

**a)**

As given by the question:

$$b_R = (X_1'X_1)^{-1}X_1'y \quad (1)$$

$$y = X_1\beta_1 + X_2\beta_2 + \varepsilon \quad (2)$$

replacing y in (1) with (2) results:

$$b_R = (X_1'X_1)^{-1}X_1'(X_1\beta_1 + X_2\beta_2 + \varepsilon) = \beta_1 + (X_1'X_1)^{-1}X_1'X_2\beta_2 + (X_1'X_1)^{-1}X_1'\varepsilon$$

as  $(X_1'X_1)^{-1}X_1'X_2 = P$ , and  $E\left((X_1'X_1)^{-1}X_1'\varepsilon\right) = (X_1'X_1)^{-1}X_1'E(\varepsilon)$ , given the assumption

that  $E(\varepsilon) = 0$ , we get  $E(b_R) = \beta_1 + P\beta_2$

**b)**

$$Var(b_R) = Var\left((X_1'X_1)^{-1}X_1'\varepsilon\right)$$

$$\text{let } (X_1'X_1)^{-1}X_1' = A$$

$$Var(b_R) = A Var(\varepsilon)A' = A\sigma^2IA' = AA'\sigma^2$$

$$AA' = (X_1'X_1)^{-1}X_1'X_1(X_1'X_1)^{-1} = (X_1'X_1)^{-1}, \text{ therefore,}$$

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

**c)**

$$y = X_1b_1 + X_2b_2 + e \quad (3)$$

replace y in (1) with (3), we get

$$b_R = (X_1'X_1)^{-1}X_1'(X_1b_1 + X_2b_2 + e) = b_1 + Pb_2 + (X_1'X_1)^{-1}X_1'e$$

compare model  $y = Xb + e_f$  with model  $y = X_1b_1 + X_2b_2 + e$ , it can be seen that

$$e_f = e$$

$$\text{therefore, } X'e_f = 0 \Rightarrow X'e = 0 \Rightarrow X_1'e = 0$$

$$\text{Because } X'e_f = X'(y - Xb) = X'y - X'X(X'X)^{-1}X'y = 0$$

$$\text{Therefore, } (X_1'X_1)^{-1}X_1'e = 0 \text{ and } b_R = b_1 + Pb_2$$

**d)**

$P = (X_1'X_1)^{-1}X_1'X_2$ , note that for model  $y = Xb + e$ , we estimate  $\beta$  as  $b = (X'X)^{-1}X'y$ . It can be seen that if we replace  $X$  with  $X_1$ ,  $y$  with  $X_2$ , for  $i_{th}$  column in  $P$ , it is the same as

the coefficients by regressing the  $i_{th}$  column in  $X_2$ , i.e.,  $i_{th}$  column in  $P$  can be obtained by regressing  $X_1'(\text{constant, Female})$  on  $i_{th}$  column in  $X_2$  respectively (Age, Educ, Parttime).

e)

According to d), we get

$$P = \begin{pmatrix} 40.05 & 2.26 & 0.20 \\ -0.11 & -0.49 & 0.25 \end{pmatrix}$$

f)

$$P = \begin{pmatrix} 40.05 & 2.26 & 0.20 \\ -0.11 & -0.49 & 0.25 \end{pmatrix}, b_2 = \begin{pmatrix} 0.031 \\ 0.233 \\ -0.365 \end{pmatrix}, b_1 = \begin{pmatrix} 3.05 \\ -0.04 \end{pmatrix} \text{ therefore}$$

$$Pb_2 = \begin{pmatrix} 1.70 \\ -0.21 \end{pmatrix}, b_1 + Pb_2 = \begin{pmatrix} 1.70 + 3.05 \\ -0.21 - 0.04 \end{pmatrix} = \begin{pmatrix} 4.75 \\ -0.25 \end{pmatrix}$$

which is almost identical to  $b_R = \begin{pmatrix} 4.73 \\ -0.25 \end{pmatrix}$  regarding the numerical values.