240A Solutions to Final Exam Winter 2007

1.(a) The usual $\widehat{\gamma} = [\mathbf{Z}'\mathbf{Z}]^{-1}\mathbf{Z}'\mathbf{y}$ with $\mathbf{Z} = \mathbf{XP}$ implies

$$\widehat{\gamma} = [(\mathbf{XP})'(\mathbf{XP})]^{-1}(\mathbf{XP})'\mathbf{y} = [\mathbf{P}'\mathbf{X}'\mathbf{XP}]^{-1}\mathbf{P}'\mathbf{X}'\mathbf{y} = \mathbf{P}^{-1}(\mathbf{X}'\mathbf{X})^{-1}(\mathbf{P}')^{-1}\mathbf{P}'\mathbf{X}'\mathbf{y} = \mathbf{P}^{-1}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}.$$

So $\widehat{\gamma} = \mathbf{P}^{-1}\widehat{\boldsymbol{\beta}}$.

(b) We have

$$\widehat{\mathbf{v}} = \mathbf{y} - \mathbf{Z}\widehat{\gamma} = \mathbf{y} - (\mathbf{XP})\mathbf{P}^{-1}\widehat{\boldsymbol{\beta}} = \mathbf{y} - \widehat{\boldsymbol{\beta}}.$$

The two residuals are identical.

(c) Here

$$\mathbf{P} = \begin{bmatrix} \mathbf{I}_{k-1} & 0 \\ 0 & 100 \end{bmatrix} \quad \Rightarrow \quad \mathbf{P}^{-1} = \begin{bmatrix} \mathbf{I}_{k-1} & 0 \\ 0 & \frac{1}{100} \end{bmatrix}.$$

So

$$\widehat{\gamma} = \begin{bmatrix} \widehat{\gamma}_{k-1} \\ \widehat{\gamma}_k \end{bmatrix} = \mathbf{P}^{-1}\widehat{\boldsymbol{\beta}} = \begin{bmatrix} \mathbf{I}_{k-1} & 0 \\ 0 & \frac{1}{100} \end{bmatrix} \begin{bmatrix} \widehat{\boldsymbol{\beta}}_{k-1} \\ \widehat{\beta}_k \end{bmatrix} = \begin{bmatrix} \widehat{\boldsymbol{\beta}}_{k-1} \\ \frac{1}{100}\widehat{\beta}_k \end{bmatrix}.$$

The first $k-1$ OLS coefficients are the same and the k^{th} is divided by 100.

2.(a) $s^2 = \widehat{\mathbf{u}}'\widehat{\mathbf{u}}/(N-k)$, so $\bar{R}^2 = 1 - s^2/\left[\sum_i (y_i - \bar{y})^2/(N-1)\right]$.
Since $\sum_i (y_i - \bar{y})^2/(N-1)$ does not change across models, $\bar{R}^2$ will not change across models if s^2 does not change.

(b) Let $s_u^2 = \widehat{\mathbf{u}}'\widehat{\mathbf{u}}/(N-k)$ be s^2 in the unrestricted model and $s_r^2 = \widetilde{\mathbf{u}}'\widetilde{\mathbf{u}}/(N-k+q)$ be s^2 in the restricted model, where we add q since q restrictions mean q more degrees of freedom.
Then since $s_u^2 = s_r^2 = s^2$ we have

$$F = \frac{(\widehat{\mathbf{u}}'\widehat{\mathbf{u}} - \widetilde{\mathbf{u}}'\widetilde{\mathbf{u}})/q}{\widehat{\mathbf{u}}'\widehat{\mathbf{u}}/(N-k)} = \frac{((N-k+q)s^2 - (N-k)s^2)/q}{s^2} = \frac{(qs^2)/q}{s^2} = 1.$$

(c) Conclude that if there is no change in $\bar{R}^2$ then $F = 1$.
[Note that $F = 1$ leads to non-rejection of restrictions at conventional levels of significance].

3.(a) Since $d1 + d2 + d3 = 1$ implies $d2 = 1 - d1 - d3$ we have

$$\begin{aligned} y &= \alpha + \beta x + \gamma d2 + \phi d3 + u \\ &= \alpha + \beta x + \gamma(1 - d1 - d3) + \phi d3 + u \\ &= (\alpha + \gamma) + \beta x - \gamma d1 + (\phi - \gamma)d3 + u \end{aligned}$$

so $\widehat{a} = \widehat{\alpha} + \widehat{\gamma}$, $\widehat{b} = \widehat{\beta}$, $\widehat{c} = -\widehat{\gamma}$, and $\widehat{d} = \widehat{\phi} - \widehat{\gamma}$.

(b) We have $\mathbf{R}\boldsymbol{\beta} = r$ where $\mathbf{R} = \begin{bmatrix} 2 & -1 \end{bmatrix}$ and $r = 0$. Then

$$\begin{aligned} \mathrm{W} &= (\mathbf{R}\widehat{\boldsymbol{\beta}} - r)' \, (\mathbf{R}\widehat{\mathrm{V}}[\widehat{\boldsymbol{\beta}}]\mathbf{R})^{-1}(\mathbf{R}\widehat{\boldsymbol{\beta}} - r) \\ &= 5 \times \left(\begin{bmatrix} 2 & -1 \end{bmatrix} \begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \end{bmatrix} \right)^{-1} \times 5 \\ &= 5 \times \left(\begin{bmatrix} 5 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \end{bmatrix} \right)^{-1} \times 5 = 5 \times (10)^{-1} \times 5 = 2.5. \end{aligned}$$

Since W$= 2.5 < \chi^2_{.05;1} = z^2_{.025} = 1.96^2 = 3.84$ we do not reject H_0.
Or can use $\sqrt{W} = t = \sqrt{2.5} < z_{0.05} = 1.96$ and we do not reject H_0.

(c) Assume instruments $\mathbf{Z}$ such that $\mathrm{E}[\mathbf{u}|\mathbf{Z}]=\mathbf{0}$ and use the IV estimator $\widehat{\boldsymbol{\beta}}_{\mathrm{IV}}=(\mathbf{Z}'\mathbf{X})^{-1}\mathbf{Z}'\mathbf{y}$.

(d) $\ln L(\boldsymbol{\beta})=\ln f(\mathbf{y})=-\frac{N}{2}\ln 2\pi-\frac{1}{2}\ln|\boldsymbol{\Sigma}|-\frac{1}{2}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})'\boldsymbol{\Sigma}^{-1}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})$. So

$$\frac{\partial \ln L(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}}=-\frac{1}{2}\left[\frac{\partial}{\partial \boldsymbol{\beta}}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})'\boldsymbol{\Sigma}^{-1}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})\right]=-\frac{1}{2}\left[-2\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{y}+2\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X}\boldsymbol{\beta}\right].$$

Setting to zero: $\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{y}-\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X}\boldsymbol{\beta}=\mathbf{0}\Rightarrow\widehat{\boldsymbol{\beta}}=(\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{y}$.

$$\widehat{\mathrm{V}}[\widehat{\boldsymbol{\beta}}]=-\left\{\mathrm{E}\left[\frac{\partial^2 \ln L(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}\partial \boldsymbol{\beta}'}\right]\right\}^{-1}=-\left\{\mathrm{E}\left[-\frac{1}{2}2\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X}\right]\right\}^{-1}=(\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}.$$

(e) FGLS is OLS of $(y_i/\sqrt{z_i})$ on $(\mathbf{x}_i/\sqrt{z_i})$.

(f) Use Cochrane-Orcutt.
OLS of y_t on $\mathbf{x}_t$ gives $\widehat{u}_t$.
OLS of $\widehat{u}_t$ on $\widehat{u}_{t-1}$ gives $\widehat{\rho}$ where AR(1) error process is $u_t=u_{t-1}+\varepsilon_t$.
OLS of $y_t-\widehat{\rho}y_{t-1}$ on $\mathbf{x}_t-\widehat{\rho}\mathbf{x}_{t-1}$ gives asymptotically efficient $\widehat{\boldsymbol{\beta}}$.

4.(a) We have $\widehat{\boldsymbol{\beta}}=(\mathbf{Z}'\mathbf{A}\mathbf{X})^{-1}\mathbf{Z}'\mathbf{A}(\mathbf{X}\boldsymbol{\beta}+\mathbf{u})=\boldsymbol{\beta}+(\mathbf{Z}'\mathbf{A}\mathbf{X})^{-1}\mathbf{Z}'\mathbf{A}\mathbf{u}$.
So $\mathrm{E}[\widehat{\boldsymbol{\beta}}]=\boldsymbol{\beta}+\mathrm{E}[(\mathbf{Z}'\mathbf{A}\mathbf{X})^{-1}\mathbf{Z}'\mathbf{A}\mathbf{u}]=\boldsymbol{\beta}+(\mathbf{Z}'\mathbf{A}\mathbf{X})^{-1}\mathbf{Z}'\mathbf{A}\mathrm{E}[\mathbf{u}]=\boldsymbol{\beta}$, as $\mathrm{E}[\mathbf{u}]=\mathbf{0}$.
And $\mathrm{V}[\widehat{\boldsymbol{\beta}}]=\mathrm{E}[(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta})(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta})']=\mathrm{E}[((\mathbf{Z}'\mathbf{A}\mathbf{X})^{-1}\mathbf{Z}'\mathbf{A}\mathbf{u})\times((\mathbf{Z}'\mathbf{A}\mathbf{X})^{-1}\mathbf{Z}'\mathbf{A}\mathbf{u})']$
$=(\mathbf{Z}'\mathbf{A}\mathbf{X})^{-1}\mathbf{Z}'\mathbf{A}\mathrm{E}[\mathbf{u}\mathbf{u}']\mathbf{Z}\mathbf{A}(\mathbf{X}'\mathbf{A}\mathbf{Z})^{-1}=(\mathbf{Z}'\mathbf{A}\mathbf{X})^{-1}\mathbf{Z}'\mathbf{A}\boldsymbol{\Sigma}\mathbf{A}\mathbf{Z}(\mathbf{X}'\mathbf{A}\mathbf{Z})^{-1}$.

(b) We have

$$\begin{aligned}\widehat{\boldsymbol{\beta}} &= \boldsymbol{\beta}+\left(N^{-1}\mathbf{Z}'\mathbf{A}\mathbf{X}\right)^{-1}N^{-1}\mathbf{Z}'\mathbf{A}\mathbf{u}\\ &\xrightarrow{p}\boldsymbol{\beta}+\left(\mathrm{plim}\,N^{-1}\mathbf{Z}'\mathbf{A}\mathbf{X}\right)^{-1}\mathrm{plim}\,N^{-1}\mathbf{Z}'\mathbf{A}\mathbf{u}\\ &\xrightarrow{p}\boldsymbol{\beta}\text{ if first plim is finite and second is zero.}\end{aligned}$$

Second is expected to be zero as if a LLN can be applied then

$$\mathrm{plim}\,N^{-1}\mathbf{Z}'\mathbf{A}\mathbf{u}=\lim N^{-1}\mathrm{E}[\mathbf{Z}'\mathbf{A}\mathbf{u}]=\mathbf{0}$$

as $\mathrm{E}_{\mathbf{Z},\mathbf{u}}[\mathbf{Z}'\mathbf{A}\mathbf{u}]=\mathrm{E}_{\mathbf{Z}}[\mathrm{E}_{\mathbf{u}|\mathbf{Z}}[\mathbf{Z}'\mathbf{A}\mathbf{u}|\mathbf{Z}]=\mathrm{E}_{\mathbf{Z}}[\mathbf{Z}'\mathbf{A}\times\mathbf{0}]$ since $\mathrm{E}_{\mathbf{u}|\mathbf{Z}}[\mathbf{u}|\mathbf{Z}]=\mathbf{0}$ is given and $\mathbf{A}$ is nonstochastic.

(c) We have

$$\begin{aligned}\sqrt{N}(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}) &= \left(N^{-1}\mathbf{Z}'\mathbf{A}\mathbf{X}\right)^{-1}\frac{1}{\sqrt{N}}\mathbf{Z}'\mathbf{A}\mathbf{u}\\ &\xrightarrow{d}\left(\mathrm{plim}\,N^{-1}\mathbf{Z}'\mathbf{A}\mathbf{X}\right)^{-1}\times\mathcal{N}[\mathbf{0},\ \mathrm{plim}\,N^{-1}\mathbf{Z}'\mathbf{A}\boldsymbol{\Sigma}\mathbf{A}\mathbf{Z}]\\ &\xrightarrow{d}\mathcal{N}\left[\mathbf{0},\ \left(\mathrm{plim}\,N^{-1}\mathbf{X}'\mathbf{A}\mathbf{X}\right)^{-1}\ \mathrm{plim}\,N^{-1}\mathbf{Z}'\mathbf{A}\boldsymbol{\Sigma}\mathbf{A}\mathbf{Z}\left(\mathrm{plim}\,N^{-1}\mathbf{X}'\mathbf{A}\mathbf{Z}\right)^{-1}\right]\end{aligned}$$

where we use $\frac{1}{\sqrt{N}}\mathbf{Z}'\mathbf{A}\mathbf{u}$ has mean $\mathbf{0}$ and variance $\mathrm{E}_{\mathbf{Z}}[\frac{1}{N}\mathbf{Z}'\mathbf{A}\boldsymbol{\Sigma}\mathbf{A}\mathbf{Z}]$ and assume that a CLT can be applied so that $\mathrm{V}[\frac{1}{\sqrt{N}}\mathbf{Z}'\mathbf{A}\mathbf{u}]^{-1/2}\frac{1}{\sqrt{N}}\mathbf{Z}'\mathbf{A}\mathbf{u}\xrightarrow{d}\mathcal{N}[\mathbf{0},\ \mathbf{I}]$.

(d) Use

$$\mathrm{V}[\widehat{\boldsymbol{\beta}}]=\mathcal{N}[\boldsymbol{\beta},(\mathbf{Z}'\mathbf{A}\mathbf{X})^{-1}\mathbf{Z}'\mathbf{A}\mathrm{Diag}[\widehat{u}_i^2]\mathbf{A}\mathbf{Z}(\mathbf{X}'\mathbf{A}\mathbf{Z})^{-1}],\ \text{where }\widehat{u}_i=y_i-\mathbf{x}_i'\widehat{\boldsymbol{\beta}}.$$

5.(a) avg_ed, pct_el, yr_rnd, and pct_emer are statistically significance at 5 percent.
All have expected sign (though for yr_rnd it is perhaps not clear a priori.)

[Aside: Year-round schools are intended to give better results due to shorter summer break, though there are other reasons for believing the opposite effect may hold.]

(b) Test $H_0 : \beta_{\text{avg\_ed}} < 50$ against $H_0 : \beta_{\text{avg\_ed}} \geq 50$.
$t = (73.8491 - 50)/1.87 = 12.75 > z_{.05} = 1.645$. So reject H_0.
Conclude that one more year of parent education is associated with a more than 50 point rise in school API.

(c) The full model as $\bar{R}^2$ is higher.

(d) After doing the full regression give command

```
test pct_meal pct_el yr_rnd pct_cred pct_emer
```

(e) Yes. The heteroskedasticity test clearly rejects the null hypothesis of homoskedastic errors. [Though in output not give it actually turns out that for these data `, robust` makes little difference].

(f) By far the most important determinant is parental education.
R^2 is already 0.835 with just this as a regressor, rising somewhat to 0.853 when all the other regressors are included.

(g) This gives $\widehat{\mathrm{E}}[\mathbf{y}|\mathbf{x} = \mathbf{x}_f]$, the predicted value of $\mathrm{E}[\mathbf{y}|\mathbf{x} = \mathbf{x}_f]$ and its standard error, where prediction is at `avg_ed=12,pct_meal=20, pct_el=20, yr_rnd=0.03, pct_cred=80, pct_emer=10.`
[It also gives $\widehat{\mathbf{y}}|\mathbf{x} = \mathbf{x}_f$ but not its standard error since that additionally needs to allow for estimating the error by zero].

	Exam / 50			Exam / 50		
$75th$ percentile	40	(80%)			B+	29 and above
Median	36	(72%)	A	39 and above	B	24 and above
$25th$ percentile	34	(68%)	A-	34 and above		

Comments

Question 1 was done poorly even though very straightforward.

Question 2 was done poorly. Many missed that degrees of freedom differ in the two models, since there were q restrictions. Only one person got the conclusion that if $\bar{R}^2$ does not change then $F = 1$ (and hence if one restriction $t = 1$). This means that choosing a larger model because $\bar{R}^2$ increases will lead to larger models than if we test at critical value 0.05. This question should be answerable given good knowledge of undergraduate econometrics (and did not use matrix algebra).

Question 3 was done well, aside from part (a) and in part (b) there were many errors though correct basic approach.

Question 4 was done well. In part (d) there was confusion about what exactly "White" standard errors do.

Question 5 was done well aside from part (g), though note that a "fundamental" conclusion is not many lines long with several conclusions.