Sketch of the answers for the Ec501 Midterm (2011)

1. (a) **(10 marks)** $E(y_i|x_{i2}) = \beta_1 + \beta_2 x_{i2}$. The conditional expectation is the function of the regressors that best predicts y_i in the sense that it minimizes the mean square error of the predictions.

   (b) **(10 marks)**

$$(X'X)^{-1} = \begin{bmatrix} 0.25 & 0 \\ 0 & 0.10 \end{bmatrix}, \quad X'y = \begin{bmatrix} 6 \\ 3 \end{bmatrix}, \quad (X'X)^{-1} X'y = \begin{bmatrix} \frac{3}{2} \\ \frac{3}{10} \end{bmatrix}$$

   (c) **(10 marks)** $e'e = y'y - b'X'y = 10 - \frac{99}{10} = 0.1$

$$s^2 = \frac{0.1}{4-2} = 0.05$$

   (d) **(10 marks)**

$$\begin{aligned} R^2 &= 1 - \frac{RSS}{TSS} \\ RSS &= (n-K)\,\sigma^2 = 2 \times 0.05 = 0.1 \\ TSS &= \sum (y_i - \bar{y})^2 = 4 \times 0.25 = 1 \\ R^2 &= 0.9 \end{aligned}$$

   (e) **(10 marks)**

$$t_2 = \frac{0.3}{\sqrt{0.005}} = 4.2426$$

   Since $|t_2| < t^{0.025}_{n-K} = 4.303$ we do not reject the null hypothesis, suggesting that there is no relation between the two variables.

   (f) **(10 marks)** $X'e = X'M\epsilon = X'(I-P)\,\epsilon = (X' - X')\,\epsilon = 0$. This implies that $n^{-1}\sum e_i = n^{-1}\sum e_i x_{i2} = 0$ and hence

$$\bar{y} = b_1 + b_2\bar{x}_2 + n^{-1}\sum e_i = b_1 + b_2\bar{x}_2$$

2. (a) **(10 marks)** $t_4 = \frac{-0.037583+0.05}{0.0035859} = 3.4627$. Since $|t_4| > t^{0.025}_{n-K} = 1.960$ we reject the null hypothesis

   (b) **(10 marks)** Since $F = 266.42 > F^{0.05}_{3,2999} = 2.60$ we reject the null hypothesis

   (c) i. **(10 marks)**

$$\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{bmatrix} = \begin{bmatrix} 0.05 \\ 0 \end{bmatrix}$$

$$(y_i - 0.05x_{i3}) = \beta_1 + \beta_2\,(x_{i2} - x_{i4}) + \epsilon_i$$

   ii. **(10 marks)**

$$\begin{aligned} F &= \frac{3003-4}{2}\,\frac{467.083386 - 0.155587488 \times (3003-4)}{0.155587488 \times (3003-4)} \\ F(2,2999) &= 1.53 \end{aligned}$$