where $\boldsymbol{\Omega}_0$ is the matrix $\boldsymbol{\Omega}(\rho)$ defined in (10.05), evaluated at ρ_0 and ω_0. Evidently, (10.07) will in general not be consistently estimated by the OLS covariance matrix estimator $s^2(\boldsymbol{X}^\top\boldsymbol{X})^{-1}$. Except in special cases, it is not possible to say whether the incorrect standard error estimates obtained using OLS will be larger or smaller than the correct ones obtained by taking the square roots of the diagonal elements of (10.07). However, analysis of special cases suggests that for values of ρ greater than 0 (the most commonly encountered case) the incorrect OLS standard errors are usually too small; see, among others, Nicholls and Pagan (1977), Sathe and Vinod (1974), and Vinod (1976).

Expression (10.07) applies to any situation in which OLS is incorrectly used in place of GLS and not merely to situations in which the errors follow an AR(1) process. So does the previous result that $\hat{\boldsymbol{\beta}}$ is unbiased if $\boldsymbol{X}$ is fixed and $E(\boldsymbol{X}^\top\boldsymbol{u}) = \mathbf{0}$. But recall from Section 9.5 that, even when these conditions are satisfied, $\hat{\boldsymbol{\beta}}$ may fail to be consistent if the errors are correlated enough among themselves. We may conclude that, when the regressors are fixed and the covariance matrix of the error terms is such that there is not too much correlation of the error terms, the OLS estimates will be consistent, but the OLS covariance matrix estimate will not be. A consistent estimate of the covariance matrix of the OLS estimator can usually be found. However, since the proof of the Gauss-Markov Theorem depended on the assumption that $E(\boldsymbol{u}\boldsymbol{u}^\top) = \sigma^2\mathbf{I}$, OLS is not the best linear unbiased estimator when this assumption does not hold.

The preceding discussion assumed that there were no lagged dependent variables among the columns of $\boldsymbol{X}$. When this assumption is dropped, the results change drastically, and OLS is seen to be both biased and inconsistent. The simplest way to see this is to think about an element of $\boldsymbol{X}^\top\boldsymbol{u}$ corresponding to the lagged dependent variable (or to one of the lagged dependent variables if there is more than one). If the dependent variable is lagged j periods, this element is

$$\sum_{t=1}^{n} y_{t-j}u_t. \tag{10.08}$$

Now recall expression (10.03), in which we expressed u_t as a function of u_{t-j} and of all the innovations between periods $t-j+1$ and t. Since y_{t-j} is equal to $\boldsymbol{X}_{t-j}\boldsymbol{\beta} + u_{t-j}$, it is clear from (10.03) that (10.08) cannot possibly have expectation zero. Thus we conclude that when $\boldsymbol{X}$ includes lagged dependent variables and u_t is serially correlated,

$$\operatorname*{plim}_{n\to\infty}\left(\frac{1}{n}\boldsymbol{X}^\top\boldsymbol{u}\right) \neq \mathbf{0}, \tag{10.09}$$

which implies that

$$\operatorname*{plim}_{n\to\infty}\bigl(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0\bigr) = \operatorname*{plim}_{n\to\infty}\left(\frac{1}{n}\boldsymbol{X}^\top\boldsymbol{X}\right)^{-1}\operatorname*{plim}_{n\to\infty}\left(\frac{1}{n}\boldsymbol{X}^\top\boldsymbol{u}\right) \neq \mathbf{0}. \tag{10.10}$$