in month t would affect the value of instruments maturing in months t, $t+1$, and $t+2$ but would not directly affect the value of instruments maturing later, because the latter would not yet have been issued. This suggests that the error term should be modeled by an MA(2) process; see Frankel (1980) and Hansen and Hodrick (1980). Moving average errors also arise when data are gathered using a survey that includes some of the same respondents in consecutive periods, such as the labor force surveys in both the United States and Canada, which are used to estimate unemployment rates; see Hausman and Watson (1985).

It is generally somewhat harder to estimate regression models with moving average errors than to estimate models with autoregressive errors. To see why, suppose that we want to estimate the model

$$y_t = x_t(\boldsymbol{\beta}) + u_t, \quad u_t = \varepsilon_t - \alpha\varepsilon_{t-1}, \quad \varepsilon_t \sim \mathrm{IID}(0, \omega^2). \tag{10.61}$$

Compared with (10.57), we have dropped the subscript from α and changed its sign for convenience; the sign change is of course purely a normalization. Let us make the asymptotically innocuous assumption that the unobserved innovation ε_0 is equal to zero (techniques that do not make this assumption will be discussed below). Then we see that

$$\begin{aligned} y_1 &= x_1(\boldsymbol{\beta}) + \varepsilon_1 \\ y_2 &= x_2(\boldsymbol{\beta}) - \alpha\big(y_1 - x_1(\boldsymbol{\beta})\big) + \varepsilon_2 \\ y_3 &= x_3(\boldsymbol{\beta}) - \alpha\big(y_2 - x_2(\boldsymbol{\beta})\big) - \alpha^2\big(y_1 - x_1(\boldsymbol{\beta})\big) + \varepsilon_3, \end{aligned} \tag{10.62}$$

and so on. By making the definitions

$$\begin{aligned} &y_0^* = 0; \quad y_t^* = y_t + \alpha y_{t-1}^*, \;\; t = 1, \ldots, n; \\ &x_0^* = 0; \quad x_t^*(\boldsymbol{\beta}, \alpha) = x_t(\boldsymbol{\beta}) + \alpha x_{t-1}^*(\boldsymbol{\beta}, \alpha), \;\; t = 1, \ldots, n, \end{aligned} \tag{10.63}$$

we can write equations (10.62) in the form

$$y_t = -\alpha y_{t-1}^* + x_t^*(\boldsymbol{\beta}, \alpha) + \varepsilon_t, \tag{10.64}$$

which makes it clear that we have a nonlinear regression model. But the regression function depends on the entire sample up to period t, since y_{t-1}^* depends on all previous values of y_t and x_t^* depends on $x_{t-i}(\boldsymbol{\beta})$ for all $i \geq 0$. In the by no means unlikely case in which $|\alpha| = 1$, the dependence of y_t on past values does not even tend to diminish as those values recede into the distant past. If we have a specialized program for estimation with MA(1) errors, or a smart nonlinear least squares program that allows us to define the regression function recursively, as in (10.63), estimating (10.64) need not be any more difficult than estimating other nonlinear regression models. But if appropriate software is lacking, this estimation can be quite difficult.