Econ 7010 - 2015 - Midterm 2 Answer Key

Part A

1. The F-test is an exact test for finite samples, and is U.M.P. The Wald test is an asymptotic test, meaning that the distribution of the Wald test statistic is only approximately chi-square distributed for large samples. However, for large samples, the F-test and Wald test are identical.

   The F-test is preferable since it is U.M.P., however, it can be invalid if the assumptions underlying the F-distribution are violated (e.g. Normality of errors), or in appropriate if the estimation methodology is asymptotic (e.g. IV estimation). In such cases, the Wald test should be used.

2. The Newton algorithm may be written as:

$$\theta_{n+1} = \theta_n - g(\theta_n) / H(\theta_n)$$

where $g(\theta_n) = 3\theta_n^2 - 3$ and $H(\theta_n) = 6\theta_n$.

Starting at $\theta_0 = 2$, we have:

$$\theta_1 = 2 - \frac{3(2)^2 - 3}{6(2)} = 1.25$$

$$\theta_2 = 1.25 - \frac{3(1.25)^2 - 3}{6(1.25)} = 1.025$$

Carrying on in this fashion we will see that θ_n converges to 1. $\theta = 1$ indeed minimizes $f(\theta)$, since $g(1) = 0$ and $H(1) = (+)$.

3.) The NLLS estimator is derived in a way similar to the OLS estimator. We still seek to minimize the residual sum-of-squares $\Sigma(y_i - \hat{y}_i)^2$, however, $\hat{y}_i$ can be the result of a non-linear equation. The optimization problem is:

$$\min_{\theta} \Sigma \left(y_i - f(X, \theta) \right)^2 ,$$

where $f(X,\theta) = \hat{y}_i$, and $f(X,\theta)$ may be a non-linear function in θ.

Since $f(X,\theta)$ is non-linear, the first order conditions from the above optimization problem will also usually be non-linear, and the solution will usually only be obtainable via numerical methods (e.g. Newton's method).

4. An estimator is M.S. consistent if its bias and variance go to $\emptyset$ as $n \to \infty$. This means that the asymptotic variance of two M.S. consistent estimators cannot be directly compared since both are zero.

It turns out that many estimators have an 'n' in the denominator, and that the variance of such estimators also have an 'n' in the denominator. In these cases, two M.S. consistent estimators have variances that are approaching zero at rate n^{-1}. We can 'control' for this by multiplying the estimators by $\sqrt{n}$, and then can compare asymptotic variances in a meaningful way.

5. If X and ε are not independent, then:

$$\text{plim}\left(\frac{X'\varepsilon}{n}\right) \neq 0. \quad (= \gamma, \text{ say})$$

Assuming $\text{plim}\left(\frac{X'X}{n}\right) = Q$:

i) $$\text{plim}(b) = \text{plim}\left(\frac{X'X}{n}\right)^{-1} \text{plim}\left(\frac{X'\varepsilon}{n}\right) + \beta = Q^{-1}\gamma + \beta \neq \beta,$$

and so OLS is inconsistent.

ii) The IV estimator is $(Z'X)^{-1}Z'y$. Assuming that:

$$\text{plim}\left(\frac{Z'X}{n}\right) = Q_{zx} \quad (\text{p.d.})$$

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

$$\begin{aligned}
\text{plim}(b_{IV}) &= \text{plim}\left(\beta + (Z'X)^{-1}Z'\varepsilon\right) \\
&= \beta + \text{plim}\left(\frac{Z'X}{n}\right)^{-1} \text{plim}\left(\frac{Z'\varepsilon}{n}\right) \\
&= \beta + Q_{zx} \cdot 0 = \beta,
\end{aligned}$$

and so IV is consistent.

6. a) $e_* = y - Xb_* = \underbrace{y - Xb}_{e} - X(X'X)^{-1}R'[R(X'X)^{-1}R']^{-1}(Rb-q)$

$$e_*'e_* = e'e - (Rb-q)'[R(X'X)^{-1}R']^{-1}\underbrace{R(X'X)^{-1}\underbrace{X'X(X'X)^{-1}}_{I}R'[R(X'X)^{-1}R']^{-1}}_{I}(Rb-q)$$

(I have used the result $X'e = e'X = 0$)

$$e_*'e_* = e'e - (Rb-q)'[R(X'X)^{-1}R']^{-1}(Rb-q)$$

b) The RLS estimator is found by minimizing $e_*'e_*$. This minimization problem can be interpreted as minimizing $e'e$ subject to constraints. Hence, the minimized value of $e_*'e_*$ cannot be less than that of $e'e$.

c) $e_*'e_*$ will be equal to $e'e$ if the constraints don't matter. That is, if $Rb = q$. That is, if OLS estimates the restrictions exactly. (This is quite different from the restrictions being true.)

d) $y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 X_{4i} + \beta_5 X_{5i} + \varepsilon_i$.

Imposing the constraints $\beta_1 + \beta_2 = 1$ and $\beta_3 = \beta_4$:

$y_i = 1 + \beta_2(X_{2i} - 1) + \beta_3(X_{3i} + X_{4i}) + \beta_5 X_{5i} + \varepsilon_i$, for example.

Or, $y_i - 1 = \beta_2(X_{2i} - 1) + \beta_3(X_{3i} + X_{4i}) + \beta_5 X_{5i} + \varepsilon_i$.

Hence, estimation of the restricted model can be accomplished by subtracting 1 from the y and X_2 data, and adding the X_3 and X_4 data, and applying OLS. Note that there is no longer an intercept.

7. a) $\bar{\varepsilon}_j = 1/n_j \sum_{j=1}^{n_j} \varepsilon_j$

$$var(\bar{\varepsilon}_j) = var(1/n_j \Sigma \varepsilon_j) = 1/n_j^2 \Sigma\, var(\varepsilon_j) = 1/n_j^2\, n_j \sigma^2 = \frac{\sigma^2}{n_j} = \sigma^2/j$$

b) $var(\bar{\varepsilon}) = \sigma^2 \Omega = \sigma^2 \begin{bmatrix} 1 & 0 & 0 & & \\ 0 & 1/2 & 0 & & 0 \\ 0 & 0 & 1/3 & & \\ & 0 & & \ddots & 1/m \end{bmatrix}$

We need to find P such that $P'P = \Omega^{-1}$

$$P = \begin{bmatrix} \sqrt{1} & 0 & 0 & & \\ 0 & \sqrt{2} & 0 & & 0 \\ 0 & 0 & \sqrt{3} & & \\ & 0 & & \ddots & \sqrt{m} \end{bmatrix}$$

Pre-multiplying the y and X data by P, and then applying OLS to the transformed data, amounts to GLS.

c) This is the same as ignoring the consequences of heteroskedasticity in general: (i) estimation of var(b) would be inconsistent and so hypothesis testing would be invalid; (ii) OLS is still unbiased and consistent, however, inefficient.

d) Either: (i) Use OLS and White's het. consistent estimator for var(b), or (ii) apply FGLS by using a proxy variable for n_j or by an iterative procedure to estimate Ω.

8. a) From the formula sheet:

$$F = \frac{(e_*'e_* - e'e)/J}{e'e/(n-k)}$$

Here, $e'e = e_M'e_M + e_F'e_F = 0.06 + 0.05 = 0.11$, and we have $e_*'e_* = 0.15$. Since the two models are estimated separately, there will be a different intercept and two different slope coefficients for each, hence the unrestricted model would contain 6 parameters while the restricted would contain 3 parameters. Therefore, $J=3$ and $K=6$. So,

$$F = \frac{(0.15 - 0.11)/3}{0.11/(1000-6)} = 120.49$$

(This is a large F-stat - we would reject the null.)

b) Instead of dividing the sample into two, we could instead use the dummy variable, and estimate the model:

$$\text{income}_i = \beta_1 + \beta_2 \text{education}_i + \beta_3 \text{age}_i + \beta_4 \text{male}_i + \beta_5 \text{male}_i \times \text{education}_i + \beta_6 \text{male}_i \times \text{age}_i + \varepsilon,$$

and then use an F-test to test the joint restrictions $H_0: \beta_4 = \beta_5 = \beta_6 = 0$.

c) If $education_i$ is correlated with the error term then the OLS estimator for β_2 is inconsistent.

d) In the 2SLS interpretation of IV estimation, we could proceed in the following way:

1. Estimate the following regression: $education_i = \beta_1 + \beta_2 distance_i + u_i$
2. Obtain the fitted values, $\widehat{education}_i$, from 1
3. Estimate the regression: $income_i = \beta_1 + \beta_2 \widehat{education}_i + \beta_3 age_i + \varepsilon_i$

$\widehat{education}_i$ is uncorrelated with ε_i, and the OLS estimator for β_2 is consistent.