4
Estimation and Asymptotic Properties

Given a function $\boldsymbol{f}$ parameterized by a vector $\boldsymbol{\psi} \in \mathbb{R}^m$, $\boldsymbol{f}(x, \boldsymbol{\psi})$, the parameters $\boldsymbol{\Gamma}_1, \ldots, \boldsymbol{\Gamma}_p$ and $\boldsymbol{\psi}$ are estimated in two stages. Set $\boldsymbol{\eta}' = [\boldsymbol{\psi}', \text{vec}(\boldsymbol{\Gamma}_1)', \ldots, \text{vec}(\boldsymbol{\Gamma}_p)']'$ and define the Nonlinear Least Squares (NLLS) estimator of $\boldsymbol{\eta}$ as

$$\widehat{\boldsymbol{\eta}} = \underset{\boldsymbol{\eta}}{\mathsf{argmin}} \mathcal{Q}_T(\boldsymbol{Y}, \boldsymbol{\eta}) = \underset{\boldsymbol{\eta}}{\mathsf{argmin}} \sum_{t=1}^{T} \boldsymbol{\epsilon}_t(\boldsymbol{\eta})' \boldsymbol{\epsilon}_t(\boldsymbol{\eta}), \tag{4.1}$$

where $\boldsymbol{\epsilon}_t(\boldsymbol{\eta}) = \Delta \boldsymbol{y}_t - \boldsymbol{f}(\boldsymbol{\beta}' \boldsymbol{y}_{t-1}, \boldsymbol{\psi}) - \sum_{i=1}^{p} \boldsymbol{\Gamma}_i \Delta \boldsymbol{y}_{t-i}$ and $\boldsymbol{Y} = (\boldsymbol{y}_1, \ldots, \boldsymbol{y}_T)'$ is a $(T \times n)$ matrix representing the dataset.

Consider the following estimation procedure:

(a) Estimate $\widehat{\boldsymbol{\beta}}$ super-consistently as discussed in Section 3.

(b) Estimate Equation (2.1) by NLLS using $\widehat{\boldsymbol{\beta}}$ instead of $\boldsymbol{\beta}$.

Define the following Jacobian:

$$\boldsymbol{J}(\boldsymbol{y}_t, \boldsymbol{\eta}) = \frac{\partial \boldsymbol{\epsilon}_t(\boldsymbol{\eta})}{\partial \boldsymbol{\eta}}$$

Proposition 2 *Under Assumptions 1–4,* $\sqrt{T}\,(\widehat{\boldsymbol{\eta}} - \boldsymbol{\eta}) \xrightarrow{d} \mathsf{N}(\mathbf{0}, \boldsymbol{\Sigma})$. *Furthermore, the matrix* $\boldsymbol{\Sigma}$ *is consistently estimated by*

$$\widehat{\boldsymbol{\Sigma}} = \left[\sum_{t=1}^{T} \boldsymbol{J}(\boldsymbol{y}_t, \widehat{\boldsymbol{\eta}})' \boldsymbol{J}(\boldsymbol{y}_t, \widehat{\boldsymbol{\eta}})\right]^{-1} \left[\sum_{t=1}^{T} \boldsymbol{J}(\boldsymbol{y}_t, \widehat{\boldsymbol{\eta}})' \boldsymbol{\epsilon}_t(\boldsymbol{\eta}) \boldsymbol{\epsilon}_t(\boldsymbol{\eta})' \boldsymbol{J}(\boldsymbol{y}_t, \widehat{\boldsymbol{\eta}})\right] \left[\sum_{t=1}^{T} \boldsymbol{J}(\boldsymbol{y}_t, \widehat{\boldsymbol{\eta}})' \boldsymbol{J}(\boldsymbol{y}_t, \widehat{\boldsymbol{\eta}})\right]^{-1}.$$

Proof: See Appendix in Section 8. ■

Again, as a consequence of the faster convergence rate of the cointegrating vector $\widehat{\boldsymbol{\beta}}$, the nonlinear least squares of the second stage has standard asymptotics. Using a maximum likelihood approach, Kristensen and Rahbek

(2010) [14] showed that the general distribution of the parameters estimator is non-standard, drawing attention to the fact that it was possible that some models would yield normal distributions, for example, the linear model. However, they did not provide a condition to that neither gave any nonlinear example.

However, having rejected the null hypothesis of linearity in the test presented in the previous section, what model should a researcher estimate? In other words, which function $\boldsymbol{f}$ to choose? In some applications, it is possible that the researcher has a specific function in mind. For example, in Kapetanios, Shin and Snell (2006) [13] it is shown that a incomplete information model may lead exactly to a model with logistic transition equation.

Yet, this will not always be the case. When there is no theoretical function available, we propose a heuristic procedure based on a semi-parametric approach. Comparing the semi-parametric estimate with the existing functions in the nonlinear literature, it is possible to choose the most adequate.

In this paper we do not provide a formal proof for the validity of this approach, even though it seems plausible to provide one, considering the Taylor expansion argument used in the demonstrations of semi-parametric estimations, c.f. Pagan and Ullah (1999) [18].