where $\boldsymbol{g}(\boldsymbol{\theta})$ denotes the gradient of Q, that is, the k–vector with typical component $\partial Q(\boldsymbol{\theta})/\partial\theta_j$. As usual, $\mathcal{H}^*$ denotes a matrix of which the elements are evaluated at the appropriate $\boldsymbol{\theta}_j^*$.

If we are to be able to deduce the asymptotic normality of $\hat{\boldsymbol{\theta}}$ from (17.21), it must be possible to apply a law of large numbers to $\mathcal{H}^*$ and a central limit theorem to $n^{1/2}\boldsymbol{g}(\boldsymbol{\theta}_0)$. We would then obtain the result that

$$n^{1/2}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) \overset{a}{=} -\Big(\operatorname*{plim}_{n\to\infty} \mathcal{H}_0\Big)^{-1} n^{1/2}\boldsymbol{g}(\boldsymbol{\theta}_0). \tag{17.22}$$

What regularity conditions do we need for (17.22)? First, in order to justify the short Taylor expansion in (17.20), it is necessary that Q be at least twice continuously differentiable with respect to $\boldsymbol{\theta}$. If so, then it follows that the Hessian of Q is $O(1)$ as $n \to \infty$. Because of this, we denote it by $\mathcal{H}_0$ rather than $\boldsymbol{H}$; see Section 8.2. Then we need conditions that allow the application of a law of large numbers and a central limit theorem. Rather formally, we may state a theorem based closely on Theorem 8.3 as follows:

Theorem 17.2. Asymptotic Normality of M-Estimators

The *M*-estimator derived from the sequence of criterion functions Q is asymptotically normal if it satisfies the conditions of Theorem 17.1 and if in addition

(i) for all n and for all $\boldsymbol{\theta} \in \Theta$, $Q^n(\boldsymbol{y}^n, \boldsymbol{\theta})$ is twice continuously differentiable with respect to $\boldsymbol{\theta}$ for almost all $\boldsymbol{y}$, and the limit function $\bar{Q}(\mu, \boldsymbol{\theta})$ is twice continuously differentiable with respect to $\boldsymbol{\theta}$ for all $\boldsymbol{\theta} \in \Theta$ and for all $\mu \in \mathbb{M}$;

(ii) for all DGPs $\mu \in \mathbb{M}$ and for all sequences $\{\boldsymbol{\theta}^n\}$ that tend in probability to $\boldsymbol{\theta}(\mu)$ as $n \to \infty$, the Hessian matrix $\mathcal{H}^n(\boldsymbol{y}^n, \boldsymbol{\theta}^n)$ of Q^n with respect to $\boldsymbol{\theta}$ tends uniformly in probability to a positive definite, finite, nonrandom matrix $\mathcal{H}(\mu)$; and

(iii) for all DGPs $\mu \in \mathbb{M}$, $n^{1/2}$ times the gradient of $Q^n(\boldsymbol{y}^n, \boldsymbol{\theta})$, or $n^{1/2}\boldsymbol{g}\big(\boldsymbol{y}^n, \boldsymbol{\theta}(\mu)\big)$, converges in distribution as $n \to \infty$ to a multivariate normal distribution with mean zero and finite covariance matrix $\boldsymbol{V}(\mu)$.

Under these conditions, the distribution of $n^{1/2}\big(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}(\mu)\big)$ tends to $N\big(\mathbf{0},\, \mathcal{H}(\mu)^{-1}\boldsymbol{V}(\mu)\mathcal{H}(\mu)^{-1}\big)$.

It is not worth spending any time on the proof of Theorem 17.2. What we must do, instead, is to return to the GMM case and investigate the conditions under which the criterion function (17.13), suitably divided by n^2, satisfies the requirements of the theorem. Without further ado, we assume that all of the contributions $f_{ti}(y_t, \boldsymbol{\theta})$ are at least twice continuously differentiable with respect to $\boldsymbol{\theta}$ for all $\boldsymbol{\theta} \in \Theta$, for all y_t, and for all allowed values of any predetermined or exogenous variables on which they may depend. Next, we