AE HW3

Chunyu Qu

Econometrics 1
Philip Shaw
Problem Set 3
Spring 2021

Chapter 4 Problems:
4.1, 4.2, 4.3, 4.5

1. OLS properties.

The baseline is $E[u] = cov(x_i, u) = 0$

$E[u|x] = 0 \overset{LIE}{\Rightarrow} E[u] = E[E[u|x]|x] = 0$

$cov(u, x) = E[u'x] - E[u]E[x]$

$\because E[ux|x] = x E[u|x] = 0$

$\therefore E[u'x] = E[E[u'x|x]|x] = 0$

$\therefore cov(u, x) = 0$

$\boxed{\text{OLS1}}$ $E[\underset{\sim}{u}'\underset{\sim}{x}] = 0$

* This is a week form of uncorrelation condition. Note $E[\underset{\sim}{u}|\underset{\sim}{x}] \Rightarrow E[\underset{\sim}{u}'\underset{\sim}{x}] = 0$

$\boxed{\text{OLS2}}$ $rank(E[x'x]) = K$, x is $1 \times K$ vector of regressors

* Since $x'x$ is automatically symmetric and non-singular, OLS2. furtherly makes it Positive Definite (PD).

This means $(x'x)^{-1}x'y$ is valid, i.e. no multicollinearity issue.

$\boxed{\text{Thm. 4.1}}$

OLS1 & OLS2 $\Rightarrow$ consistent $\hat{\beta}$ for β.

$\Rightarrow$ identification ✓

$\boxed{\text{OLS3.}}$ $E[u^2\underset{\sim}{x}'\underset{\sim}{x}] = \sigma^2 E[\underset{\sim}{x}'\underset{\sim}{x}]$

* OLS3. is the weakest form of HMSKD. Note $Var(u|x) = E[u^2|x] = \sigma^2$

$\Downarrow$

OLS 3.

- Find the AST-distribution

(i) When only OLS1, OLS2 hold.

$$\sqrt{N}(\hat{\beta} - \beta) \overset{a}{\sim} A^{-1} N(0, B)$$

$$A = E[\underset{\sim}{x}'\underset{\sim}{x}], \quad B = E[u^2\underset{\sim}{x}'\underset{\sim}{x}]$$

$$Avar\sqrt{N}(\hat{\beta} - \beta) = (A^{-1})'BA^{-1} = A^{-1}BA^{-1}$$

as A^{-1} is symetric, PD.

(ii) when OLS1 - OLS3 all hold.

We apply Theorem 4.2.

$$\sqrt{N}(\hat{\beta} - \beta) \overset{a}{\sim} N(0, \sigma^2 A^{-1})$$

$$Avar\sqrt{N}(\hat{\beta} - \beta) = \sigma^2 A^{-1}$$

$$Avar(\hat{\beta}) = \frac{\sigma^2 A^{-1}}{N}.$$

$$ASE(\hat{\beta}) = \sigma\sqrt{\frac{A^{-1}}{N}}$$

2. Endogeneity

- x_j is endogenous if $cov(x_j, u) \neq 0$

  exogenous if $cov(x_j, u) = 0$

- Sources of endg.

① Omitted variable

② measurement error $cov(x^*, u) \neq 0$

③ Simultaneity $x \leftrightarrow y$

4.1. Consider a standard log(*wage*) equation for men under the assumption that all explanatory variables are exogenous:

$$\log(wage) = \beta_0 + \beta_1 married + \beta_2 educ + \mathbf{z}\boldsymbol{\gamma} + u \tag{4.49}$$

$$\mathrm{E}(u \mid married, educ, \mathbf{z}) = 0$$

where $\mathbf{z}$ contains factors other than marital status and education that can affect wage. When β_1 is small, $100 \cdot \beta_1$ is approximately the ceteris paribus percentage difference in wages between married and unmarried men. When β_1 is large, it is preferable to use the exact percentage difference in $\mathrm{E}(wage \mid married, educ, \mathbf{z})$. Call this θ_1.

a. Show that, if u is independent of all explanatory variables in equation (4.49), then $\theta_1 = 100 \cdot [\exp(\beta_1) - 1]$. [Hint: Find $\mathrm{E}(wage \mid married, educ, \mathbf{z})$ for $married = 1$ and $married = 0$, and find the percentage difference.] A natural, consistent, estimator of θ_1 is $\hat{\theta}_1 = 100 \cdot [\exp(\hat{\beta}_1) - 1]$, where $\hat{\beta}_1$ is the OLS estimator from equation (4.49).

Step 1. ② obtain $E[wage \mid married, educ, z]$

(4.49) $\Rightarrow$ $wage = \exp\{\beta_0 + \beta_1 married + \beta_2 educ + \underset{\sim}{z}\gamma + u\}$

$= e^{u} \cdot \exp\{\beta_0 + \beta_1 married + \beta_2 educ + \underset{\sim}{z}\gamma\}$.

As given $E[u \mid \underset{\sim}{x}] = 0 \Rightarrow E[u] = 0,\ cov(u, x_j) = 0$

By $E[g(u)] = E[g(u) \mid \underset{\sim}{x}]$ $g \in C(\mathbb{R})$ $E[e^u] = E[e^u \mid x] \equiv \delta_0$

which is an arbitrary real value.

Thus, $E[wage \mid married, educ, z] = E[e^u \exp\{\underset{\sim}{x}\} \mid \underset{\sim}{x}] = E[e^u \mid \underset{\sim}{x}] \exp\{\underset{\sim}{x}\}$

$= \delta_0 \exp\{\beta_0 + \beta_1 married + \beta_2 educ + \underset{\sim}{z}\gamma\}$

$$= \begin{cases} \delta_0 \exp\{\beta_0 + \beta_1 + \beta_2 educ + \underset{\sim}{z}\gamma\}, & \text{for } married = 1 \\ \delta_0 \exp\{\beta_0 + \beta_2 educ + z\gamma\} & \text{for } married = 0 \end{cases}$$

Step 2. obtain θ_1

① $$\theta_1 = \frac{E[wage \mid married = 1, educ, z] - E[wage \mid married = 0, educ, z]}{E[wage \mid married = 0, educ, z]} \times 100$$

$= (e^{\beta_1} - 1) \times 100$

b. Use the delta method (see Section 3.5.2) to show that asymptotic standard error of $\hat{\theta}_1$ is $[100 \cdot \exp(\hat{\beta}_1)] \cdot \mathrm{se}(\hat{\beta}_1)$.

Recap $\quad Ase(\hat{\gamma}_N) = [c'(\hat{\theta}_N)] \, Ase(\hat{\theta}_N)$

where $\quad \gamma_N = c(\theta_N) \qquad c \in C(\mathbb{R})$

Compute $c'(\theta_N)$

$$\hat{\theta}_1 \equiv c(\hat{\beta}_1) = (e^{\hat{\beta}_1} - 1) \times 100$$

$$\therefore \quad c'(\hat{\theta}_1) = 100 e^{\hat{\beta}_1}$$

$$\therefore \quad Ase(\hat{\theta}_1) = 100 \, e^{\hat{\beta}_1} \, Ase(\hat{\beta}_1)$$

c. Repeat parts a and b by finding the exact percentage change in E(*wage* | *married*, *educ*, **z**) for any given change in *educ*, $\Delta educ$. Call this θ_2. Explain how to estimate θ_2 and obtain its asymptotic standard error.

Take edu as an example.

$$E[wage \mid married, educ, \underset{\sim}{z}] = \begin{cases} \delta_0 \exp\{\beta_0 + \beta_1 married + \beta_2 educ + \underset{\sim}{z}\gamma\} \\ \delta_0 \exp\{\beta_0 + \beta_1 married + \beta_2 (educ + \Delta educ) + \underset{\sim}{z}\gamma\} \end{cases}$$

$$\theta_2 = \frac{E[wage \mid married, educ + \Delta educ, \underset{\sim}{z}] - E[wage \mid married, educ, \underset{\sim}{z}]}{E[wage \mid married, educ, \underset{\sim}{z}]} \cdot 100$$

$$= (e^{\Delta educ \beta_2} - 1) \cdot 100$$

$$\therefore \quad Ase(\hat{\theta}_2) = 100 \cdot \Delta educ \cdot e^{\Delta educ \cdot \beta_2} \, Ase(\hat{\beta}_2)$$

d. Use the data in NLS80.RAW to estimate equation (4.49), where **z** contains the remaining variables in equation (4.29) (except ability, of course). Find $\hat{\theta}_1$ and its standard error: find $\hat{\theta}_2$ and its standard error when $\Delta educ = 4$.

d. For the estimated version of equation (4.29), $\hat{\beta}_1 = .199$, $\mathrm{se}(\hat{\beta}_1) = .039$, $\hat{\beta}_2 = .065$, and $\mathrm{se}(\hat{\beta}_2) = .006$. Therefore, $\hat{\theta}_1 = 22.01$ and $\mathrm{se}(\hat{\theta}_1) = 4.76$. For $\hat{\theta}_2$ we set $\Delta educ = 4$. Then $\hat{\theta}_2 = 29.7$ and $\mathrm{se}(\hat{\theta}_2) = 3.11$.

4.2. a. Show that, under random sampling and the zero conditional mean assumption $E(u \mid \mathbf{x}) = 0$, $E(\hat{\boldsymbol{\beta}} \mid \mathbf{X}) = \boldsymbol{\beta}$ if $\mathbf{X}'\mathbf{X}$ is nonsingular. (Hint: Use Property CE.5 in the appendix to Chapter 2.)

Pf: With OLS.1 and OLS.2, we have that

$\because \; \hat{\underset{\sim}{\beta}} = (X'X)^{-1}X'Y, \quad Y = X\beta + U$

$\therefore \; \hat{\underset{\sim}{\beta}} = (X'X)^{-1}X'X\beta + (X'X)^{-1}X'u$

$= \underset{\sim}{\beta} + (X'X)^{-1}X'u$

$\therefore \; E[\hat{\underset{\sim}{\beta}} \mid X] = E[\underset{\sim}{\beta} \mid X] + E[(X'X)^{-1}X'u \mid X]$

$= \beta + (X'X)^{-1}X' E[u \mid X] \qquad (1)$

$\because \; E(u \mid \underset{\sim}{x}) = 0$ i.e. u_i and x_i are independent for any i

Thus, by CE.5 as follows.

PROPERTY CE.5: If the vector $(\mathbf{u}, \mathbf{v})$ is independent of the vector $\mathbf{x}$, then $E(\mathbf{u} \mid \mathbf{x}, \mathbf{v}) = E(\mathbf{u} \mid \mathbf{v})$.

(u, x_i) is independent of all other $x_1, \cdots x_{i-1}, x_{i+1} \cdots x_n$.

$\therefore \; E[u \mid [x_1, \cdots, x_N]] = E[u \mid x_i]$

$\Rightarrow E[u \mid X] = E[u \mid \underset{\sim}{x}] = 0$

$\therefore (1) \Rightarrow \quad E[\hat{\beta} \mid X] = \beta$

b. In addition to the assumptions from part a, assume that $\text{Var}(u \mid \mathbf{x}) = \sigma^2$. Show that $\text{Var}(\hat{\boldsymbol{\beta}} \mid \mathbf{X}) = \sigma^2(\mathbf{X}'\mathbf{X})^{-1}$.

pf: we need to absorb $Var(u|\underset{\sim}{x}) = \sigma^2$ into $Var(\hat{\beta}|X)$

One way to do that is use $\hat{\beta} = \beta + (X'X)^{-1}X'U$

$$Var(\hat{\beta}|X) = Var(\beta + (X'X)^{-1}X'U|X)$$

$$= Var((X'X)^{-1}X'U|X)$$

$E[u_i, u_j|X] = E[u_i, u_j|x_i, x_j]$

$E[u_i|x_i, u_i, x_j] = E[u_i|x_i] = 0$

$\Rightarrow E[u_i u_j|x_i, u_j, x_j]$

$= u_j E[u_i|x_i, u_j, x_j]$

$= 0$ $\rightarrow$ CE.5

$$= (X'X)^{-1}X' \, Var(U|X) \, ((X'X)^{-1}X')'$$

$$((X'X)^{-1}X')' = X[(X'X)']^{-1} = X(X'X)^{-1}$$

$$= (X'X)^{-1}X' Var(U|X) \, X(X'X)^{-1}$$

$cov(u_i, u_j|X)$

$= E[u_i u_j|X] - E[u_i|X]E[u_j|X]$, with $E[u_i u_j|X] = E[u_i u_j|x_i, u_j, x_j]$, $E[u_i|X] = 0$, $E[u_j|X] = 0$

$= 0$

$$Var(U|X) = \begin{bmatrix} Var(u_1|X) & \underset{= 0}{cov(u_1, u_2|X)} & \cdots & \underset{= 0}{cov(u_1, u_N|X)} \\ & \ddots & & \\ & & \ddots & \\ & & & Var(u_N|X) \end{bmatrix}$$

$$\because Var(U|X) = E[U'U|X] - (\underset{= 0}{E[U|X]})^2$$

i.e. $Var[u_i|x_i] = E[u_i^2|x_i] = \sigma^2$

$\therefore Var(U|X) = \sigma^2 I$

$$\therefore Var(\hat{\beta}|X) = [(X'X)^{-1}X'X(X'X)^{-1}]\sigma^2 = \sigma^2(X'X)^{-1}$$

4.3. Suppose that in the linear model (4.5), $\mathrm{E}(\mathbf{x}'u) = \mathbf{0}$ (where $\mathbf{x}$ contains unity), $\mathrm{Var}(u \mid \mathbf{x}) = \sigma^2$, but $\mathrm{E}(u \mid \mathbf{x}) \neq \mathrm{E}(u)$.

$\Downarrow$ $E[u|\underset{\sim}{x}] = 0$

a. Is it true that $\mathrm{E}(u^2 \mid \mathbf{x}) = \sigma^2$?

$$E(u^2|\underset{\sim}{x}) = Var(u|\underset{\sim}{x}) - (E[u|\underset{\sim}{x}])^2$$

$$= \sigma^2 - (E[u|\underset{\sim}{x}])^2 \neq \sigma^2$$

Since if $E[u^2|\underset{\sim}{x}] = 0$, $\Rightarrow E[u] = 0$, but $E[u|\underset{\sim}{x}] \neq E[u]$

$\therefore E[u^2|\underset{\sim}{x}] \neq 0$

b. What relevance does part a have for OLS estimation?

As denoted $E[x'u]$ – OLS 1, Full rank — OLS 2

HMSKD – OLS 3.

Since $V(u|x) = \sigma^2$ is a stronger version than OLS3 in our text.

we simply consider as HMSKD.

(1) First, we know that $\hat{\beta}_{OLS}$ is identified

(2) We cannot infer if $\hat{\beta}_{OLS}$ is unbiased or not. But with $E[u|x]$, we infer $\hat{\beta}_{OLS}$ is unbiased.

(3) One key property is that $E[u|\underset{\sim}{x}] = 0 \Rightarrow$ HMSKD, and $E[u|\underset{\sim}{x}] = 0$ is a strong condition that implies $E[u'x] = 0$, $E[u] = 0$. but not vice versa. Thus, with $E[x'u] = 0$, unless we are given HMSKD, we cannot say anything about the variance.

(4) $E[u|\underset{\sim}{x}] = 0$ is also necessary for consistency, but we only have $E[u'x] = 0$, so we cannot garantee consistency.

4.5. Let y and z be random scalars, and let $\mathbf{x}$ be a $1 \times K$ random vector, where one element of $\mathbf{x}$ can be unity to allow for a nonzero intercept. Consider the population model

$$\mathrm{E}(y \mid \mathbf{x}, z) = \mathbf{x}\boldsymbol{\beta} + \gamma z \tag{4.50}$$

$$\mathrm{Var}(y \mid \mathbf{x}, z) = \sigma^2 \tag{4.51}$$

where interest lies in the $K \times 1$ vector $\boldsymbol{\beta}$. To rule out trivialities, assume that $\gamma \neq 0$. In addition, assume that $\mathbf{x}$ and z are orthogonal in the population: $\mathrm{E}(\mathbf{x}'z) = \mathbf{0}$.

Consider two estimators of $\boldsymbol{\beta}$ based on N independent and identically distributed observations: (1) $\hat{\boldsymbol{\beta}}$ (obtained along with $\hat{\gamma}$) is from the regression of y on $\mathbf{x}$ and z; (2) $\tilde{\boldsymbol{\beta}}$ is from the regression of y on $\mathbf{x}$. Both estimators are consistent for $\boldsymbol{\beta}$ under equation (4.50) and $\mathrm{E}(\mathbf{x}'z) = \mathbf{0}$ (along with the standard rank conditions).

a. Show that, without any additional assumptions (except those needed to apply the law of large numbers and central limit theorem), Avar $\sqrt{N}(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta})$ − Avar $\sqrt{N}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})$ is always positive semidefinite (and usually positive definite). Therefore—from the standpoint of asymptotic analysis—it is always better under equations (4.50) and (4.51) to include variables in a regression model that are uncorrelated with the variables of interest.

Pf: We know that $E[\underset{\sim}{x}'\underset{\sim}{x}]$ is PSD, intuitively, we use theorem 4.2 to estimate asymtotic dist'n.

Step 1. Estimate $\sqrt{N}(\hat{\beta} - \beta)$ with $y = x\beta + \gamma z + u$

Check: OLS 1. $E(y \mid \underset{\sim}{x}, z) = x\underset{\sim}{\beta} + \gamma z \Rightarrow E[u] = 0$, $E[x'u] = \underset{\substack{\| \\ 0}}{\mathrm{Cov}(x', u)}^{\text{iid}} - E[x']\underset{\substack{\| \\ 0}}{E[u]} = 0$ ✓

OLS 3. $\mathrm{Var}(y \mid \underset{\sim}{x}, z) = \sigma^2$ is a stronger HMSKD than $E[u^2 x'x] = \sigma^2 E[\underset{\sim}{x}'\underset{\sim}{x}]$ ✓

OLS 2. holds unless special cases such as perfect collinearity ✓

THEOREM 4.2 (Asymptotic Normality of OLS): Under Assumptions OLS.1–OLS.3,

$$\sqrt{N}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \overset{a}{\sim} \mathrm{Normal}(0, \sigma^2 \mathbf{A}^{-1}) \tag{4.9}$$

Let $E(y \mid \underset{\sim}{x}, \underset{\sim}{z}) \equiv E[y \mid \underset{\sim}{w}] = \underset{\sim}{w}\underset{\sim}{\delta}$ where $\underset{\sim}{\delta} = \begin{bmatrix} \underset{\sim}{\beta} \\ \gamma \end{bmatrix}$ $\underset{\sim}{W} = [\underset{1\times K}{\underset{\sim}{x}}, \underset{1\times 1}{z}]$

$\mathrm{Var}[y \mid \underset{\sim}{w}] = \sigma^2$

By theorem 4.2

$\Rightarrow \mathrm{Avar}\sqrt{N}(\hat{\delta} - \delta) = \sigma^2 A^{-1}$

where $A = E[W'W] = E\left[\begin{bmatrix}\underset{\sim}{X} \\ z\end{bmatrix}[\underset{\sim}{X}\ z]\right]$

$$= E\left[\begin{bmatrix} \underbrace{\underset{\sim}{X}'X}_{k\times k} & \underbrace{z\underset{\sim}{X}'}_{k\times 1} \\ \underbrace{z\underset{\sim}{X}}_{1\times k} & \underbrace{z^2}_{1\times 1}\end{bmatrix}_{(k+1)\times(k+1)}\right] \overset{E(X'z)=0}{=} \begin{bmatrix} E[X'X] & 0 \\ 0 & z^2\end{bmatrix}$$

By the property of inverse of block matrix

$$A^{-1} = \begin{bmatrix} E[X'X]^{-1}_{k\times k} & 0 \\ 0 & \frac{1}{z^2}\end{bmatrix}$$

Thus $\sqrt{N}(\hat{\delta} - \delta) = \sqrt{N}\begin{bmatrix}\hat{\beta}-\beta \\ \hat{\gamma}-\gamma\end{bmatrix} \sim N\left(0,\ \sigma^2\begin{bmatrix} E[X'X]^{-1} & 0 \\ 0 & \frac{1}{z^2}\end{bmatrix}\right)$

$\Rightarrow \sqrt{N}(\hat{\beta}-\beta) \sim N(0,\ \sigma^2 E[X'X]^{-1})$ ①.

$\sqrt{N}(\hat{\gamma}-\gamma) \sim N\left(0,\ \sigma^2 \frac{1}{z^2}\right)$

Step 2. Estimate $\sqrt{N}(\tilde{\beta} - \beta)$ with $y = x\beta + V$

$\tilde{\beta}$ is obtained with a short form treating γz as an augmented error term rewrite as.

$y = \underset{\sim}{x}\underset{\sim}{\beta} + \gamma z + u = x\underset{\sim}{\beta} + V$. with $V = \gamma z + u$

and $u = y - E[y \mid \underset{\sim}{x}, z]$

Check. OLS1. $E[x'V] = E[x'(\gamma z + u)] = \gamma \overset{0}{\overset{\shortparallel}{E(x'z)}} + E[x'u] = 0$ ✓. OLS2 ✓

OLS3. W.T.S. $E[V^2\underset{\sim}{x}'\underset{\sim}{x}] = E[V^2] \cdot E[\underset{\sim}{x}'x]$.

where $E[V^2] = E[\gamma^2 z^2] + 2\gamma E[\gamma z u] + E[u^2] = \gamma^2 z^2 + 2\gamma^2 E[zu] + \sigma^2$

i.e. $E[V^2\underset{\sim}{x}'\underset{\sim}{x}] \overset{?}{=} [\gamma^2 z^2 + 2\gamma^2 E[zu] + \sigma^2]\ E[x'x]$ (A)

$\because E[V^2\underset{\sim}{x}'\underset{\sim}{x}] \overset{LIE}{=} E[E(V^2\underset{\sim}{x}'\underset{\sim}{x} \mid \underset{\sim}{x})] = E[x'x\, E[V^2 \mid \underset{\sim}{x}]]$

$= E\left[E[\underbrace{\gamma^2 z^2}_{} + \underbrace{2\gamma z u}_{} + u^2 \mid x]\, x'x\right]$

$= \gamma^2 E[E[z^2 \mid x] \cdot x'x] + 2\gamma E[E[zu \mid x]\, x'x] + E[\overbrace{E[u^2 \mid x]}^{\sigma^2}\, x'x]$

$= \gamma^2 E[R_x^{+} x'x] + \boxed{\text{By iid } \underset{\overset{\shortparallel}{0}}{E[zu \mid x]} = E[z \mid x]E[u \mid x]} + \sigma^2 E[x'x]$ (B)

As $\gamma^2 E[R_x^+ \underset{\sim}{x}' \underset{\sim}{x}] \geq 0$

Thus, equation (A) (B) are not necessarily equal. OLS 3 is not guaranteed.

In this case, we adopt the general form of theorem 4.2, as

$$\boxed{\text{OLS 1 \& OLS 2} \Rightarrow \sqrt{N}(\hat{\beta}-\beta) \sim N(0, (A^{-1})' B A^{-1}) \quad B \equiv E[u^2 \underset{\sim}{x}' \underset{\sim}{x}] \quad A = E[\underset{\sim}{x}' \underset{\sim}{x}]}$$

$$\therefore \sqrt{N}(\hat{\beta}-\beta) \sim N\left(0,\ E[\underset{\sim}{x}' x]^{-1} E[V^2 \underset{\sim}{x}' \underset{\sim}{x}] E[\underset{\sim}{x}' \underset{\sim}{x}]^{-1}\right)$$

Step 3. $\text{Avar}\sqrt{N}(\tilde{\beta}-\beta) - \text{Avar}\sqrt{N}(\hat{\beta}-\beta)$

$$= E[x'x]^{-1} E[V^2 \underset{\sim}{x}' \underset{\sim}{x}] E[\underset{\sim}{x}' \underset{\sim}{x}]^{-1} - \sigma^2 E[\underset{\sim}{x}' \underset{\sim}{x}]^{-1}$$

$$= E[\underset{\sim}{x}' \underset{\sim}{x}]^{-1} \underline{E[V^2 \underset{\sim}{x}' \underset{\sim}{x}]} E[\underset{\sim}{x}' \underset{\sim}{x}]^{-1} - \sigma^2 E[\underset{\sim}{x}' \underset{\sim}{x}]^{-1} \underline{E[\underset{\sim}{x}' \underset{\sim}{x}]} E[\underset{\sim}{x}' \underset{\sim}{x}]^{-1}$$

$$= E[\underset{\sim}{x}' \underset{\sim}{x}]^{-1} \left[E[V^2 \underset{\sim}{x}' \underset{\sim}{x}] - \sigma^2 E[x'x]\right] E[\underset{\sim}{x}' \underset{\sim}{x}]^{-1}$$

By B. $E[V^2 x' \underset{\sim}{x}] - \sigma^2 E[x'x]] = \gamma^2 E[R_x^+ \underset{\sim}{x}' \underset{\sim}{x}]$

$\therefore$ PSD $\times$ PSD $\times$ PSD $\Rightarrow$ PSD,

$\therefore \text{Avar}\sqrt{N}(\tilde{\beta}-\beta) - \text{Avar}\sqrt{N}(\hat{\beta}-\beta)$ is PSD.

b. Consider the special case where $z = (x_K - \mu_K)^2$, where $\mu_K \equiv E(x_K)$, and x_K is symetrically distributed: $E[(x_K - \mu_K)^3] = 0$. Then β_K is the partial effect of x_K on $E(y \mid \mathbf{x})$ evaluated at $x_K = \mu_K$. Is it better to estimate the average partial effect with or without $(x_K - \mu_K)^2$ included as a regressor?

$$y = x\beta + \gamma z + u$$

$$\begin{bmatrix} y \\ \downarrow \end{bmatrix}_{N\times 1} = \begin{bmatrix} x_1 & x_2 \ldots & x_k \\ \downarrow & \downarrow & \downarrow \end{bmatrix}_{N\times k} \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_k \end{bmatrix}_{k\times 1} + \hat{\gamma} \cdot \begin{bmatrix} z \\ \downarrow \end{bmatrix}_{N\times 1} + \begin{bmatrix} u \\ \downarrow \end{bmatrix}_{N\times 1}$$

Now $z = (x_k - E[x_k])^2$ which is the 2nd-centeral moment

$\beta_k = \dfrac{\partial E[y|\underset{\sim}{x}]}{\partial x_k}$, when $x_k = \mu_k$, $z = 0$, $y = x\beta + u$.

$$E_{x_k}\left[\frac{\partial E[y|\underset{\sim}{x}]}{\partial x_k}\right] = \beta_k + E_{x_k}\left[\frac{\partial (x_k - \mu_k)^2}{\partial x_k}\right] + \underset{\substack{\shortparallel \\ 0}}{E_{x_k}[u]} = \beta_k + E_{x_k}[2(x_k - \mu_k)] = 0$$

with $E[(x_k - \mu_k)^3] = 0$, we know $\text{cov}((x_k - \mu_k)^2, x_k) = 0$

So $E[zx] = 0$ still holds. (OLS1 ✓) (see prob 2.2 (c))

It is better since it allows us to compute PE of x_k at any value. We can obtain PE at $x_k = \mu_k$ with $y = x\beta + \gamma z + u$ (z is scalar), but not at other values of x_k.

c. Under the setup in Problem 2.3, with $\mathrm{Var}(y \mid \mathbf{x}) = \sigma^2$, is it better to estimate β_1 and β_2 with or without $x_1 x_2$ in the regression?

2.3. Suppose that

$$\mathrm{E}(y \mid x_1, x_2) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2 \tag{2.47}$$

a. Write this expectation in error form (call the error u), and describe the properties of u.

b. Suppose that x_1 and x_2 have zero means. Show that β_1 is the expected value of $\partial \mathrm{E}(y \mid x_1, x_2)/\partial x_1$ (where the expectation is across the population distribution of x_2). Provide a similar interpretation for β_2.

c. Now add the assumption that x_1 and x_2 are independent of one another. Show that the linear projection of y on $(1, x_1, x_2)$ is

$$\mathrm{L}(y \mid 1, x_1, x_2) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 \tag{2.48}$$

The same reason as b. Since HMSKD is guaranteed, with $x_1 x_2$ we can have more flexibility on evaluating PE on x_1, x_2.

Since $V[y \mid \underset{\sim}{x}] = V[y \mid x_1, x_2, x_1 x_2] = \sigma^2$. recap in a, $V(y \mid x_1, x_2)$ cannot be σ^2 if x_1 x_2 are not independent, most of the time, and so HMSKD is violated if we take away $x_1 x_2$, which will cause the loss of series of nice properties.