Econometrics 710 Problem Set #7

March 21, 2015

TA: Cristian Hernandez
cahernandez3@wisc.edu

Exercise 7.1

$$\begin{aligned}
\widetilde{\beta} &= \underset{\beta:\, \beta_2=0}{\arg\min}\, SSR(\beta) = (y - X_1\beta_1 - X_2\beta_2)'(y - X_1\beta_1 - X_2\beta_2) \\
&= \underset{\beta:\, \beta_2=0}{\arg\min}\, SSR(\beta_1, 0) = (y - X_1\beta_1)'(y - X_1\beta_1) \\
&= \begin{pmatrix} (X_1'X_1)^{-1}X_1'y \\ 0 \end{pmatrix}.
\end{aligned}$$

The CLS estimate of β subject to $\beta_2 = 0$ is just an OLS regression of y on x_1.

Exercise 7.2

$$\begin{aligned}
\tilde{\beta} &= \underset{\beta:\, \beta_1=c}{\arg\min}\, SSR(\beta) = (y - X_1\beta_1 - X_2\beta_2)'(y - X_1\beta_1 - X_2\beta_2) \\
&= \underset{\beta:\, \beta_1=c}{\arg\min}\, SSR(c, \beta_2) = (y - X_1c - X_2\beta_2)'(y - X_1c - X_2\beta_2) \\
&= \underset{\beta:\, \beta_1=c}{\arg\min}\, SSR(c, \beta_2) = (z - X_2\beta_2)'(z - X_2\beta_2) \\
&= \begin{pmatrix} (X_2'X_2)^{-1}X_2'z \\ 0 \end{pmatrix} \\
&= \begin{pmatrix} (X_2'X_2)^{-1}X_2'(y - X_1c) \\ 0 \end{pmatrix}.
\end{aligned}$$

, where $z_i = y_i - x_ic$. Therefore, the CLS estimate of β subject to $\beta_1 = c$ is just an OLS regression of $y - x_1c$ on x_2.

Exercise 7.3

$$\begin{aligned}
\tilde{\beta} &= \underset{\beta:\, \beta_1+\beta_2=0}{\arg\min}\, SSR(\beta) = (y - X_1\beta_1 - X_2\beta_2)'(y - X_1\beta_1 - X_2\beta_2) \\
&= \underset{\beta:\, \beta_1+\beta_2=0}{\arg\min}\, SSR(\beta_1, -\beta_1) = (y - (X_1 - X_2)\beta_1)'(y - (X_1 - X_2)\beta_1) \\
&= \begin{pmatrix} ((X_1 - X_2)'(X_1 - X_2))^{-1}(X_1 - X_2)'y \\ -((X_1 - X_2)'(X_1 - X_2))^{-1}(X_1 - X_2)'y \end{pmatrix}.
\end{aligned}$$

Exercise 7.4

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

Since $R'(X'X)^{-1}R[R'(X'X)^{-1}R]^{-1}$ is equal to the identity matrix of size q (number of constraints).

Exercise 7.5

Under the restriction (7.1), $R'\beta = c$

(a) Note that $\hat{\beta} = \beta + (X'X)^{-1}X'e$,

$$R'\hat{\beta} - c = R'(\hat{\beta} - \beta) = R'(X'X)^{-1}X'e$$

(b) From (7.9),

$$\begin{aligned}
\tilde{\beta}_{cls} - \beta &= (\beta + (X'X)^{-1}X'e) - (X'X)^{-1}R[R'(X'X)^{-1}R]^{-1}(R'\hat{\beta} - c) - \beta \\
&= (X'X)^{-1}X'e - (X'X)^{-1}R(R'(X'X)^{-1}R)^{-1}R'(X'X)^{-1}X'e
\end{aligned}$$

(c)

$$\begin{aligned}
\tilde{e} &= y - X\tilde{\beta}_{cls} = X\beta + e - X\tilde{\beta}_{cls} \\
&= e - X((X'X)^{-1}X'e - (X'X)^{-1}R(R'(X'X)^{-1}R)^{-1}R'(X'X)^{-1}X'e) \quad \text{(from } (b)\text{)} \\
&= (I - P + A)e
\end{aligned}$$

where $A = X(X'X)^{-1}R(R'(X'X)^{-1}R)^{-1}R'(X'X)^{-1}X'$

(d)

$$\begin{aligned}
A' &= \left(X(X'X)^{-1}R(R'(X'X)^{-1}R)^{-1}R'(X'X)^{-1}X'\right)' \\
&= X(X'X)^{-1}R(R'(X'X)^{-1}R)^{-1}R'(X'X)^{-1}X' = A
\end{aligned}$$

$$\begin{aligned}
A^2 &= X(X'X)^{-1}R(R'(X'X)^{-1}R)^{-1}R'(X'X)^{-1}X'X(X'X)^{-1}R(R'(X'X)^{-1}R)^{-1}R'(X'X)^{-1}X' \\
&= X(X'X)^{-1}R(R'(X'X)^{-1}R)^{-1}R'(X'X)^{-1}R(R'(X'X)^{-1}R)^{-1}R'(X'X)^{-1}X' \\
&= X(X'X)^{-1}R(R'(X'X)^{-1}R)^{-1}R'(X'X)^{-1}X' = A
\end{aligned}$$

$$\begin{aligned}
tr(A) &= tr\left(X(X'X)^{-1}R(R'(X'X)^{-1}R)^{-1}R'(X'X)^{-1}X'\right) \\
&= tr\left((X'X)^{-1}R(R'(X'X)^{-1}R)^{-1}R'(X'X)^{-1}X'X\right) \\
&= tr\left(R(R'(X'X)^{-1}R)^{-1}R'(X'X)^{-1}\right) \\
&= tr\left((R'(X'X)^{-1}R)^{-1}R'(X'X)^{-1}R\right) = tr(I_q) = q
\end{aligned}$$

$$\begin{aligned}
PA &= (I - M)A = (I - X(X'X)^{-1}X')[X(X'X)^{-1}R(R'(X'X)^{-1}R)^{-1}R'(X'X)^{-1}X'] \\
&= A - X(X'X)^{-1}R(R'(X'X)^{-1}R)^{-1}R'(X'X)^{-1}X' = A - A = 0
\end{aligned}$$

Exercise 7.6

$\mathbb{E}(s^2_{cls}|X) = \frac{1}{n-k+q}\mathbb{E}(\tilde{e}'\tilde{e}|X)$. Also note that $(I - P + A)'(I - P + A) = I - P + A$ by exercise 7.5 (d).

$$\begin{aligned}
\mathbb{E}(\tilde{e}'\tilde{e}|X) &= \mathbb{E}(e'(I - P + A)'(I - P + A)e|X) \quad \text{(by exercise 7.5}(c)\text{)} \\
&= \mathbb{E}(e'(I - P + A)e|X) = \mathbb{E}(tr((I - P + A)ee')|X) \\
&= tr((I - P + A)\mathbb{E}(ee'|X)) = \sigma^2 tr(I - P + A) \quad \text{(homoskedastic regression assumption)} \\
&= (n - k + q)\sigma^2 \quad \text{(by (3.24) and exercise 7.5}(d)\text{)}
\end{aligned}$$

Thus, $\mathbb{E}(s^2_{cls}|X) = \sigma^2$

Exercise 7.7

Since W_n is positive definite, it is invertible and symmetric. The Lagrangian is

$$\begin{aligned}\mathcal{L}(\beta,\lambda) &= \frac{1}{2}J_n(\beta,W_n)+\lambda'(R'\beta-c)\\ &= \frac{1}{2}n(\hat{\beta}-\beta)'W_n(\hat{\beta}-\beta)+\lambda'(R'\beta-c)\end{aligned}$$

The FOC with respect to β and λ is

$$\begin{aligned}\frac{\partial\mathcal{L}(\tilde{\beta}_{md},\tilde{\lambda}_{md})}{\partial\beta}: &\qquad nW_n\tilde{\beta}_{md}-nW_n\hat{\beta}+R\tilde{\lambda}_{md}=0,\\ \frac{\partial\mathcal{L}(\tilde{\beta}_{md},\tilde{\lambda}_{md})}{\partial\lambda}: &\qquad R'\tilde{\beta}_{md}-c=0\end{aligned}$$

Premultiplying $R'W_n^{-1}$ on the first equation, we get

$$(R'W_n^{-1}R)\tilde{\lambda}_{md}=n(R'\hat{\beta}-R'\tilde{\beta}_{md})$$

Imposing FOC for λ and solving for λ gives us (7.21)

$$\tilde{\lambda}_{md}=n(R'W_n^{-1}R)^{-1}(R'\hat{\beta}-c).$$

Plugging $\tilde{\lambda}_{md}$ into the FOC for β gives (7.22)

$$\tilde{\beta}_{md}=\hat{\beta}-W_n^{-1}R(R'W_n^{-1}R)^{-1}(R'\hat{\beta}-c)$$

When $W_n=\hat{Q}_{xx}=\frac{1}{n}(X'X)$, $\tilde{\beta}_{md}$ specializes to $\tilde{\beta}_{cls}$.

Exercise 7.8

Under Assumption 6.1.1, and 7.5.1 we know that $\hat{\beta}\xrightarrow{p}\beta$. Under Assumption 7.5.1,

$$\begin{aligned}\tilde{\beta}_{md} &= \hat{\beta}-W_n^{-1}R(R'W_nR)^{-1}(R'\hat{\beta}-c)\\ \tilde{\beta}_{md} &= \hat{\beta}-W_n^{-1}R(R'W_nR)^{-1}R'(\hat{\beta}-\beta)\\ &\xrightarrow{p} \beta-W^{-1}R(R'WR)^{-1}R'(\beta-\beta)=\beta.\end{aligned}$$

Exercise 7.9

Under the assumption 7.5.1,

$$\begin{aligned}\sqrt{n}(\tilde{\beta}_{md}-\beta) &= \sqrt{n}(\hat{\beta}-\beta)-\sqrt{n}W_n^{-1}R(R'W_nR)^{-1}(R'\hat{\beta}-c)\\ &= \sqrt{n}(\hat{\beta}-\beta)-W_n^{-1}R(R'W_nR)^{-1}R'\sqrt{n}(\hat{\beta}-\beta)\\ &= A_n\sqrt{n}(\hat{\beta}-\beta)\end{aligned}$$

where $A_n=I-W_n^{-1}R(R'W_nR)^{-1}R'$ is $k\times k$ matrix.

By an assumption 7.5.2 and the continuous mapping theorem, $A_n \xrightarrow{p} A = I - W^{-1}R(R'W^{-1}R)^{-1}R'$. We also know that under assumption 6.1.2, $\sqrt{n}(\hat{\beta}-\beta) \xrightarrow{d} N(0, V_\beta)$, where $V_\beta = Q_{xx}^{-1}\Omega Q_{xx}^{-1}$. Thus,

$$\sqrt{n}(\tilde{\beta}_{md}-\beta) \xrightarrow{d} N(0, V_\beta(W))$$

where

$$\begin{aligned} V_\beta(W) &= AV_\beta A' = (I - W^{-1}R(R'W^{-1}R)^{-1}R')V_\beta(I - W^{-1}R(R'W^{-1}R)^{-1}R')' \\ &= V_\beta - W^{-1}R(R'W^{-1}R)^{-1}R'V_\beta - V_\beta R(R'W^{-1}R)^{-1}R'W^{-1} \\ &+ W^{-1}R(R'W^{-1}R)^{-1}R'V_\beta R(R'W^{-1}R)^{-1}R'W^{-1} \end{aligned}$$

Exercise 7.10

From exercise 7.7, $\tilde{\beta}_{cls}$ is a special case of $\tilde{\beta}_{md}$ when $W_n = \hat{Q}_{xx}$, thus Theorem 7.5.3 is a special case of Theorem 7.5.2 by replacing $W_n = \hat{Q}_{xx} = \frac{1}{n}\sum_{i=1}^n x_i x_i' \xrightarrow{p} W = Q_{xx}$.

The result in exercise 7.9 implies that the asymptotic variance of $\tilde{\beta}_{cls}$ is

$$\begin{aligned} V_{cls} = V_\beta(Q_{xx}) &= V_\beta - Q_{xx}^{-1}R(R'Q_{xx}^{-1}R)^{-1}R'V_\beta - V_\beta R(R'Q_{xx}^{-1}R)^{-1}R'Q_{xx}^{-1} \\ &+ Q_{xx}^{-1}R(R'Q_{xx}^{-1}R)^{-1}R'V_\beta R(R'Q_{xx}^{-1}R)^{-1}R'Q_{xx}^{-1} \end{aligned}$$

Exercise 7.11

From exercise 7.9,

$$\begin{aligned} V_\beta^* = V_\beta(V_\beta^{-1}) &= V_\beta - V_\beta R(R'V_\beta R)^{-1}R'V_\beta - V_\beta R(R'V_\beta R)^{-1}R'V_\beta \\ &+ V_\beta R(R'V_\beta R)^{-1}R'V_\beta R(R'V_\beta R)^{-1}R'V_\beta \\ &= V_\beta - V_\beta R(R'V_\beta R)^{-1}R'V_\beta \end{aligned}$$

Exercise 7.12

To show $V_\beta^* \leq V_\beta$, it is enough to show that $V_\beta - V_\beta^* = V_\beta R(R'V_\beta R)^{-1}R'V_\beta \geq 0$.

Under assumption 7.5.1, $V_\beta = (\mathbb{E}x_i x_i')^{-1}(\mathbb{E}x_i x_i' e_i^2)(\mathbb{E}x_i x_i')^{-1}$ is positive definite, thus $(R'V_\beta R)$ is also positive definite. By the spectral decomposition, $R'V_\beta R = H\Lambda H'$, where $H'H = I$ with strictly positive diagonal elements of Λ. Thus $V_\beta - V_\beta^* = V_\beta R(R'V_\beta R)^{-1}R'V_\beta = G'G \geq 0$ where $G = \Lambda^{-1/2}HR'V_\beta$

Exercise 7.13

(7.29)

Note that $\tilde{\beta}_{1,cls} = \hat{Q}_{11}^{-1}\hat{Q}_{1y} = \beta_1 + \left(\frac{1}{n}X_1'X_1\right)^{-1}\frac{1}{n}X_1 e$. Thus under the assumption 6.1.2 and $\beta_2 = 0$,

$$\sqrt{n}(\tilde{\beta}_{1,cls} - \beta_1) \to_d N(0, Q_{11}^{-1}\Omega_{11}Q_{11}^{-1}),$$

where $Q_{11} = \mathbb{E}(x_{1i}x_{1i}')$ and $\Omega_{11} = \mathbb{E}(x_{1i}x_{1i}'e_i^2)$. Thus, $\mathrm{avar}(\tilde{\beta}_{1,cls}) = Q_{11}^{-1}\Omega_{11}Q_{11}^{-1}$.

(7.30)

$$\begin{aligned}
\tilde{\beta}_{md} &= \hat{\beta} - \begin{pmatrix} \hat{V}_{11} & \hat{V}_{12} \\ \hat{V}_{21} & \hat{V}_{22} \end{pmatrix} \begin{bmatrix} 0 \\ I \end{bmatrix} \left([0 \quad I] \begin{pmatrix} \hat{V}_{11} & \hat{V}_{12} \\ \hat{V}_{21} & \hat{V}_{22} \end{pmatrix} \begin{bmatrix} 0 \\ I \end{bmatrix} \right)^{-1} ([0 \quad I]\hat{\beta}) \\
&= \hat{\beta} - \begin{bmatrix} \hat{V}_{12} \\ \hat{V}_{22} \end{bmatrix} (\hat{V}_{22}^{-1})\hat{\beta}_2 \\
&= \begin{pmatrix} \hat{\beta}_1 - \hat{V}_{12}\hat{V}_{22}^{-1}\hat{\beta}_2 \\ 0 \end{pmatrix}
\end{aligned}$$

Thus, $\tilde{\beta}_{1,md} = \hat{\beta}_1 - \hat{V}_{12}\hat{V}_{22}^{-1}\hat{\beta}_2$

(7.31) With theorem 7.6.1 and similar calculations above with $R = [0, I]'$, we get

$$\begin{aligned}
V_\beta^* &= V_\beta - V_\beta R (R' V_\beta R)^{-1} R' V_\beta \\
&= \begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix} - \begin{pmatrix} V_{12} \\ V_{22} \end{pmatrix} V_{22}^{-1} \begin{pmatrix} V_{21} & V_{22} \end{pmatrix} \\
&= \begin{pmatrix} V_{11} - V_{12}V_{22}^{-1}V_{21} & 0 \\ 0 & 0 \end{pmatrix}
\end{aligned}$$

Thus, $\mathrm{avar}(\tilde{\beta}_{1,md}) = V_{11} - V_{12}V_{22}^{-1}V_{21}$.

Exercise 7.14

$$\begin{aligned}
\mathrm{avar}(\hat{\beta}_1) &= Q_{11\cdot 2}^{-1}(\Omega_{11} - Q_{12}Q_{22}^{-1}\Omega_{21} - \Omega_{12}Q_{22}^{-1}Q_{21} + Q_{12}Q_{22}^{-1}\Omega_{22}Q_{22}^{-1}Q_{21})Q_{11\cdot 2}^{-1} \\
&= \frac{4}{3}\left(1 - \frac{1}{2}\cdot\frac{7}{8} - \frac{7}{8}\cdot\frac{1}{2} + \frac{1}{2}\cdot\frac{1}{2}\right)\frac{4}{3} = \frac{2}{3}
\end{aligned}$$

$$\mathrm{avar}(\tilde{\beta}_{1,cls}) = Q_{11}^{-1}\Omega_{11}Q_{11}^{-1} = 1$$

By matrix calculations, we get

$$V_\beta = \begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix} = \begin{pmatrix} \frac{2}{3} & \frac{1}{6} \\ \frac{1}{6} & \frac{2}{3} \end{pmatrix}.$$

$$\mathrm{avar}(\tilde{\beta}_{1,md}) = V_{11} - V_{12}V_{22}^{-1}V_{21} = \frac{2}{3} - \frac{1}{6}\cdot\frac{3}{2}\cdot\frac{1}{6} = \frac{5}{8}$$

Exercise 7.15

(a)

I dropped $married_2$ since it has only one value different than zero in the sample. We don't have enough variation in that variable to learn about its coefficient. Table 1 summarizes the results. Note that the reported $\widetilde{R}^2$ corresponds to the leave-one-out cross validation R-squared, a much better measure of fit than the standard measures R-squared and Theil's R-squared adjusted.

Variable	OLS estimate	Std. error
education	0.0874	0.0029
experience	0.0284	0.0028
experience^2/100	-0.0371	0.0055
married$_1$	0.1822	0.0250
married$_3$	-0.0385	0.0556
widowed	0.2375	0.1817
divorced	0.0748	0.0453
separated	0.0177	0.0531
constant	1.1906	0.0462
$\widetilde{R}^2 = 0.244$		
$n = 4{,}230$		

Table 1: OLS estimates for the log(Wage) equation using sub-samples of white male hispanics ($n = 4230$). Standard errors are heteroskedastic-robust (Horn-Horn-Duncan formula).

(b), (c)

For constrained least-squares estimator (CLS) and efficient minimum distance estimator (EMD), I use the following estimators and asymptotic standard errors.

$$
\begin{aligned}
\tilde{\beta}_{cls} &= \hat{\beta} - (X'X)^{-1}R(R'(X'X)^{-1}R)^{-1}(R'\hat{\beta} - c) \\
\tilde{\beta}_{emd} &= \hat{\beta} - \hat{V}_\beta R(R'\hat{V}_\beta R)^{-1}(R'\hat{\beta} - c) \\
\hat{V}_{cls} &= \hat{V}_\beta - Q_{xx}^{-1}R(R'Q_{xx}^{-1}R)^{-1}R'\hat{V}_\beta - \hat{V}_\beta R(R'Q_{xx}^{-1}R)^{-1}R'Q_{xx}^{-1} \\
&\quad + \quad Q_{xx}^{-1}R(R'Q_{xx}^{-1}R)^{-1}R'\hat{V}_\beta R(R'Q_{xx}^{-1}R)^{-1}R'Q_{xx}^{-1} \\
\hat{V}_{emd} &= \hat{V}_\beta - \hat{V}_\beta R(R'\hat{V}_\beta R)^{-1}R'\hat{V}_\beta,
\end{aligned}
$$

where $\hat{\beta}$ is OLS estimates from (a), $R = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 \end{pmatrix}'$, $c = (0,0)'$ and I use $\hat{V}_\beta = \bar{V}_\beta = n\bar{V}_{\hat{\beta}} = n(X'X)^{-1}(\sum_{i=1}^n (1-h_{ii})^{-1}x_i x_i' \hat{e}_i^2)(X'X)^{-1}$ with OLS residuals $\hat{e}_i = y_i - x_i'\beta$. (One can alternatively use CLS, EMD residuals as well.)

Variable	$\widehat{\beta}$	$se(\widehat{\beta})$	$\widehat{\beta}_{cls}$	$se(\widehat{\beta}_{cls})$	$\widehat{\beta}_{emd}$	$se(\widehat{\beta}_{emd})$
education	0.0874	0.0029	0.0875	0.0029	0.0876	0.0029
experience	0.0284	0.0028	0.0285	0.0028	0.0286	0.0028
experience^2/100	-0.0371	0.0055	-0.0373	0.0055	-0.0375	0.0055
married$_1$	0.1822	0.0250	0.1818	0.0250	0.1817	0.0250
married$_3$	-0.0385	0.0556	-0.0388	0.0556	-0.0388	0.0556
widowed	0.2375	0.1817	0.1818	0.0250	0.1817	0.0250
divorced	0.0748	0.0453	0.0564	0.0382	0.0518	0.0379
separated	0.0177	0.0531	0.0564	0.0382	0.0518	0.0379
constant	1.1906	0.0462	1.1881	0.0460	1.1873	0.0460

Table 2: Estimates for the log(Wage) equation using constrained least-squares and efficient minimum distance estimation subject to $R'\beta = 0$. Standard errors are heteroskedastic-robust (Horn-Horn-Duncan formula).

(d) Assume experience is non-negative

$$\beta_2 + \frac{\beta_3}{50} experience \geq 0 \quad \text{for} \quad experience \in [0, 50] \Longleftrightarrow \beta_2 \geq 0, \beta_2 + \beta_3 \geq 0$$

(Due to the way we defined *experience*, there are 4 observations with negative experience in the data; the minimum value is -2. Allowing experience to be negative is not a big issue, since we could just replace the inequality constraints for $\beta_2 - 1/25\beta_3 \geq 0$ and $\beta_2 + \beta_3 \geq 0$. This modification doesn't affect the following numerical results.)

(e)

Consider following minimum distance estimation with inequality constraints and with general weight matrix W_n. $\hat{\beta}$ denotes OLS estimates.

$$\begin{aligned}
\tilde{\beta}(W_n) &= \underset{R'\beta=c, A'\beta\leq b}{\arg\min} \; n(\hat{\beta}-\beta)'W_n(\hat{\beta}-\beta) \\
&= \underset{R'\beta=c, A'\beta\leq b}{\arg\min} \; \beta'(nW_n)\beta - 2(nW_n\hat{\beta})'\beta \\
&= \underset{R'\beta=c, A'\beta\leq b}{\arg\min} \; \frac{1}{2}\beta'H\beta + f'\beta
\end{aligned}$$

where $H = nW_n$, $f = -nW_n\hat{\beta}$, and $(R'\beta = c, A\beta \leq b) \Leftrightarrow (\beta_4 = \beta_7, \beta_8 = \beta_9, -\beta_2 \leq 0, -\beta_2 - \beta_3 \leq 0)$. This is a *quadratic programming* problem that can be solved with MATLAB (see the functions `quadprog`—quadratic programming—and `lsqlin`—constrained least squares). I used `quadprog` and `lsqlin` and got the same results. To use `lsqlin` for the efficient minimum distance estimator we have to decompose the weight matrix using the Cholesky decomposition, say find a matrix C_n such that $W_n = C_n'C_n$ and

$$\begin{aligned}
\tilde{\beta}(W_n) &= \underset{R'\beta=c, A'\beta\leq b}{\arg\min} \; n(\hat{\beta}-\beta)'W_n(\hat{\beta}-\beta) \\
&= \underset{R'\beta=c, A'\beta\leq b}{\arg\min} \; n(\hat{\beta}-\beta)'C_n'C_n(\hat{\beta}-\beta) \\
&= \underset{R'\beta=c, A'\beta\leq b}{\arg\min} \; n(C_n\hat{\beta}-C_n\beta)'(C_n\hat{\beta}-C_n\beta) \\
&= \underset{R'\beta=c, A'\beta\leq b}{\arg\min} \; n(Z-S\beta)'(Z-S\beta)
\end{aligned}$$

, which is solving the constrained least squares problem of $Z = C_n\widehat{\beta}$ on $S = C_n$. Table 3 shows the estimated parameters for the CLS estimator and the efficient minimum distance estimator.

	$\widehat{\beta}_{cls}$ (lsqlin)	$\widehat{\beta}_{cls}$ (quadprog)	$\widehat{\beta}_{emd}$ (lsqlin)	$\widehat{\beta}_{emd}$ (quadprog)
education	0.0889	0.0889	0.0884	0.0884
experience	0.0216	0.0216	0.0216	0.0216
experience^2/100	-0.0216	-0.0216	-0.0216	-0.0216
married$_1$	0.1906	0.1906	0.1930	0.1930
married$_3$	-0.0316	-0.0316	-0.0321	-0.0321
widowed	0.1906	0.1906	0.1930	0.1930
divorced	0.0642	0.0642	0.0671	0.0671
separated	0.0642	0.0642	0.0671	0.0671
constant	1.2181	1.2181	1.2222	1.2222

Table 3: Estimation results for log(Wage) equation with inequality constraints.