Final Exam - Econ 7010 2013 - Answer Key

1.) FALSE

The Chow test is equivalent to testing the joint significance of the coefficients on dummy variables which interact with all X variables, not just the intercept.

2.) Models are "nested" when it is possible to get one of the models by imposing restrictions on the other model. The restricted model must have a higher sum of squared residuals, since we are minimizing subject to a constraint (when compared to the unrestricted model).

3.) We could implement FGLS by transforming the model, and then applying OLS. The transformed model is:

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

where P is based on s_1^2, s_2^2 and s_3^2. P has to be such that $var(P\varepsilon) = \sigma^2 I_n$. One possible choice for P is:

$$P = \begin{array}{cl} \left[\begin{array}{c} 1/s_1 \\ \vdots \\ 1/s_1 \\ \hline 1/s_2 \\ \vdots \\ 1/s_2 \\ \hline 1/s_3 \\ \vdots \\ 1/s_3 \end{array}\right] & \begin{array}{l} \\ \\ n_1 \\ \\ \\ n_2 \\ \\ \\ \\ \end{array} \end{array}$$

4.) The RLS estimator is:

$$b_* = b - (X'X)^{-1}R'[R(X'X)^{-1}R']^{-1}(Rb - q).$$ Taking expectations:

$$E(b_*) = \beta - (X'X)^{-1}R'[R(X'X)^{-1}R']^{-1}(R\beta - q),$$

since $E(b) = \beta$.

Hence, the only way that b_* is unbiased is if the 2nd term in the expression of $E(b_*)$ is equal to zero. This will happen when $R\beta = q$, i.e., when the restrictions we have imposed are true.

5.) (i) We can rephrase this as: wrongly excluding a relevant regressor (the intercept), or falsely imposing a restriction ($\beta_0 = 0$). In this case, OLS is biased and inconsistent, but efficient compared to the model which does include an intercept.

(ii) In this case we have included an irrelevant regressor, or failed to impose a valid restriction. OLS is still unbiased and consistent, but is inefficient compared to the model without an intercept.

6.) No, you do not. A variable, X, may have non-linear effects on y, which can ~~be modelled~~ sometimes be modelled by specifying a model that is still linear in the parameters, but non-linear in the variables. For example, we frequently see models that take the logs of the LHS, RHS, or both. This is a way of modelling non-linear effects.

For the example provided, a suitable population model would be:

$$y = \beta_0 + \beta_1 X + \frac{\beta_2}{2} X^2 + \ldots + \varepsilon$$

This way, $\frac{\partial y}{\partial X} = \beta_1 + \beta_2 X$, as required.

7.) Sometimes, when a model is non-linear in the parameters, we can make it linear by taking a suitable transformation. One example of this which we have seen in class is the Cobb-Douglas production function.

If there is no way to transform the model to make it linear, then we can use NLLS, where we still seek to minimize the sum of squares. This objective function ($e'e$) however, is non-linear. Solving for the NLLS estimator entails taking derivatives of $e'e$, which will likely be non-linear themselves. Hence, the FOC's do not have an analytic solution in this case, and must be solved ~~analytically~~ numerically. The Newton-Raphson algorithm is one possible choice.

The algorithm searches for $\hat{\theta}$ such that $\left.\frac{\partial e'e}{\partial \theta}\right|_{\hat{\theta}} = 0$. It starts with an initial guess (θ_0), and uses the first and second derivatives of $e'e$, evaluated at θ_0, to inform the next guess.

8.) In White's test for heteroskedasticity, the null and alternative hypotheses are:

$H_0: var(\varepsilon) = \sigma^2 I_n$

$H_A: var(\varepsilon) \neq \sigma^2 I_n$

This is a deconstructive test. That is, rejection of the null does not imply the form of heteroskedasticity.

To perform the test, we estimate our model by OLS, obtaining $e'e$. We then seek to explain variation in $e'e$ by regressing it on all X variables, squares of X variables, cross-products, etc. If we are able to explain a significant amount of the variation in $e'e$ in this fashion, we should reject the null. The appropriate test statistic is $nR^2 \sim \chi^2_k$, where R^2 is obtained from the auxilliary regression, and k is the number of parameters included in the auxilliary regression.

9.) $b = (X'X)^{-1}X'y = (X'X)^{-1}X'(X\beta + \varepsilon) = \beta + (X'X)^{-1}X'\varepsilon$

$$= \beta + \left(\frac{X'X}{n}\right)^{-1}\left(\frac{X'\varepsilon}{n}\right)$$

Assuming that $\text{plim}\left(\frac{X'X}{n}\right) = Q$, a finite, p.s.d. matrix,

and that $\text{plim}\left(\frac{X'\varepsilon}{n}\right) = 0$ (the X's and ε's are independent),

and by applying Slutsky's theorem repeatedly, we have;

$$\text{plim}(b) = \beta + Q^{-1} \cdot 0 = \beta,$$

and b is consistent for β.

10.) We can derive the formula for the over-identified b_{IV} estimator by two-stage least squares:

1. Regress the endogenous variable on the set of instruments and obtain $\hat{X}$.

   Fit: $X = Z\gamma + u$, where $\hat{\gamma} = (Z'Z)^{-1}Z'X$.
   Get: $\hat{X} = Z\hat{\gamma} = Z(Z'Z)^{-1}Z'X$.

2. Regress y on $\hat{X}$.

   Fit: $y = \hat{X}\beta + \varepsilon$, where $b_{IV} = (\hat{X}'\hat{X})^{-1}\hat{X}'y$

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

   $$= (X'Z(Z'Z)^{-1}Z'X)^{-1}X'Z(Z'Z)^{-1}Z'y,$$

which is the formula for the IV estimator of β, in the over-identified case.

II.) A.1 i) The assumption is violated if our pop. model is non-linear in the parameters.

ii) This model can't be estimated by OLS.

iii) Some other estimation methodology must be used, such as NLLS, or MLE.

A.2 i) If one or more of the X variables is linearly related to one or more of the other X^s. This most commonly arises through the use of dummy variables (dummy variable trap).

ii) Estimation will fail. The $(X'X)$ matrix can't be inverted, hence b can't be calculated.

iii) In the case of the D.V. trap, one of the dummies must be dropped.

A.3 i) This assumption is likely violated fairly often, however, if an intercept is included in the model, it is as if we have this assumption.

ii) OLS will be biased and inconsistent.

iii) Include an intercept.

A.4 i) When there is heteroskedasticity, autocorrelation, or both, this assumption will be violated.

ii) OLS will still be unbiased and consistent, but will be inefficient. The estimator for cov(b), however, will be inconsistent. This is because we will be assuming the wrong formula for cov(b). F-stats, t-stats, associated p-values will all be wrong.

iii) There are several things we can do.

- We can stick with OLS, but at least estimate cov(b) correctly. This involves using the residuals from OLS. This is called White's heteroskedastic consistent (or robust) standard errors. OLS is still inefficient.

- If the form of var(ε) is known, we can use GLS. This provides an efficient estimator.

- If var(ε) is unknown, as will typically be the case in practice, the unknown form of heteroskedasticity must be estimated. This can be done by using the residuals from OLS. This is called FGLS, or when only het. is present, weighted least squares.

A.5 i) This assumption could be violated, for example, if there are unobservable or omitted ~~effects~~ variables which affect y, and are ~~inclu~~ correlated with the included variables.

ii) Endogeneity will cause a severe effect; the OLS estimators will be inconsistent.

iii) We can use instrumental variables estimation instead. This involves finding suitable instruments - variables which are correlated with the endogenous variables but uncorrelated with the error term. This will recover consistency.

A.6 i) The central limit theorem provides fairly good justification for this assumption. If the CLT fails, however, the ε^s could follow any distribution.

ii) OLS will have it's usual properties, however, t-stats will not be t-distributed and F-stats won't be F-distributed.

iii) We can use asymptotic approximations for the test statistics, e.g. use the z-test instead of t and the Wald test instead of F.

12.) a) $L(\sigma^2 | y) = \frac{1}{\sigma^n (\sqrt{2\pi})^n} e^{-\Sigma \frac{y_i^2}{2\sigma^2}}$

$$\log L = -n \log \sigma - n \log \sqrt{2\pi} - \Sigma \frac{y_i^2}{2\sigma^2}$$

$$\frac{\partial \log L}{\partial \sigma} = \frac{-n}{\sigma} + 2\Sigma \frac{y_i^2}{2\sigma^3} = 0$$

$$\hookrightarrow \hat{\sigma}^2 = \frac{\Sigma y_i^2}{n}$$

b) It would be biased if we had to estimate μ as well, but we don't. There is no degrees of freedom distortion here. It is unbiased.

c) We can use the invariance property of MLE^s to get:

$$\hat{\sigma}^4 = (\hat{\sigma}^2)^2$$

d) The LM test starts with the fact that the FOC^s for the unrestricted model, when evaluated at the MLE^s, will be equal to zero by definition. If the restrictions are true, then evaluating the same ~~[illegible]~~ FOC^s ~~[illegible]~~ at the restricted MLE^s will also give zero. If the restrictions are false, however, the FOC^s will not be satisfied, and we will reject the restrictions.

13.)

a) $\hat{\beta}_1 = (X_1'\Omega^{-1}X_1)^{-1}X_1'\Omega^{-1}y$

$var(\hat{\beta}_1) = (X_1'\Omega^{-1}X_1)^{-1}X_1'\Omega^{-1}\, var(y)\, \Omega^{-1}X_1(X_1'\Omega^{-1}X_1)^{-1}$

$= (X_1'\Omega^{-1}X_1)^{-1}X_1'\Omega^{-1}\sigma^2\Omega^{-1}X_1(X_1'\Omega^{-1}X_1)^{-1}$

$= \sigma^2(X_1'\Omega^{-1}X_1)^{-1}X_1'\Omega^{-2}X_1(X_1'\Omega^{-1}X_1)^{-1}$

b) $E(\hat{\beta}_1) = \beta_1 + (X_1'\Omega^{-1}X_1)^{-1}X_1'\Omega^{-1}X_2\beta_2 \neq \beta_1$

The bias will be zero if $X_2\beta_2 = 0$ (nothing has been omitted from the model) or if $X_1'\Omega^{-1}X_2 = 0$ (the vectors X_1 and X_2 are orthogonal after transformation).

c) $b_1 = (X_1'M_2X_1)^{-1}X_1'M_2y$

$var(b_1) = (X_1'M_2X_1)^{-1}X_1'M_2\, var(y)\, M_2X_1(X_1'M_2X_1)^{-1}$

$= \sigma^2(X_1'M_2X_1)^{-1}X_1'M_2X_1(X_1'M_2X_1)^{-1}$

$= \sigma^2(X_1'M_2X_1)^{-1}$

d) Variance could go either way. Compared to b_1 from model (1), $\hat{\beta}_1$ from model (2) will have smaller variance due to the restrictions which have been imposed. However, we know that when $var(\varepsilon) = \sigma^2 I$, OLS estimator is efficient, and applying GLS will increase variance.

14.) a) $R^2 = 1 - \dfrac{e'e}{\Sigma(y_i - \bar{y})^2}$

$e'e = 9716547$

S.D. dependent var $= \sqrt{\frac{1}{n-1} \Sigma(y_i - \bar{y})^2}$

So, $\Sigma(y_i - \bar{y})^2 = (242.1072)^2 \times (204) = 11957642.84$

and $R^2 = 0.1874$.

The model explains 18.74% of the sample variation in y.

b) $10.39204 \pm 1.960 \times (3.982739)$

So the interval is $[2.58587 ; 18.19821]$

If we constructed such intervals many times, 95% of them would cover the true value of the coefficient. This particular interval may or may not cover it.

c) $H_0: \beta_2 = 10$ vs. $H_A: \beta_2 \neq 10$.

$$t = \frac{17.47498 - 10}{6.184717} = 1.2086$$

Using a critical value of $t_c = 1.96$, we do not reject H_0 at the 5% sig. level.

d) The relevant p-value can be read from the Eviews output:

Prob (F-statistic) = 0.000001