seem to have introduced, and then reintroduced, it to econometricians. The theorem is much more general, and much more generally useful, than a casual reading of those papers might suggest, however. Among other things, it almost totally eliminates the need to invert partitioned matrices when one is deriving many standard results about ordinary (and nonlinear) least squares.

The FWL Theorem applies to any regression where there are two or more regressors, and these can logically be broken up into two groups. The regression can thus be written as

$$\boldsymbol{y} = \boldsymbol{X}_1\boldsymbol{\beta}_1 + \boldsymbol{X}_2\boldsymbol{\beta}_2 + \text{residuals}, \tag{1.18}$$

where $\boldsymbol{X}_1$ is $n \times k_1$ and $\boldsymbol{X}_2$ is $n \times k_2$, with $\boldsymbol{X} \equiv [\boldsymbol{X}_1 \quad \boldsymbol{X}_2]$ and $k = k_1 + k_2$. For example, $\boldsymbol{X}_1$ might be seasonal dummy variables or trend variables and $\boldsymbol{X}_2$ genuine economic variables. This was in fact the type of situation dealt with by Frisch and Waugh (1933) and Lovell (1963). Another possibility is that $\boldsymbol{X}_1$ might be regressors, the joint significance of which we desire to test, and $\boldsymbol{X}_2$ might be other regressors that are not being tested. Or $\boldsymbol{X}_1$ might be regressors that are known to be orthogonal to the regressand, and $\boldsymbol{X}_2$ might be regressors that are not orthogonal to it, a situation which arises very frequently when we wish to test nonlinear regression models; see Chapter 6.

Now consider another regression,

$$\boldsymbol{M}_1\boldsymbol{y} = \boldsymbol{M}_1\boldsymbol{X}_2\boldsymbol{\beta}_2 + \text{residuals}, \tag{1.19}$$

where $\boldsymbol{M}_1$ is the matrix that projects off $\mathcal{S}(\boldsymbol{X}_1)$. In (1.19) we have first regressed $\boldsymbol{y}$ and each of the k_2 columns of $\boldsymbol{X}_2$ on $\boldsymbol{X}_1$ and then regressed the vector of residuals $\boldsymbol{M}_1\boldsymbol{y}$ on the $n \times k_2$ matrix of residuals $\boldsymbol{M}_1\boldsymbol{X}_2$. The FWL Theorem tells us that the residuals from regressions (1.18) and (1.19), and the OLS estimates of $\boldsymbol{\beta}_2$ from those two regressions, will be *numerically identical.* Geometrically, in regression (1.18) we project $\boldsymbol{y}$ directly onto $\mathcal{S}(\boldsymbol{X}) \equiv \mathcal{S}(\boldsymbol{X}_1, \boldsymbol{X}_2)$, while in regression (1.19) we first project $\boldsymbol{y}$ and all of the columns of $\boldsymbol{X}_2$ off $\mathcal{S}(\boldsymbol{X}_1)$ and then project the residuals $\boldsymbol{M}_1\boldsymbol{y}$ onto the span of the matrix of residuals, $\mathcal{S}(\boldsymbol{M}_1\boldsymbol{X}_2)$. The FWL Theorem tells us that these two apparently rather different procedures actually amount to the same thing.

The FWL Theorem can be proved in several different ways. One standard proof is based on the algebra of partitioned matrices. First, observe that the estimate of $\boldsymbol{\beta}_2$ from (1.19) is

$$\left(\boldsymbol{X}_2^{\top}\boldsymbol{M}_1\boldsymbol{X}_2\right)^{-1}\boldsymbol{X}_2^{\top}\boldsymbol{M}_1\boldsymbol{y}. \tag{1.20}$$

This simple expression, which we will make use of many times, follows immediately from substituting $\boldsymbol{M}_1\boldsymbol{X}_2$ for $\boldsymbol{X}$ and $\boldsymbol{M}_1\boldsymbol{y}$ for $\boldsymbol{y}$ in expression (1.04) for the vector of OLS estimates. The algebraic proof would now use results on the inverse of a partitioned matrix (see Appendix A) to demonstrate that the OLS estimate from (1.18), $\hat{\boldsymbol{\beta}}_2$, is identical to (1.20) and would then go