Exo2

4- Montrez que
$y\,'X\beta = \beta'X'y$

Preuve 1
$y\,'X\beta = y\,'(X\beta) =$

$$= \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}' \left(\begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1K} \\ x_{21} & x_{22} & \cdots & x_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nK} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_K \end{bmatrix} \right) = \begin{bmatrix} y_1 & y_2 & \cdots & y_n \end{bmatrix} \begin{bmatrix} \sum_{k=1}^{K} x_{1k}\beta_k \\ \sum_{k=1}^{K} x_{2k}\beta_k \\ \vdots \\ \sum_{k=1}^{K} x_{nk}\beta_k \end{bmatrix} = \sum_{i=1}^{n} y_i \left(\sum_{k=1}^{K} x_{ik}\beta_k \right)$$

$\beta'X'y = (\beta'X')y = (X\beta)'y$

$$= \left(\begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1K} \\ x_{21} & x_{22} & \cdots & x_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nK} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_K \end{bmatrix} \right)' \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} \sum_{k=1}^{K} x_{1k}\beta_k \\ \sum_{k=1}^{K} x_{2k}\beta_k \\ \vdots \\ \sum_{k=1}^{K} x_{nk}\beta_k \end{bmatrix}' \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \sum_{i=1}^{n} \left(\sum_{k=1}^{K} x_{ik}\beta_k \right) y_i = \sum_{i=1}^{n} y_i \left(\sum_{k=1}^{K} x_{ik}\beta_k \right)$$

c.q.f.d.

Preuve 2 (rapido presto)
On sait que la transposée d'une matrice 1x1 $a_{1\times 1}$ est $\left(a_{1\times 1}\right)' = a_{1\times 1}$

Alors on sait que
$\underbrace{y\,'X\beta}_{1\times 1} = \underbrace{\left(y\,'X\beta\right)'}_{1\times 1} = X\,'\beta\,'y$
à partir du résultat $\left(a_{1\times 1}\right)' = a_{1\times 1}$

5-We have the following linear regression model for $K = 2$ with a constant:

$$y_i = \beta_1 + \beta_2 x_{i2} + \varepsilon_i \qquad \forall i = 1,...,n$$

In matrix form we have

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{12} \\ 1 & x_{22} \\ \vdots & \vdots \\ 1 & x_{n2} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

With the estimator $\hat{\beta} = (X'X)^{-1}X'y$ in matrix form we can write and expand thereafter into summations:

$$\hat{\beta} \equiv \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} = (X'X)^{-1}X'y = \left(\begin{bmatrix} 1 & 1 & 1 & 1 \\ x_{12} & x_{22} & \dots & x_{n2} \end{bmatrix} \begin{bmatrix} 1 & x_{12} \\ 1 & x_{22} \\ \vdots & \vdots \\ 1 & x_{n2} \end{bmatrix} \right)^{-1} \begin{bmatrix} 1 & 1 & 1 & 1 \\ x_{12} & x_{22} & \dots & x_{n2} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

The first part is

$$(X'X)^{-1} = \left(\begin{bmatrix} 1 & 1 & 1 & 1 \\ x_{12} & x_{22} & \dots & x_{n2} \end{bmatrix} \begin{bmatrix} 1 & x_{12} \\ 1 & x_{22} \\ \vdots & \vdots \\ 1 & x_{n2} \end{bmatrix} \right)^{-1} = \begin{bmatrix} \sum_{i=1}^{n} 1 & \sum_{i=1}^{n} 1 \times x_{i2} \\ \sum_{i=1}^{n} x_{i2} \times 1 & \sum_{i=1}^{n} x_{i2} \times x_{i2} \end{bmatrix}^{-1} = \begin{bmatrix} n & \sum_{i=1}^{n} x_{i2} \\ \sum_{i=1}^{n} x_{i2} & \sum_{i=1}^{n} x_{i2}^2 \end{bmatrix}^{-1}$$

The second part is

$$X'y = \begin{bmatrix} 1 & 1 & 1 & 1 \\ x_{12} & x_{22} & \dots & x_{n2} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} 1 \times y_i \\ \sum_{i=1}^{n} x_{i2} \times y_i \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} x_{i2} y_i \end{bmatrix}$$

Remember that if

$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ its inverse is given by $A^{-1} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$

From that we have that

$$(X'X)^{-1} = \frac{1}{\left(n\sum_{i=1}^{n} x_{i2}^2 - \sum_{i=1}^{n} x_{i2} \sum_{i=1}^{n} x_{i2} \right)} \begin{pmatrix} \sum_{i=1}^{n} x_{i2}^2 & -\sum_{i=1}^{n} x_{i2} \\ -\sum_{i=1}^{n} x_{i2} & n \end{pmatrix}$$

Hence,

$$\hat{\beta} = (X'X)^{-1}X'y = \frac{1}{\left(n\sum_{i=1}^{n} x_{i2}^2 - \left(\sum_{i=1}^{n} x_{i2}\right)^2\right)} \begin{pmatrix} \sum_{i=1}^{n} x_{i2}^2 & -\sum_{i=1}^{n} x_{i2} \\ -\sum_{i=1}^{n} x_{i2} & n \end{pmatrix} \begin{bmatrix} \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} x_{i2} y_i \end{bmatrix}$$

$$= \frac{1}{\left(n\sum_{i=1}^{n} x_{i2}^2 - \left(\sum_{i=1}^{n} x_{i2}\right)^2\right)} \begin{bmatrix} \left(\sum_{i=1}^{n} x_{i2}^2 \sum_{i=1}^{n} y_i\right) - \left(\sum_{i=1}^{n} x_{i2} \sum_{i=1}^{n} x_{i2} y_i\right) \\ \left(n\sum_{i=1}^{n} x_{i2} y_i\right) - \left(\sum_{i=1}^{n} x_{i2} \sum_{i=1}^{n} y_i\right) \end{bmatrix}$$

We can rewrite the denominator as[1]

$$\left(n\sum_{i=1}^{n} x_{i2}^2 - \left(\sum_{i=1}^{n} x_{i2}\right)^2\right) = n\underbrace{\left(\sum_{i=1}^{n} x_{i2}^2 - \frac{1}{n}\left(\sum_{i=1}^{n} x_{i2}\right)^2\right)}_{n\sigma_{x_2 x_2}} = nn\sigma_{x_2 x_2} = n^2 \sigma_{x_2}^2 = n^2 \operatorname{var}(x_2)$$

Thus we have

$$\hat{\beta} = \frac{1}{n^2 \sigma_{x_2 x_2}} \begin{bmatrix} \left(\sum_{i=1}^{n} x_{i2}^2 \sum_{i=1}^{n} y_i\right) - \left(\sum_{i=1}^{n} x_{i2} \sum_{i=1}^{n} x_{i2} y_i\right) \\ \left(n\sum_{i=1}^{n} x_{i2} y_i\right) - \left(\sum_{i=1}^{n} x_{i2} \sum_{i=1}^{n} y_i\right) \end{bmatrix} = \frac{1}{\sigma_{x_2 x_2}} \begin{bmatrix} \frac{1}{n^2}\left(\sum_{i=1}^{n} x_{i2}^2 \sum_{i=1}^{n} y_i\right) - \frac{1}{n^2}\left(\sum_{i=1}^{n} x_{i2} \sum_{i=1}^{n} x_{i2} y_i\right) \\ \frac{1}{n^2}\left(n\sum_{i=1}^{n} x_{i2} y_i\right) - \frac{1}{n^2}\left(\sum_{i=1}^{n} x_{i2} \sum_{i=1}^{n} y_i\right) \end{bmatrix}$$

$$= \left(\frac{1}{\sigma_{x_2 x_2}}\right) \begin{bmatrix} \left(\left(\frac{\sum_{i=1}^{n} x_{i2}^2}{n}\right)\left(\frac{\sum_{i=1}^{n} y_i}{n}\right)\right) - \left(\left(\frac{\sum_{i=1}^{n} x_{i2}}{n}\right)\left(\frac{\sum_{i=1}^{n} x_{i2} y_i}{n}\right)\right) \\ \left(\left(\frac{\sum_{i=1}^{n} x_{i2} y_i}{n}\right)\right) - \left(\left(\frac{\sum_{i=1}^{n} x_{i2}}{n}\right)\left(\frac{\sum_{i=1}^{n} y_i}{n}\right)\right) \end{bmatrix}$$

[1] Recall that

$$\operatorname{var}(x) = E(x^2) - (E(x))^2 = \left(\sum_{i=1}^{n} x_i^2 \big/ n\right) - \left(\sum_{i=1}^{n} x_i \big/ n\right)^2 = \left(\sum_{i=1}^{n} x_i^2 \big/ n\right) - \left(\frac{1}{n^2}\right)\left(\sum_{i=1}^{n} x_i\right)^2$$

$$= \left(\frac{1}{\sigma_{x_2x_2}}\right)\left[\frac{\left(\left(\frac{\sum_{i=1}^{n} x_{i2}^2}{n}\right)\overline{y}\right) - \left(\overline{x}_2\left(\frac{\sum_{i=1}^{n} x_{i2}y_i}{n}\right)\right)}{\left(\left(\frac{\sum_{i=1}^{n} x_{i2}y_i}{n}\right)\right) - \left(\left(\frac{\sum_{i=1}^{n} x_{i2}}{n}\right)\left(\frac{\sum_{i=1}^{n} y_i}{n}\right)\right)}\right]$$

From the variance and the covariance[2] equation we can write:

$$= \left(\frac{1}{\sigma_{x_2x_2}}\right)\left[\frac{\left(\left(\sigma_{x_2x_2} + \left(\frac{\sum_{i=1}^{n} x_{i2}}{n}\right)^2\right)\overline{y}\right) - \left(\overline{x}_2\left(\sigma_{x_2y} + \left(\frac{\sum_{i=1}^{n} y_i}{n}\right)\left(\frac{\sum_{i=1}^{n} x_{i2}}{n}\right)\right)\right)}{\left(\left(\frac{\sum_{i=1}^{n} x_{i2}y_i}{n}\right)\right) - \left(\left(\frac{\sum_{i=1}^{n} x_{i2}}{n}\right)\left(\frac{\sum_{i=1}^{n} y_i}{n}\right)\right)}\right]$$

$$= \left(\frac{1}{\sigma_{x_2x_2}}\right)\left[\frac{\left(\left(\sigma_{x_2x_2} + \left(\overline{x}_2\right)^2\right)\overline{y}\right) - \left(\overline{x}_2\left(\sigma_{x_2y} + \overline{y}\cdot\overline{x}_2\right)\right)}{\left(\left(\frac{\sum_{i=1}^{n} x_{i2}y_i}{n}\right)\right) - \left(\left(\frac{\sum_{i=1}^{n} x_{i2}}{n}\right)\left(\frac{\sum_{i=1}^{n} y_i}{n}\right)\right)}\right]$$

$$= \left(\frac{1}{\sigma_{x_2x_2}}\right)\left[\frac{\sigma_{x_2x_2}\overline{y} + \left(\overline{x}_2\right)^2\overline{y} - \overline{x}_2\sigma_{x_2y} - \overline{y}\cdot\left(\overline{x}_2\right)^2}{\left(\left(\frac{\sum_{i=1}^{n} x_{i2}y_i}{n}\right)\right) - \left(\left(\frac{\sum_{i=1}^{n} x_{i2}}{n}\right)\left(\frac{\sum_{i=1}^{n} y_i}{n}\right)\right)}\right]$$

---

[2] $\operatorname{cov}(x_2, y) = \sigma_{x_2y} = \left(\frac{\sum_{i=1}^{n} x_{i2}y_i}{n}\right) - \left(\frac{\sum_{i=1}^{n} y_i}{n}\right)\left(\frac{\sum_{i=1}^{n} x_{i2}}{n}\right) = \frac{\sum_{i=1}^{n} x_{i2}y_i}{n} - \overline{y}\cdot\overline{x}_2$

With the subtraction we have

$$= \left(\frac{1}{\sigma_{x_2 x_2}} \right) \left[\begin{matrix} \sigma_{x_2 x_2} \overline{y} - \overline{x}_2 \sigma_{x_2 y} \\ \left(\left(\frac{\sum_{i=1}^{n} x_{i2} y_i}{n} \right) \right) - \left(\left(\frac{\sum_{i=1}^{n} x_{i2}}{n} \right) \left(\frac{\sum_{i=1}^{n} y_i}{n} \right) \right) \end{matrix} \right]$$

And from the covariance equation again we can write:

$$= \left(\frac{1}{\sigma_{x_2 x_2}} \right) \left[\begin{matrix} \left(\sigma_{x_2 x_2} \overline{y} \right) - \left(\overline{x}_2 \sigma_{x_2 y} \right) \\ \left(\sigma_{x_2 y} \right) \end{matrix} \right] = \left[\begin{matrix} \left(\frac{\sigma_{x_2 x_2} \overline{y}}{\sigma_{x_2 x_2}} - \frac{\sigma_{x_2 y} \overline{x}_2}{\sigma_{x_2 x_2}} \right) \\ \left(\frac{\sigma_{x_2 y}}{\sigma_{x_2 x_2}} \right) \end{matrix} \right] = \left[\begin{matrix} \overline{y} - \hat{\beta}_2 \overline{x}_2 \\ \left(\frac{\sigma_{x_2 y}}{\sigma_{x_2 x_2}} \right) \end{matrix} \right] = \left[\begin{matrix} \overline{y} - \hat{\beta}_2 \overline{x}_2 \\ \left(\frac{\mathrm{cov}(x_2, y)}{\mathrm{var}(x_2)} \right) \end{matrix} \right] = \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix}$$

$$= \hat{\beta}$$

Thus for the model with one constant and one varying regressor we have:

$$\hat{\beta}_2 = \left(\frac{\sigma_{x_2 y}}{\sigma_{x_2 x_2}} \right) = \frac{\mathrm{cov}(x_2, y)}{\mathrm{var}(x_2)}$$

And from that we can compute

$$\hat{\beta}_1 = \overline{y} - \hat{\beta}_2 \overline{x}_2$$

This is a big mess, this is why we prefer the matrix notation and the more sexy and talkative matrix algebra $\hat{\beta} = (X'X)^{-1} X'y$ expression.