf{x}_t'\hat{\boldsymbol{\beta}}_{MLE}$

c.) $a = r = k + 1$ and $\mathbf{h}(\boldsymbol{\theta}, \mathbf{w}_t) = \mathbf{s}_t(\boldsymbol{\theta})$

d.)

$$\hat{\mathbf{D}}' = T^{-1}\sum_{t=1}^{T}\frac{\partial \ell_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'}\bigg|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} = T^{-1}\sum_{t=1}^{T}\begin{bmatrix} \frac{-\mathbf{x}_t\mathbf{x}_t'}{\hat{\sigma}^2} & \frac{-\mathbf{x}_t\hat{u}_t}{\hat{\sigma}^4} \\ \frac{-\hat{u}_t\mathbf{x}_t'}{\hat{\sigma}^4} & \frac{1}{2\hat{\sigma}^4} - \frac{\hat{u}_t^2}{\hat{\sigma}^6} \end{bmatrix}$$

$$= \begin{bmatrix} -\hat{\sigma}^{-2}T^{-1}\sum_{t=1}^{T}\mathbf{x}_t\mathbf{x}_t' & \mathbf{0} \\ \mathbf{0}' & -1/(2\hat{\sigma}^4) \end{bmatrix} \xrightarrow{p} \begin{bmatrix} -\sigma_0^{-2}\mathbf{Q} & \mathbf{0} \\ \mathbf{0}' & -1/(2\sigma_0^4) \end{bmatrix}$$

for $\mathbf{Q} = E(\mathbf{x}_t\mathbf{x}_t')$

e.)

$$\hat{\mathbf{S}} = T^{-1}\sum_{t=1}^{T}\begin{bmatrix} \frac{\hat{u}_t^2\mathbf{x}_t x_t'}{\hat{\sigma}^4} & \frac{\hat{u}_t\mathbf{x}_t}{\hat{\sigma}^2}\left(\frac{-1}{2\hat{\sigma}^2} + \frac{\hat{u}_t^2}{2\hat{\sigma}^4}\right) \\ \frac{\hat{u}_t\mathbf{x}_t'}{\hat{\sigma}^2}\left(\frac{-1}{2\hat{\sigma}^2} + \frac{\hat{u}_t^2}{2\hat{\sigma}^4}\right) & \left(\frac{-1}{2\hat{\sigma}^2} + \frac{\hat{u}_t^2}{2\hat{\sigma}^4}\right)^2 \end{bmatrix}$$

$$= T^{-1}\sum_{t=1}^{T}\begin{bmatrix} \frac{\hat{u}_t^2\mathbf{x}_t x_t'}{\hat{\sigma}^4} & \frac{\hat{u}_t\mathbf{x}_t}{\hat{\sigma}^2}\left(\frac{\hat{u}_t^2}{2\hat{\sigma}^4}\right) \\ \frac{\hat{u}_t\mathbf{x}_t'}{\hat{\sigma}^2}\left(\frac{\hat{u}_t^2}{2\hat{\sigma}^4}\right) & \frac{(\hat{u}_t^2-\hat{\sigma}^2)^2}{4\hat{\sigma}^8} \end{bmatrix} \xrightarrow{p} \begin{bmatrix} \sigma_0^{-2}\mathbf{Q} & \mathbf{0} \\ \mathbf{0}' & 1/(2\sigma_0^4) \end{bmatrix}.$$

Note $\mathbf{S} = -\mathbf{D}$

f.) $\mathbf{V} = \mathbf{S}^{-1} = \begin{bmatrix} \sigma_0^2\mathbf{Q}^{-1} & \mathbf{0} \\ \mathbf{0}' & (2\sigma_0^4) \end{bmatrix}$. The first block says $\sqrt{T}(\hat{\boldsymbol{\beta}}_{MLE} - \boldsymbol{\beta}_0) \xrightarrow{L} N(\mathbf{0}, \sigma_0^2\mathbf{Q}^{-1})$ which is the usual OLS result.

g.) If we set the off-diagonal elements of $\mathbf{S}$ to their plim of zero (though this is not imposed by the properties of the estimate $\hat{\mathbf{S}}$ itself), the upper-left block of $\hat{\mathbf{V}}$ is

$$\left[\left(-\hat{\sigma}^{-2}T^{-1}\sum_{t=1}^{T}\mathbf{x}_t\mathbf{x}_t'\right)\left(T^{-1}\frac{\sum_{t=1}^{T}\hat{u}_t^2\mathbf{x}_t x_t'}{\hat{\sigma}^4}\right)^{-1}\left(-\hat{\sigma}^{-2}T^{-1}\sum_{t=1}^{T}\mathbf{x}_t\mathbf{x}_t'\right)\right]^{-1}$$

$$= T\left[\left(\sum_{t=1}^{T}\mathbf{x}_t\mathbf{x}_t'\right)\left(\sum_{t=1}^{T}\hat{u}_t^2\mathbf{x}_t x_t'\right)^{-1}\left(\sum_{t=1}^{T}\mathbf{x}_t\mathbf{x}_t'\right)\right]^{-1}.$$

Since this is the asymptotic variance of $\sqrt{T}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0)$, to get the variance of $\hat{\boldsymbol{\beta}}$ we divide by T. This is the identical matrix as when we calculate White standard errors.

h.)

$$T\mathbf{q}(\hat{\boldsymbol{\theta}})'\left[\mathbf{G}(\hat{\boldsymbol{\theta}})\hat{\mathbf{V}}\mathbf{G}(\hat{\boldsymbol{\theta}})'\right]^{-1}\mathbf{q}(\hat{\boldsymbol{\theta}}) \xrightarrow{p} \chi^2(m)$$

$$\mathbf{G}(\hat{\boldsymbol{\theta}}) = \frac{\partial \mathbf{q}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\bigg|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}}$$

i.) Minimize $Q(\boldsymbol{\theta}) = T[\mathbf{g}(\boldsymbol{\theta};\mathbf{Y}_T)]'\hat{\mathbf{S}}^{-1}[\mathbf{g}(\boldsymbol{\theta};\mathbf{Y}_T)]$ with respect to $\boldsymbol{\theta}$ for $\hat{\mathbf{S}}$ as given in (e) subject to the constraint that $\mathbf{q}(\boldsymbol{\theta}) = \mathbf{0}$. Compare the minimized value of $Q(\hat{\boldsymbol{\theta}}_{\text{restricted}})$ with $\chi^2(m)$.