Bibliography

[1] Apostol, T. *Calculus.* Wiley, New York, 1967.

[2] Balcombe, N., Bailey, A., and Brooks, T. Threshold effects in price transmission: The case of Brazilian wheat, maize and soya prices. *American Journal of Agricultural Economics 89* (2007), 308–323.

[3] Bec, F., and Rahbek, A. Vector equilibrium correction models with non-linear discontinuous adjustments. *Econometrics Journal 7* (2004), 628–651.

[4] Chung, H., Ho, T.-W., and Wei, L.-J. The dynamic relationship between the prices of ADRs and their underlying stocks: Evidence from the threshold vector error correction model. *Applied Economics 37* (2005), 2387–2394.

[5] Davies, R. Hypothesis testing when a nuisance parameter is present only under the alternative. *Biometrika 74* (1987), 33–43.

[6] Engle, R., and Granger, C. Co-integration and error correction: Representation, estimation, and testing. *Econometrica 55* (1987), 251–276.

[7] Gonzalo, J., and Pitarakis, J.-Y. *Threshold Effects in Multivariate Error Correction Models.* 2006.

[8] Granger, C., and Terasvirta, T. *Modelling Nonlinear Dynamic Relationships.* Oxford University Press, Oxford, 1993.

[9] Hansen, B. Inference when a nuisance parameter is not identified under the null hypothesis. *Econometrica 64* (1996), 413–430.

[10] Hansen, B., and Seo, B. Testing for two-regime threshold cointegration in vector error-correction models. *Journal of Econometrics 110* (2002), 293–318.

[11] Ibragimov, R., and Phillips, P. Regression asymptotics using martingale convergence methods. *Econometric Theory 24* (2008), 888–947.

[12] Johansen, S. Estimation and hypothesis testing of cointegration vectors in gaussian vector autoregressive models. *Econometrica 59* (1991), 1551–1580.

[13] Kapetanios, G., Shin, Y., and Snell, A. Testing for cointegration in nonlinear smooth transition error correction models. *Econometric Theory 22* (2006), 279–303.

[14] Kristensen, D., and Rahbek, A. Likelihood-based inference in nonlinear error-correction models. *Journal of Econometrics* (2009), Forthcoming.

[15] Lo, M., and Zivot, E. Threshold cointegration and nonlinear adjustment to the law of one price. *Macroeconomic Dynamics 5* (2001), 533–576.

[16] Luukkonen, R., Saikkonen, P., and Teräsvirta, T. Testing linearity against smooth transition autoregressive models. *Biometrika 75* (1988), 491–499.

[17] Michael, P., Nobay, A., and Peel, D. Transaction costs and nonlinear adjustment in real exchange rates: An empirical investigation. *Journal of Political Economy 105* (1997), 862–879.

[18] Pagan, A., and Ullah, A. *Nonparametric Econometrics*. Wiley, Cambridge, 1999.

[19] Saikkonen, P. Stability results for nonlinear error correction models. *Journal of Econometrics 127* (2005), 69–81.

[20] Saikkonen, P. Stability of regime switching error correction models under linear cointegration. *Econometric Theory 24* (2008), 294–318.

[21] Saikkonen, P., and Choi, I. Cointegrating smooth transition regression. *Econometric Theory 20* (2004), 301–340.

[22] Seo, B. Testing for smooth transition nonlinearity in partially nonstationary vector autoregressions. *Journal of the Korean Statistical Society 36* (2007), 257–274.

[23] Seo, M. H. Bootstrap testing for the null of no cointegration in a threshold vector error correction model. *Journal of Econometrics 134* (2006), 129–150.

[24] Suárez-Fariñas, M., Pedreira, C., and Medeiros, M. Local global neural networks: A new approach for nonlinear time series modeling. *Journal of the American Statistical Association 99* (2004), 1092–1107.

[25] Teräsvirta, T. Specification, estimation, and evaluation of smooth transition autoregressive models. *Journal of the American Statistical Association 89* (1994), 208–218.

[26] van Dijk, D., Teräsvirta, T., and Franses, P. Smooth transition autoregressive models - a survey of recent developments. *Econometric Reviews 21* (2002), 1–47.

[27] Wooldridge, J. M. *Econometric Analysis of Cross Section and Panel Data*. MIT Press, Cambridge, 2001.

8 Appendix

The proofs to Propositions 1 and 2 are in the first section, while the proofs of all lemmas used are in the second section of this appendix.

8.1 Propositions

Proof. (Proposition 1) Suppose $\mathcal{H}_0$ is true. Without loss of generality and or the sake of clarity, consider a system of two variables, only one lag of only one of the variables and no constant. Under $\mathcal{H}_0$ and the knowledge of β all the regressors are stationary. Therefore, the asymptotic distributions would be standard. We will show that using super-consistent $\widehat{\boldsymbol{\beta}}$ yields the same limiting distributions. Consider the first equation of the system:

$$\Delta y_{1t} = \theta_{11}(\widehat{z}_{t-1}) + \theta_{21}(\widehat{z}_{t-1})^2 + \theta_{31}(\widehat{z}_{t-1})^3 + \gamma\Delta y_{1t-1} + \widetilde{\epsilon}_{1t}, \tag{8.1}$$

where $z_t = \boldsymbol{\beta}'\boldsymbol{y}_{t-1}$, $\widehat{z}_t = \widehat{\boldsymbol{\beta}}'\boldsymbol{y}_{t-1}$ and $\widetilde{\epsilon}_{1t} = \epsilon_{1t}+\theta_{11}(\widehat{\beta}_2-\beta_2)y_{2t-1}+\theta_{21}[2z_{t-1}(\widehat{\beta}_2-\beta_2)y_{2t-1}+3(\widehat{\beta}_2-\beta_2)^2y_{2t-1}^2]+\theta_{31}[3z_{t-1}^2(\widehat{\beta}_2-\beta_2)y_{2t-1}-3z_{t-1}(\widehat{\beta}_2-\beta)^2y_{2t-1}^2+5(\widehat{\beta}_2-\beta_2)^3y_{2t-1}^3]$. We will prove the result for this equation, the extension to both equations being straightforward but involving much longer and tedious manipulations.

First, note that

$$\sqrt{T}\begin{pmatrix}\widehat{\gamma}_1-\gamma_1\\ \widehat{\theta}_{11}-\theta_{11}\\ \widehat{\theta}_{21}-\theta_{21}\\ \widehat{\theta}_{31}-\theta_{31}\end{pmatrix}=$$

$$T\begin{pmatrix}\sum_{t=1}^{T}\Delta y_{1t-1}^2 & \sum_{t=1}^{T}\Delta y_{1t-1}\widehat{z}_{t-1} & \sum_{t=1}^{T}\Delta y_{1t-1}\widehat{z}_{t-1}^2 & \sum_{t=1}^{T}\Delta y_{1t-1}\widehat{z}_{t-1}^3\\ \sum_{t=1}^{T}\Delta y_{1t-1}\widehat{z}_{t-1} & \sum_{t=1}^{T}\widehat{z}_{t-1}^2 & \sum_{t=1}^{T}\widehat{z}_{t-1}^3 & \sum_{t=1}^{T}\widehat{z}_{t-1}^4\\ \sum_{t=1}^{T}\Delta y_{1t-1}\widehat{z}_{t-1}^2 & \sum_{t=1}^{T}\widehat{z}_{t-1}^3 & \sum_{t=1}^{T}\widehat{z}_{t-1}^4 & \sum_{t=1}^{T}\widehat{z}_{t-1}^5\\ \sum_{t=1}^{T}\Delta y_{1t-1}\widehat{z}_{t-1}^3 & \sum_{t=1}^{T}\widehat{z}_{t-1}^4 & \sum_{t=1}^{T}\widehat{z}_{t-1}^5 & \sum_{t=1}^{T}\widehat{z}_{t-1}^6\end{pmatrix}^{-1}$$

$$\times\frac{1}{\sqrt{T}}\begin{pmatrix}\sum_{t=1}^{T}\Delta y_{1t-1}\widetilde{\epsilon}_t\\ \sum_{t=1}^{T}\widehat{z}_{t-1}\widetilde{\epsilon}_t\\ \sum_{t=1}^{T}\widehat{z}_{t-1}^2\widetilde{\epsilon}_t\\ \sum_{t=1}^{T}\widehat{z}_{t-1}^3\widetilde{\epsilon}_t\end{pmatrix}.$$

Hence, to establish our result, it suffices to show that:

(a) $\text{plim}\,T^{-1}\sum_{t=1}^{T}\Delta y_{1t-1}\widehat{z}_{t-1}^l=\text{plim}\,T^{-1}\sum_{t=1}^{T}\Delta y_{1t-1}z_{t-1}^l,\ \forall l=1,2,3.$

(b) $\text{plim}\,T^{-1}\sum_{t=1}^{T}\widehat{z}_{t-1}^k=\text{plim}\,T^{-1}\sum_{t=1}^{T}z_{t-1}^k,\forall k=2,3,4,5,6.$

(c) $T^{-1/2}\sum_{t=1}^{T}\Delta y_{1t-1}\widetilde{\epsilon_t}$ has the same asymptotic distribution of $T^{-1/2}\sum_{t=1}^{T}\Delta y_{1t-1}\epsilon_t$.

(d) $T^{-1/2}\sum_{t=1}^{T}\widehat{z}_{t-1}^h\widetilde{\epsilon_t}$ has the same asymptotic distribution of $T^{-1/2}\sum_{t=1}^{T}z_{t-1}^h\epsilon_t$ for h=1,2,3.

(a) and (b) follow directly from Lemma 3.

To prove (d), note that the expression $\widehat{z}_{t-1}^2\widetilde{\epsilon}_{1t}$ equals $\widehat{z}_{t-1}^2\epsilon_{1t}$ plus a number of terms in the form

$$(\widehat{\beta}_2-\beta_2)^i\sum_{t=1}^{T}z_{t-1}^k y_{2t-1}^i.$$

Following what was shown in Lemma 3, as long as $i \geq 1$ and $k > 0$, the expression is $o_p(1)$, i.e., the limit when $T \to \infty$ is zero. If $k = 0$, we need $i \geq 2$. In this case, the inequality is respected in all expressions.

Hence, $T^{-1/2}\sum_{t=1}^{T} \widehat{z}_{t-1}^{h}\widetilde{\epsilon}_{1t} = T^{-1/2}\sum_{t=1}^{T} \widehat{z}_{t-1}^{h}\epsilon_{1t} + o_p(1)$. Again from Lemma 3, $T^{-1/2}\sum_{t=1}^{T} \widehat{z}_{t-1}^{h}\epsilon_{1t} = T^{-1/2}\sum_{t=1}^{T} z_{t-1}^{h}\epsilon_{1t} + o_p(1)$, and from here the result follows.

Proof of claim (c) is analogous to the proof to claim (d).

Now, suppose $\mathcal{H}_A$ is true. To prove the consistency we will show that, under $\mathcal{H}_A$, the F-statistic diverges to infinity. Under the alternative, Δy_{1t} follows Equation (8.1) except for the error, which becomes $\widetilde{\epsilon}_{1t}^{*} = \widetilde{\epsilon}_{1t} + \frac{1}{6}f^{(4)}(k_t, \psi)(z_{t-1})^4$ for some fixed $k_t \in \mathbb{R}$.

Let $\boldsymbol{Z}_t = \left(\widehat{z}_{t-1}, \widehat{z}_{t-1}^{2}, \widehat{z}_{t-1}^{3}, \Delta y_{1t-1}\right)'$ and $\boldsymbol{\mathcal{Z}}_T = (\boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_T)'$. Then, $\mathsf{plim}\left(\frac{1}{T}\boldsymbol{\mathcal{Z}}_T\boldsymbol{\mathcal{Z}}_T'\right)^{-1} = \boldsymbol{\Omega}$ is unchanged whether $\mathcal{H}_0$ is true or not. We will show that under $\mathcal{H}_A$, $T^{-1/2}(\widehat{\theta}_{21} - 0)$ diverges. Let $\widetilde{\boldsymbol{\Omega}}$ be the relevant partition of $\boldsymbol{\Omega}$.

Note that

$$\begin{aligned} T^{-1/2}\widehat{\theta}_{21} &= \widetilde{\boldsymbol{\Omega}}T^{-1/2}\sum_{t=1}^{T}\widehat{z}_{t-1}^{2}\Delta y_{1t} \\ &= T^{1/2}\theta_{21} + \widetilde{\boldsymbol{\Omega}}T^{-1/2}\sum_{t=1}^{T}\widehat{z}_{t-1}^{2}\widetilde{\epsilon}_{1t}^{*} \\ &= T^{1/2}\theta_{21} + \widetilde{\boldsymbol{\Omega}}T^{-1/2}\sum_{t=1}^{T}\left[\widehat{z}_{t-1}^{2}\widetilde{\epsilon}_{1t} + \widehat{z}_{t-1}^{2}\frac{1}{6}f^{(4)}(k_t,\psi)(z_{t-1})^4\right]. \end{aligned}$$

We know, from Assumptions 1 and 4, that $\frac{1}{6}f^{(4)}(k_t, \psi)$ is bounded. So, we can write

$$-K\sum_{t=1}^{T}\widehat{z}_{t-1}^{2}(z_{t-1})^4 < \sum_{t=1}^{T}\widehat{z}_{t-1}^{2}\frac{1}{6}f^{(4)}(k_1,\psi)(z_{t-1})^4 < K\sum_{t=1}^{T}\widehat{z}_{t-1}^{2}(z_{t-1})^4.$$

Pre-multiplying by $T^{-1/2}$, taking limits, and using the results in Lemma 3, we get

$$O_p(1) < \sum_{t=1}^{T}\widehat{z}_{t-1}^{2}\frac{1}{6}f^{(4)}(k_1,\psi)(z_{t-1})^4 < O_p(1).$$

From (d) we know that $T^{-1/2}\sum_{t=1}^{T}\widehat{z}_{t-1}^{2}\widetilde{\epsilon}_{1t}$ is $O_p(1)$. Therefore, we have two limited terms plus $T^{1/2}\theta_{21}$, which will diverge to ∞, giving us the result. The same argument applies to the F-test, only with lengthier calculations.

Finally, from (b) it is easy to see that $\mathsf{plim}T^{-1}(\widehat{\epsilon}_{1t}^{2})$ exists.

■

Proof. (Proposition 2) Again, for the sake of simplicity, let us consider only one lag of only one variable. In addition, without loss of generality, we will assume ψ is scalar. The NLLS problem is

$$\min \frac{1}{2} T^{-1} \sum_{t=1}^{T} \left[\Delta y_{1t} - f(\widehat{\boldsymbol{\beta}}' \boldsymbol{y}_{t-1}, \psi) - \gamma \Delta y_{1t-1} \right]^2 .$$

The first order conditions are:

$$T^{-1} \sum_{t=1}^{T} \widehat{\boldsymbol{s}}_t(\widehat{\psi}, \widehat{\gamma}) = 0,$$

where

$$\widehat{\boldsymbol{s}}_t(\widehat{\psi}, \widehat{\gamma}) = \left[\Delta y_{1t} - f(\widehat{\boldsymbol{\beta}}' \boldsymbol{y}_{t-1}, \widehat{\psi}) - \widehat{\gamma} \Delta y_{1t-1} \right] \begin{bmatrix} f_\psi(\widehat{\boldsymbol{\beta}}' \boldsymbol{y}_{t-1}, \widehat{\psi}) \\ \Delta y_{1t-1} \end{bmatrix} .$$

We will always use the hat to make clear whether the function is calculated with $\widehat{\boldsymbol{\beta}}' \boldsymbol{y}_{t-1}$ or $\boldsymbol{\beta}' \boldsymbol{y}_{t-1}$.

We can make a mean-value expansion around (ψ, γ):

$$\sum_{t=1}^{T} \widehat{\boldsymbol{s}}_t(\psi, \gamma) + \sum_{t=1}^{T} \widehat{\boldsymbol{H}}_t(\widetilde{\psi}, \widetilde{\gamma}) \begin{pmatrix} \psi - \widehat{\psi} \\ \gamma - \widehat{\gamma} \end{pmatrix} = 0,$$

where

$$\widehat{\boldsymbol{H}}_t(\psi, \gamma) = \begin{bmatrix} \frac{\partial \widehat{\boldsymbol{s}}_t(\psi, \gamma)}{\partial \psi} \\ \frac{\partial \widehat{\boldsymbol{s}}_t(\psi, \gamma)}{\partial \gamma} \end{bmatrix}',$$

$(\widetilde{\psi}, \widetilde{\gamma}) = (\widehat{\psi}, \widehat{\gamma}) + t(\psi, \gamma)$, for some $t \in (0, 1)$.

From Lemma 4

$$\text{plim} T^{-1} \sum_{t=1}^{T} \widehat{\boldsymbol{H}}_t(\widetilde{\psi}, \widetilde{\gamma}) = \text{plim} T^{-1} \sum_{t=1}^{T} \boldsymbol{H}_t(\psi, \gamma) = \boldsymbol{H}(\psi, \gamma),$$

for some fixed $\boldsymbol{H}(\psi, \gamma)$.

Therefore,

$$T^{1/2} \begin{pmatrix} \psi - \widehat{\psi} \\ \gamma - \widehat{\gamma} \end{pmatrix} = [\boldsymbol{H}_t(\psi, \gamma)]^{-1} \left(-T^{-1/2} \right) \sum_{t=1}^{T} \widehat{\boldsymbol{s}}_t(\psi, \gamma) + o_p(1).$$

All we have to show now is that

$$(T^{-1/2}) \sum_{t=1}^{T} \widehat{\boldsymbol{s}}_t(\psi, \gamma) = (T^{-1/2}) \sum_{t=1}^{T} \boldsymbol{s}_t(\psi, \gamma) + o_p(1).$$

It is sufficient to show that

(a) $T^{-1/2}\sum_{t=1}^{T}\Delta y_{1t} f_\psi(\widehat{\boldsymbol{\beta}}'\boldsymbol{y}_{t-1},\widehat{\psi}) = T^{-1/2}\sum_{t=1}^{T}\Delta y_{1t} f_\psi(\boldsymbol{\beta}'\boldsymbol{y}_{t-1},\widehat{\psi}) + o_p(1)$

(b) $T^{-1/2}\sum_{t=1}^{T} f(\widehat{\boldsymbol{\beta}}'\boldsymbol{y}_{t-1},\widehat{\psi}) f_\psi(\widehat{\boldsymbol{\beta}}'\boldsymbol{y}_{t-1},\widehat{\psi}) =$
$T^{-1/2}\sum_{t=1}^{T} f(\boldsymbol{\beta}'\boldsymbol{y}_{t-1},\widehat{\psi}) f_\psi(\boldsymbol{\beta}'\boldsymbol{y}_{t-1},\widehat{\psi}) + o_p(1)$

Claims (a) and (b) follow directly from Lemma 3 and Assumption 4.

As to the covariance matrix estimator, the proof is standard. Since we only need to use the Law of Large Numbers, the non-stationarity of the variables does not bring any extra complications Wooldridge (2001) [27].

■

8.2 Lemmas

Lemma 1 *Suppose f is a function which is n times continuously differentiable on the closed interval [a - r, a + r] and n + 1 times differentiable on the open interval (a-r,a+r). If there exists a positive real constant M_n such that $|f^{(n+1)}(x)| < M_n, \forall x \in (a-r, a+r)$, then*

$$f(x) = f(a) + f'(a)\frac{(x-a)}{1!} + f''(a)\frac{(x-a)^2}{2!} + \ldots + f^{(n)}(b)\frac{(x-a)^n}{n!}$$

for some $b \in (a, x)$.

Proof. See Apostol (1967) [1].

■

From Ibragimov and Phillips (2008) [11]:

Theorem 1 *Let $f : \mathbb{R} \to \mathbb{R}$ be a twice continuously differentiable function such that f' satisfies the growth condition $|f'(x)| \leq K(1 + |x|^\alpha), \forall x \in \boldsymbol{R}$ for some constants $K > 0$ and $\alpha < 0$. Suppose that u_t and v_t are two linear processes $u_t = \sum_{j=1}^{\infty}\gamma_j\epsilon_{t-j}$ and $v_t = \sum_{j=1}^{\infty}\delta_j\epsilon_{t-j}$ where $\sum_{j=1}^{\infty} j|\gamma_j| < \infty$, $\sum_{j=1}^{\infty} j|\delta_j| < \infty$ and $(\epsilon_t)_{t\in\boldsymbol{Z}}$ are zero-mean i.i.d. random variables with $E[\epsilon_0^2] < \infty$ and $E[|\epsilon_0|^p] < \infty$ for $p \geq max(6, 4\alpha)$. Then*

$$\frac{1}{\sqrt{T}}\sum_{t=2}^{[Tr]} f\left(\frac{1}{\sqrt{T}}\sum_{i=1}^{t-1} u_i\right) v_t \xrightarrow{d} \lambda_{uv}\int_0^r f'(\omega_u W(v))dv + \omega_v\int_0^r f(\omega_u W(v))d(W(v),$$

where $\omega_u = E[u_t^2]$, $\omega_v = E[v_t^2]$ and $\lambda_{uv} = \sum_{j=1}^{\infty} E[u_0 v_0]$.

The exact form of the limiting distribution is not relevant for our results. What we need is the following corollary.

Lemma 2 *Under the conditions of Theorem 1,*

$$T^{-1/2}\sum_{t=2}^{T} f\left(T^{-1/2}\sum_{i=1}^{t-1} u_i\right) v_t = O_p(1)$$

.

Note that the derivatives of any polynomial function satisfy the growth condition.

Lemma 3 *Let v_t be a stationary process, $\boldsymbol{y}_t$ be an $I(1)$ cointegrated vector, with cointegration vector $\boldsymbol{\beta}$ and $\widehat{\boldsymbol{\beta}}$ a super-consistent estimate of $\boldsymbol{\beta}$. Let also, for some $d < \infty$, $f : \mathbb{R} \to \mathbb{R}$ be d times continuously differentiable and the d-th derivative of f be limited. Then, $T^{-1/2}\sum_{t=1}^{T} f(\boldsymbol{\beta}'\boldsymbol{y}_t)v_t = T^{-1/2}\sum_{t=1}^{T} f(\widehat{\boldsymbol{\beta}}'\boldsymbol{y}_t)v_t + o_p(1)$, and $\text{plim}\, T^{-1}\sum_{t=1}^{T} f(\boldsymbol{\beta}'\boldsymbol{y}_t) = \text{plim}\, T^{-1}\sum_{t=1}^{T} f(\widehat{\boldsymbol{\beta}}'\boldsymbol{y}_t)$.*

Proof. For the first result, consider first a two dimensional case $\boldsymbol{\beta}'\boldsymbol{y}_t = y_{1t} + \beta_2 y_{2t}$. Using Lemma 1 to expand f around $\beta_2 y_{2t}$,

$$f(\widehat{\boldsymbol{\beta}}'\boldsymbol{y}_t) = f(\boldsymbol{\beta}'\boldsymbol{y}_t) + f'(\boldsymbol{\beta}'\boldsymbol{y}_t)(\widehat{\beta}_2 - \beta_2)y_{2t} + \cdots + \frac{f^{(d-1)}(\boldsymbol{\beta}'\boldsymbol{y}_t)(\widehat{\beta}_2 - \beta_2)^{d-1}y_{2t}^{d-1}}{(d-1)!} + \frac{f^d(\widetilde{\boldsymbol{\beta}'\boldsymbol{y}_t})(\widehat{\beta}_2 - \beta_2)^d y_{2t}^d}{d!}$$

for some $\widetilde{\boldsymbol{\beta}'\boldsymbol{y}_t} \in \left(\widehat{\boldsymbol{\beta}}'\boldsymbol{y}_t, \boldsymbol{\beta}'\boldsymbol{y}_t\right)$. Taking the k-th term, such that $3 \le k \le d-1$, we have, by Lemma 2,

$$\begin{aligned}(\widehat{\beta}_2 - \beta_2)^k \sum_{t=1}^{T} f^k(\boldsymbol{\beta}'\boldsymbol{y}_t)v_t y_{2t}^k/k! &= T^{-(k-1)/2}\left[T(\widehat{\beta}_2 - \beta_2)\right]^k \\ &\quad\times\left[T^{-(k+1)/2}\frac{\sum_{t=1}^{T} f^k(\boldsymbol{\beta}'\boldsymbol{y}_t)v_t y_{2t}^k}{k!}\right] \\ &= T^{-(k-1)/2}O_p(1)O_p(1) = o_p(1).\end{aligned}$$

For $k = 2$ we get a $O_p(1)$, but it will be further divided by $T^{1/2}$, giving us an $o_p(1)$. Since the d-th derivative is limited, for some $M \in \mathbb{R}$, the sum of the d-th term is bounded by

$$\begin{aligned}\pm M\sum_{t=1}^{T}(\widehat{\beta}_2 - \beta_2)^d v_t y_{2t}^d &= \pm M T^{-(d-1)/2}\left[T(\widehat{\beta}_2 - \beta_2)\right]^d\left[T^{-(d+1)/2}\sum_{t=1}^{T} v_t y_{2t}^d\right] \\ &= T^{-(d-1)/2}O_p(1)O_p(1).\end{aligned}$$

Again, if $d \ne 1$ we have an $o_p(1)$ expression, if $d = 1$, we are back to the $k = 2$ case. Therefore, the only remaining term is the first, which gives us the result.

For the multidimensional case, just repeat the reasoning for each dimension of $\boldsymbol{\beta}'$.

The second result is proven by the same line of reasoning. The only difference is that in the end of each expression we will have $T^{-(i-2)/2}O_p(1)O_p(1)$, giving us an $o_p(1)$ except for $i = 1, 2$. But since the expression will be divided by T, we will have $o_p(1)$ for every i.

■

Lemma 4 $\text{plim}\, T^{-1}\sum_{t=1}^{T}\widehat{\boldsymbol{H}}_t(\widetilde{\psi},\widetilde{\gamma}) = \text{plim}\, T^{-1}\sum_{t=1}^{T}\boldsymbol{H}_t(\psi,\gamma)$.

Proof. $\widehat{\boldsymbol{H}}_t(\widetilde{\psi},\widetilde{\gamma})$ equals

$$\begin{bmatrix} L & \Delta y_{1t-1} f_{\psi}(\widehat{\boldsymbol{\beta}}'\boldsymbol{y}_{t-1},\widetilde{\psi}) \\ \Delta y_{1t-1} f_{\psi}(\widehat{\boldsymbol{\beta}}'\boldsymbol{y}_{t-1},\widetilde{\psi}) & \Delta y_{1t-1}^2 \end{bmatrix}$$

Where $L = f_{\psi\psi}(\widehat{\boldsymbol{\beta}}'\boldsymbol{y}_{t-1},\widetilde{\psi})(-\Delta\boldsymbol{y}_{1t} + f(\widehat{\boldsymbol{\beta}}'\boldsymbol{y}_{t-1},\widetilde{\psi}) + \widetilde{\gamma}\Delta y_{1t-1}) + f_{\psi}(\widehat{\boldsymbol{\beta}}'\boldsymbol{y}_{t-1},\widetilde{\psi})^2$. From Lemma 3 and Assumption 4:

(a) $\text{plim}\, T^{-1}\sum_{t=1}^{T}\Delta y_{1t-1} f_{\psi}(\widehat{\boldsymbol{\beta}}'\boldsymbol{y}_{t-1},\widetilde{\psi}) = \text{plim}\, T^{-1}\sum_{t=1}^{T}\Delta y_{1t-1} f_{\psi}(\boldsymbol{\beta}'\boldsymbol{y}_{t-1},\widetilde{\psi})$

(b) $\text{plim}\, T^{-1}\sum_{t=1}^{T} f_{\psi\psi}(\widehat{\boldsymbol{\beta}}'\boldsymbol{y}_{t-1},\widetilde{\psi})\left[-\Delta y_{1t} + f(\widehat{\boldsymbol{\beta}}'\boldsymbol{y}_{t-1},\widetilde{\psi}) + \widetilde{\gamma}\Delta y_{1t-1}\right] =$
$\text{plim}\, T^{-1}\sum_{t=1}^{T} f_{\psi\psi}(\boldsymbol{\beta}'\boldsymbol{y}_{t-1},\widetilde{\psi})\left[-\Delta y_{1t} + f(\boldsymbol{\beta}'\boldsymbol{y}_{t-1},\widetilde{\psi}) + \widetilde{\gamma}\Delta y_{1t-1}\right]$

(c) $\text{plim}\, T^{-1}\sum_{t=1}^{T} f_{\psi}(\widehat{\boldsymbol{\beta}}'\boldsymbol{y}_{t-1},\widetilde{\psi})^2 = \text{plim}\, T^{-1}\sum_{t=1}^{T} f_{\psi}(\boldsymbol{\beta}'\boldsymbol{y}_{t-1},\widetilde{\psi})^2$

Therefore, we have established that $\text{plim}\; T^{-1}\sum_{t=1}^{T}\widehat{\boldsymbol{H}}_t(\widetilde{\psi},\widetilde{\gamma}) = \text{plim}\; T^{-1}\sum_{t=1}^{T}\boldsymbol{H}_t(\widetilde{\psi},\widetilde{\gamma})$. Usual nonlinear least squares approach, c.f. Wooldridge (2001) [27], may easily be used to establish $\text{plim}\, T^{-1}\sum_{t=1}^{T}\boldsymbol{H}_t(\widetilde{\psi},\widetilde{\gamma}) = \text{plim}\, T^{-1}\sum_{t=1}^{T}\boldsymbol{H}_t(\psi,\gamma)$. These two equalities give us the result.

■