CHAPTER IX

SOLUTIONS TO THE EXERCISES

9.2 Chapter 2

1. We have

$$LML' = (I - AX)U(I - AX)' + AVA'$$

which is formula (2.2.8) for the risk as the sum of the bias and the variance defined by (2.2.4). The minimum risk is given by (since $LK = I$)

$$LK(K'M^{-1}K)^{-1}K'L' = (U^{-1} + X'V^{-1}X)^{-1}.$$

Equality is attained when

$$L = LK(K'M^{-1}K)^{-1}K'M^{-1} = (U^{-1} + X'V^{-1}X)^{-1}[U^{-1}, X'V^{-1}].$$

From the definition $L = [I - AX, A]$, this implies that

$$A = (U^{-1} + X'V^{-1}X)^{-1}X'V^{-1},$$

which is formula (1.7.8). □

3. The restrictions (2.9.3) and (2.9.4) may be written in the form $XZ' = 0$ where

$$Z = \begin{bmatrix} l & 0 & 1 & -1 & 0 \\ m & \rho & 0 & 0 & -1 \end{bmatrix},$$

and where $l = \log\lambda$ and $m = \log\mu$. For $\psi\beta$ to be estimable (where ψ is a 1×5 vector) we must have $\psi = aX$ for some $1 \times n$ vector a. Since Z has rank 2 and is orthogonal to X, and since X is assumed to have maximal rank subject to (2.9.3) and (2.9.4), therefore X has rank 3. Thus Z is polar to X, hence $\psi Z' = 0$. That is, if ψ is orthogonal to the space spanned by the rows of Z, it is in the row space of X; i.e., $\psi Z' = 0$ implies $\psi = aX$ for some a. The estimable functionals ψ are therefore the vectors ψ for which $\psi Z' = 0$.

(a) Choosing ψ to be the coordinate vector δ^j with 1 in jth place and 0s elsewhere, δ^j is estimable if and only if the jth column of Z is zero. This is clearly not the case for $j = 2, \ldots, 5$; thus for $j = 2, \ldots, 5$, β_j is not estimable. For $j = 1$, column 1 of Z is zero if only if $l = m = 0$ (i.e., $\lambda = \mu = 1$), in which case β_1 is estimable; otherwise β_1 is not estimable.

(b) Yes, since

$$\psi Z' = (0, 0, 1, 1, 0) \begin{bmatrix} l & m \\ 0 & \rho \\ 1 & 0 \\ -1 & 0 \\ 0 & -1 \end{bmatrix} = (0, 0).$$

(c) Yes, since

$$\psi Z' = (1,\ n+1,\ x_{n+1,3},\ l + x_{n+1,3},\ m + \rho(n+1)) \begin{bmatrix} l & m \\ 0 & \rho \\ 1 & 0 \\ -1 & 0 \\ 0 & -1 \end{bmatrix} = (0,0).$$

Thus, $E\{y_{n+1}\}$ is estimable; that is, if the restrictions (2.9.3) and (2.9.4) continue to hold in period $n+1$, it is possible to predict the mean of the dependent variable y_t, even though none of the individual coefficients β_j is estimable.

4. **(a)** Step (1) gives

$$b_2^* = (X_2'X_2)^{-1}X_2'y \quad \text{and} \quad B_2^* = (X_2'X_2)^{-1}X_2'X_1. \tag{9.2.1}$$

Defining H_2 as in the Hint and substituting (9.2.1) in (2.9.7), Step (2) gives

$$y^* = (I - H_2)y \quad \text{and} \quad X_1^* = (I - H_2)X_1. \tag{9.2.2}$$

Substitution of (9.2.2) in (2.9.6) gives, by Step (3),

$$b_1^* = [X_1'(I - H_2)X_1]^{-1}X_1'(I - H_2)y. \tag{9.2.3}$$

We need to show that this is the same as the ordinary least-squares estimator of β_1 from (2.9.6). Defining $I_k = [\Phi_1', \Phi_2']$ as a partition of the $k \times k$ identity matrix into its first k_1 and last k_2 columns, we have

$$X_1 = X\Phi_1' = [X_1\ X_2]\begin{bmatrix} I_{k_1} \\ 0 \end{bmatrix} \quad \text{and} \quad X_2 = X\Phi_2' = [X_1\ X_2]\begin{bmatrix} 0 \\ I_{k_2} \end{bmatrix}, \tag{9.2.4}$$

and since we see easily that $H_iH = HH_i = H_i$ for $i = 1, 2$, it follows that

$$(I - H_i)(I - H) = (I - H)(I - H_i) = I - H \quad (i = 1, 2). \tag{9.2.5}$$

Consequently, from (9.2.4) and the definition of H (from the Hint), we have

$$(I - H)X_i = (I - H)X\Phi_i' = 0, \quad \text{and} \quad (I - H_i)X_i = 0 \quad (i = 1, 2). \tag{9.2.6}$$

Now, writing the residual vector from the least-squares regression of (2.9.6) as

$$e = y - Xb = (I - H)y, \tag{9.2.7}$$

we obtain the identity

$$y = Xb + e = X_1b_1 + X_2b_2 + (I - H)y. \tag{9.2.8}$$

Now we observe from (9.2.6), (9.2.5), and (9.2.4) that premultiplication of (9.2.8) by $X_1'(I - H_2)$ annihilates the last two terms on the right, leaving

$$X_1'(I - H_2)y = X_1'(I - H_2)X_1b_1. \tag{9.2.9}$$

Premultiplying both sides by $[X_1'(I - H_2)X_1]^{-1}$ we obtain

$$b_1 = [X_1'(I - H_2)X_1]^{-1}X_1'(I - H_2)y. \tag{9.2.10}$$

But this is the same as b_1^* in (9.2.3), hence $b_1 = b_1^*$.

(b) This follows from (9.2.2) and the fact that

$$\begin{aligned} e^* &= (I - H_2)(y - X_1 b_1) && (\text{since } b_1^* = b_1) \\ &= (I - H_2)(y - X_1 b_1 - X_2 b_2) && (\text{since } (I - H_2)X_2 = 0) \\ &= (I - H_2)(y - Xb) && \\ &= (I - H_2)(I - H)y && \\ &= (I - H)y && (\text{from } (9.2.5)) \\ &= e && (\text{from } (9.2.7)). \end{aligned}$$

The above result (both parts (a) and (b)) is known as the Frisch-Waugh theorem, for which references are given below. See Frisch & Waugh (1933), Reiersöl (1945), Lovell (1963), Davidson & MacKinnon (1933, pp. 19–24), Fiebig, Bartels, & Krämer (1996), and Chipman (1998, pp. 58–108).