3. THE PARTITIONED REGRESSION MODEL

Consider taking a regression equation in the form of

$$y = [\, X_1 \quad X_2 \,] \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \varepsilon = X_1\beta_1 + X_2\beta_2 + \varepsilon. \tag{1}$$

Here, $[X_1, X_2] = X$ and $[\beta_1', \beta_2']' = \beta$ are obtained by partitioning the matrix X and vector β of the equation $y = X\beta + \varepsilon$ in a conformable manner. The normal equations $X'X\beta = X'y$ can be partitioned likewise. Writing the equations without the surrounding matrix braces gives

$$X_1'X_1\beta_1 + X_1'X_2\beta_2 = X_1'y, \tag{2}$$

$$X_2'X_1\beta_1 + X_2'X_2\beta_2 = X_2'y. \tag{3}$$

From (2), we get the equation $X_1'X_1\beta_1 = X_1'(y - X_2\beta_2)$, which gives an expression for the leading subvector of $\hat{\beta}$:

$$\hat{\beta}_1 = (X_1'X_1)^{-1}X_1'(y - X_2\hat{\beta}_2). \tag{4}$$

To obtain an expression for $\hat{\beta}_2$, we must eliminate β_1 from equation (3). For this purpose, we multiply equation (2) by $X_2'X_1(X_1'X_1)^{-1}$ to give

$$X_2'X_1\beta_1 + X_2'X_1(X_1'X_1)^{-1}X_1'X_2\beta_2 = X_2'X_1(X_1'X_1)^{-1}X_1'y. \tag{5}$$

When the latter is taken from equation (3), we get

$$\Big\{X_2'X_2 - X_2'X_1(X_1'X_1)^{-1}X_1'X_2\Big\}\beta_2 = X_2'y - X_2'X_1(X_1'X_1)^{-1}X_1'y. \tag{6}$$

On defining

$$P_1 = X_1(X_1'X_1)^{-1}X_1', \tag{7}$$

can we rewrite (6) as

$$\Big\{X_2'(I - P_1)X_2\Big\}\beta_2 = X_2'(I - P_1)y, \tag{8}$$

whence

$$\hat{\beta}_2 = \Big\{X_2'(I - P_1)X_2\Big\}^{-1}X_2'(I - P_1)y. \tag{9}$$

Now let us investigate the effect that conditions of orthogonality amongst the regressors have upon the ordinary least-squares estimates of the regression parameters. Consider a partitioned regression model, which can be written as

$$y = [\,X_1, X_2\,] \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \varepsilon = X_1\beta_1 + X_2\beta_2 + \varepsilon. \tag{10}$$

It can be assumed that the variables in this equation are in deviation form. Imagine that the columns of X_1 are orthogonal to the columns of X_2 such that $X_1'X_2 = 0$. This is the same as assuming that the empirical correlation between variables in X_1 and variables in X_2 is zero.

The effect upon the ordinary least-squares estimator can be seen by examining the partitioned form of the formula $\hat{\beta} = (X'X)^{-1}X'y$. Here, we have

$$X'X = \begin{bmatrix} X_1' \\ X_2' \end{bmatrix} [\,X_1 \quad X_2\,] = \begin{bmatrix} X_1'X_1 & X_1'X_2 \\ X_2'X_1 & X_2'X_2 \end{bmatrix} = \begin{bmatrix} X_1'X_1 & 0 \\ 0 & X_2'X_2 \end{bmatrix}, \tag{11}$$

where the final equality follows from the condition of orthogonality. The inverse of the partitioned form of $X'X$ in the case of $X_1'X_2 = 0$ is

$$(X'X)^{-1} = \begin{bmatrix} X_1'X_1 & 0 \\ 0 & X_2'X_2 \end{bmatrix}^{-1} = \begin{bmatrix} (X_1'X_1)^{-1} & 0 \\ 0 & (X_2'X_2)^{-1} \end{bmatrix}. \tag{12}$$

We also have

$$X'y = \begin{bmatrix} X_1' \\ X_2' \end{bmatrix} y = \begin{bmatrix} X_1'y \\ X_2'y \end{bmatrix}. \tag{13}$$

On combining these elements, we find that

$$\begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} = \begin{bmatrix} (X_1'X_1)^{-1} & 0 \\ 0 & (X_2'X_2)^{-1} \end{bmatrix} \begin{bmatrix} X_1'y \\ X_2'y \end{bmatrix} = \begin{bmatrix} (X_1'X_1)^{-1}X_1'y \\ (X_2'X_2)^{-1}X_2'y \end{bmatrix}. \tag{14}$$

In this special case, the coefficients of the regression of y on $X = [X_1, X_2]$ can be obtained from the separate regressions of y on X_1 and y on X_2.

It should be understood that this result does not hold true in general. The general formulae for $\hat{\beta}_1$ and $\hat{\beta}_2$ are those which we have given already under (4) and (9):

$$\begin{aligned} &\hat{\beta}_1 = (X_1'X_1)^{-1}X_1'(y - X_2\hat{\beta}_2), \\ &\hat{\beta}_2 = \left\{X_2'(I - P_1)X_2\right\}^{-1}X_2'(I - P_1)y, \quad P_1 = X_1(X_1'X_1)^{-1}X_1'. \end{aligned} \tag{15}$$

It can be confirmed easily that these formulae do specialise to those under (14) in the case of $X_1'X_2 = 0$.

The purpose of including X_2 in the regression equation when, in fact, interest is confined to the parameters of β_1 is to avoid falsely attributing the explanatory power of the variables of X_2 to those of X_1.

Let us investigate the effects of erroneously excluding X_2 from the regression. In that case, the estimate will be

$$\begin{aligned}\tilde{\beta}_1 &= (X_1'X_1)^{-1}X_1'y \\ &= (X_1'X_1)^{-1}X_1'(X_1\beta_1 + X_2\beta_2 + \varepsilon) \\ &= \beta_1 + (X_1'X_1)^{-1}X_1'X_2\beta_2 + (X_1'X_1)^{-1}X_1'\varepsilon.\end{aligned} \tag{16}$$

On applying the expectations operator to these equations, we find that

$$E(\tilde{\beta}_1) = \beta_1 + (X_1'X_1)^{-1}X_1'X_2\beta_2, \tag{17}$$

since $E\{(X_1'X_1)^{-1}X_1'\varepsilon\} = (X_1'X_1)^{-1}X_1'E(\varepsilon) = 0$. Thus, in general, we have $E(\tilde{\beta}_1) \neq \beta_1$, which is to say that $\tilde{\beta}_1$ is a biased estimator. The only circumstances in which the estimator will be unbiased are when either $X_1'X_2 = 0$ or $\beta_2 = 0$. In other circumstances, the estimator will suffer from a problem which is commonly described as *omitted-variables bias.*

The Regression Model with an Intercept

Now consider again the equations

$$y_t = \alpha + x_{t.}\beta + \varepsilon_t, \quad t = 1, \ldots, T, \tag{18}$$

which comprise T observations of a regression model with an intercept term α and with k explanatory variables in $x_{t.} = [x_{t1}, x_{t2}, \ldots, x_{tk}]$. These equations can also be represented in a matrix notation as

$$y = \iota\alpha + Z\beta + \varepsilon. \tag{19}$$

Here, the vector $\iota = [1, 1, \ldots, 1]'$, which consists of T units, is described alternatively as the dummy vector or the summation vector, whilst $Z = [x_{tj}; t = 1, \ldots T; j = 1, \ldots, k]$ is the matrix of the observations on the explanatory variables.

Equation (19) can be construed as a case of the partitioned regression equation of (1). By setting $X_1 = \iota$ and $X_2 = Z$ and by taking $\beta_1 = \alpha$, $\beta_2 = \beta_z$ in equations (4) and (9), we derive the following expressions for the estimates of the parameters α, β_z:

$$\hat{\alpha} = (\iota'\iota)^{-1}\iota'(y - Z\hat{\beta}_z), \tag{20}$$

$$\hat{\beta}_z = \big\{Z'(I - P_\iota)Z\big\}^{-1} Z'(I - P_\iota)y, \quad \text{with} \tag{21}$$
$$P_\iota = \iota(\iota'\iota)^{-1}\iota' = \frac{1}{T}\iota\iota'.$$

To understand the effect of the operator P_ι in this context, consider the following expressions:

$$\iota' y = \sum_{t=1}^{T} y_t,$$
$$(\iota'\iota)^{-1}\iota' y = \frac{1}{T}\sum_{t=1}^{T} y_t = \bar{y}, \tag{22}$$
$$P_\iota y = \iota\bar{y} = \iota(\iota'\iota)^{-1}\iota' y = [\bar{y}, \bar{y}, \ldots, \bar{y}]'.$$

Here, $P_\iota y = [\bar{y}, \bar{y}, \ldots, \bar{y}]'$ is simply a column vector containing T repetitions of the sample mean. From the expressions above, it can be understood that, if $x = [x_1, x_2, \ldots x_T]'$ is vector of T elements, then

$$x'(I - P_\iota)x = \sum_{t=1}^{T} x_t(x_t - \bar{x}) = \sum_{t=1}^{T}(x_t - \bar{x})x_t = \sum_{t=1}^{T}(x_t - \bar{x})^2. \tag{23}$$

The final equality depends upon the fact that $\sum(x_t - \bar{x})\bar{x} = \bar{x}\sum(x_t - \bar{x}) = 0$.

The Regression Model in Deviation Form

Consider the matrix of cross-products in equation (1.22). This is

$$Z'(I - P_\iota)Z = \{(I - P_\iota)Z\}'\{Z(I - P_\iota)\} = (Z - \bar{Z})'(Z - \bar{Z}). \tag{24}$$

Here, $\bar{Z} = [(\bar{x}_j; j = 1, \ldots, k)_t; t = 1, \ldots, T]$ is a matrix in which the generic row $(\bar{x}_1, \ldots, \bar{x}_k)$, which contains the sample means of the k explanatory variables, is repeated T times. The matrix $(I - P_\iota)Z = (Z - \bar{Z})$ is the matrix of the deviations of the data points about the sample means, and it is also the matrix of the residuals of the regressions of the vectors of Z upon the summation vector ι. The vector $(I - P_\iota)y = (y - \iota\bar{y})$ may be described likewise.

It follows that the estimate of β_z is precisely the value which would be obtained by applying the technique of least-squares regression to a meta-equation

$$\begin{bmatrix} y_1 - \bar{y} \\ y_2 - \bar{y} \\ \vdots \\ y_T - \bar{y} \end{bmatrix} = \begin{bmatrix} x_{11} - \bar{x}_1 & \ldots & x_{1k} - \bar{x}_k \\ x_{21} - \bar{x}_1 & \ldots & x_{2k} - \bar{x}_k \\ \vdots & & \vdots \\ x_{T1} - \bar{x}_1 & \ldots & x_{Tk} - \bar{x}_k \end{bmatrix} \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} \varepsilon_1 - \bar{\varepsilon} \\ \varepsilon_2 - \bar{\varepsilon} \\ \vdots \\ \varepsilon_T - \bar{\varepsilon} \end{bmatrix}, \tag{25}$$

which lacks an intercept term. In summary notation, the equation may be denoted by

$$y - \iota\bar{y} = [Z - \bar{Z}]\beta_z + (\varepsilon - \bar{\varepsilon}). \tag{26}$$

Observe that it is unnecessary to take the deviations of y. The result is the same whether we regress y or $y - \iota\bar{y}$ on $[Z - \bar{Z}]$. The result is due to the symmetry and idempotency of the operator $(I - P_\iota)$ whereby $Z'(I - P_\iota)y = \{(I - P_\iota)Z\}'\{(I - P_\iota)y\}$.

Once the value for $\hat{\beta}$ is available, the estimate for the intercept term can be recovered from the equation (1.21) which can be written as

$$\begin{aligned} \bar{\alpha} &= \bar{y} - \bar{Z}\hat{\beta}_z \\ &= \bar{y} - \sum_{j=1}^{k} \bar{x}_j \hat{\beta}_j. \end{aligned} \tag{27}$$