and employing a specification-consistent estimator. Even if the latter provides consistent estimates of some parameters, it may say nothing about the most interesting ones and may allow serious specification errors to go undetected. Another reason is that there is little evidence concerning the properties of (17.66) in finite samples. The results of Cragg (1983) and Tauchen (1986) for related estimators suggest that these may sometimes be poor.

One important practical problem is how to choose the lag truncation parameter p. Theory is signally unhelpful here. As we mentioned earlier, there are results establishing rates at which p may tend to infinity as the sample size tends to infinity. But if an investigator has a sample of precisely 136 observations, what value of p should be chosen? Andrews (1991b) confronts this problem directly and provides data-dependent methods for choosing p, based on the estimation of an optimal value of a parameter he defines. It is fair to say that none of his methods is elementary, and we cannot discuss them here. Perhaps the most encouraging outcome of his investigations is that, in the neighborhood of the optimal value of p, variations in p have little influence on the performance of the HAC estimator.

Andrews (1991b) also provides valuable evidence about HAC covariance matrix estimators, (17.64) and others, from Monte Carlo experiments. Perhaps the most important finding is that *none* of the HAC estimators he considers is at all reliable for sample sizes up to 250 or so if the errors follow an AR(1) process with autocorrelation parameter greater than 0.9. This disappointing result is related to the fact that AR(1) processes with parameters near unity are close to having what is called a **unit root**. This phenomenon is studied in Chapter 20, and we will see that unit roots throw most conventional econometric theory into confusion.

If we stay away from unit roots and near-unit roots, things are more orderly. We saw in Chapter 16 that it is possible to use HCCMEs even in the presence of homoskedasticity with little loss of accuracy, provided that one of the better HCCMEs is used. It appears that much the same is true for HAC estimators. With an ordinary regression model with serially uncorrelated, homoskedastic errors, the loss of precision due to the use of the Newey-West estimator, say, as opposed to the usual OLS estimator, $\hat{\sigma}^2(\boldsymbol{X}^\top\boldsymbol{X})^{-1}$, is small. With some of the other HAC estimators considered by Andrews, the loss is smaller still, which implies that the Newey-West estimator is generally not the best available. Similarly, if the errors are heteroskedastic but still serially uncorrelated, then an HCCME is much better than the OLS estimator but only very slightly better than the HAC estimator.

If the errors are autocorrelated at order one and homoskedastic, both the OLS estimator and the HCCME are dominated not only by the HAC estimator, as one would expect, but also by the straightforward estimator computed by estimating the autocorrelation parameter ρ and using the covariance matrix estimator of a feasible GLS procedure. This last estimator is in these circumstances preferable to the HAC ones. In fact, it is only when the errors are