if $\boldsymbol{X}$ simply consisted of a constant vector, $\boldsymbol{e}_t^\top \boldsymbol{P}_X \boldsymbol{e}_t$ would equal $1/n$. Even when there is no constant term, h_t can never be 0 unless every element of $\boldsymbol{X}_t$ is 0. However, it is evidently quite possible for h_t to equal 1. Suppose, for example, that one column of $\boldsymbol{X}$ is the dummy variable $\boldsymbol{e}_t$. In that case, $h_t = \boldsymbol{e}_t^\top \boldsymbol{P}_X \boldsymbol{e}_t = \boldsymbol{e}_t^\top \boldsymbol{e}_t = 1$.

It is interesting to see what happens when we add a dummy variable $\boldsymbol{e}_t$ to a regression. It turns out that $\hat{u}_t$ will equal zero and that the t^{th} observation will have no effect at all on any coefficient except the one corresponding to the dummy variable. The latter simply takes on whatever value is needed to make $\hat{u}_t = 0$, and the remaining coefficients are those that minimize the SSR for the remaining $n - 1$ observations. These results are easily established by using the FWL Theorem.

Consider the following two regressions, where for ease of notation the data have been ordered so that observation t is the last observation, and $\boldsymbol{y}_{(t)}$ and $\boldsymbol{X}_{(t)}$ denote the first $n - 1$ rows of $\boldsymbol{y}$ and $\boldsymbol{X}$, respectively:

$$\begin{bmatrix} \boldsymbol{y}_{(t)} \\ y_t \end{bmatrix} = \begin{bmatrix} \boldsymbol{X}_{(t)} \\ \boldsymbol{X}_t \end{bmatrix} \boldsymbol{\beta} + \text{residuals}, \tag{1.43}$$

and

$$\begin{bmatrix} \boldsymbol{y}_{(t)} \\ y_t \end{bmatrix} = \begin{bmatrix} \boldsymbol{X}_{(t)} & \boldsymbol{0} \\ \boldsymbol{X}_t & 1 \end{bmatrix} \begin{bmatrix} \boldsymbol{\beta} \\ \alpha \end{bmatrix} + \text{residuals}. \tag{1.44}$$

Regression (1.43) is simply the regression of $\boldsymbol{y}$ on $\boldsymbol{X}$, which yields parameter estimates $\hat{\boldsymbol{\beta}}$ and least squares residuals $\hat{\boldsymbol{u}}$. Regression (1.44) is regression (1.43) with $\boldsymbol{e}_t$ as an additional regressor. By the FWL Theorem, the estimate of $\boldsymbol{\beta}$ from (1.44) must be identical to the estimate of $\boldsymbol{\beta}$ from the regression

$$\boldsymbol{M}_t \begin{bmatrix} \boldsymbol{y}_{(t)} \\ y_t \end{bmatrix} = \boldsymbol{M}_t \begin{bmatrix} \boldsymbol{X}_{(t)} \\ \boldsymbol{X}_t \end{bmatrix} \boldsymbol{\beta} + \text{residuals}, \tag{1.45}$$

where $\boldsymbol{M}_t$ is the matrix that projects orthogonally onto $\mathcal{S}^\perp(\boldsymbol{e}_t)$. Multiplying any vector by $\boldsymbol{M}_t$ merely annihilates the last element of that vector. Thus regression (1.45) is simply

$$\begin{bmatrix} \boldsymbol{y}_{(t)} \\ 0 \end{bmatrix} = \begin{bmatrix} \boldsymbol{X}_{(t)} \\ \boldsymbol{0} \end{bmatrix} \boldsymbol{\beta} + \text{residuals}. \tag{1.46}$$

The last observation, in which the regressand and all regressors are zero, obviously has no effect at all on parameter estimates. Regression (1.46) is therefore equivalent to regressing $\boldsymbol{y}_{(t)}$ on $\boldsymbol{X}_{(t)}$ and so must yield OLS estimates $\hat{\boldsymbol{\beta}}^{(t)}$. For regression (1.46), the residual for observation t is clearly zero; the FWL Theorem then implies that the residual for observation t from regression (1.44) must likewise be zero, which implies that $\hat{\alpha}$ must equal $y_t - \boldsymbol{X}_t \hat{\boldsymbol{\beta}}^{(t)}$.

These results make it easy to derive the results (1.40) and (1.41), which were earlier stated without proof. Readers who are not interested in the