Alastair Hall

ECON61001:
Econometric Methods

Semester 1, 2020-21

Solutions to Problem Set for Tutorial 6

1.(a) If both equations (1) and (2) on the Problem set hold then $v_g = \bar{u}_g = n_g^{-1}\sum_{i=N_{g-1}+1}^{N_g} u_i$. Define $H_g = \{h_i;\ i = N_{g-1}+1,\ N_{g-1}+2, \ldots, N_g)$. Using the LIE-II in the hint, we have:

$$E[v_g|\bar{h}_g] \;=\; E[\bar{u}_g|\bar{h}_g] \;=\; E\left[\, E[\bar{u}_g|H_g, \bar{h}_g] \,|\, \bar{h}_g \right], \tag{1}$$

Using the hint, it follows that

$$E[\bar{u}_g|H_g, \bar{h}_g] \;=\; E[\bar{u}_g|H_g] \;=\; E[\sum_{i=N_{g-1}+1}^{N_g} u_i/n_g|H_g] \;=\; \sum_{i=N_{g-1}+1}^{N_g} E[u_i|H_g]/n_g. \tag{2}$$

From Assumption *CS2*, we have that u_i and h_j are independent for all $i \neq j$ and so $E[u_i|H_g] = E[u_i|h_i]$ which equals zero via Assumption *CS4*. Therefore, it follows from (2) that $E[v_g|H_g] = 0$, which in turn implies $E[v_g|\bar{h}_g] = 0$ via (1).

Now consider $Var[v_g|\bar{h}_g]$. Since $E[v_g|\bar{h}_g] = 0$, it follows that $Var[v_g|\bar{h}_g] = E[v_g^2|\bar{h}_g]$. Using LIE-II and the suggestion in the hint,

$$E[v_g^2|\bar{h}_g] = E\left[\, E[v_g^2|H_g] \,|\, \bar{h}_g \right]. \tag{3}$$

Multiplying out, we have

$$E[v_g^2|H_g] \;=\; E[\sum_{i=N_{g-1}+1}^{N_g}\sum_{j=N_{g-1}+1}^{N_g} u_i u_j/n_g^2 \,|\, H_g] \;=\; n_i^{-2}\sum_{i=N_{g-1}+1}^{N_g}\sum_{j=N_{g-1}+1}^{N_g} E[u_i u_j|H_g]. \tag{4}$$

Assumption *CS2* states that (u_i, h_i) and (u_j, h_j) are independent for all $i \neq j$. Therefore, if $i \neq j$ then $E[u_i u_j|H_i] = E[u_i|h_i]E[u_j|h_j] = 0$ from Assumption *CS4*. From Assumptions *CS2* and *CS5*, it follows that $E[u_i^2|H_i] = E[u_i^2|h_i] = \sigma_0^2$. Using these results in (4), we obtain:

$$E[v_i^2|H_g] \;=\; n_g^{-2}\sum_{i=N_{g-1}+1}^{N_g} E[u_i^2|H_g] \;=\; n_g^{-2}\sum_{i=N_{g-1}+1}^{N_g} E[u_i^2|h_i] \;=\; \sigma_0^2/n_g.$$

1.(b) Assumption *CS2* states that (u_i, h_i) is independent of (u_j, h_j) for all $i \neq j$. So $E[v_g|\bar{X}] = E[v_g|\bar{h}_g] = 0$, from part (a). Since this holds for all i, we obtain $E[v|\bar{X}] = 0$.

Similarly, $Var[v_g|\bar{X}] = Var[v_g|\bar{h}_g] = \sigma_0^2/n_g$. Assumption *CS2* states that (u_i, h_i) is independent of (u_j, h_j) and so $(v_g, \bar{h}_g')$ is independent of $(v_\ell, \bar{h}_\ell')'$ which implies $Cov[v_g, v_\ell|\bar{h}_g, \bar{h}_\ell] = 0$

for all $g \neq \ell$. Again using Assumption *CS2*, it follows that $Cov[v_g, v_\ell|\bar{X}] = 0$. Combining these results, we obtain:

$$Var[v|\bar{X}] \;=\; \sigma_0^2 \begin{bmatrix} n_1^{-1} & 0 & 0 & \dots & 0 \\ 0 & n_2^{-1} & 0 & \dots & 0 \\ \vdots & & \ddots & & \vdots \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & \dots & 0 & n_{\mathcal{G}}^{-1} \end{bmatrix}.$$

1.(c) If $n_g \neq n$, say, for all g then $\{v_g\}$ are heteroscedastic. This means that the OLS estimator based on equation (1) in the question is unbiased but inefficient.

1.(d) Let $\hat{\beta}_{GLS}$ be the GLS estimator of β. From (a), it follows that the regression model can be written as

$$\bar{y} \;=\; \bar{X}\beta \;+\; v$$

where $\bar{y}$ is the $\mathcal{G} \times 1$ vector with g^{th} element $\bar{y}_g$, $\bar{X}$ is the $\mathcal{G} \times 2$ matrix with g^{th} row $[1, \bar{h}_g]$ and v is the $\mathcal{G} \times 1$ vector with g^{th} element v_g. From part (a) it follows that $Var[v] = \Sigma = \sigma_0^2 V$ where $V = diag[n_1^{-1}, n_2^{-1}, \dots n_{\mathcal{G}}^{-1}]$. From Lecture 5, it can be recalled that the GLS estimator is given by

$$\hat{\beta}_{GLS} \;=\; (\bar{X}'\Sigma^{-1}\bar{X})^{-1}\bar{X}'\Sigma^{-1}\bar{y} \;=\; (\bar{X}'V^{-1}\bar{X})^{-1}\bar{X}'V^{-1}\bar{y},$$

as the multiple of σ_0^2 in Σ cancels out. This estimator is feasible - provided that $\{n_g; i = 1, 2 \dots \mathcal{G}\}$ are known.

2. If $\beta_i|x_i \sim N(\beta_0, \sigma_0^2 I_K)$ then we can write $\beta_i = \beta_0 + v_i$ where $v_i|x_i \sim N(0, \sigma_0^2 I_K)$. This means that the model for y_i can be written as

$$y_i \;=\; x_i'(\beta_0 \;+\; v_i) \;=\; x_i'\beta_0 \;+\; u_i$$

where $u_i = x_i'v_i$. Notice that $E[u_i|x_i] = E[x_i'v_i|x_i] = x_i'E[v|x_i] = 0$ and $Var[u_i|x_i] = Var[x_i'v_i|x_i] = \sigma_0^2 x_i'x_i$. Therefore one possible justification for the heteroscedsatic linear model is *parameter variation* in the mean of a linear model.

3.(a) Substituting for y, we obtain:

$$\hat{\beta}_W \;=\; (X'W_2X)^{-1}X'W_2(X\beta_0 \;+\; u) \;=\; \beta_0 \;+\; (X'W_2X)^{-1}X'W_2u.$$

3.(b) Using part (a), we have

$$E\left[\hat{\beta}_W\right] \;=\; \beta_0 \;+\; E\left[(X'W_2X)^{-1}X'W_2u\right].$$

Since X and W_2 are constants, it follows that:

$$E\left[(X'W_2X)^{-1}X'W_2u\right] \;=\; (X'W_2X)^{-1}X'W_2\,E\left[u\right] \;=\; 0 \text{ using } CA4.$$

3.(c) Using parts (a) and (b), we have:

$$\begin{aligned} Var[\hat{\beta}_W] &= E\Big[\,(\hat{\beta}_W - \beta_0)(\hat{\beta}_W - \beta_0)'\,\Big] \\ &= E\Big[\,(X'W_2X)^{-1}X'W_2uu'W_2X(X'W_2X)^{-1}\,\Big] \\ &= (X'W_2X)^{-1}X'W_2E[uu']W_2X(X'W_2X)^{-1}, \text{ as both } X \text{ and } W_2 \text{ are constants,} \\ &= (X'W_2X)^{-1}X'W_2\Sigma W_2X(X'W_2X)^{-1}, \text{ as } Var[u] = \Sigma. \end{aligned}$$

3.(d) Using part(a), *CA6*, and X and W_2 constant, $\hat{\beta}_W$ is a linear combination of $u \sim N(0, \Sigma)$. Therefore, it follows from Lemma 2.1 in the Lecture Notes that $\hat{\beta}_W \sim N\Big(\beta_0, Var[\hat{\beta}_W]\Big)$.

3.(e) Using the LIE, we have:

$$E\Big[\,\hat{\beta}_W\,\Big] = \beta_0 + E\Big[\,E\Big[\,(X'W_2X)^{-1}X'W_2u \,|\, X\Big]\Big]$$

Since W_2 is constant, we have

$$E\Big[\,E\Big[\,(X'W_2X)^{-1}X'W_2u \,|\, X\Big]\Big] = E\Big[\,(X'W_2X)^{-1}X'W_2\,E\,[u \,|\, X]\Big] = 0 \text{ using } \textit{SR4}.$$

Therefore, we have $E\Big[\,\hat{\beta}_W\,\Big] = \beta_0$ and so $\hat{\beta}_W$ is an unbiased estimator of β_0.

4.(a) This follows directly because the WLS estimator is OLS applied to the regression model $(w_i * y_i) = (w_i * x_i)'\beta_0 + (w_i * u_i)$.

4.(b) Since $\{(x_i', u_i)\}_{i=1}^N$ is an i.i.d. sequence and $\{w_i\}_{i=1}^N$ are constants, it follows that $(\breve{x}_i', \breve{u}_i)$ and $(\breve{x}_j', \breve{u}_j)$ are independent for all $i \neq j$. However, $(\breve{x}_i', \breve{u}_i)$ and $(\breve{x}_i', \breve{u}_j)$ are not identically distributed: to see this, note that $E[\breve{x}_i\breve{x}_i'] = w_i^2E[x_ix_i'] = w_i^2Q$ and so changes with i (assuming the weights are not independent of i).

4.(c)(i) Using Question 3(a), we have:

$$\hat{\beta}_W = \beta_0 + (N^{-1}X'W_2X)^{-1}N^{-1}X'W_2u. \tag{5}$$

We are given in the question that,

$$N^{-1}X'W_2X \xrightarrow{p} Q_w, \text{ a positive definite matrix,}$$

and so via Slutsky's Theorem,

$$(N^{-1}X'W_2X)^{-1} \xrightarrow{p} Q_w^{-1}. \tag{6}$$

We are also given that

$$N^{-1/2}X'W_2u \xrightarrow{d} N(0, \Omega_w),$$

form which it follows that

$$N^{-1}X'W_2u = N^{-1/2}\Big\{N^{-1/2}X'W_2u\Big\} \xrightarrow{p} 0. \tag{7}$$

Combining (5)-(7) and using Slutsky's Theorem, we obtain:

$$\hat{\beta}_W \xrightarrow{p} \beta_0 + Q_w^{-1} \times 0 = \beta_0,$$

and so $\hat{\beta}_W$ is a consistent estimator for β_0.

4.(ii) Using Question 3(a), we have:

$$N^{1/2}(\hat{\beta}_W - \beta_0) = (N^{-1}X'W_2X)^{-1}N^{-1/2}X'W_2u. \tag{8}$$

Using the limit theorems given in the question, it then follows from Lemma 3.5 in the Lecture Notes (in Section 3.1) that:

$$N^{1/2}(\hat{\beta}_W - \beta_0) \xrightarrow{d} N(0, Q_w^{-1}\Omega_w Q_w^{-1}).$$

4.(c)(iii) A suitable test statistic is:

$$S_N = N(R\hat{\beta}_W - r)'[R\hat{V}_W R']^{-1}(R\hat{\beta}_W - r),$$

where $\hat{V}_W = \hat{Q}_w^{-1}\hat{\Omega}_w\hat{Q}_w^{-1}$, $\hat{Q}_w = N^{-1}X'W_2X$, $\hat{\Omega}_w = N^{-1}X'\hat{M}_wX$ and $\hat{M}_w = diag(\hat{u}_1^2w_1^4, \hat{u}_2^2w_2^4, \ldots, \hat{u}_N^2w_N^4)$ and $\hat{u}_i = y_i - x_i'\hat{\beta}_W$. It can be shown under the conditions here that $\hat{V}_W \xrightarrow{p} V_W = Q_w^{-1}\Omega_w Q_w^{-1}$ and so it follows from Lemma 3.6 in the Lecture Notes (in Section 3.1) that under H_0 $S_N \xrightarrow{d} \chi^2_{n_r}$.