BGPE Frontiers in Econometrics 2008

A. Colin Cameron, U.C.-Davis

Solutions to Final Exam

1.(a) We have $\text{E}[\widetilde{\widehat{\boldsymbol{\beta}}}] = \text{E}[\boldsymbol{\beta} + (\mathbf{Z}'\mathbf{A}\mathbf{X})^{-1}\mathbf{Z}'\mathbf{A}\mathbf{u}] = \boldsymbol{\beta} + (\mathbf{Z}'\mathbf{A}\mathbf{X})^{-1}\mathbf{Z}'\mathbf{A}\text{E}[\mathbf{u}] = \boldsymbol{\beta}$ using $\mathbf{Z}$, $\mathbf{A}$ and nonstochastic and $\text{E}[\mathbf{u}] = \mathbf{0}$.
And $\text{V}[\widehat{\boldsymbol{\beta}}] = \text{E}[(\widehat{\boldsymbol{\beta}} - \text{E}[\widehat{\boldsymbol{\beta}}])(\widehat{\boldsymbol{\beta}} - \text{E}[\widehat{\boldsymbol{\beta}}])'] = \text{E}[(\mathbf{Z}'\mathbf{A}\mathbf{X})^{-1}\mathbf{Z}'\mathbf{A}\mathbf{u}((\mathbf{Z}'\mathbf{A}\mathbf{X})^{-1}\mathbf{Z}'\mathbf{A}\mathbf{u})']$
$= (\mathbf{Z}'\mathbf{A}\mathbf{X})^{-1}\mathbf{Z}'\mathbf{A}\Sigma\mathbf{A}\mathbf{Z}(\mathbf{X}'\mathbf{A}\mathbf{Z})^{-1}$.

(b) 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

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

assuming first plim is finite non-zero and second is zero (which essentially requires $\text{E}[\mathbf{u}|\mathbf{Z}] = \mathbf{0}$).

(c) We have

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

where it is assumed that relevant LLN and CLT can be applied, where

$$\mathbf{B} = \lim \text{V}\left[\frac{1}{\sqrt{N}}\mathbf{Z}'\mathbf{A}\mathbf{u}\right] = \lim \text{E}\left[N^{-1}\mathbf{Z}'\mathbf{A}\mathbf{u}\mathbf{u}'\mathbf{A}\mathbf{Z}\right] = \lim N^{-1}\mathbf{Z}'\mathbf{A}\Sigma\mathbf{A}\mathbf{Z}.$$

(d) Given heteroskedastic errors the White approach can be applied so

$$\text{plim}\, N^{-1}\mathbf{Z}'\mathbf{A}\text{Diag}[\widehat{u}_i^2]\mathbf{A}\mathbf{Z} = \lim N^{-1}\mathbf{Z}'\mathbf{A}\Sigma\mathbf{A}\mathbf{Z}, \text{ where } \widehat{u}_i = y_i - \mathbf{x}_i'\widehat{\boldsymbol{\beta}},$$

and

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

(e) Given $\mathbf{A}$ is diagonal symmetric $\mathbf{A}^{1/2} = \text{Diag}[\sqrt{a_{ii}}]$ satisfies $\mathbf{A}^{1/2}\mathbf{A}^{1/2} = \mathbf{A}$, so

$$\begin{aligned}
\widehat{\boldsymbol{\beta}} &= [\mathbf{X}'\mathbf{A}\mathbf{X}]^{-1}\mathbf{X}'\mathbf{A}\mathbf{y} = [(\mathbf{X}'\mathbf{A}^{1/2})(\mathbf{A}^{1/2}\mathbf{X})]^{-1}(\mathbf{X}'\mathbf{A}^{1/2})(\mathbf{A}^{1/2}\mathbf{y}) \\
&= [(\mathbf{A}^{1/2}\mathbf{X})'(\mathbf{A}^{1/2}\mathbf{X})]^{-1}(\mathbf{A}^{1/2}\mathbf{X})'(\mathbf{A}^{1/2}\mathbf{y})
\end{aligned}$$

which is OLS of $\mathbf{A}^{1/2}\mathbf{y}$ on $\mathbf{A}^{1/2}\mathbf{X}$, or OLS of $\sqrt{a_{ii}}y_i$ on $\sqrt{a_{ii}}\mathbf{x}_i$.

2.(a) Here

$$\begin{aligned} \ln f(y) &= 2\ln\lambda + \ln y - \lambda y \text{ and } \lambda = 2\exp(-\mathbf{x}'\boldsymbol{\beta}) \text{ and } \ln\lambda = \ln 2 - \mathbf{x}'\boldsymbol{\beta} \\ &\Rightarrow \ln L(\boldsymbol{\beta}) = \sum\nolimits_i \ln f(y_i) = \sum\nolimits_i \{2\ln 2 - 2\mathbf{x}_i'\boldsymbol{\beta} + \ln y_i - 2\exp(-\mathbf{x}'\boldsymbol{\beta})y_i\}. \end{aligned}$$

(b) Here

$$\begin{aligned} \frac{\partial \ln L(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} &= \sum\nolimits_i (-2\mathbf{x}_i + 2y_i\exp(-\mathbf{x}_i'\boldsymbol{\beta})\mathbf{x}_i) \\ &= \sum\nolimits_i 2\times\{y_i\exp(-\mathbf{x}_i'\boldsymbol{\beta}) - 1\}\mathbf{x}_i \quad \text{rearranging} \\ &= \sum\nolimits_i 2\times\frac{y_i - \exp(\mathbf{x}_i'\boldsymbol{\beta})}{\exp(\mathbf{x}_i'\boldsymbol{\beta})}\mathbf{x}_i \quad \text{multiplying by } \frac{\exp(\mathbf{x}_i'\boldsymbol{\beta})}{\exp(\mathbf{x}_i'\boldsymbol{\beta})} \end{aligned}$$

(c) Essential condition is $\mathrm{E}[y_i|\mathbf{x}_i] = \exp(\mathbf{x}_i'\boldsymbol{\beta}_0)$ as then $\mathrm{E}\left[\frac{y_i-\exp(\mathbf{x}_i'\boldsymbol{\beta})}{\exp(\mathbf{x}_i'\boldsymbol{\beta})}\mathbf{x}_i\right] = \mathbf{0}$ at $\boldsymbol{\beta} = \boldsymbol{\beta}_0$ so then $\mathrm{E}\left[\left.\frac{\partial \ln L(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}}\right|_{\boldsymbol{\beta}_0}\right] = \mathbf{0}$.

(d) For the MLE

$$\sqrt{N}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0) \xrightarrow{d} \mathcal{N}[\mathbf{0}, \mathbf{A}_0^{-1}]$$

where $\mathbf{A}_0 = -\lim \mathrm{E}\left[\left.\frac{1}{N}\frac{\partial^2 \ln L(\boldsymbol{\beta})}{\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}'}\right|_{\boldsymbol{\beta}_0}\right]$ where

$$\begin{aligned} \frac{\partial^2 \ln L(\boldsymbol{\beta})}{\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}'} &= \sum\nolimits_i -2y_i\exp(-\mathbf{x}_i'\boldsymbol{\beta})\mathbf{x}_i\mathbf{x}_i' \\ \mathrm{E}\left[\left.\frac{\partial^2 \ln L(\boldsymbol{\beta})}{\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}'}\right|_{\boldsymbol{\beta}_0}\right] &= \sum\nolimits_i -2\exp(\mathbf{x}_i'\boldsymbol{\beta}_0)\exp(-\mathbf{x}_i'\boldsymbol{\beta}_0)\mathbf{x}_i\mathbf{x}_i' = -\sum\nolimits_i 2\mathbf{x}_i\mathbf{x}_i' \end{aligned}$$

so

$$\sqrt{N}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0) \xrightarrow{d} \mathcal{N}\left[\mathbf{0}, \left(\lim 2N^{-1}\sum\nolimits_i \mathbf{x}_i\mathbf{x}_i'\right)^{-1}\right].$$

(e) Use Newton-Raphson

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

and can cancel the 2's. Or can take expected value of the hessian and use

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

3. **(a)** Here u_{it} is causing problems, so do instrumental variables of y_{it} on $\mathbf{x}_{it}$ with instruments $\mathbf{z}_{it}$. Here $\mathbf{z}_{it}$ should be uncorrelated with u_{it}(and α_i) and correlated with α_i.

(b) Actually same solution as part (a). Or can first diefference to get rid of α_i and then do IV.

(c) Estimation is based on $\mathrm{E}[\mathbf{z}_i(y_i - \Lambda(\mathbf{x}_i'\boldsymbol{\beta}))] = \mathbf{0}$.
If just-identified the estimator solves $\sum_i \mathbf{z}_i(y_i - \Lambda(\mathbf{x}_i'\boldsymbol{\beta})) = \mathbf{0}$.
If over-identified GMM estimator minimizes $[\sum_i \mathbf{z}_i(y_i - \Lambda(\mathbf{x}_i'\boldsymbol{\beta}))]' \, \mathbf{W} \, [\sum_i \mathbf{z}_i(y_i - \Lambda(\mathbf{x}_i'\boldsymbol{\beta}))]$.

(d) Density for i^{th} observation is $F_1(\mathbf{x}_i'\boldsymbol{\beta}_1)^{y_{1i}} \times F_2(\mathbf{x}_i'\boldsymbol{\beta}_2)^{y_{2i}} \times F_3(\mathbf{x}_i'\boldsymbol{\beta}_3)^{y_{3i}}$.
Log-likelihood is $\sum_i [y_{1i} \ln F_1(\mathbf{x}_i'\boldsymbol{\beta}_1) + y_{2i} \ln F_2(\mathbf{x}_i'\boldsymbol{\beta}_2) + y_{3i} \ln F_3(\mathbf{x}_i'\boldsymbol{\beta}_3)]$.

(e) For nonlinear estimators the oprimal estimator has smallest asymptotic variance matrix among all consistent estimators.

4.(a) For logit, sign of coefficient gives sign of the marginal effect.
Expect `fairpoor` down with increase in `income`; up with increase in `age`; up with increase in `chronic`.
So the signs of the coefficients are as expected.

(b) Yes. Individually at level 0.05 as $p < 0.05$ using either one-sided or two-sided test. And jointly as overall Wald test (`chi(2)`) has $p = 0.00$.

(c) A $1,000 increase in income is a one-unit increase in income.
Using $\partial \ln \Pr[y = 1|\mathbf{x}]/\partial \mathbf{x} = \Lambda(\mathbf{x}'\boldsymbol{\beta})(1-\Lambda(\mathbf{x}'\boldsymbol{\beta}))\boldsymbol{\beta}$ evaluated at $\Lambda(\mathbf{x}'\boldsymbol{\beta}) = \bar{y}$ we have that probability of `fairpoor` falls by $.097 \times .903 \times .0236 = 0.0021$.
Or use rule of thumb. Falls by $0.025 \times .0236 = 0.0059$.
Or log-odds ratio falls by $-.0236$. Or odds-ratio is $e^{-.0236} = 0.978$ times that without income change.

(d) Consistency requires that $\Pr[y_i = 1|\mathbf{x}_i] = \exp(\mathbf{x}_i'\boldsymbol{\beta})/[1 + \exp(\mathbf{x}_i'\boldsymbol{\beta})]$.
Efficiency requires correct density. Here density for each observations is necessarily Bernoulli, but we do additionally require independence over i for correct joint density of all observations.

(e) Difference is in the standard errors. The robust se's here are actually quite different for `income`, increasing from 0.0029 to 0.0037, a more than 30% increase.

(f) The model fit of logit is better than probit as the log-likelihood is $(-1249.3 - -1253.9) = 4.6$ higher for logit. [This is a large difference actually, as a likelihood ratio test with one degree of freedom favors the more general model at level 0.05 if critical value is 3.84 or likelihood ratio change exceeds $3.84/2 = 1.92$.]
Also using the pseudo-R^2 logit is favored as has higher pseudo-R^2.