# Chapter 1
## Econometrics

There are no exercises or applications in Chapter 1.

# Chapter 2
## The Linear Regression Model

There are no exercises or applications in Chapter 2.

# Chapter 3
## Least Squares

### ■ Exercises

1. Let $\mathbf{X} = \begin{bmatrix} 1 & x_1 \\ \dots & \dots \\ 1 & x_n \end{bmatrix}$.

   a. The normal equations are given by (3-12), $\mathbf{X}'\mathbf{e} = \mathbf{0}$ (we drop the minus sign), hence for each of the columns of **X**, $\mathbf{x}_k$, we know that $\mathbf{X}_k'\mathbf{e} = \mathbf{0}$. This implies that $\Sigma_{i=1}^n e_i = 0$ and $\Sigma_{i=1}^n x_i e_i = 0$.

   b. Use $\Sigma_{i=1}^n e_i$ to conclude from the first normal equation that $a = \bar{y} - b\bar{x}$.

   c. We know that $\Sigma_{i=1}^n e_i = 0$ and $\Sigma_{i=1}^n x_i e_i = 0$. It follows then that $\Sigma_{i=1}^n (x_i - \bar{x})e_i = 0$ because $\Sigma_{i=1}^n \bar{x}e_i = \bar{x}\Sigma_{i=1}^n e_i = 0$. Substitute $e_i$ to obtain $\Sigma_{i=1}^n (x_i - \bar{x})(y_i - a - bx_i) = 0$. or $\Sigma_{i=1}^n (x_i - \bar{x})(y_i - \bar{y} - b(x_i - \bar{x})) = 0$.

   Then, $\Sigma_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) = b\Sigma_{i=1}^n (x_i - \bar{x})(x_i - \bar{x}))$ so $b = \dfrac{\Sigma_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\Sigma_{i=1}^n (x_i - \bar{x})^2}$.

   d. The first derivative vector of $\mathbf{e}'\mathbf{e}$ is $-2\mathbf{X}'\mathbf{e}$. (The normal equations.) The second derivative matrix is $\partial^2(\mathbf{e}'\mathbf{e})/\partial\mathbf{b}\partial\mathbf{b}' = 2\mathbf{X}'\mathbf{X}$. We need to show that this matrix is positive definite. The diagonal elements are $2n$ and $2\Sigma_{i=1}^n x_i^2$ which are clearly both positive. The determinant is $[(2n)(2\Sigma_{i=1}^n x_i^2)] - (2\Sigma_{i=1}^n x_i)^2 = 4n\Sigma_{i=1}^n x_i^2 - 4(n\bar{x})^2 = 4n[(\Sigma_{i=1}^n x_i^2) - n\bar{x}^2] = 4n[(\Sigma_{i=1}^n (x_i - \bar{x})^2]$. Note that a much simpler proof appears after (3-6).

2. Write **c** as **b** + (**c** − **b**). Then, the sum of squared residuals based on **c** is

$$\begin{aligned}(\mathbf{y} - \mathbf{X}\mathbf{c})'(\mathbf{y} - \mathbf{X}\mathbf{c}) &= [\mathbf{y} - \mathbf{X}(\mathbf{b} + (\mathbf{c} - \mathbf{b}))]'[\mathbf{y} - \mathbf{X}(\mathbf{b} + (\mathbf{c} - \mathbf{b}))] \\ &= [(\mathbf{y} - \mathbf{X}\mathbf{b}) + \mathbf{X}(\mathbf{c} - \mathbf{b})]'[(\mathbf{y} - \mathbf{X}\mathbf{b}) + \mathbf{X}(\mathbf{c} - \mathbf{b})] \\ &= (\mathbf{y} - \mathbf{X}\mathbf{b})'(\mathbf{y} - \mathbf{X}\mathbf{b}) + (\mathbf{c} - \mathbf{b})'\mathbf{X}'\mathbf{X}(\mathbf{c} - \mathbf{b}) + 2(\mathbf{c} - \mathbf{b})'\mathbf{X}'(\mathbf{y} - \mathbf{X}\mathbf{b}).\end{aligned}$$

   But, the third term is zero, as $2(\mathbf{c} - \mathbf{b})'\mathbf{X}'(\mathbf{y} - \mathbf{X}\mathbf{b}) = 2(\mathbf{c} - \mathbf{b})\mathbf{X}'\mathbf{e} = \mathbf{0}$. Therefore,

$$(\mathbf{y} - \mathbf{X}\mathbf{c})'(\mathbf{y} - \mathbf{X}\mathbf{c}) = \mathbf{e}'\mathbf{e} + (\mathbf{c} - \mathbf{b})'\mathbf{X}'\mathbf{X}(\mathbf{c} - \mathbf{b})$$

   or

$$(\mathbf{y} - \mathbf{X}\mathbf{c})'(\mathbf{y} - \mathbf{X}\mathbf{c}) - \mathbf{e}'\mathbf{e} = (\mathbf{c} - \mathbf{b})'\mathbf{X}'\mathbf{X}(\mathbf{c} - \mathbf{b}).$$

   The right-hand side can be written as $\mathbf{d}'\mathbf{d}$ where $\mathbf{d} = \mathbf{X}(\mathbf{c} - \mathbf{b})$, so it is necessarily positive. This confirms what we knew at the outset, least squares is least squares.

3. In the regression of **y** on **i** and **X**, the coefficients on **X** are $\mathbf{b} = (\mathbf{X}'\mathbf{M}^0\mathbf{X})^{-1}\mathbf{X}'\mathbf{M}^0\mathbf{y}$. $\mathbf{M}^0 = \mathbf{I} - \mathbf{i}(\mathbf{i}'\mathbf{i})^{-1}\mathbf{i}'$ is the matrix which transforms observations into deviations from their column means. Since $\mathbf{M}^0$ is idempotent and symmetric we may also write the preceding as $[(\mathbf{X}'\mathbf{M}^{0\prime})(\mathbf{M}^0\mathbf{X})]^{-1}(\mathbf{X}'\mathbf{M}^{0\prime})(\mathbf{M}^0\mathbf{y})$ which implies that the regression of $\mathbf{M}^0\mathbf{y}$ on $\mathbf{M}^0\mathbf{X}$ produces the least squares slopes. If only **X** is transformed to deviations, we would compute $[(\mathbf{X}'\mathbf{M}^{0\prime})(\mathbf{M}^0\mathbf{X})]^{-1}(\mathbf{X}'\mathbf{M}^{0\prime})\mathbf{y}$, but, of course, this is identical. However, if only **y** is transformed, the result is $(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{M}^0\mathbf{y}$, which is likely to be quite different.

4. What is the result of the matrix product $\mathbf{M}_1\mathbf{M}$ where $\mathbf{M}_1$ is defined in (3-19) and **M** is defined in (3-14)?

$$\mathbf{M}_1\mathbf{M} = (\mathbf{I} - \mathbf{X}_1(\mathbf{X}_1'\mathbf{X}_1)^{-1})(\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}) = \mathbf{M} - \mathbf{X}_1(\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1'\mathbf{M}$$

There is no need to multiply out the second term. Each column of $\mathbf{M}\mathbf{X}_1$ is the vector of residuals in the regression of the corresponding column of $\mathbf{X}_1$ on all of the columns in **X**. Since that **x** is one of the columns in **X**, this regression provides a perfect fit, so the residuals are zero. Thus, $\mathbf{M}\mathbf{X}_1$ is a matrix of zeroes which implies that $\mathbf{M}_1\mathbf{M} = \mathbf{M}$.

5. The original **X** matrix has *n* rows. We add an additional row, $\mathbf{x}_s'$. The new **y** vector likewise has an additional element. Thus, $\mathbf{X}_{n,s} = \begin{bmatrix}\mathbf{X}_n \\ \mathbf{x}_s'\end{bmatrix}$ and $\mathbf{y}_{n,s} = \begin{bmatrix}\mathbf{y}_n \\ y_s\end{bmatrix}$. The new coefficient vector is

$\mathbf{b}_{n,s} = (\mathbf{X}_{n,s}'\,\mathbf{X}_{n,s})^{-1}(\mathbf{X}_{n,s}'\mathbf{y}_{n,s})$. The matrix is $\mathbf{X}_{n,s}'\,\mathbf{X}_{n,s} = \mathbf{X}_n'\mathbf{X}_n + \mathbf{x}_s\mathbf{x}_s'$. To invert this, use (A-66);

$(\mathbf{X}_{n,s}'\mathbf{X}_{n,s})^{-1} = (\mathbf{X}_n'\mathbf{X}_n)^{-1} - \dfrac{1}{1+\mathbf{x}_s'(\mathbf{X}_n'\mathbf{X}_n)^{-1}\mathbf{x}_s}(\mathbf{X}_n'\mathbf{X}_n)^{-1}\mathbf{x}_s\mathbf{x}_s'(\mathbf{X}_n'\mathbf{X}_n)^{-1}$. The vector is

$(\mathbf{X}_{n,s}'\,\mathbf{y}_{n,s}) = (\mathbf{X}_n'\,\mathbf{y}_n) + \mathbf{x}_s y_s$. Multiply out the four terms to get

$$(\mathbf{X}_{n,s}'\,\mathbf{X}_{n,s})^{-1}(\mathbf{X}_{n,s}'\mathbf{y}_{n,s}) =$$

$$\mathbf{b}_n - \frac{1}{1+\mathbf{x}_s'(\mathbf{X}_n'\mathbf{X}_n)^{-1}\mathbf{x}_s}(\mathbf{X}_n'\mathbf{X}_n)^{-1}\mathbf{x}_s\mathbf{x}_s'\mathbf{b}_n + (\mathbf{X}_n'\mathbf{X}_n)^{-1}\,\mathbf{x}_s y_s - \frac{1}{1+\mathbf{x}_s'(\mathbf{X}_n'\mathbf{X}_n)^{-1}\mathbf{x}_s}(\mathbf{X}_n'\mathbf{X}_n)^{-1}\mathbf{x}_s\mathbf{x}_s'(\mathbf{X}_n'\mathbf{X}_n)^{-1}\mathbf{x}_s y_s$$

$$= \mathbf{b}_n + (\mathbf{X}_n'\mathbf{X}_n)^{-1}\,\mathbf{x}_s y_s - \frac{\mathbf{x}_s'(\mathbf{X}_n'\mathbf{X}_n)^{-1}\mathbf{x}_s}{1+\mathbf{x}_s'(\mathbf{X}_n'\mathbf{X}_n)^{-1}\mathbf{x}_s}(\mathbf{X}_n'\mathbf{X}_n)^{-1}\mathbf{x}_s y_s - \frac{1}{1+\mathbf{x}_s'(\mathbf{X}_n'\mathbf{X}_n)^{-1}\mathbf{x}_s}(\mathbf{X}_n'\mathbf{X}_n)^{-1}\mathbf{x}_s\mathbf{x}_s'\mathbf{b}_n$$

$$\mathbf{b}_n + \left[1 - \frac{\mathbf{x}_s'(\mathbf{X}_n'\mathbf{X}_n)^{-1}\mathbf{x}_s}{1+\mathbf{x}_s'(\mathbf{X}_n'\mathbf{X}_n)^{-1}\mathbf{x}_s}\right](\mathbf{X}_n'\mathbf{X}_n)^{-1}\mathbf{x}_s y_s - \frac{1}{1+\mathbf{x}_s'(\mathbf{X}_n'\mathbf{X}_n)^{-1}\mathbf{x}_s}(\mathbf{X}_n'\mathbf{X}_n)^{-1}\mathbf{x}_s\mathbf{x}_s'\mathbf{b}_n$$

$$\mathbf{b}_n + \frac{1}{1+\mathbf{x}_s'(\mathbf{X}_n'\mathbf{X}_n)^{-1}\mathbf{x}_s}(\mathbf{X}_n'\mathbf{X}_n)^{-1}\mathbf{x}_s y_s - \frac{1}{1+\mathbf{x}_s'(\mathbf{X}_n'\mathbf{X}_n)^{-1}\mathbf{x}_s}(\mathbf{X}_n'\mathbf{X}_n)^{-1}\mathbf{x}_s\mathbf{x}_s'\mathbf{b}_n$$

$$\mathbf{b}_n + \frac{1}{1+\mathbf{x}_s'(\mathbf{X}_n'\mathbf{X}_n)^{-1}\mathbf{x}_s}(\mathbf{X}_n'\mathbf{X}_n)^{-1}\mathbf{x}_s(y_s - \mathbf{x}_s'\mathbf{b}_n).$$

6. Define the data matrix as follows: $\mathbf{X} = \begin{bmatrix}\mathbf{i} & \mathbf{x} & \mathbf{0} \\ 1 & 0 & 1\end{bmatrix} = \begin{bmatrix}\mathbf{X}_1, & \mathbf{0} \\ & 1\end{bmatrix} = \begin{bmatrix}\mathbf{X}_1 & \mathbf{X}_2\end{bmatrix}$ and $\mathbf{y} = \begin{bmatrix}\mathbf{y}_o \\ y_m\end{bmatrix}$.

(The subscripts on the parts of **y** refer to the "observed" and "missing" rows of **X**.)
We will use Frish-Waugh to obtain the first two columns of the least squares coefficient vector. $\mathbf{b}_1 = (\mathbf{X}_1'\mathbf{M}_2\mathbf{X}_1)^{-1}(\mathbf{X}_1'\mathbf{M}_2\mathbf{y})$. Multiplying it out, we find that $\mathbf{M}_2$ = an identity matrix save for the last diagonal element that is equal to 0.

$\mathbf{X}_1'\mathbf{M}_2\mathbf{X}_1 = \mathbf{X}_1'\mathbf{X}_1 - \mathbf{X}_1'\begin{bmatrix}\mathbf{0} & \mathbf{0} \\ \mathbf{0}' & 1\end{bmatrix}\mathbf{X}_1$. This just drops the last observation. $\mathbf{X}_1'\mathbf{M}_2\mathbf{y}$ is computed likewise.

Thus, the coefficients on the first two columns are the same as if $y_0$ had been linearly regressed on $\mathbf{X}_1$.

The denominator of $R^2$ is different for the two cases (drop the observation or keep it with zero fill and the dummy variable). For the first strategy, the mean of the $n - 1$ observations should be different from the mean of the full $n$ unless the last observation happens to equal the mean of the first $n - 1$.

For the second strategy, replacing the missing value with the mean of the other $n - 1$ observations, we can deduce the new slope vector logically. Using Frisch-Waugh, we can replace the column of $x$'s with deviations from the means, which then turns the last observation to zero. Thus, once again, the coefficient on the $x$ equals what it is using the earlier strategy. The constant term will be the same as well.

7. For convenience, reorder the variables so that $\mathbf{X} = [\mathbf{i}, \mathbf{P}_d, \mathbf{P}_n, \mathbf{P}_s, \mathbf{Y}]$. The three dependent variables are $\mathbf{E}_d$, $\mathbf{E}_n$, and $\mathbf{E}_s$, and $\mathbf{Y} = \mathbf{E}_d + \mathbf{E}_n + \mathbf{E}_s$. The coefficient vectors are

$$\mathbf{b}_d = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{E}_d,$$
$$\mathbf{b}_n = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{E}_n, \text{ and}$$
$$\mathbf{b}_s = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{E}_s.$$

The sum of the three vectors is

$$\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'[\mathbf{E}_d + \mathbf{E}_n + \mathbf{E}_s] = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}.$$

Now, $\mathbf{Y}$ is the last column of $\mathbf{X}$, so the preceding sum is the vector of least squares coefficients in the regression of the last column of $\mathbf{X}$ on all of the columns of $\mathbf{X}$, including the last. Of course, we get a perfect fit. In addition, $\mathbf{X}'[\mathbf{E}_d + \mathbf{E}_n + \mathbf{E}_s]$ is the last column of $\mathbf{X}'\mathbf{X}$, so the matrix product is equal to the last column of an identity matrix. Thus, the sum of the coefficients on all variables except income is 0, while that on income is 1.

8. Let $\overline{R}_K^2$ denote the adjusted $R^2$ in the full regression on $K$ variables including $\mathbf{x}_k$, and let $\overline{R}_1^2$ denote the adjusted $R^2$ in the short regression on $K$-1 variables when $\mathbf{x}_k$ is omitted. Let $R_K^2$ and $R_1^2$ denote their unadjusted counterparts. Then,

$$R_K^2 = 1 - \mathbf{e}'\mathbf{e}/\mathbf{y}'\mathbf{M}^0\mathbf{y}$$
$$R_1^2 = 1 - \mathbf{e}_1'\mathbf{e}_1/\mathbf{y}'\mathbf{M}^0\mathbf{y}$$

where $\mathbf{e}'\mathbf{e}$ is the sum of squared residuals in the full regression, $\mathbf{e}_1'\mathbf{e}_1$ is the (larger) sum of squared residuals in the regression which omits $\mathbf{x}_k$, and $\mathbf{y}'\mathbf{M}^0\mathbf{y} = \Sigma_i (y_i - \overline{y})^2$.

Then, $\overline{R}_K^2 = 1 - [(n - 1)/(n - K)](1 - R_K^2)$

and $\overline{R}_1^2 = 1 - [(n - 1)/(n-(K - 1))](1 - R_1^2)$.

The difference is the change in the adjusted $R^2$ when $\mathbf{x}_k$ is added to the regression,

$$\overline{R}_K^2 - \overline{R}_1^2 = [(n - 1)/(n - K + 1)][\mathbf{e}_1'\mathbf{e}_1/\mathbf{y}'\mathbf{M}^0\mathbf{y}] - [(n - 1)/(n - K)][\mathbf{e}'\mathbf{e}/\mathbf{y}'\mathbf{M}^0\mathbf{y}].$$

The difference is positive if and only if the ratio is greater than 1. After cancelling terms, we require for the adjusted $R^2$ to increase that $\mathbf{e}_1'\mathbf{e}_1/(n - K + 1)]/[(n - K)/\mathbf{e}'\mathbf{e}] > 1$. From the previous problem, we have that $\mathbf{e}_1'\mathbf{e}_1 = \mathbf{e}'\mathbf{e} + b_K^2 (\mathbf{x}_k'\mathbf{M}_1\mathbf{x}_k)$, where $\mathbf{M}_1$ is defined above and $b_k$ is the least squares coefficient in the full regression of $\mathbf{y}$ on $\mathbf{X}_1$ and $\mathbf{x}_k$. Making the substitution, we require $[(\mathbf{e}'\mathbf{e} + b_K^2 (\mathbf{x}_k'\mathbf{M}_1\mathbf{x}_k))$ $(n - K)]/[(n - K)\mathbf{e}'\mathbf{e} + \mathbf{e}'\mathbf{e}] > 1$. Since $\mathbf{e}'\mathbf{e} = (n - K)s^2$, this simplifies to $[\mathbf{e}'\mathbf{e} + b_K^2 (\mathbf{x}_k'\mathbf{M}_1\mathbf{x}_k)]/$ $[\mathbf{e}'\mathbf{e} + s^2] > 1$. Since all terms are positive, the fraction is greater than one if and only $b_K^2 (\mathbf{x}_k'\mathbf{M}_1\mathbf{x}_k) > s^2$ or $b_K^2 (\mathbf{x}_k'\mathbf{M}_1\mathbf{x}_k/s^2) > 1$. The denominator is the estimated variance of $b_k$, so the result is proved.

9. This $R^2$ must be lower. The sum of squares associated with the coefficient vector which omits the constant term must be higher than the one which includes it. We can write the coefficient vector in the regression without a constant as $\mathbf{c} = (0,\mathbf{b}^*)$ where $\mathbf{b}^* = (\mathbf{W}'\mathbf{W})^{-1}\mathbf{W}'\mathbf{y}$, with $\mathbf{W}$ being the other $K - 1$ columns of $\mathbf{X}$. Then, the result of the previous exercise applies directly.

10. We use the notations 'Var[.]' and 'Cov[.]' to indicate the sample variances and covariances. Our information is

$$\text{Var}[N] = 1,\ \text{Var}[D] = 1,\ \text{Var}[Y] = 1.$$

Since $C = N + D$, $\text{Var}[C] = \text{Var}[N] + \text{Var}[D] + 2\text{Cov}[N, D] = 2(1 + \text{Cov}[N, D])$.

From the regressions, we have

$\text{Cov}[C, Y]/\text{Var}[Y] = \text{Cov}[C, Y] = 0.8.$

But, $\text{Cov}[C, Y] = \text{Cov}[N, Y] + \text{Cov}[D, Y]$.

Also, $\text{Cov}[C, N]/\text{Var}[N] = \text{Cov}[C, N] = 0.5$,

but, $\text{Cov}[C, N] = \text{Var}[N] + \text{Cov}[N, D] = 1 + \text{Cov}[N, D]$, so $\text{Cov}[N, D] = -0.5$,

so that $\text{Var}[C] = 2(1 + -0.5) = 1$.

And, $\text{Cov}[D, Y]/\text{Var}[Y] = \text{Cov}[D, Y] = 0.4$.

Since $\text{Cov}[C, Y] = 0.8 = \text{Cov}[N, Y] + \text{Cov}[D, Y]$, $\text{Cov}[N, Y] = 0.4$.

Finally, $\text{Cov}[C, D] = \text{Cov}[N, D] + \text{Var}[D] = -0.5 + 1 = 0.5$.

Now, in the regression of $C$ on $D$, the sum of squared residuals is $(n - 1)\{\text{Var}[C] - (\text{Cov}[C,D]/\text{Var}[D])^2\text{Var}[D]\}$ based on the general regression result $\Sigma e^2 = \Sigma(y_i - y)^2 - b^2\Sigma\,(x_i - \bar{x})^2$. All of the necessary figures were obtained above. Inserting these and $n - 1 = 20$ produces a sum of squared residuals of 15.

11. The relevant submatrices to be used in the calculations are

| | Investment | Constant | GNP | Interest |
|---|---|---|---|---|
| Investment | * | 3.0500 | 3.9926 | 23.521 |
| Constant | | 15 | 19.310 | 111.79 |
| GNP | | | 25.218 | 148.98 |
| Interest | | | | 943.86 |

The inverse of the lower right $3 \times 3$ block is $(\mathbf{X}'\mathbf{X})^{-1}$,

$$(\mathbf{X}'\mathbf{X})^{-1} = \begin{matrix} 7.5874 & & \\ -7.41859 & 7.84078 & \\ .27313 & -.598953 & .06254637 \end{matrix}$$

The coefficient vector is $\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = (-.0727985, .235622, -.00364866)'$. The total sum of squares is $\mathbf{y}'\mathbf{y} = .63652$, so we can obtain $\mathbf{e}'\mathbf{e} = \mathbf{y}'\mathbf{y} - \mathbf{b}'\mathbf{X}'\mathbf{y}$. $\mathbf{X}'\mathbf{y}$ is given in the top row of the matrix. Making the substitution, we obtain $\mathbf{e}'\mathbf{e} = .63652 - .63291 = .00361$. To compute $R^2$, we require $\Sigma_i (y_i - \bar{y})^2 = .63652 - 15(3.05/15)^2 = .01635333$, so $R^2 = 1 - .00361/.0163533 = .77925$.

12. The results cannot be correct. Since $\log S/N = \log S/Y + \log Y/N$ by simple, exact algebra, the same result must apply to the least squares regression results. That means that the second equation estimated must equal the first one plus $\log Y/N$. Looking at the equations, that means that all of the coefficients would have to be identical save for the second, which would have to equal its counterpart in the first equation, plus 1. Therefore, the results cannot be correct. In an exchange between Leff and Arthur Goldberger that appeared later in the same journal, Leff argued that the difference was a simple rounding error. You can see that the results in the second equation resemble those in the first, but not enough so that the explanation is credible. Further discussion about the data themselves appeared in a subsequent discussion. [See Goldberger (1973) and Leff (1973).]

## Application

```
?=============================================================================
? Chapter 3 Application 1
?=============================================================================
Read $
(Data appear in the text.)
Namelist ; X1 = one,educ,exp,ability$
Namelist ; X2 = mothered,fathered,sibs$
?=============================================================================
? a.
?=============================================================================
Regress  ; Lhs = wage ; Rhs = x1$
+----------------------------------------------------+
| Ordinary    least squares regression               |
| LHS=WAGE     Mean                 =   2.059333     |
|              Standard deviation   =   .2583869     |
| WTS=none     Number of observs.   =         15     |
| Model size   Parameters           =          4     |
|              Degrees of freedom   =         11     |
| Residuals    Sum of squares       =   .7633163     |
|              Standard error of e  =   .2634244     |
| Fit          R-squared            =   .1833511     |
|              Adjusted R-squared   =  -.3937136E-01 |
| Model test   F[  3,    11] (prob) =    .82 (.5080) |
+----------------------------------------------------+
+--------+--------------+----------------+--------+--------+----------+
|Variable| Coefficient  | Standard Error |t-ratio |P[|T|>t]| Mean of X|
+--------+--------------+----------------+--------+--------+----------+
 Constant|    1.66364000       .61855318     2.690   .0210
 EDUC    |     .01453897       .04902149      .297   .7723   12.8666667
 EXP     |     .07103002       .04803415     1.479   .1673   2.80000000
 ABILITY |     .02661537       .09911731      .269   .7933    .36600000
?=============================================================================
? b.
?=============================================================================
Regress  ; Lhs = wage ; Rhs = x1,x2$
+----------------------------------------------------+
| Ordinary    least squares regression               |
| LHS=WAGE     Mean                 =   2.059333     |
|              Standard deviation   =   .2583869     |
| WTS=none     Number of observs.   =         15     |
| Model size   Parameters           =          7     |
|              Degrees of freedom   =          8     |
| Residuals    Sum of squares       =   .4522662     |
|              Standard error of e  =   .2377673     |
| Fit          R-squared            =   .5161341     |
|              Adjusted R-squared   =   .1532347     |
| Model test   F[  6,     8] (prob) =   1.42 (.3140) |
+----------------------------------------------------+
+--------+--------------+----------------+--------+--------+----------+
|Variable| Coefficient  | Standard Error |t-ratio |P[|T|>t]| Mean of X|
+--------+--------------+----------------+--------+--------+----------+
 Constant|     .04899633       .94880761      .052   .9601
 EDUC    |     .02582213       .04468592      .578   .5793   12.8666667
 EXP     |     .10339125       .04734541     2.184   .0605   2.80000000
 ABILITY |     .03074355       .12120133      .254   .8062    .36600000
 MOTHERED|     .10163069       .07017502     1.448   .1856   12.0666667
 FATHERED|     .00164437       .04464910      .037   .9715   12.6666667
 SIBS    |     .05916922       .06901801      .857   .4162   2.20000000
```

```
?=========================================================================
? c.
?=========================================================================
Regress  ; Lhs = mothered ; Rhs = x1 ; Res = meds $
Regress  ; Lhs = fathered ; Rhs = x1 ; Res = feds $
Regress  ; Lhs = sibs     ; Rhs = x1 ; Res = sibss $
Namelist ; X2S = meds,feds,sibss $
Matrix   ; list ; Mean(X2S) $
Matrix Result   has  3 rows and  1 columns.
               1
        +--------------
       1| -.1184238D-14
       2|  .1657933D-14
       3| -.5921189D-16
The means are (essentially) zero.  The sums must be zero, as these new variables are
orthogonal to the columns of X1. The first column in X1 is a column of ones, so this
means that these residuals must sum to zero.
?=========================================================================
? d.
?=========================================================================
Namelist ; X = X1,X2 $
Matrix   ; i = init(n,1,1) $
Matrix   ; M0 = iden(n) - 1/n*i*i' $
Matrix   ; b12 = <X'X>*X'wage$
Calc     ; list ; ym0y =(N-1)*var(wage) $
Matrix   ; list ; cod = 1/ym0y * b12'*X'*M0*X*b12 $
Matrix COD      has  1 rows and  1 columns.
               1
        +--------------
       1|     .51613
Matrix   ; e = wage - X*b12 $
Calc     ; list ; cod = 1 - 1/ym0y * e'e $
+-------------------------------------+
 COD     =       .516134
The R squared is the same using either method of computation.
Calc     ; list ; RsqAd = 1 - (n-1)/(n-col(x))*(1-cod)$
+-------------------------------------+
 RSQAD   =       .153235
? Now drop the constant
Namelist ; X0 = educ,exp,ability,X2 $
Matrix   ; i = init(n,1,1) $
Matrix   ; M0 = iden(n) - 1/n*i*i' $
Matrix   ; b120 = <X0'X0>*X0'wage$
Matrix   ; list ; cod = 1/ym0y * b120'*X0'*M0*X0*b120 $
Matrix COD      has  1 rows and  1 columns.
               1
        +--------------
       1|     .52953
Matrix   ; e0 = wage - X0*b120 $
Calc     ; list ; cod = 1 - 1/ym0y * e0'e0 $
+-------------------------------------+
| Listed Calculator Results          |
+-------------------------------------+
 COD     =       .515973
The R squared now changes depending on how it is computed.  It also goes up, completely
artificially.
```

```
?=========================================================================
? e.
?=========================================================================
The R squared for the full regression appears immediately below.
? f.
Regress ; Lhs = wage ; Rhs = X1,X2 $
+----------------------------------------------------+
| Ordinary    least squares regression               |
| WTS=none     Number of observs.   =         15     |
| Model size   Parameters           =          7     |
|              Degrees of freedom   =          8     |
| Fit          R-squared            =   .5161341     |
+----------------------------------------------------+
+--------+-------------+----------------+--------+--------+----------+
|Variable| Coefficient  | Standard Error |t-ratio |P[|T|>t]| Mean of X|
+--------+-------------+----------------+--------+--------+----------+
 Constant|     .04899633       .94880761      .052   .9601
 EDUC    |     .02582213       .04468592      .578   .5793   12.8666667
 EXP     |     .10339125       .04734541     2.184   .0605   2.80000000
 ABILITY |     .03074355       .12120133      .254   .8062    .36600000
 MOTHERED|     .10163069       .07017502     1.448   .1856   12.0666667
 FATHERED|     .00164437       .04464910      .037   .9715   12.6666667
 SIBS    |     .05916922       .06901801      .857   .4162   2.20000000
Regress ; Lhs = wage ; Rhs = X1,X2S $
+----------------------------------------------------+
| Ordinary    least squares regression               |
| WTS=none     Number of observs.   =         15     |
| Model size   Parameters           =          7     |
|              Degrees of freedom   =          8     |
| Fit          R-squared            =   .5161341     |
|              Adjusted R-squared   =   .1532347     |
+----------------------------------------------------+
+--------+-------------+----------------+--------+--------+----------+
|Variable| Coefficient  | Standard Error |t-ratio |P[|T|>t]| Mean of X|
+--------+-------------+----------------+--------+--------+----------+
 Constant|    1.66364000       .55830716     2.980   .0176
 EDUC    |     .01453897       .04424689      .329   .7509   12.8666667
 EXP     |     .07103002       .04335571     1.638   .1400   2.80000000
 ABILITY |     .02661537       .08946345      .297   .7737    .36600000
 MEDS    |     .10163069       .07017502     1.448   .1856 -.118424D-14
 FEDS    |     .00164437       .04464910      .037   .9715  .165793D-14
 SIBSS   |     .05916922       .06901801      .857   .4162 -.592119D-16
```

In the first set of results, the first coefficient vector is $\mathbf{b}_1 = (\mathbf{X}_1'\mathbf{M}_2\mathbf{X}_1)^{-1}\mathbf{X}_1'\mathbf{M}_2\mathbf{y}$ and $\mathbf{b}_2 = (\mathbf{X}_2'\mathbf{M}_1\mathbf{X}_2)^{-1}\mathbf{X}_2'\mathrm{M}_1\mathbf{y}$.

In the second regression, the second set of regressors is $\mathbf{M}_1\mathbf{X}_2$, so $\mathbf{b}_1 = (\mathbf{X}_1'\mathbf{M}_{12}\,\mathbf{X}_1)^{-1}\mathbf{X}_1'\mathbf{M}_{12}\mathbf{y}$ where $\mathbf{M}_{12} = \mathbf{I} - (\mathbf{M}_1\mathbf{X}_2)[(\mathrm{M}_1\mathbf{X}_2)'(\mathbf{M}_1\mathbf{X}_2)]^{-1}(\mathbf{M}_1\mathbf{X}_2)'$.

Thus, because the "M" matrix is different, the coefficient vector is different. The second set of coefficients in the second regression is $\mathbf{b}_2 = [(\mathbf{M}_1\mathbf{X}_2)'\mathbf{M}_1(\mathbf{M}_1\mathbf{X}_2)]^{-1}\,(\mathbf{M}_1\mathbf{X}_2)\mathbf{M}_1\mathbf{y} = (\mathbf{X}_2'\mathbf{M}_1\mathbf{X}_2)^{-1}\mathbf{X}_2'\mathbf{M}_1\mathbf{y}$ because $\mathbf{M}_1$ is idempotent.