**Econometrics 710 Problem Set #8**

March 23, 2015

TA: Cristian Hernandez
cahernandez3@wisc.edu

## Exercise 8.1

Consider the model

$$\begin{aligned} y_i &= x_i'\beta + z_i\theta + e_i \\ \mathbb{E}(e_i x_i) &= 0 \\ \mathbb{E}(e_i z_i) &= 0 \end{aligned}$$

, where $y_i \in \mathbb{R}$, $x_i \in \mathbb{R}^k$, and $z_i \in \mathbb{R}$. Now let $SSR(\widehat{\beta})$ denote the sum of squared residuals from the OLS regression of $y_i$ on $x_i$ and $z_i$; $SSR(\widehat{\beta}_{cls})$ denote the sum of squared residuals from the OLS regression of $y_i$ on $x_i$ only—this is the exclusion restriction $\theta = 0$—; and $SST = \sum_{i=1}^n (y_i - \bar{y})^2$. Let $\bar{R}_k^2$ denote Theil's adjusted R-squared for the constrained model and $\bar{R}_{k+1}^2$ denote Theil's adjusted R-squared for the unconstrained model. Then,

$$\begin{aligned} \bar{R}_{k+1}^2 > \bar{R}_k^2 &\iff 1 - \frac{1}{n-k-1}\frac{SSR(\widehat{\beta})}{SST} > 1 - \frac{1}{n-k}\frac{SSR(\widehat{\beta}_{cls})}{SST} \\ &\iff F_n = \frac{SSR(\widehat{\beta}_{cls}) - SSR(\widehat{\beta})}{SSR(\widehat{\beta})/(n-k-1)} > 1 \end{aligned}$$

Note that $F_n$ is the $F$-statistic for the unconstrained model with $k+1$ explanatory variables and the exclusion restriction $\theta = 0$. Since we have only one restriction $q = 1$, it follows that

$$F_n = J_n^0/q = J_n^0 = W_n^0 = t_{k+1}^2$$

, where $J_n^0$ is the minimum distance statistic statistic under homoskedasticity, $W_n^0$ is the Wald statistic under homoskedasticity, and $t_{k+1}$ is the $t$-ratio of $\widehat{\theta}$ with the homoskedastic standard error. Then, $F_n > 1 \iff t_{k+1}^2 > 1 \iff |t_{k+1}| > 1$.

## Exercise 8.2

(a) By the asymptotic properties of the OLS estimator for each sample, we know that

$$\sqrt{n}(\hat{\beta}_1 - \beta_1) \xrightarrow{d} N(0, V_1)$$

$$\sqrt{n}(\hat{\beta}_2 - \beta_2) \xrightarrow{d} N(0, V_2)$$

where,

$$V_1 = (\mathbb{E}\mathbf{x}_{1i}\mathbf{x}_{1i}')^{-1}(\mathbb{E}\mathbf{x}_{1i}\mathbf{x}_{1i}'e_{1i}^2)(\mathbb{E}\mathbf{x}_{1i}\mathbf{x}_{1i}')^{-1}, V_2 = (\mathbb{E}\mathbf{x}_{2i}\mathbf{x}_{2i}')^{-1}(\mathbb{E}\mathbf{x}_{2i}\mathbf{x}_{2i}'e_{2i}^2)(\mathbb{E}\mathbf{x}_{2i}\mathbf{x}_{2i}')^{-1},$$

Independence of the samples implies that,

$$\begin{pmatrix} \sqrt{n}(\hat{\beta}_1 - \beta_1) \\ \sqrt{n}(\hat{\beta}_2 - \beta_2) \end{pmatrix} \to_d N\left( \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} V_1 & 0 \\ 0 & V_2 \end{pmatrix} \right).$$

By the CMT,

$$\sqrt{n}((\hat{\beta}_1 - \hat{\beta}_2) - (\beta_1 - \beta_2)) \xrightarrow{d} N(0, V_1 + V_2).$$

(b)

From (a), an appropriate test statistic for $H_0 : \beta_2 = \beta_1$ is the Wald statistic.

$$W_n = n(\hat{\beta}_1 - \hat{\beta}_2)'(\hat{V}_1 + \hat{V}_2)^{-1}(\hat{\beta}_1 - \hat{\beta}_2)$$

where $\hat{V}_1, \hat{V}_2$ are consistent estimators of $V_1, V_2$.

(c) From (a) and (b), $\sqrt{n}((\hat{\beta}_1 - \hat{\beta}_2) - (\beta_1 - \beta_2)) \to_d N(0, V_1 + V_2)$, and $\hat{V}_1 \xrightarrow{p} V_1, \hat{V}_2 \xrightarrow{p} V_2$. Under $H_0$, $W_n \xrightarrow{d} \chi^2(k)$ as $n \to \infty$

## Exercise 8.3

(a) Regression equation : $I_i = \beta_1 Q_i + \beta_2 C_i + \beta_3 D_i + \beta_4 + e_i$

The following equation reports the OLS estimates with Horn-Horn-Duncan heteroskedastic-robust standard errors in parenthesis

$$\hat{I}_i = -\underset{(0.391)}{0.421} Q_i + \underset{(1.686)}{7.229} C_i + \underset{(0.472)}{0.736} D_i + \underset{(0.521)}{7.510}$$

(b)

Asymptotic confidence intervals for the coefficients: $\hat{\beta}_j \pm 1.96 se(\hat{\beta}_j)$

$CI(\beta_1)$ : [ -1.187 0.345], $CI(\beta_2)$ : [ 3.925 10.534], $CI(\beta_3)$ : [ -0.189 1.662], $CI(\beta_4)$ : [ 6.487 8.531]

(c)

(1) We set the null and alternative hypotheses as follows

$$H_0 : R'\beta = 0, \qquad H_1 : R'\beta \neq 0$$

where

$$\beta = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{pmatrix}, R = \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$$

The Wald statistic is $W_n = n(R'\hat{\beta})'(R'\hat{V}_\beta R)^{-1}R'\hat{\beta} = 21.027$. Since we have to test two restrictions, the asymptotic distribution of $W_n$ is the chi-square distribution with 2 degrees of freedom. The critical value for the chi-square distribution with 2 degrees of freedom is $\chi^2_{0.95}(2) = 5.99 < W_n$ for $\alpha = 0.05$. Therefore, we reject $H_0$ at 5% significance level.

(2) For testing the coefficient of $Q_i$ are zero, the null hypothesis is $H_0 : R'\beta = 0$, where $R = (1, 0, 0, 0)'$

The Wald statistic is $W_n = n\hat{\beta}'R(R'\hat{V}_\beta R)^{-1}R'\hat{\beta} = \frac{n\hat{\beta}_1^2}{[\hat{V}_\beta]_{1,1}} = t_n^2$ where $t_n = \frac{|\hat{\beta}_1|}{s.e.(\hat{\beta}_1)}$

Under $H_0$, $W_n$ converges to chi-square distribution with 1 degree of freedom. (or $t_n \xrightarrow{d} |Z|$, $Z \sim N(0,1)$). Since $W_n = 1.163 < 3.84 = \chi^2_{0.95}(1)$ (or $t_n = 1.078 < 1.96$) , we fail to reject $H_0$ under 5 % significance level.

From (1) and (2), it is not likely that coefficient of C and D in regression equations are both zero, and we fail to reject the hypothesis that coefficient of Q is zero under 5 % significance level. These results are inconsistent with the predictions of the theory.

(d)

Regression equation:

$$I_i = \beta_1 Q_i + \beta_2 C_i + \beta_3 D_i + \beta_4 Q_i^2 + \beta_5 C_i^2 + \beta_6 D_i^2 + \beta_7 Q_i C_i + \beta_8 Q_i D_i + \beta_9 C_i D_i + \beta_{10} + e_i$$

The following equation reports the OLS estimates with Horn-Horn-Duncan heteroskedastic-robust standard errors in parenthesis

$$\begin{aligned} \hat{I}_i &= \underset{(0.392)}{0.406} Q_i + \underset{(3.313)}{9.958} C_i + \underset{(1.163)}{1.736} D_i - \underset{(0.045)}{0.101} Q_i^2 - \underset{(4.200)}{12.620} C_i^2 \\ &\quad - \underset{(0.648)}{0.869} D_i^2 + \underset{(0.757)}{2.098} Q_i C_i - \underset{(0.419)}{0.213} Q_i D_i + \underset{(3.524)}{5.738} C_i D_i + \underset{(0.570)}{5.930} \end{aligned}$$

To test $H_0 : R'\beta = 0$ where $R = [0_{6\times 3}\ \ I_{6\times 6}\ \ 0_{6\times 1}]'$, I use the Wald statistic $W_n$. $W_n = 36.774 > 12.59 = \chi^2_{0.95}(6)$. We reject the null hypothesis at 5% significance level.

## Exercise 8.4

(a) The following equation reports the OLS estimates with Horn-Horn-Duncan heteroskedastic-robust standard errors in parenthesis

$$\begin{aligned} \widehat{\log TC_i} &= -3.527 + 0.720 \log Q_i + 0.436 \log PL_i - 0.220 \log PK_i + 0.427 \log PF_i \\ &\quad (1.741) \quad (0.033) \qquad\quad (0.248) \qquad\quad (0.328) \qquad\quad (0.076) \end{aligned}$$

(b) If the null hypothesis $H_0 : \beta_3 + \beta_4 + \beta_5 = 1$ is true, that means that the total cost function exhibits constant returns to scale with respect to the price of labor, capital and fuel.

(c) CLS estimates and consistent asymptotic covariance matrix estimates are

$$\tilde{\beta}_{cls} = \hat{\beta} - (X'X)^{-1} R[R'(X'X)^{-1}R]^{-1}(R'\hat{\beta} - 1)$$

where $R = [0, 0, 1, 1, 1]'$ and

$$\begin{aligned} \hat{V}_{cls} &= \hat{V}_\beta - \hat{Q}_{xx}^{-1} R(R'\hat{Q}_{xx}^{-1}R)^{-1} R'\hat{V}_\beta - \hat{V}_\beta R(R'\hat{Q}_{xx}^{-1}R)^{-1} R'\hat{Q}_{xx}^{-1} \\ &+ \hat{Q}_{xx}^{-1} R(R'\hat{Q}_{xx}^{-1}R)^{-1} R'\hat{V}_\beta R(R'\hat{Q}_{xx}^{-1}R)^{-1} R'\hat{Q}_{xx}^{-1} \end{aligned}$$

where $\hat{V}_\beta$ is a consistent estimator of $V_\beta$, and $\hat{Q}_{xx} = \frac{1}{n} X'X$. The results are

$$\begin{aligned} \widehat{\log TC_i} &= -4.691 + 0.721 \log Q_i + 0.593 \log PL_i - 0.007 \log PK_i + 0.415 \log PF_i \\ &\quad (0.818) \quad (0.033) \qquad\quad (0.169) \qquad\quad (0.156) \qquad\quad (0.074) \end{aligned}$$

(d) The efficient minimum distance estimator and consistent asymptotic covariance matrix estimator are $\tilde{\beta}_{emd} = \hat{\beta} - \hat{V}_\beta R[R'\hat{V}_\beta R]^{-1}(R'\hat{\beta} - 1)$ The following equation summarizes the results

$$\hat{V}_{emd} = \hat{V}_\beta - \hat{V}_\beta R(R'\hat{V}_\beta R)^{-1} R'\hat{V}_\beta$$

$$\begin{aligned}\widehat{\log TC_i} &= \underset{(0.815)}{-4.746} + \underset{(0.033)}{0.720} \log Q_i + \underset{(0.169)}{0.580} \log PL_i + \underset{(0.154)}{0.009} \log PK_i + \underset{(0.073)}{0.410} \log PF_i\end{aligned}$$

Note that these results are similar to the ones with CLS.

(e) The Wald statistic for testing $H_0 : \beta_1 + \beta_2 + \beta_3 = 1$ is $W_n = n(R'\hat{\beta} - 1)'(R'\hat{V}_\beta R)^{-1}(R'\hat{\beta} - 1) = \frac{n(\hat{\beta}_1+\hat{\beta}_2+\hat{\beta}_3-1)^2}{(R'\hat{V}_\beta R)} = 0.629 < 3.84 = \chi^2_{0.95}(1)$. Thus, we fail to reject $H_0$ under 5 % significance level.

(f) Since the null hypothesis is linear in the parameters, the efficient minimum distance statistic is algebraically equivalent to the Wald statistic.

$$J_n^* = J_n(\tilde{\beta}_{emd}) = n(\hat{\beta} - \tilde{\beta}_{emd})'\hat{V}_\beta^{-1}(\hat{\beta} - \tilde{\beta}_{emd})$$

Thus, test results are same as in (e).

### Exercise 9.1

From the asymptotic distribution of the NLLS estimator, we know

$$\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \xrightarrow{d} N(0, \boldsymbol{V_\theta}),$$

where

$$\boldsymbol{V_\theta} = \left(\mathbb{E}(\boldsymbol{g_{\theta i} g'_{\theta i}})\right)^{-1} \left(\mathbb{E}(\boldsymbol{g_{\theta i} g'_{\theta i}} e_i^2)\right)^{-1} \left(\mathbb{E}(\boldsymbol{g_{\theta i} g'_{\theta i}})\right)^{-1}, \boldsymbol{g_{\theta i}} = \frac{\boldsymbol{\partial}}{\boldsymbol{\partial \theta}} \boldsymbol{g(x_i, \theta)}$$

$\hat{\boldsymbol{V}}$ is the estimator of $\text{var}(\hat{\boldsymbol{\theta}})$, which is approximately $\frac{1}{n}\hat{\boldsymbol{V}}_{\boldsymbol{\theta}}$

For any fixed $\mathbf{x}$, by the Delta-Method,

$$\sqrt{n}(g(\mathbf{x}, \hat{\boldsymbol{\theta}}) - g(\mathbf{x}, \boldsymbol{\theta})) \xrightarrow{d} N(0, \boldsymbol{G' V_\theta G})$$

where $\boldsymbol{G} = \frac{\partial g(\mathbf{x}, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}}$.

A 95% asymptotic confidence interval for $g(\mathbf{x})$ is $\left[g(\mathbf{x}, \hat{\boldsymbol{\theta}}) \pm 1.96 \cdot \sqrt{\hat{\boldsymbol{G}}'\hat{\boldsymbol{V}}\hat{\boldsymbol{G}}}\right]$, where $\hat{\boldsymbol{G}} = \frac{\partial g(\mathbf{x}, \hat{\boldsymbol{\theta}})}{\partial \boldsymbol{\theta}}$

### Exercise 9.2

(a) Consider the following two specifications of the model

$$\begin{aligned}(\text{Model 1}) : \log TC_i &= \beta_1 + \beta_2 \log Q_i + \beta_3 \log PL_i + \beta_4 \log PK_i + \beta_5 \log PF_i + e_i \\ (\text{Model 2}) : \log TC_i &= \beta_1 + \beta_2 \log Q_i + \beta_3 \log PL_i + \beta_4 \log PK_i + \beta_5 \log PF_i + \beta_6 (\log Q_i)^2 + e_i\end{aligned}$$

If we want to pick the model or judge new specification whether include variable $(\log Q_i)^2$ or not, you may want to test $H_0 : \beta_6 = 0$ in Model 2. Since the t-statistic is $t_n = 5.450 > 1.96$, based on this test, adding a variable $(\log Q_i)^2$ to the Model 1 seems a reasonable thing to do.

It is reasonable to think that firm's marginal cost changes as output changes. We know that standard total cost function has S-shape, which implies increasing returns to scale for low-output firms, and decreasing returns to scale for high-output firms.

Note that there exist more structured techniques to compare models. For example, you could estimate both models and report the leave-one-out cross validation $\widetilde{R}^2$ (you have Matlab code for that, check the solution from problem set 7) and choose the model that has the highest $\widetilde{R}^2$.

(b) Now, try a different non-linear specification:

$$\begin{aligned} \log TC_i &= \beta_1 + \beta_2 \log Q_i + \beta_3 \log PL_i + \beta_4 \log PK_i + \beta_5 \log PF_i \\ &\quad + \beta_6 \log Q_i (1 + \exp(-(\log Q_i - \beta_7)))^{-1} + e_i. \end{aligned} \tag{1}$$

From the data, we need to choose $\beta_7$ so that appropriate number of observations of $\log Q_i$ are distributed both above and below over $\beta_7$. So we choose reasonable range for $\beta_7$ as 16th minimum and maximum values of $\log Q_i$ from the data, which is [4.143, 8.645].

(c) I picked 1000 values from 4.143 to 8.643, equally spaced within the interval, and computed the sum of squared residuals that result from the CLS procedure for each value of $\beta_7$. SSE is minimized

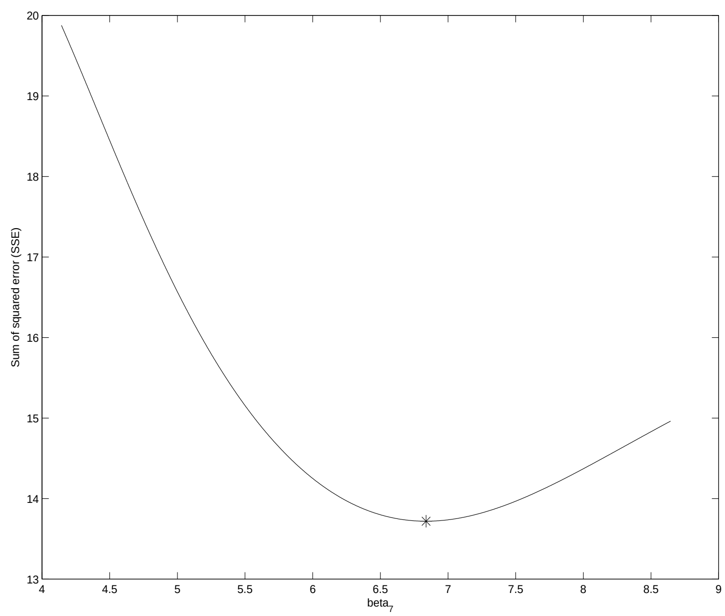


Figure 1: Sum of squared errors for each values of $\beta_7 \in [4.143, 8.643]$

at $\hat{\beta}_7 = 6.838$. The following equation summarizes estimation results

$$\begin{aligned} \widehat{\log TC_i} &= -4.043 + 0.433 \log Q_i + 0.479 \log PL_i + 0.046 \log PK_i + 0.475 \log PF_i \\ &+ 0.228 \frac{\log Q_i}{1 + \exp(-(\log Q_i - 6.838))} \end{aligned}$$

(d) I think that the easiest way to find the standard errors for the constrained non-linear least squares problem is to replace the constraint in the model, re-arrange the equation, re-estimate the model, and then report the usual standard errors for the NLLS estimator. This procedure yields the right standard errors as long as the constraint is true.

Then, we replace the constraint $\beta_5 = 1 - \beta_5 - \beta_4$ into the equation and get the following equation

$$\log \frac{TC_i}{PF_i} = \beta_1 + \beta_2 \log Q_i + \beta_3 \log \frac{PL_i}{PF_i} + \beta_4 \log \frac{PK_i}{PF_i} + \beta_6 \boldsymbol{z}_i + e_i.$$

Now let $m(\boldsymbol{x}, \theta) = \beta_1 + \beta_2 x_{1i} + \beta_3 x_{2i} + \beta_4 x_{3i} + \beta_6 \frac{x_{1i}}{1+\exp(\beta_7 - x_{1i})}$, where $\theta = (\beta_1, \beta_2, \beta_3, \beta_4, \beta_6, \beta_7)'$, $y_i = \log \frac{TC_i}{PF_i}$, $x_{1i} = \log Q_i$, $x_{2i} = \log \frac{PL_i}{PF_i}$, and $x_{3i} = \log \frac{PK_i}{PF_i}$. Then, the asymptotic variance of the NLLS estimator $\hat{\theta}$ is

$$\hat{V}_\theta = \left( \frac{1}{n} \sum_{i=1}^{n} \hat{m}_{\theta i} \hat{m}'_{\theta i} \right)^{-1} \left( \frac{1}{n} \sum_{i=1}^{n} \hat{m}_{\theta i} \hat{m}'_{\theta i} \hat{e}_i^2 \right) \left( \frac{1}{n} \sum_{i=1}^{n} \hat{m}_{\theta i} \hat{m}'_{\theta i} \right)^{-1}$$

where

$$\begin{aligned} \hat{m}_{\theta i} &= \frac{\partial m(x_i, \hat{\theta})}{\partial \theta} = \left( 1, x_{1i}, x_{2i}, x_{3i}, \frac{x_{1i}}{1 + \exp(\hat{\beta}_7 - x_{1i})}, -\hat{\beta}_6 \frac{x_{1i} \exp(\hat{\beta}_7 - x_{1i})}{(1 + \exp(\hat{\beta}_7 - x_{1i}))^2} \right)' \\ \hat{e}_i &= y_i - m(x_i, \hat{\theta}), \end{aligned}$$

Using the estimates $(\hat{\beta}_1, ..., \hat{\beta}_7)'$ from (c), the standard errors for all parameters estimates $\hat{\theta} = (\hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3, \hat{\beta}_4, \hat{\beta}_6, \hat{\beta}_7)'$ are $0.702, 0.097, 0.138, 0.116, 0.054$ and $0.345$ respectively. The standard error of $\hat{\beta}_5 = 1 - \hat{\beta}_3 - \hat{\beta}_4$ is $0.069$.

Note that there is another way to solve for the constrained non-linear estimator. Given that the non-linear estimator is asymptotically normal, and that the constraint is linear, we can use the efficient minimum distance estimator using the unconstrained NLLS result, along with the estimated covariance matrix.