Econ 7010 - Final Exam 2016 - Answer Key

Part A - Short Answer

2.) For the strong consistency of an estimator, sufficient conditions are for the bias of the estimator and the variance of the estimator to go to zero as the sample size goes to infinity.

Since s^2 is unbiased, we just have to check its variance in the limit:

$$\lim_{n \to \infty} \frac{2\sigma^4}{n-k} \to 0$$

Now, $\hat{\sigma}^2$ is biased, but the bias disappears as n increases:

$$\lim_{n \to \infty} \frac{n-k}{n} \sigma^2 \to \sigma^2,$$

and variance goes to zero:

$$\lim_{n \to \infty} \frac{2\sigma^4}{n} \to 0,$$

so that both estimators are consistent.

3.) Maximum Likelihood Estimators are BAN (best asymptotically normal). That is, they are consistent, asymptotically efficient, and asymptotically normally distributed. This is all obtained provided that the likelihood function has been specified correctly. In addition, MLEs also have an invariance property.

4.) i) $R = \begin{bmatrix} 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$; $q = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

ii) Since ε is not Normal, the F-statistic would not follow the F-distribution. The distribution of the F-statistic would depend on the distribution of ε. Instead of the F-test, the Wald test could be used. The Wald test would be asymptotically chi-square distributed, meaning that the chi-square distribution would be used as an approximation for this testing problem.

5.) See Topic 1, pg. 11.

6.) There were two situations discussed in class where the algorithm may fail to converge: when the Hessian is singular, and when the algorithm cycles (see pg. 21 and 22 of "New Slides"). In addition, the algorithm may find a local min or a max, instead of the global min.

7.) See many other answer keys.

8.) A p-value is the probability of calculating a test statistic that is more adverse to the null hypothesis than the one just calculated, provided the null hypothesis is true. If a p-val is low, this means that a rare occurence has been witnessed, provided the null is true. Alternatively, the null is false and should be rejected.

PART B

1. a) $y^* = \begin{bmatrix} y \\ \hdashline q \end{bmatrix}$; $X^* = \begin{bmatrix} X \\ \hdashline R \end{bmatrix}$; $\varepsilon^* = \begin{bmatrix} \varepsilon \\ \hdashline v \end{bmatrix}$

This has been obtained from the two equations:

$$y = X\beta + \varepsilon$$
$$q = R\beta + v$$

b) $X^{*\prime}X^* = [X' \vdots R'] \begin{bmatrix} X \\ \hdashline R \end{bmatrix} = [X'X + R'R]$

$X^{*\prime}y = [X' \vdots R'] \begin{bmatrix} y \\ \hdashline q \end{bmatrix} = X'y + R'q$

$\hat{\beta} = (X^{*\prime}X^*)^{-1}X^{*\prime}y^* = [X'X + R'R]^{-1}[X'y + R'q]$

c) $E[\hat{\beta}] = [X'X + R'R]^{-1}[X'(X\beta + E(\varepsilon)) + R'(R\beta + E(v))]$
$= [X'X + R'R]^{-1}[X'X\beta + R'R\beta] = [X'X + R'R]^{-1}[X'X + R'R]\beta$
$= \beta$ (unbiased)

d) $\hat{\beta} = [X'X + R'R]^{-1}X'X\beta + [X'X + R'R]^{-1}R'R\beta + [X'X + R'R]^{-1}X'\varepsilon + [X'X + R'R]^{-1}R'v$

$$V(\hat{\beta}) = [X'X + R'R]^{-1}X'\sigma^2 I X[X'X + R'R]^{-1} + [X'X + R'R]^{-1}R'VR[X'X + R'R]^{-1}$$

$$= [X'X + R'R]^{-1}(\sigma^2 X'X + R'VR)[X'X + R'R]^{-1}$$

e) If $V = \sigma^2 I$, then:

$$V(\hat{\beta}) = [X'X + R'R]^{-1}(\sigma^2[X'X + R'R])[X'X + R'R]^{-1}$$

$$= \sigma^2[X'X + R'R]^{-1}$$

f) We would have to prove that $V(b) - V(\hat{\beta}) = \sigma^2[(X'X)^{-1} - (X'X + R'R)^{-1}]$ is positive definite. Since R has full rank, $R'R$ is positive definite, and $\hat{\beta}$ is more efficient than b.

g) Use $t_{n-k} = \dfrac{\hat{\beta}_i - \beta_0}{\sqrt{s^2(X'X + R'R)^{-1}_{ii}}}$, where $s^2 = (y - X\hat{\beta})'(y - X\hat{\beta})/(n-k)$.

Part C

3. a) $\bar{\varepsilon}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} \varepsilon_j$

$$\text{var}(\bar{\varepsilon}_i) = \frac{1}{n_i^2} n_i \, \text{var}(\varepsilon_j) = \frac{\sigma^2}{n_i}$$

Other than being heteroskedastic, $\bar{\varepsilon}_i$ satisfies all of the usual assumptions.

b) For this problem, the Ω matrix, where $V(\varepsilon) = \sigma^2 \Omega$, is:

$$\Omega = \begin{bmatrix} 1/n_1 & & 0 \\ & 1/n_2 & \\ 0 & & \ddots \; 1/n_m \end{bmatrix}$$

The P matrix, where $P'P = \Omega^{-1}$, is:

$$P = \begin{bmatrix} \sqrt{n_1} & & 0 \\ & \sqrt{n_2} & \\ 0 & & \ddots \; \sqrt{n_m} \end{bmatrix}$$

The WLS, or GLS estimator, is obtained by applying OLS to the transformed model:

$$Py = PX\beta + P\varepsilon$$

This estimator is preferable due to its relative efficiency.

c) $P = \begin{bmatrix} 4 & & & 0 \\ & 2 & & \\ & & 3 & \\ 0 & & & 2 \end{bmatrix}$; $y^* = Py = \begin{bmatrix} 12 \\ 4 \\ 15 \\ 2 \end{bmatrix}$; $X^* = PX = \begin{bmatrix} 4 & 8 \\ 2 & 6 \\ 3 & 3 \\ 2 & 8 \end{bmatrix}$

$$X^{*\prime}X^* = \begin{bmatrix} 4 & 2 & 3 & 2 \\ 8 & 6 & 3 & 8 \end{bmatrix} \begin{bmatrix} 4 & 8 \\ 2 & 6 \\ 3 & 3 \\ 2 & 8 \end{bmatrix} = \begin{bmatrix} 33 & 69 \\ 69 & 173 \end{bmatrix}$$

$$(X^{*\prime}X^*)^{-1} = \begin{bmatrix} 0.1825 & -0.0728 \\ -0.0728 & 0.0348 \end{bmatrix}$$

$$X^{*\prime}y^* = \begin{bmatrix} 105 \\ 181 \end{bmatrix}$$

$$(X^{*\prime}X^*)^{-1}X^{*\prime}y^* = \begin{bmatrix} 5.99 \\ -1.34 \end{bmatrix}$$

d) If group size is unknown, then Ω is unknown and must be estimated. One possibility is to first estimate OLS, and obtain the residuals e_i. The OLS residuals are consistent estimators for ε since OLS is still consistent in the presence of heteroskedasticity. A consistent estimator for Ω is then:

$$\hat{\Omega} = \begin{bmatrix} e_1^2 & & & 0 \\ & e_2^2 & & \\ & & \ddots & \\ 0 & & & e_n^2 \end{bmatrix}$$

The FGLS estimator is then $(X'\hat{\Omega}^{-1}X)^{-1}X'\hat{\Omega}^{-1}y$.

5. a) i) The instruments and the error term must be independent:

$$\text{plim}\left(\frac{Z'\varepsilon}{n}\right) = 0$$

ii) The instruments must be correlated with the endogenous regressors:

$$\text{plim}\left(\frac{Z'X}{n}\right) = Q_{ZX} \quad \text{(a finite matrix)}$$

iii) The $Z'Z$ matrix must be finite in the limit:

$$\text{plim}\left(\frac{Z'Z}{n}\right) = Q_{ZZ} \quad \text{(a finite matrix)}$$

b) The IV estimator is:

$$\hat{\beta}_{IV} = [X'Z(Z'Z)^{-1}Z'X]^{-1}X'Z(Z'Z)^{-1}Z'y$$

$$= \left[\frac{X'Z}{n}\left(\frac{Z'Z}{n}\right)^{-1}\frac{Z'X}{n}\right]^{-1}\frac{X'Z}{n}\left(\frac{Z'Z}{n}\right)^{-1}\frac{Z'y}{n}$$

Using the above assumptions, and Slutsky's theorem:

$$\text{plim}(\hat{\beta}_{IV}) = [Q_{XZ}Q_{ZZ}^{-1}Q_{ZX}]^{-1}Q_{XZ}Q_{ZZ}^{-1}(Q_{ZX}\beta + 0)$$

$$= \beta$$

c) $$\text{plim}(\hat{\beta}_{IV}) = [Q_{XZ}Q_{ZZ}^{-1}Q_{ZX}]^{-1}Q_{XZ}Q_{ZZ}^{-1}(Q_{ZX}\beta + Q_{ZW}\gamma + 0)$$

$$= \beta + [Q_{XZ}Q_{ZZ}^{-1}Q_{ZX}]^{-1}Q_{XZ}Q_{ZZ}^{-1}Q_{ZW}\gamma \neq \beta$$

6. parts (a) and (b) - see "MLE Practice Questions", pg. 1-2.

c) The MLE for λ is:

$$\tilde{\lambda} = \frac{1}{n} \Sigma y_i$$

$$E[\tilde{\lambda}] = \frac{1}{n} \Sigma E(y_i) = \frac{1}{n} n\lambda = \lambda \quad \text{(unbiased)}$$

Note that we are told that the mean and variance of y are equal to λ.

d) A regression model can be obtained by "linking" the mean of each y_i to a set of regressors:

$$\lambda_i = \exp(X_i'\beta),$$

where the exponent of the linear function $X_i'\beta$ is used in order to ensure that λ_i remains positive. The above "link" function is then substituted into the log-likelihood, and maximization takes place with respect to β insted of λ.

e) Not relevant for the 2018 exam.