Testing Self-Exciting Threshold Autoregressive Models Against Structural Change Models: A Nonnested Approach (First Draft)*

M. Rodrigo Dupleich Ulloa †

May 20, 2005

Abstract

This paper studies the asymptotic distribution theory for typical Non-nested tests applied to two type of time series nonlinear models. Our competing models are self-exciting threshold autoregressive (SETAR) and structural break (SB) model. We consider Non-nested tests based on Davidson and McKinnon [1981], Cox [1961] and Cox [1962]. However, contrary to standard asymptotic results in Non-nested tests, our asymptotic limit results are non-standard, combining the tools of Bai [1995] and Bai and Perron [1998]; and Hansen [1996] and Hansen [2000], for SB and SETAR models, respectively.

JEL Classification: C12; C22; C52.

*I would like to acknowledge ideas, comments and critics of Michael Clements, Jeremy Smith, Richard J. Smith and Paulo Parente. Obviously, any errors remains mine

†Address: Department of Economics, University of Warwick, Coventry-UK, CV4 7AL. email: M.R.Dupleich-Ulloa@warwick.ac.uk

1 Introduction

The recent interest in modelling nonlinearities in economic time series has been matched by the development of procedures for testing for the presence of non-linearities. Typically tests fall into broad campus: "portmanteau" tests of the null hypothesis of linearity against an unspecified form on non-linearity (e.g. Neural Network tests as Lee et al. [1993]), and tests designed to have power against an unspecified form of non-linearity. Tong [1990] provide a good review. Test against an unspecified non-linear alternative are generally uninformative about the type of non-linearity present if the null is rejected, and give little indication as to the class of non-linear model to use.

Tests designed to have power against a specific non-linear alternative may nevertheless reject the null in the presence of other forms of non-linearity, raising the possibility of adopting an inappropriate class of non-linear model. We consider two classes of nonlinear time series models, Self-exciting threshold autoregressive (SETAR) models (analyzed in Tong [1990] and Hansen [1996], among others) and Structural break (SB) models (studied in Bai [1995], and Davies et al. [1995]) It might be reasonable to ask whether a economic or financial variables are subject to changes on their statistical structure due to significant shocks. Or else, the process shows a complex dynamics that results from non-linear structures within the underlying DGP.

In order to ask this question, in this paper we investigate standard non-nested hypothesis tests derived by Cox [1961] and Cox [1962], which generalizes the likelihood ratio procedure used in the case of nested hypothesis. The main idea is to "mean adjust" the Likelihood ratio statistic with a measure of closeness between to densities (e.g. Kullback-Leibler measure). In parallel, nonnested hypothesis tests designed by Atkinson [1970] and Davidson and McKinnon [1981] for which local approximations of artificial nesting in a neighborhood of the null hypothesis are also considered.

Under, correct specification of a SETAR or Structural Break (SB) process, the behavior of the parameter estimates are govern by nonstandard distribution, in particular the threshold parameter and the point break, for SETAR and SB respectively (see Hansen [2000] and Bai [1995]). As a consequence, we find that the asymptotic null distribution of the our non-nested tests considered are also nonstandard, differing from the typical procedures for testing separate families of hypothesis.

The plan of the paper is as follows. In section 2 we briefly outline the SETAR and SC models. This Section describes the Hansen [1996] and Andrews [1993] asymptotic distribution theory of the SETAR and SB models, respectively, We set up the standard non-nested tests within our competing models in section 3, and asymptotic theory for these tests is developed in the present context in Section 4. Finally, we end offer some concluding remarks in section 5. Mathematical proofs of the Propositions and Corollaries are presented in the Appendix.

2 Asymptotic results on Structural Break and SETAR models

2.1 Self-excited Threshold Autoregressive model

A $SETAR(q)$ model represents our alternative nonlinear model, and we specify this model similar to Hansen [2000]. such that,

$$\begin{aligned} y_t &= g\left(\mathbf{x}_t, \theta_S\right) + \epsilon_t \\ &= \mathbf{x}_t' \alpha_1 \cdot I(y_{t-d} \leq \gamma) + \mathbf{x}_t' \alpha_2 \cdot I(y_{t-d} > \gamma) + \varepsilon_t \end{aligned} \tag{1}$$

Where, $\alpha_i = (\alpha_{io}, \alpha_{i1}, ..., \alpha_{iq})'$ and $\mathbf{x}_t = (1, y_{t-1}, ..., y_{t-q})'$. Similar to Hansen [2000], the above model can be written as,

$$y_t = \mathbf{x}_t' \alpha_2 + \mathbf{x}_t'\left(\gamma\right) \alpha_T + \varepsilon_t$$

In which, $\alpha_T = \alpha_2 - \alpha_1$ and $\mathbf{x}_t\left(\gamma\right) = \mathbf{x}_t \cdot I(y_{t-d} \leq \gamma)$. In a matrix form,

$$\mathbf{Y} = \mathbf{X}\alpha_2 + \mathbf{X}_\gamma \alpha_T + \varepsilon \tag{2}$$

And,

$$\mathbf{Y} = \mathbf{X}_\gamma^* \alpha^* + \varepsilon \tag{3}$$

Here, $\mathbf{X} = \left(\mathbf{x}_1', ..., \mathbf{x}_T'\right)', \mathbf{X}_\gamma = \left(\mathbf{x}_1'\left(\gamma\right), ..., \mathbf{x}_T\left(\gamma\right)\right)'$, $\mathbf{X}_\gamma^* = \left[\mathbf{X}, \mathbf{X}_\gamma\right]$, $\alpha^* = (\alpha_2', \alpha_T')'$ and $\epsilon = (\epsilon_1, ..., \epsilon_T)$. Analogously,

$$\mathbf{Y} = \mathbf{X}_{1\gamma}\alpha_1 + \mathbf{X}_{2\gamma}\alpha_2 + \varepsilon \tag{4}$$

And,

$$\mathbf{Y} = \mathbf{X}_\gamma^{**}\alpha + \varepsilon \tag{5}$$

Where , $\alpha = (\alpha_1', \alpha_2')'$. and $\mathbf{X}_\gamma^{**} = [\mathbf{X}_{1\gamma}, \mathbf{X}_{2\gamma}]$, with $\mathbf{X}_{i\gamma} = \left(\mathbf{x}_{i,1}'\left(\gamma\right), ..., \mathbf{x}_{i,T}\left(\gamma\right)\right)'$, $\mathbf{x}_{1,t}\left(\gamma\right) = \mathbf{x}_t \cdot I(y_{t-d} \leq \gamma)$ and $\mathbf{x}_{2,t}\left(\gamma\right) = \mathbf{x}_t \cdot I(y_{t-d} > \gamma)$. As noted in Gonzalo and Pitarakis [2002], we have $\mathbf{X} = \sum_{i=1}^{2} \mathbf{X}_{i\gamma}$ and the regressor are such that $\mathbf{X}_{1\gamma}'\mathbf{X}_{2\gamma} = 0$.Again, assumptions on $\{\epsilon_t, \, t \geq 0\}$ are setup in the conditions stated bellow. Threshold models are extensively studied in Tong [1990] and Franses and Van Dijk [2000], among others. Detailed analysis on consistency and limiting distribution is developed in Chan [1993] and Hansen [2000]. Moreover, Chan [1990] and Hansen [1996] consider linearity tests in a threshold framework. Similar to Hansen [1996], the following regularity conditions will be required later,

Condition 2.1. *Let the following assumptions hold,*

C.1 $\{\varepsilon_t\}$ is iid sequence, with zero mean, nonzero variance σ_ε^2 and $\mathbb{E}\,|\varepsilon_t|^{4r} < \infty$ for some $r > 1$. The density of ε_t is bounded and continuous and independent of $\{y_{t-1}, y_{t-2,...}\}$.

C.2 The stability condition, $\max(\alpha_1'\iota, \alpha_2'\iota) < 1$ for which $\iota = [0, 1, 1, ..., 1]'$ a $(q+1)$ vector.

C.3 The autoregressive function is discontinuous, that is $(\alpha_1 - \alpha_2) \cdot \mathbf{x}^* \neq 0$, where $\mathbf{x}^* = (1, y_{q-1}, y_{q-2}, ..., y_0)$ and $y_{q-d} = \gamma$.

Condition (C.1) entails that the absolute continuity of $\pi(\cdot)$ and its *pdf* bounded away from 0 and ∞ over each bounded set. (C.3) shows the nonlinear nature of our problem and γ can be considered as the "location parameter" of the discontinuity. The stochastic stability of the SETAR process is ensured by (C.2), which implies that all roots of the polynomials $\Theta_i(z) = 1 - \sum_{j=1}^{q} \alpha_{ij} z^i$ lie outside the unit circle of the complex plane, for $i = 1, 2$.

Consequently, under (C.1)-(C.3), the SETAR process in (1) is geometric ergodic, which refers to the rate of convergence to the "invariant" distribution $\pi(\cdot)$. Moreover if the chain is started from the initial distribution $\pi(\cdot)$, the process is strict stationary. This result follows from Chan and Tong [1985] once an appropriate drift condition is selected. See also Cline and Pu [1999] for a discussion between the links between exponential stability of a dynamic system and geometric stability of a time series, for which the approach of Chan and Tong [1985] relies. Condition (C.1) can be relaxed to mixing conditions, and $\{y_t\}_{t=0}^{\infty}$ is a Near epoch dependent (NED) process, under (C.2) and (C.3) (see Davidson [2002] for these result), but for simplicity we might keep these framework.

In a previous work, under similar conditions, Chan [1993] obtains an asymptotic distribution for which $T(\widehat{\gamma} - \gamma_0)$ converges weakly to a random variable M_-, where $[M_-, M_+)$ is the unique random interval over which a compound Poisson process attains its global minimum. However, for the practical point of view, these results are not useful since the limit process depends on nuisance parameters. Hansen [2000] tackles this problem using a different approach based on controlling the rate at α_T shirks to zero such that $T(\widehat{\gamma} - \gamma_0)$ weakly converges to a nonstandard but free of nuisance parameters. Hence, the following Proposition summarizes the main results of Hansen [2000], in terms of our purposes, under conditional homoskedasticity,

Proposition 2.2. *Suppose that Assumption 1 in Hansen [2000] holds. Then,*

1. $\widehat{\gamma} \xrightarrow{p} \gamma^0$, and for $\varepsilon > 0$ $P\left(T^{1-2\alpha}\left|\widehat{\gamma} - \gamma^0\right| > \overline{v}\right) \leq \varepsilon$ (Lemma A.5 and Lemma A.9).

2. $\sqrt{T}\left(\widehat{\alpha}-\alpha^{0}\right) \xrightarrow{d} N\left(0, \mathbf{V}_{\alpha}^{\dagger}\right)$ (Lemma A.12).

3. $T^{1-2\alpha}\left(\widehat{\gamma}-\gamma^{0}\right) \Rightarrow \frac{\sigma_{\epsilon}^{2}}{(\mathbf{c}'\mathbf{D}\mathbf{c})f} \arg\max_{-\infty<r<\infty}\left[-\frac{1}{2}|r|+\mathbf{W}(r)\right]$, such that the two sided Brownian motion $\mathbf{W}(r)$ on the real line is,

$$\mathbf{W}(r)=\mathbf{W}_{1}(-r)\cdot I(r<0)+\mathbf{W}_{2}(r)\cdot I(r>0)$$

   Where $\mathbf{W}_{i}(r)$ are independent standard Brownian motions on $[0,\infty)$.

Here $\mathbf{D}(\gamma)=\mathbb{E}\left(\mathbf{x}_{t}\mathbf{x}_{t}' \;/\; y_{t-d}=\gamma\right)$ we define $c<\infty$, such that $\alpha_{T}=\mathbf{c}T^{-\alpha}$ with $0<\alpha<1/2$. Part (1), analyzed in Lemma A.9 of Hansen [2000], determines the convergence rate of the threshold estimate, it is decreasing in α since the speed equals to $T^{1-2\alpha}$. As α increases, the precision in the estimation of γ reduces due to a smallest threshold effect.α_{T}. Lemmas A.11 and A.12 proof the result stated in Proposition 2.2.

Proposition's part 3, considers the weak convergence of the threshold estimate $\widehat{\gamma}$ to a functional of a two-sided Wiener process on $C[0,\infty]$ As showed in Chan [1993], the two sided distribution results from the discontinuity created by the "location" parameter γ of the autoregressive function, this parameter cuts with an hyperplane the Wiener measure in $\mathbf{W}(r)$.

Lower the value of the ratio $\sigma_{\epsilon}^{2}/\left(\mathbf{c}'\mathbf{D}\mathbf{c}\right)f$, the asymptotic distribution of $\widehat{\gamma}$ becomes less dispersed. That is, when we have a small σ_{ϵ}^{2}, a large threshold effect $|c|$ or a high value of $f\left(\gamma_{0}\right)$ (i.e. many observations are concentrated around the location parameter). The distribution of $\arg\max_{-\infty<r<\infty}\left[-\frac{1}{2}|r|+\mathbf{W}(r)\right]$ is analyzed in Bhattacharya and Brockwell [1976]. Chan [1993] obtains a type of similar results in which $T\left(\widehat{\gamma}-\gamma^{0}\right)$ converges weakly to a distribution based on a Compound Poisson process.

2.2 Autoregressive Structural Break model

A standard autoregressive process with one structural break can be written as follows,

$$\begin{aligned} y_{t} &= f\left(\mathbf{z}_{t},\theta_{B}\right)+u_{t} \\ &= \mathbf{z}_{t}'\delta_{1}\cdot I(t\leq T_{1})+\mathbf{z}_{t}'\delta_{2}\cdot I(t>T_{1})+u_{t} \end{aligned} \tag{6}$$

Here, $\delta_{i}=\left(\delta_{io},\delta_{i1},...,\delta_{iq}\right)'$ and $\mathbf{z}_{t}=\left(1,y_{t-1},...,y_{t-q}\right)'$. Similar to Bai [1995] and Bai and Perron [1998], the model given above can be rewritten as,

$$\mathbf{Y}=\mathbf{Z}_{1}\delta_{1}+\mathbf{Z}_{2}\delta_{2}+\mathbf{u} \tag{7}$$

Or,

$$\mathbf{Y}=\overline{\mathbf{Z}}\delta+\mathbf{u} \tag{8}$$

Where, $\mathbf{Y}=(y_1,...,y_T)$, $\mathbf{u}=(u_1,...,u_T)$, $\delta=(\delta_1',\delta_2')'$, $\overline{Z}=[\mathbf{Z}_1,\mathbf{Z}_2]$ a $(T\times(2(q+1)))$ matrix with $\mathbf{Z}_1=\left(\mathbf{z}_1',...,\mathbf{z}_{T_1}',0,...,0\right)'$ and $\mathbf{Z}_2=\left(0,...,0,\mathbf{z}_{T_1+1}',...,\mathbf{z}_T'\right)'$. Additionally,

$$\mathbf{Y}=\mathbf{Z}\delta_2+\mathbf{Z}_1\delta_T+\mathbf{u} \tag{9}$$

And,

$$\mathbf{Y}=\overline{\mathbf{Z}}^*\delta^*+\mathbf{u} \tag{10}$$

Where, $\delta^*=(\delta_1',\delta_T')'$, $\overline{Z}^*=[\mathbf{Z},\mathbf{Z}_1]$, a $(T\times 2p)$ matrix (hereafter $p=q+1$). In (9), $\mathbf{Z}=(\mathbf{z}_1',...,\mathbf{z}_T')'$ and the magnitude of switch is given by $\delta_T=\delta_1-\delta_2$. Through, the paper either the re-parametrization in (8) or (10) are used. Note that $\mathbf{Z}=\sum_{i=1}^2\mathbf{Z}_i$ and $\mathbf{Z}_1'\mathbf{Z}_2=0$. Equation (8) characterize a pure change model as specified in Bai [1995] in which all the autoregressive parameters switch with the break. Throughout, true value of a parameter is denoted with a 0 superscript, such that $\delta^0=\left(\delta_1^{0\prime},\delta_2^{0\prime}\right)'$ and T_1^0 (i.e. $T_1^0=[\pi_0T]$, where $\pi_0\in(0,1)$ and $[\cdot]$ is the greatest integer function).For our purposes, similar to Davies et al. [1995], we establish the following conditions,

Condition 2.3. *, The following assumptions are assumed to hold,*

C.3 $\{u_t\}$ is iid sequence, with zero mean, nonzero variance σ_ε^2 and $\mathbb{E}\left|u_t\right|^{4r}<\infty$ for some $r>1$. The density of u_t is bounded and continuous and independent of $\{y_{t-1},y_{t-2,...}\}$.

C.4 The stability condition, $\max\left(\delta_1'\iota,\delta_2'\iota\right)<1$ for which $\iota=[0,1,1,...,1]'$ a $(q+1)$ vector.

C.5 The autoregressive function is discontinuous, that is $(\delta_1-\delta_2)\cdot\mathbf{z}^*\neq 0$, where $\mathbf{z}^*=(1,y_{q-1,}y_{q-2,...,}y_0)$ and $[\pi_0T]$, where $\pi_0\in(0,1)$.

Our assumptions are similar Bai [1995] and Bai and Perron [1998], and would imply for instance the framework in GMM analyzed in (Andrews [1993]). As in Davies et al. [1995], the process in (6) is piecewise strict stationary and strong mixing. Under a suitable drift criteria, the chain converges to its stationary distribution at a uniform geometric rate. As the previous case, we summarized the main results of Bai [1995] and Bai and Perron [1998], in the following proposition,

Proposition 2.4. *Under Assumptions A1-A7 of given Bai and Perron [1998], we have,*

1. $\widehat{\pi}\xrightarrow{p}\pi^0$ and every $\eta>0$ there exists a $C<\infty$ such that for all large T, $P\left(Tv_t^2\left(\widehat{\pi}-\pi^0\right)>C\right)<\eta$.
2. $\sqrt{T}\left(\widehat{\delta}-\delta^0\right)\xrightarrow{d}N\left(0,\mathbf{V}_\delta^\dagger\right)$
3. $\frac{(\Delta'Q_1\Delta)}{\sigma_u^2}v_t^2\left(\widehat{T}_1-T_1^0\right)\Rightarrow\arg\max_s Z(s)$.

Similar to Hansen's results, $Z(s)$ is a two-sided Wiener process with a drift such that,

$$Z(s) = \begin{cases} W_1(-s) - |s|/2, \text{ if } s \leq 0 \\ \sqrt{\xi} W_2(s) - \xi |s|/2, \text{ if } s > 0 \end{cases}$$

And, $W_i(s)$, $i = 1, 2$ are independent Wiener processes defined on $[0, \infty]$ Moreover,

$$\xi = \frac{\Delta' \mathbf{C}_2 \Delta}{\Delta' \mathbf{C}_1 \Delta}$$

From Bai and Perron's Assumption (A.7), we know that

$$\left(\Delta T_i^0\right)^{-1} \sum_{t=T_{i-1}^0+1}^{T_{i-1}^0+\left[s\Delta T_{i-1}^0\right]} \mathbf{z}_t \mathbf{z}_t' \xrightarrow{p} s\mathbf{C}_i$$

uniformly in $s \in [0, 1]$, for $i = 1, 2$. ($T_0 = 0$ and $T_2 = T$). Assumption (A.6) establishes that $\delta_T = v_t \Delta$ for some Δ independent of T, where $v_t \to 0$ and $T^{\frac{1}{2}-\alpha} v_t \to \infty$ for some $\alpha \in (0, 1/2)$. The former assumption would imply that even if the magnitude of the shift decrease as T grows unbounded, the estimated point of break remains consistent, result established in the second part of Proposition (2.4). As a result the convergence rate of $\widehat{\pi}$ is $T v_t^2$ rather than T. As explained in Bai and Perron [1998], this result permits central limit theorems to operate since we concentrate the analysis in the neighborhood of length C/v_t^2 around the true break date, as v_t shrinks this open ball increases.

The first part of Proposition 2.4 establishes the consistency of the break point estimator $\widehat{\pi}$. It also considers the convergence of the estimates to the true value π_0. This result permits the $\sqrt{T}$ asymptotic normality of the estimated coefficients $\widehat{\delta}$ obtained in part 2 of Proposition 2.4 (See Corollary 1 in Bai [1995]).

3 Non-nested tests

This section is interested in the comparison of different hypothesis of competing models, in our approach our results are derived from papers by Cox [1961]; and Atkinson [1970]. The first author derives the asymptotic distribution of a tests statistic based on the Neyman-Pearson likelihood-ratio, which generalizes the likelihood procedure used in standard nested tests. The second type of literature focuses in introducing a third model, called artificial nesting model that contains both competing models. This approach consists of embedding the alternatives in a general combined model. This procedure is extended in Davidson and McKinnon [1981], they derive a test by considering some local approximations of artificial nesting models in the neighborhood of the hypothesis H_0.

There two other approaches that our paper does not consider, one procedure tests non-nested hypothesis is to take a Bayesian point of view. The second approach derives a Wald tests of the encompassing hypothesis which modify a standard Wald tests for non-nested hypothesis, taking into account the fact that both of the competing models do not necessary belong to the true distribution. The idea behind the encompassing principle relies on whether a particular null model can predict the salient features of the competing alternative model, see Mizon and Richard [1986].

3.1 Davidson and McKinnon P-test

Similar to Davidson and McKinnon [1981], we can define a convex combination of separate family of parametric models specified in (6) and (2) that "artificially" nests both models Lets consider the two competing hypothesis, H_{SB} : $\mathbf{Y} = \overline{\mathbf{Z}}\delta + \mathbf{u}$ and $H_{TR} : \mathbf{Y} = \mathbf{X}^{*}_{\gamma}\alpha^{*} + \epsilon$, the basic idea is to introduce a third hypothesis in which both H_{SB} and H_{TR} are nested and characterize by a common constraint. Therefore, the associated mixture is given by,

$$\mathbf{Y} = \omega \mathbf{X}^{*}_{\gamma}\alpha^{*} + (1-\omega)\,\overline{\mathbf{Z}}\delta + \varsigma \tag{11}$$

Where $\omega \in [0, 1]$ and $H_{SB} : \{\omega = 0\}$ and $H_{TR} : \{\omega = 1\}$. As pointed out in Gourieroux and Monfort [1994], the procedure consists in testing $\{\omega = 0\}$ against $\{\omega \geq 0\}$ and $\{\omega = 1\}$ against $\{\omega \leq 0\}$. The drawback of this procedure is that $(\alpha, \gamma, \sigma_{\varepsilon})$ are unidentifiable under H_{SB} (and (δ, π, σ_u) under H_{TR}). To circumvent the unidentifiability problem, Davidson and McKinnon [1981] by replacing the nuisance parameter $(\alpha, \gamma, \sigma_{\varepsilon})$ by its estimator under H_{TR}. As a result, a pseudo-nested model is given by,

$$\mathbf{Y} = \omega \mathbf{X}^{*}_{\widetilde{\gamma}}\widetilde{\alpha}^{*}(\widetilde{\gamma}) + (1-\omega)\,\overline{\mathbf{Z}}\delta + \varsigma^{*} \tag{12}$$

Similar to Davidson and McKinnon [1993], equation (12) can be pre-multiplied by $\mathbf{M}_{\overline{Z}}$, an orthogonal projection matrix on to the space generated by the columns of $\overline{Z}$, such that,

$$\mathbf{M}_{\overline{\mathbf{Z}}}\mathbf{Y} = \omega \mathbf{M}_{\overline{\mathbf{Z}}}\mathbf{X}^{*}_{\widetilde{\gamma}}\widetilde{\alpha}^{*}(\widetilde{\gamma}) + \mathbf{M}_{\overline{\mathbf{Z}}}\varsigma^{*}$$

Where $\mathbf{M}_{\overline{\mathbf{Z}}} = \mathbf{I}_T - \mathbf{P}_{\overline{\mathbf{Z}}}$, such that $\mathbf{P}_{\overline{Z}} = \overline{\mathbf{Z}}\left(\overline{\mathbf{Z}}'\overline{\mathbf{Z}}\right)^{-1}\overline{\mathbf{Z}}'$ and $\mathbf{M}_{\overline{\mathbf{Z}}}\overline{\mathbf{Z}} = 0$, $\mathbf{P}_{\overline{\mathbf{Z}}}$ projects vectors of compatible dimension on to the column space of $\overline{\mathbf{Z}}$. Therefore, the OLS estimator of ω is,

$$\widehat{\omega}(\widehat{\pi}, \widetilde{\gamma}) = \left(\widetilde{\alpha}^{*}(\widetilde{\gamma})'\,\mathbf{X}^{*\prime}_{\widetilde{\gamma}}\mathbf{M}_{\overline{\mathbf{Z}}}\mathbf{X}^{*}_{\widetilde{\gamma}}\widetilde{\alpha}^{*}(\widetilde{\gamma})\right)^{-1}\widetilde{\alpha}^{*}(\widetilde{\gamma})'\,\mathbf{X}^{*\prime}_{\widetilde{\gamma}}\mathbf{M}_{\overline{Z}}\mathbf{Y} \tag{13}$$

As $\mathbf{X}^{*\prime}_{\widetilde{\gamma}}\widetilde{\alpha}^{*}(\widetilde{\gamma}) = \mathbf{P}_{\mathbf{X}^{*}_{\widetilde{\gamma}}}\mathbf{Y}$. The estimator given in (13) can be rewritten as,

$$\widehat{\omega} = \left(\mathbf{Y}'\mathbf{P}_{\mathbf{X}^{*}_{\widetilde{\gamma}}}\mathbf{M}_{\overline{\mathbf{Z}}}\mathbf{P}_{\mathbf{X}^{*}_{\widetilde{\gamma}}}\mathbf{Y}\right)^{-1}\mathbf{Y}'\mathbf{P}_{\mathbf{X}^{*}_{\widetilde{\gamma}}}\mathbf{M}_{\overline{Z}}\mathbf{Y}$$

Under the null hypothesis $H_0 : \omega = 0$ (i.e. $H_{SB} : \mathbf{Y} = \overline{\mathbf{Z}}^0 \delta^0 + \mathbf{u}$). Similar to Bai [1995], $\mathbf{Y} = \overline{\mathbf{Z}}\delta^0 + \mathbf{u}^*$, where $\mathbf{u}^* = \mathbf{u} + (\overline{\mathbf{Z}}^0 - \overline{\mathbf{Z}})\delta^0$. Thus, (13) equals to,

$$\widehat{\omega}\left(\widehat{\pi}, \widetilde{\gamma}\right) = \left(\widetilde{\alpha}^*\left(\widetilde{\gamma}\right)' \mathbf{X}_{\widetilde{\gamma}}^{*\prime} \mathbf{M}_{\overline{\mathbf{Z}}} \mathbf{X}_{\widetilde{\gamma}}^{*} \widetilde{\alpha}^*\left(\widetilde{\gamma}\right)\right)^{-1} \widetilde{\alpha}^*\left(\widetilde{\gamma}\right)' \mathbf{X}_{\widetilde{\gamma}}^{*\prime} \mathbf{M}_{\overline{Z}} \mathbf{u}^* \tag{14}$$

The associated t-ratio statistic is given by,

$$T_{1\omega}\left(\widetilde{\gamma}\right) = \frac{\widetilde{\alpha}^*\left(\widetilde{\gamma}\right)' \mathbf{X}_{\widetilde{\gamma}}^{*\prime} \mathbf{M}_{\overline{Z}} \mathbf{u}^*}{\widehat{\sigma}_{\varsigma^*}\left[\widetilde{\alpha}^*\left(\widetilde{\gamma}\right)' \mathbf{X}_{\widetilde{\gamma}}^{*\prime} \mathbf{M}_{\overline{\mathbf{Z}}} \mathbf{X}_{\widetilde{\gamma}}^{*} \widetilde{\alpha}^*\left(\widetilde{\gamma}\right)\right]^{1/2}} \tag{15}$$

Where,

$$\widehat{\sigma}_{\varsigma^*} = \frac{1}{T}\left\|\mathbf{Y} - \widehat{\omega}\mathbf{X}_{\widetilde{\gamma}}^{*}\widetilde{\alpha}^*\left(\widetilde{\gamma}\right) - \left(1-\widehat{\omega}\right)\overline{\mathbf{Z}}\widehat{\delta}\right\|^2$$

Analogously, if the change point model in (12) is fixed, such that $\mathbf{Y} = \omega \mathbf{X}_{\gamma}^{**}\alpha + \left(1-\omega\right)\overline{\mathbf{Z}}^{*}\widetilde{\delta}^*\left(\widetilde{\pi}\right) + \varsigma^{**}$, and we premultiply by $\mathbf{M}_{\mathbf{X}_{\gamma}^{**}} = \mathbf{I}_T - \mathbf{P}_{\mathbf{X}_{\gamma}^{**}}$ (where $\mathbf{P}_{\mathbf{X}_{\gamma}^{**}} = \mathbf{X}_{\gamma}^{**}\left(\mathbf{X}_{\gamma}^{**\prime}\mathbf{X}_{\gamma}^{**}\right)^{-1}\mathbf{X}_{\gamma}^{**\prime}$), the LS estimator of ω is given by,

$$\left(1-\widehat{\omega}\right) = \left(\widetilde{\delta}^{*\prime}\widetilde{\mathbf{Z}}^{*\prime}\mathbf{M}_{\mathbf{X}_{\widetilde{\gamma}}^{**}}\widetilde{\mathbf{Z}}^{*}\widetilde{\delta}^{*}\right)^{-1}\widetilde{\delta}^{*\prime}\widetilde{\mathbf{Z}}^{*\prime}\mathbf{M}_{\mathbf{X}_{\widetilde{\gamma}}^{**}}\mathbf{Y}$$

The above formulation helps us to test the null $H_0 : \omega = 1$ (i.e. H_{TR}: $\mathbf{Y} = \mathbf{X}_{\gamma^0}^{**}\alpha^0 + \epsilon$). Therefore,

$$\left(1-\widehat{\omega}\right) = \left(\widetilde{\delta}^{*\prime}\widetilde{\mathbf{Z}}^{*\prime}\mathbf{M}_{\mathbf{X}_{\widetilde{\gamma}}^{**}}\widetilde{\mathbf{Z}}^{*}\widetilde{\delta}^{*}\right)^{-1}\widetilde{\delta}^{*\prime}\widetilde{\mathbf{Z}}^{*\prime}\mathbf{M}_{\mathbf{X}_{\widetilde{\gamma}}^{**}}\varepsilon^* \tag{16}$$

Where, as before $\varepsilon^* = \varepsilon + (\mathbf{X}_{\gamma^0}^{**} - \mathbf{X}_{\widetilde{\gamma}}^{**})\alpha^0$ and the associated t-ratio statistic for this case,

$$T_{2\omega}\left(\widetilde{\pi}\right) = \frac{\widetilde{\delta}^{*\prime}\widetilde{\mathbf{Z}}^{*\prime}\mathbf{M}_{\mathbf{X}_{\widetilde{\gamma}}^{**}}\varepsilon^*}{\widehat{\sigma}_{\varsigma^{**}}\left[\widetilde{\delta}^{*\prime}\widetilde{\mathbf{Z}}^{*\prime}\mathbf{M}_{\mathbf{X}_{\widetilde{\gamma}}^{**}}\widetilde{\mathbf{Z}}^{*}\widetilde{\delta}^{*}\right]^{1/2}} \tag{17}$$

For which the standard errors are defined as,

$$\widehat{\sigma}_{\varsigma^{**}} = \frac{1}{T}\left\|\mathbf{Y} - \widehat{\omega}\mathbf{X}_{\widetilde{\gamma}}^{**}\widehat{\alpha} - \left(1-\widehat{\omega}\right)\widetilde{\mathbf{Z}}^{*}\widetilde{\delta}^{*}\right\|^2$$

As pointed out by Gourieroux and Monfort [1995] for the $P-$tests, "P", stands for the projection matrices $\mathbf{M}_{(\cdot)}$ appearing in (15) and (17). In order to solve the identification problem appearing our mixtures regressions, as suggested by Davies [1977], applied in Chan [1990] and Hansen [1996] for a nonlinearity test in a SETAR framework; and Andrews [1993] and Bai and Perron [1998] (among others) in a change-point test, our competing models. Since, the parameter $\widetilde{\gamma}$ (resp $\widetilde{\pi}$). in (11) (resp 12) only appears under the alternative hypothesis (i.e. H_{TR} (resp (H_{SB})), we can adopt the the union-intersection principle, and consider the test statistic under H_{SB} (resp. H_{TR}) of the form,

$$T_{1\omega}^{\dagger} = \sup_{\widetilde{\gamma}\in\Gamma} T_{1\omega}\left(\widetilde{\gamma}\right) \text{ and } T_{2\omega}^{\dagger} = \sup_{\widetilde{\pi}\in\Pi} T_{2\omega}\left(\widetilde{\pi}\right) \tag{18}$$

Here,$T_{i\omega}\left(\widetilde{\gamma}\right)$ can be seen as a point-wise t-test. Although the paper concentrates on statistics of the form (18). Our results also apply to more general transformations $g(\cdot)$ on $\widetilde{\gamma}\in\Gamma$. Based on superior local power, tests statistics of the form $ave\ T_{1\omega}^{\dagger}=\int_{\Gamma}T_{1\omega}\left(\widetilde{\gamma}\right)dW\left(\gamma\right)$ and $\exp\ T_{1\omega}^{\dagger}=\ln(\int_{\Gamma}\exp\left(1/2\ T_{1\omega}\left(\widetilde{\gamma}\right)\right)dW\left(\gamma\right))$. All three statistics are $g(T_{1\omega})$ maps functionals on $\widetilde{\gamma}\in\Gamma$; each function g is continuous with respect to the uniform metric, monotonic. Similar transformation can be constructed for $T_{2\omega}^{\dagger}$ on $\widetilde{\pi}\in\Pi$.

This approach is subject to four main drawbacks, if the conclusions of such tests indicate that either both of the rival models should be accepted or that both of them should be rejected our approach based on a comprehensive model will fail to provide any useful conclusion. Secondly, remarked by Pesaran [1974], another source of weakness of this procedure relies on the possible high degree of multicollinearity across the explanatory regressors of the competing models

Thirdly, in this artificial nesting approach, there is a degree of arbitrarily in the way a comprehensive model is constructed. Finally, in our formulation in (15), $\widehat{\alpha}^{*\prime}\mathbf{X}_{\widehat{\gamma}}^{\prime*}$ depends on the data through the estimator $\widehat{\alpha}^{*}$ thus they are correlated with the error term. However, Davidson and McKinnon [1981] omit these difficulties and study directly the asymptotic properties of (15). As a consequence, the next section set up our testing problem in Cox's theory framework.

3.2 Cox test

Following Cox [1961] and Cox [1962], a maximum-likelihood ratio procedure proposed by Cox. The idea is to consider, firstly, the usual likelihood ratio (LR) tests statistic and study its asymptotic distributional properties. Since the LR statistic divided by T tends to a non-zero limit under the null. Therefore the tests statistic cannot be equivalent to the generalized Wald and Score test statistics, analyzed in Gourieroux and Monfort [1994] and Gourieroux and Monfort [1995]. Thus, in the second step of the Cox's procedure,a correction is applied and consequently the test statistic converges to zero under the null as $T\uparrow\infty$, namely,

$$C_{1T}(\pi,\gamma)=\widehat{L}_{RS}-\mathbb{E}_{(\widehat{\delta},\widehat{\pi})}(\widehat{L}_{SR}) \tag{19}$$

Where $\widehat{L}_{RS}=(S_{T}^{TR}\left(\widehat{\alpha},\widehat{\gamma}\right)-S_{T}^{SB}(\widehat{\delta},\widehat{\pi}))$, the conditional expectation can be seen as consistent estimator of $\mathbb{E}_{(\delta_{0},\pi_{0})}(S_{T}^{TR}\left(\alpha(\pi_{0}),\gamma(\pi_{0})\right)-S_{T}^{SB}(\delta_{0},\pi_{0}))$, a Kullback-Leibler measured of the closeness of the densities of the SETAR and SB under $H_{SB}:\mathbf{Y}=\overline{\mathbf{Z}}^{0}\delta^{0}+\mathbf{u}$, Thus,

$$\mathbb{E}_{(\widehat{\delta},\widehat{\pi})}(\widehat{L}_{RS})=T\cdot\{\underset{T\rightarrow\infty}{p\lim}(\frac{\widehat{L}_{RS}}{T})\}_{(\widehat{\delta},\widehat{\pi})} \tag{20}$$

If $S_{T}^{SB}(\widehat{\delta},\widehat{\pi})$ and $S_{T}^{TR}\left(\widehat{\alpha},\widehat{\gamma}\right)$ are the estimated LS criterion functions for H_{SB} and H_{TR}, respectively, similar to Walker [1967] and Pesaran [1974], we have that .

$$S_T^{SB}(\delta,\pi) = \left\| \mathbf{Y}-\overline{Z}\delta \right\|^2 \text{ and } S_T^{TR}(\alpha,\gamma) = \left\| \mathbf{Y}-\mathbf{X}_{\gamma}^{**}\alpha \right\|^2$$

Thus,

$$\begin{aligned} S_T^{TR}(\widehat{\alpha},\widehat{\gamma}) - S_T^{SB}(\widehat{\delta},\widehat{\pi}) &= \left\| \mathbf{M}_{\mathbf{X}_{\widehat{\gamma}}^{**}}\mathbf{Y} \right\|^2 - \left\| \mathbf{M}_{\widehat{\mathbf{Z}}}\mathbf{Y} \right\|^2 \\ &= T(\widehat{\sigma}_{\varepsilon}^2(\widehat{\gamma}) - \widehat{\sigma}_u^2(\widehat{\pi})) \end{aligned} \tag{21}$$

The previous difference are the sum of squares of observed residuals obtained when the models (SETAR) and (SB) are fitted to the data. Note, our tests is directly using the difference between objective functions of the LS method, slightly similar to the GMM framework studied in Smith [1992]. It is easy to show that under H_{SB},

$$\left[\underset{T\to\infty}{p\lim} T^{-1} S_T^{SB}(\widehat{\delta},\widehat{\pi}) \right]_{(\delta_0,\pi_0)} = \sigma_u^2(\pi_0) \tag{22}$$

Usually is difficult to determine $p\lim\ T^{-1}S_T^{TR}(\alpha,\gamma)$ under H_{SB}. Similar to Walker [1967] and Pesaran and Deaton [1978], in order to estimates $\sigma_{\varepsilon\pi}^2$ and α_π (for a fixed γ_π),

$$\begin{aligned} \widehat{\alpha}_\pi &= (\mathbf{X}_{\gamma_\pi}^{**\prime}\mathbf{X}_{\gamma_\pi}^{**})^{-1}\mathbf{X}_{\gamma_\pi}^{**}\mathbf{Y} \\ \widehat{\sigma}_{\varepsilon\pi}^2 &= T^{-1} \left\| \mathbf{M}_{\mathbf{X}_{\gamma_\pi}^{**}}\mathbf{Y} \right\|^2 \end{aligned}$$

Which given that H_{SB} is true become, $\alpha_\pi = (\mathbf{X}_{\gamma_\pi}^{**\prime}\mathbf{X}_{\gamma_\pi}^{**})^{-1}\mathbf{X}_{\gamma_\pi}^{**}(\overline{\mathbf{Z}}^0\delta^0+\mathbf{u})$ and $\sigma_{\varepsilon\pi}^2 = T^{-1}\left\| \mathbf{M}_{\mathbf{X}_{\gamma_\pi}^{**}}(\overline{\mathbf{Z}}^0\delta^0+\mathbf{u}) \right\|^2$. Hence, in the view of (20) we have,

$$\underset{T\to\infty}{p\lim}\sigma_{\varepsilon\pi}^2 = \sigma_u^2(\pi_0)+\frac{1}{T}\delta^{0\prime}\overline{\mathbf{Z}}^{0\prime}\mathbf{M}_{\mathbf{X}_{\gamma_\pi}^{**}}\overline{\mathbf{Z}}^0\delta \tag{23}$$

Thus for a fixed γ_π, we have an estimate of $\sigma_{\varepsilon\pi}^2$,

$$\widehat{\sigma}_{\varepsilon\widehat{\pi}}^2 = \widehat{\sigma}_u^2(\widehat{\pi})+\frac{1}{T}\widehat{\delta}'\overline{\mathbf{Z}}'\mathbf{M}_{\mathbf{X}_{\gamma_\pi}^{**}}\overline{\mathbf{Z}}\widehat{\delta} \tag{24}$$

Hence, if γ_π is known, using (21), (22) and (24), the statistical test in (19),

$$\begin{aligned} C_{1T}(\widehat{\pi},\widehat{\gamma})_{\gamma_\pi} &= T(\widehat{\sigma}_{\varepsilon}^2(\widehat{\gamma}) - \widehat{\sigma}_u^2(\widehat{\pi})) - \widehat{\delta}'\overline{\mathbf{Z}}'\mathbf{M}_{\mathbf{X}_{\gamma_\pi}^{**}}\widehat{\overline{\mathbf{Z}}\delta}' \\ &= \left\| \mathbf{M}_{\mathbf{X}_{\widehat{\gamma}}^{**}}\mathbf{Y} \right\|^2 - \left\| \mathbf{M}_{\overline{\mathbf{Z}}}\mathbf{Y} \right\|^2 - \left\| \mathbf{M}_{\mathbf{X}_{\gamma_\pi}^{**}}\mathbf{P}_{\overline{\mathbf{Z}}}\mathbf{Y} \right\|^2 \end{aligned} \tag{25}$$

From the previous statistic, it is easy to see that $C_{1T}(\widehat{\pi},\widehat{\gamma})_{\gamma_\pi} = T(\widehat{\sigma}_{\varepsilon}^2(\widehat{\gamma}) - \widehat{\sigma}_{\varepsilon\widehat{\pi}}^2(\gamma_\pi))$. Hence, (25) may be seen as the difference between the actual residual variance of the SETAR model under H_{TR} and the estimate of the residual variance of H_{TR} under H_{SB}. Furthermore, From our Cox test is asymptotically equivalent to a generalized Wald test of the difference between the estimator $\widehat{\sigma}^2_\varepsilon(\widehat{\gamma})$ under the H_{TR} and the estimator of the pseudo parameter value $\widehat{\sigma}^2_\varepsilon(\widehat{\gamma}_\pi)$ under H_{SB}.

A considerable difference in (25) leads to the rejection of H_{SB} because the SETAR model is performing significantly well for H_{SB} regarded as true. Similar to Smith [1992] and Gourieroux and Monfort [1994], under some conditions we can get that,

$$\begin{aligned} C_{1T}(\widehat{\pi},\widehat{\gamma})_{\gamma_\pi} &= \left\|\mathbf{P}_{\mathbf{X}^{**}_{\gamma_\pi}}\mathbf{P}_{\overline{\mathbf{Z}}}\mathbf{Y}\right\|^2 - \left\|\mathbf{P}_{\mathbf{X}^{**}_{\widehat{\gamma}}}\mathbf{Y}\right\|^2 \\ &= -2\delta^{0\prime}\overline{\mathbf{Z}}'\left[\mathbf{P}_{\mathbf{X}^{**}_{\widehat{\gamma}}}\mathbf{M}_{\overline{\mathbf{Z}}}\mathbf{u}^*\right] + 2\delta^{0\prime}\overline{\mathbf{Z}}'\left[\mathbf{P}_{\mathbf{X}^{**}_{\gamma_\pi}} - \mathbf{P}_{\mathbf{X}^{**}_{\widehat{\gamma}}}\right]\mathbf{u}^* + o_p(1) \quad (26) \end{aligned}$$

Here $\mathbf{u}^* = \mathbf{u} + (\overline{\mathbf{Z}}^0 - \overline{\mathbf{Z}})\delta^0$.as before. Since is not fixed γ_π, our approach faces two main problems. First, we require an analytic derivation of the pseudo-true estimator γ_π that enters in (26), although a close form solution for γ is also not available for the well specified SETAR). In theory, we should be able to derive analytically an expression for γ_{π_0}, and in the second stage $\widehat{\gamma}_\pi$ is computed by evaluating γ_{π_0} at $\pi_0 = \widehat{\pi}$.

Secondly, in the general case, Walker [1967] and Pesaran and Deaton [1978], it is assumed that the asymptotic optimization problem for the pseudo parameter has a unique root $\widehat{\gamma}_\pi$ wpa1 Therefore, a consistency analysis on γ_π cannot be applied straightforwardly.

Pesaran and Pesaran [1993] tackles this problem by evaluating the integral in (20) with Monte Carlo integration method. Recently, Kapetanios and Weeks [2003] considers boostrap procedures to estimate the Kullback-Leibler measure in (20).

4 Distribution Theory

Similar to Carrasco [2002], the following subsections are devoted to analyze the asymptotics of the misspecified models then we proceed to analyze our nonnested tests, based on these preliminary results.

4.1 Asymptotic results under the SC model

Under the structural break model, we could directly use the results of Bai [1995], Csörgo and Horváth [1996] and Bai and Perron [1998]. However, our statistics would contain the threshold of the misspecified model (i.e. the indicator function). Therefore, the proposition given bellow analyzes the asymptotics of the following empirical covariances,

Proposition 4.1. *Under the SB model given in (10) and Conditions (C4)-(C.5), the following results hold,*

1. $T^{-1}\sum_{t=1}^{[sT]}\mathbf{x}_t\mathbf{x}_t' \xrightarrow{p} s\mathbf{C}_1(1)$ *and* $T^{-1}\sum_{t=[sT]+1}^{T}\mathbf{x}_t\mathbf{x}_t' \xrightarrow{p} (1-s)\mathbf{C}_2(1)$*, uniformly on* $(s,u_i)\in[0,1]^2$*, for* $\mathbf{C}_i(\cdot)$ *some* $(p\times p)$ *positive definite matrix,* $i=1,2$.
2. $T^{-1}\sum_{t=1}^{[sT]+d}\mathbf{x}_t(\gamma)\mathbf{x}_t(\gamma)' \xrightarrow{p} s\mathbf{C}_1(u_1)$ *and* $T^{-1}\sum_{t=[sT]+d+1}^{T}\mathbf{x}_t(\gamma)\mathbf{x}_t(\gamma)' \xrightarrow{p} (1-s)\mathbf{C}_2(u_2)$*, uniformly on* $(s,u_i)\in[0,1]^2$*,* $i=1,2$.
3. *Moreover, we have,*

$$T^{-1}\overline{\mathbf{Z}}_1'\overline{\mathbf{Z}}_1 \xrightarrow{p} \pi\mathbf{C}_1(1)\,,\; T^{-1}\overline{\mathbf{Z}}_2'\overline{\mathbf{Z}}_2 \xrightarrow{p} (1-\pi)\mathbf{C}_2(1)$$
$$T^{-1}\mathbf{X}'\mathbf{X} \xrightarrow{p} \mathbf{C}^*(1) \text{ and } T^{-1}\mathbf{X}_\gamma'\mathbf{X}_\gamma \xrightarrow{p} \mathbf{C}^*(u)$$

uniformly on $(\pi,u_i)\in[0,1]^2$*. Here* $\mathbf{C}^*(1)=\pi\mathbf{C}_1(1)+(1-\pi)\mathbf{C}_2(1)$ *and* $\mathbf{C}^*(u)=\pi\mathbf{C}_1(u_1)+(1-\pi)\mathbf{C}_2(u_2)$ *and* $u_i=\mathbb{P}(y_{t-d}\le\gamma)=p_1(\gamma)I(t\le\pi T)+p_2(\gamma)I(t>\pi T)$.

As in Chan [1990] and Tong [1990], we define $u_i:\mathbb{R}\to[0,1]$ such that $u_i(\gamma)=\mathbb{P}_i(y_t\le\gamma)$.Through the paper, we shall use the following matrices with the following asymptotic properties, such that,

Remark 4.2. *If Proposition (4.1) holds, then* $T^{-1}\mathbf{X}_{1\gamma}'\mathbf{X}_{1\gamma} \xrightarrow{p} \mathbf{C}^*(u)$ *and* $T^{-1}\mathbf{X}_{2\gamma}'\mathbf{X}_{2\gamma} \xrightarrow{p} \mathbf{C}^*(1)-\mathbf{C}^*(u)$ *uniformly on* $(\pi,u_i)\in[0,1]^2$.

Based on the Proposition given above, the results of Carrasco [2002] can be seen as a special case, such that,

Corollary 4.3. *If Proposition (4.1) holds and* $q=0$*, then,*

1. $T^{-1}\sum_{t=1}^{[sT]}x_t^2 \xrightarrow{p} s$ *and* $T^{-1}\sum_{t=[sT]+1}^{T}x_t^2 \xrightarrow{p} (1-s)$.
2. $T^{-1}\sum_{t=1}^{[sT]+d}x_t(\gamma)^2 \xrightarrow{p} s\cdot p_1(\gamma)$ *and* $T^{-1}\sum_{t=[sT]+d+1}^{T}x_t(\gamma)^2 \xrightarrow{p} (1-s)p_2(\gamma)$*, uniformly on* $(s,u_i)\in[0,1]^2$*. Here,* $x_t=1$ *and* $x_t(\gamma)=I(y_{t-d}\le\gamma)$.

Note that for the special case $q=0$, $\pi_0\mathbf{C}_1(1)$ and $(1-\pi_0)\mathbf{C}_2(1)$ equals to π_o and $(1-\pi_0)$, respectively. Additionally, $\pi_0\mathbf{C}_1(u_1)$ and $(1-\pi_0)\mathbf{C}_2(u_2)$ equals to $\pi_0 p_1(\gamma)$ and $(1-\pi_0)p_2(\gamma)$, respectively. Similar to Carrasco [2002], the difference between $p_1(\gamma)$ and $p_2(\gamma)$ mainly results because under the structural break,

$$\mathbb{P}(y_t\le\gamma)=p_1(\gamma)I(t\le\pi_0T)+p_2(\gamma)I(t>\pi_0T)$$

In terms Carrasco [2002] notation, if $q = 0$ and $\{u_t; t \geq 0\}$ is an *iid* sequence with a $N(0, \sigma_u^2)$,we have,

$$p_1(\gamma) = \Phi\left(\frac{\gamma - \delta_1^0}{\sigma_u}\right) \text{ and } p_2(\gamma) = \Phi\left(\frac{\gamma - \delta_2^0}{\sigma_u}\right)$$

Where $\Phi(\cdot)$ is the cumulative distribution function of the standard normal. In Caner and Hansen [2001] and Gonzalo and Montesinos [2002], unit roots with threshold effects are analyzed simultaneously. Thus, the asymptotic tools are employed permitted to analyze double-indexed empirical process that weakly converges to a two-parameter Brownian motion (see Bickel and Wichura [1971] and Hida [1980] for the multiparameter stochastic processes). Similar techniques are used in Bai [1996] and Csörgo and Horváth [1996] for analyzing change-point tests (e.g. Kolmogorov-Smirnov type tests based on sequential empirical processes).

As a result, our analysis combines those asymptotic techniques to analyze simultaneously threshold and change-point effects. Let $\mathbf{B}(s, u)$ be a Gaussian process on $[0,1]^2$ with zero mean and Covariance function $\mathbb{E}\left[\mathbf{B}(s,u)\mathbf{B}(r,v)'\right] = (s \wedge r)(u \wedge v)$, which we shall call a two-parameter Brownian motion on $[0,1]^2$. For the existence and properties of this process we refer to Hida [1980] and for weak convergence properties see Bickel and Wichura [1971].

The symbol denotes $\Longrightarrow$ weak convergence for sequences of (measurable) random elements of a space of bounded cadlag functions on $[0,1]^2$, that is $D(T) \times D(T) \times ... \times D(T)$ where $T = [0,1]^2$, the product metric space of all vector functions on $[0,1]^2$. Furthermore, we endow each component space $D(T)$ with the Skorohod metric and the $\sigma-$field generated by closed balls under this metric (more technical details are given in Billingsley [1968] and Pollard [1985]), therefore we have the following result, similar to Caner and Hansen [2001] and Gonzalo and Montesinos [2002],

Proposition 4.4. *Under Conditions (C3)-(C5),*

$$T^{-1/2}\sum_{t=1}^{[sT]} \mathbf{x}_t(\gamma)\, u_t \quad \Rightarrow \quad \sigma_u \mathbf{B}_1(s, u_1),\; T^{-1/2}\sum_{t=[sT]+1}^{T} \mathbf{x}_t(\gamma)\, u_t \Rightarrow \sigma_u \mathbf{B}_2(s, u_2)$$

$$T^{-1/2}\sum_{t=[sT]+1}^{T} \mathbf{z}_t u_t \quad \Rightarrow \quad \sigma_u \mathbf{B}_1(s, 1),\; \textit{and}\; T^{-1/2}\sum_{t=[sT]+1}^{T} \mathbf{z}_t u_t \Rightarrow \sigma_u \mathbf{B}_2(s, 1)$$

where $\mathbf{B}_i(s)$ *is a 2-parameter Brownian motion on* $[0,1]^2$ *with dimension p and covariance kernel* $\mathbb{E}\left[\mathbf{B}_i(s,w)\mathbf{B}_i(r,v)'\right] = \sigma_u^2 (s \wedge r)\, \mathbf{C}_i (w \wedge v)$.

Similar Brownian functionals are analyzed in Caner and Hansen [2001] and Gonzalo and Montesinos [2002] for developing asymptotic theory for threshold processes with unit root. Moreover, Bai [1996] and Csörgo and Horváth [1996] based on empirical processes techniques use this type of processes to analyze changing point tests.

Corollary 4.5. *Let $q = 0$, then the following results from Proposition (4.4) holds, then empirical processes weakly converges such that,*

$$
\begin{aligned}
T^{-1/2}\sum_{t=1}^{[sT]} I(y_{t-d} &\leq \gamma)u_t \Rightarrow \sigma_u B_1(s,u_1) \\
T^{-1/2}\sum_{t=[sT]+1}^{T} I(y_{t-d} &\leq \gamma)u_t \Rightarrow \sigma_u B_2(s,u_2)
\end{aligned}
$$

where $B_i(s)$ is a 2-parameter Brownian motion on $[0,1]^2$ with covariance $\mathbb{E}\left[B_i(s,w)B_i(r,v)\right] = \sigma_u^2\left(s\wedge r\right)\left(w\wedge v\right)$.

The section employs our results given above to analyze the limit behavior of the misspecified estimators and the nonnested tests.

4.1.1 Convergence of the SETAR pseudo parameters under Structural break model.

Proposition 9 and 10 in Carrasco [2002] show the asymptotic distribution of the parameters of a misspecified SETAR model, in which the data is generated by a MSA model. These results assume that both models have $q = 0$. That is, the models only depend on a switching intercept. As a result, we have that,

$$\widehat{\alpha} \xrightarrow{p} \alpha_a(\delta^0, T_1^0, \gamma_a)$$

And,

$$\sqrt{T}\left(\widehat{\alpha} - \alpha_a\right) \xrightarrow{d} N\left(0, K_3\right)$$

Where K_3 is given in Lemma A.3. in Carrasco [2002]. Similar to the analysis for the Cox test, in order to obtain the associated pseudo-true values of $(\alpha, \gamma, \sigma_\varepsilon^2)$ under H_{SB}, we need to measure the proximity between the two hypothesis H_{SB} and H_{TR} . Similar to Gourieroux and Monfort [1994], based on Kullback-Leibler information criterion (KLIC), and similar to the semiparametric case, we can define,

$$D_{SB}\left(\pi,\gamma\right) = \mathbb{E}_{SB}\left\|\overline{\mathbf{Z}}^0\delta^0 + \mathbf{u} - \mathbf{X}_\gamma^{**}\alpha\right\|^2$$

Where, $\mathbb{E}_{SB}\left(\cdot\right)$ is a conditional expectation with respect to the probability measure attached to the H_{SB}. The empirical counter part would be $D_{SB}\left(\pi_0, \widehat{\gamma}\right) = T^{-1}\left\|\overline{Z}\delta - \mathbf{X}_\gamma^{**}\alpha\right\|^2$. Since for a fixed γ, $\widetilde{\alpha}_{\pi_0} = (\mathbf{X}_\gamma^{**\prime}\mathbf{X}_\gamma^{**})^{-1}\mathbf{X}_\gamma^{**}(\overline{\mathbf{Z}}^0\delta^0 + \mathbf{u})$, we define the associated asymptotic pseudo-true value as the solution of the optimization problem,

$$\widetilde{\gamma}\left(\pi\right) = \operatorname{argmin}_{\gamma\in\Gamma}\ \mathbb{E}_{SB}\left\|\mathbf{M}_{\mathbf{X}_\gamma^{**}}(\mathbf{Z}^0\delta^0 + \mathbf{u})\right\|^2 \tag{27}$$

Normally, For the general case, when $q \geq 1$, the nonrandom vector of pseudo true values can be computed,

Proposition 4.6. *Under the SC model given in (10) and Conditions (C.3)-(C.5). The pseudo parameter of the misspecified SETAR model follows,*

$$\widetilde{\alpha}(\widetilde{\gamma}) \xrightarrow{p} \alpha_a(\pi_0, \gamma) = [\alpha_{1a}, \alpha_{2a}]'$$
$$\xrightarrow{p} \begin{bmatrix} \pi_0 \mathbf{C}^*(u)^{-1} \mathbf{C}_1(u_1) & (1-\pi_0)\mathbf{C}^*(u)^{-1}\mathbf{C}_2(u_2) \\ \pi_0 [\mathbf{C}^*(1) - \mathbf{C}^*(u)]^{-1}[\mathbf{C}_1(1) - \mathbf{C}_1(u_1)] & (1-\pi_0)[\mathbf{C}^*(1) - \mathbf{C}^*(u)]^{-1}[\mathbf{C}_2(1) - \mathbf{C}_2(u_2)] \end{bmatrix} \delta^0$$

uniformly on $(s,u) \in [0,1]^2$. *Here* $\mathbf{C}^*(1) = \pi_0 \mathbf{C}_1(1) + (1-\pi_0)\mathbf{C}_2(1)$ *and* $\mathbf{C}^*(u) = \pi_0 \mathbf{C}_1(u_1) + (1-\pi_0)\mathbf{C}_2(u_2)$.

Following Corollary the similarities of our Proposition 4.6 above with the one presented by Carrasco [2002],Proposition 5, given above, has similar results as part (i) Proposition 10 of Carrasco [2002].

Corollary 4.7. *Let* $q = 0$*, then the following results from Proposition (4.6) holds, then,*

$$\widetilde{\alpha}(\widetilde{\gamma}) \xrightarrow{p} \begin{bmatrix} \pi_0 \frac{p_1(\gamma)}{p^*(\gamma)} & (1-\pi_0)\frac{p_2(\gamma)}{p^*(\gamma)} \\ \pi_0 \frac{(1-p_1(\gamma))}{(1-p^*(\gamma))} & (1-\pi_0)\frac{(1-p_2(\gamma))}{(1-p^*(\gamma))} \end{bmatrix} \begin{bmatrix} \delta_1^0 \\ \delta_2^0 \end{bmatrix}$$

uniformly on $(s,u) \in [0,1]^2$. *and* $p^*(\gamma) = \pi_0 p_1(\gamma) + (1-\pi_0) p_2(\gamma)$.

4.2 Asymptotic results under the SETAR model

By symmetry of our nonnested testing problem, we shall compute the asymptotic results for the case which the data is generated by a SETAR process whereas the research fits a change of point analysis, such that,

Proposition 4.8. *If Conditions (C.1)-(C.2) and the underlying process follows a SETAR process as given in (5), then*

1. $T^{-1}\sum_{t=1}^{T} \mathbf{z}_t \mathbf{z}_t' \xrightarrow{p} \mathbf{Q}_1(1,u) + \mathbf{Q}_2(1,u)$ *uniformly over* $(s,u) \in [0,1]^2$, *for* $\mathbf{Q}_i(\cdot)$ *some* $(p \times p)$ *positive definite matrix.*
2. $T^{-1}\sum_{t=1}^{[sT]} \mathbf{z}_t \mathbf{z}_t' \xrightarrow{p} \mathbf{Q}_1(s,u) + \mathbf{Q}_2(s,u)$, *uniformly on* $(s,u) \in [0,1]^2$.
3. *That is, if* u *is fixed at* u^o *(i.e.*$\mathbb{P}(y_{t-d} \leq \gamma^0)$*), we have,*

$$T^{-1}\mathbf{X}_{1\gamma}'\mathbf{X}_{1\gamma} \xrightarrow{p} \mathbf{Q}_1(1,u) \text{ , } T^{-1}\mathbf{X}_{2\gamma}'\mathbf{X}_{2\gamma} \xrightarrow{p} \mathbf{Q}_2(1,u)$$
$$T^{-1}\mathbf{Z}_1'\mathbf{Z}_1 \xrightarrow{p} \mathbf{Q}^*(\pi) \text{ and } T^{-1}\mathbf{Z}'\mathbf{Z} \xrightarrow{p} \mathbf{Q}^*(1)$$

uniformly on $(\pi,u) \in [0,1]^2$. *Here* $\mathbf{Q}^*(1) = \mathbf{Q}_1(1,u) + \mathbf{Q}_2(1,u)$ *and* $\mathbf{Q}^*(\pi) = \mathbf{Q}_1(\pi,u) + \mathbf{Q}_2(\pi,u)$.

If the parametrization given in (7), we have,

Remark 4.9. *Proposition (4.8) implies that* $T^{-1}\mathbf{Z}_2'\mathbf{Z}_2 \xrightarrow{p} \mathbf{Q}^*(1) - \mathbf{Q}^*(\pi)$ *uniformly on* $(\pi, u) \in [0,1]^2$.

Furthermore, based on the results given above and following Carrasco [2002], they lead to,

Corollary 4.10. *If* $z_t = 1$ *(i.e.* $q = 0$*), under Proposition (4.8), we have that,* $T^{-1}\sum_{t=1}^{[sT]} z_t^2 \xrightarrow{p} s$ *and* $T^{-1}\sum_{t=1}^{[sT]+d} x_t(\gamma)^2 \xrightarrow{p} s \cdot p(\gamma)$ *uniformly on* $(s,u) \in [0,1]^2$*. Here* $p(\gamma) = u = \mathbb{P}(y_{t-d} \leq \gamma)$.

Analogously to Proposition (4.4) obtained before above, we have the weak convergence of the following double-indexed sequential empirical process, such that,

Proposition 4.11. *As* $T \uparrow \infty$*, under Conditions (C.1)-(C.3),*

$$\begin{aligned} T^{-1/2}\sum_{t=1}^{[sT]} \mathbf{x}_{1t}(\gamma)\,\varepsilon_t &\Rightarrow \sigma_\varepsilon \mathbf{B}_1(s,u) \\ T^{-1/2}\sum_{t=1}^{[sT]} \mathbf{x}_{2t}(\gamma)\,\varepsilon_t &\Rightarrow \sigma_\varepsilon \mathbf{B}_2(s,u) \end{aligned}$$

where $\mathbf{B}_i(s)$ *is a 2-parameter Brownian motion on* $[0,1]^2$ *with dimension* p *and covariance kernel* $\mathbb{E}\left[\mathbf{B}_i(s,w)\mathbf{B}_i(r,v)'\right] = \sigma_\varepsilon^2 \mathbf{Q}_i(s \wedge r, w \wedge v)$

Once more, the special case of $q = 0$ analyzed in Carrasco [2002] is implied by the Proposition 4.11above.

Corollary 4.12. *If* $q = 0$ *in Proposition (4.11), then,*

$$\begin{aligned} T^{-1/2}\sum_{t=1}^{[sT]} I(y_{t-d} &\leq \gamma)\varepsilon_t \Rightarrow \sigma_\varepsilon B_1(s,u) \\ T^{-1/2}\sum_{t=1}^{[sT]} I(y_{t-d} &> \gamma)\varepsilon_t \Rightarrow \sigma_\varepsilon B_2(s,u) \end{aligned}$$

where $\mathbf{B}_i(s)$ *is a 2-parameter Brownian motion on* $[0,1]^2$ *with covariance* $\mathbb{E}[B_i(s,w)B_i(r,v)] = \sigma_\varepsilon^2 (s \wedge r)(w \wedge v)$.

Similarly, an analysis of the misspecified SC can be done based on the asymptotic results given above.

4.2.1 Convergence of the Structural break pseudo parameters under SETAR model.

Analogously, Carrasco's Proposition 6 obtains the asymptotic distribution of $\widehat{\delta}$ when in fact the model is generated by a threshold model Carrasco's Proposition assumes $q = 0$ and normal errors, such that,

$$\sqrt{T}\left(\begin{array}{c}\widehat{\delta}_1 - \alpha_1^0 - p(\gamma)\alpha_2^0 \\ \widehat{\delta}_2\end{array}\right) \xrightarrow{d} N\left(\left(\begin{array}{c}0\\0\end{array}\right), K_2\left[\begin{array}{cc}1/\pi & -1/\pi \\ -1/\pi & 1/(\pi(1-\pi)\end{array}\right]\right)$$

And,

$$\widehat{\sigma}_u^2 \xrightarrow{d} \sigma_{ua}^2 = \sigma_\epsilon^2 + p(\gamma)(1-p(\gamma))(\alpha_2^0)^2$$

Here, K_2 depends on $\sigma_\epsilon^2, \alpha^0, \gamma^0$ and the pseudo-true parameter π (details are given in Carrasco [2002]). Nonetheless, the results given above do not provide full discussion regarding the convergence of the threshold parameter $\widehat{\gamma}$ or and the break point date $\widehat{T}_1^0$ to their pseudo-true parameters. By symmetry of the problem, the following proposition provides the criterias in which existence and consistency of the misspecified structural model can be obtained,

Proposition 4.13. *Under the SC model given in (10) and Conditions (C.1)-(C.3). The pseudo parameter of the misspecified SETAR model follows,*

$$\widetilde{\delta}(\widetilde{\pi}) \xrightarrow{p} \delta_a\left(\pi, \gamma^0\right) = [\delta_{1a}, \delta_{2a}]'$$
$$\xrightarrow{p} \left[\begin{array}{cc}\mathbf{Q}^*(\pi)^{-1}\mathbf{Q}_1(\pi) & \mathbf{Q}^*(\pi)^{-1}\mathbf{Q}_2(\pi) \\ \left[\mathbf{Q}^*(1) - \mathbf{Q}^*(\pi)\right]^{-1}\left[\mathbf{Q}_1(1) - \mathbf{Q}_1(\pi)\right] & \left[\mathbf{Q}^*(1) - \mathbf{Q}^*(\pi)\right]^{-1}\left[\mathbf{Q}_2(1) - \mathbf{Q}_2(\pi)\right]\end{array}\right]\alpha^0$$

uniformly on $(\pi, u^0) \in [0,1]^2$.

Similar to Proposition 6 in Carrasco [2002], we must have,

Corollary 4.14. *Let* $q = 0$*, then Proposition (4.13) can be rewritten as,*

$$\widetilde{\delta}(\widetilde{\pi}) \xrightarrow{p} \left[\begin{array}{cc}p\left(\gamma^0\right) & \left(1-p\left(\gamma^0\right)\right) \\ p\left(\gamma^0\right) & \left(1-p\left(\gamma^0\right)\right)\end{array}\right]\left[\begin{array}{c}\alpha_1^0 \\ \alpha_2^0\end{array}\right]$$

uniformly on $(s, u) \in [0,1]^2$.

As in Carrasco [2002], the shift of the pseudo parameters at the limit $\delta_{2a} - \delta_{1a}$ equals to zero, i.e. δ_{ia}, $i = 1, 2$, is a simple weighted average of linear parameter α_1^0 and α_2^0 under the SETAR model. Since the asymptotic behavior of both the correctly specified and misspecified models are set up, next we shall analyze the distribution theory of the "artificial" nested tests.

4.3 Asymptotic convergence of the test statistics

This section consider the asymptotic limits of our non-nested tests for competing models suggested in the previous section.

4.3.1 Convergence of the P-test T_ω

Based on the asymptotic results and the limit behavior of the competing models, we are able to study the limit process of $\widehat{\omega}$, given in (14) and (16), and the associated T_ω statistic, specified in (15) and (17).

As Shorack and Wellner [1986], Bai [1996] and Csörgo and Horváth [1996], we say that $K(s,u)$ be a Gaussian process on a unit cube $[0,1]^2$ with $\mathbb{E}\left[K(s,u)\right]=0$ and Covariance function $\mathbb{E}\left[K(s,u)K(r,v)'\right]=(u\wedge v)\left((s\wedge r)-sr\right)$, defined as a Kiefer process on $[0,1]^2$ (see also Bickel and Wichura [1971] for more technical details). This type of stochastic process has been extendedly used in the statistics literature (e.g. Csörgo and Horváth [1996] and Shorack and Wellner [1986]). However, to our knowledge these processes has been applied in the econometrics literature by Bai [1996] and Gonzalo and Montesinos [2002].

Intuitively speaking, a Kiefer process can be seen as a two-dimensional Brownian bridge process as $K(s,u)=B(s,u)-uB(s,1)$, and its existence and properties are developed in Csörgo and Révész [1981].

Asymptotic distribution of the P-tests under SC model. The following Proposition shows limit process of our Davidson and MacKinnon's type test based on (14), (15) and (18), such that,

Proposition 4.15. *Under Propositions (4.1), (4.4) and (4.6) and $H_0:\omega=0$ (i.e. $H_{SB}:\mathbf{Y}=\overline{Z}^0\delta^0+\mathbf{u}$, the SC model), we have,*

$$T_{1\omega}\Rightarrow\left[\alpha_{Ta}'\mathbf{A}\left(\pi_0,u\right)\alpha_{Ta}\right]^{-1/2}\alpha_{Ta}'\sum_{i=1}^{2}\mathbf{K}_i(\pi_0,u) \tag{28}$$

$$\sup_{0<u<1}T_{1\omega}\left(u\right)\Rightarrow\sup_{0<u<1}\left[\alpha_{Ta}'\mathbf{A}\left(\pi_0,u\right)\alpha_{Ta}\right]^{-1/2}\alpha_{Ta}'\sum_{i=1}^{2}\mathbf{K}_i(\pi_0,u) \tag{29}$$

Where,

$$\mathbf{A}\left(\pi_0,u\right)=\left[\pi_0\left(\mathbf{I}_p-\mathbf{C}_1\left(u_1\right)\mathbf{C}_1\left(1\right)^{-1}\right)\mathbf{C}_1\left(u_1\right)+\left(1-\pi_0\right)\left(\mathbf{I}_p-\mathbf{C}_2\left(u_2\right)\mathbf{C}_2\left(1\right)^{-1}\right)\mathbf{C}_2\left(u_2\right)\right] \tag{30}$$

$$\mathbf{K}_i(\pi_0,u_i)=\left(\mathbf{B}_i(\pi_0,u_i)-\mathbf{C}_i\left(u_i\right)\mathbf{C}_i\left(1\right)^{-1}\mathbf{B}_i(\pi_0,1)\right) \tag{31}$$

where $\mathbf{K}_i(\pi_0,u_i)$ is a Kiefer process on the unit cube $[0,1]^2$ with dimension p and covariance kernel

$$\mathbb{E}\left[\mathbf{K}_i(s,w)\mathbf{K}_i(r,v)'\right]=\sigma_u^2\left(s\wedge r\right)\left(\mathbf{C}_i\left(w\wedge v\right)-\mathbf{C}_i\left(w\right)\mathbf{C}_i\left(1\right)^{-1}\mathbf{C}_i\left(v\right)\right) \tag{32}$$

Proposition (4.15) gives the asymptotic distribution of the test statistic sequence $T_{1\omega}$ Our result given in (28) takes a similar form to that found for change-point in Bai [1996] Theorem 3 (i) for which general type of changes in the error distribution functions are considered under fixed alternatives. The asymptotic distribution of $T_{1\omega}$ is less dispersed when α'_{Ta} is small, i.e. when a reduced "threshold effect" is observed in the in the misspecified SETAR model. As suggested in Bai [1996], the sum of the two Kiefer processes in (28) is uniformly bounded in probability since each $\mathbf{K}_i(\cdot,\cdot)$ is also bounded in probability, for $i = 1, 2$ (See Theorem A.2.3 in Csörgo and Horváth [1996]). Next, we present, a special version of 4.15, such that,

Corollary 4.16. *If $q = 0$, equations (30)-(32) in Proposition (4.15), can be rewritten as,*

$$A(\pi_0, u) = [\pi_0 p_1(\gamma)(1 - p_1(\gamma)) + (1 - \pi_0) p_2(\gamma)(1 - p_2(\gamma))] \tag{33}$$

$$K_i(\pi_0, u_i) = (B_i(\pi_0, u_i) - p_1(\gamma) B_i(\pi_0, 1)) \tag{34}$$

And,

$$\mathbb{E}\left[K_i(s, w) K_i(r, v)'\right] = \sigma_u^2 (s \wedge r)((w \wedge v) - wv) \tag{35}$$

Note that $u_i = p_i(\gamma)$, for $i = 1, 2$. Subsequently, we begin considering the inverse situation when a SETAR process generates the data.

Asymptotic distribution of the P-tests under SETAR model. Analogously to Proposition (4.15), the asymptotic distribution of the test statistic sequence $T_{2\omega}$ given in (17) and (18),

Proposition 4.17. *Under Propositions (4.8), (4.11) and (4.13) and the null hypothesis of $H_0 : \omega = 1$ (i.e. H_{TR}: $\mathbf{Y} = \mathbf{X}^*_{\gamma^0}\alpha^0 + \epsilon$, the SETAR model), we have,*

$$T_{2\omega} \Rightarrow \left[\delta'_{Ta} \mathbf{G}\left(\pi, u^0\right) \delta_{Ta}\right]^{-1/2} \delta'_{Ta} \sum_{i=1}^{2} \mathbf{K}_i(\pi, u^0) \tag{36}$$

$$\sup_{0<\pi<1} T_{2\omega}(\pi) \Rightarrow \sup_{0<\pi<1} \left[\delta'_{Ta} \mathbf{G}\left(\pi, u^0\right) \delta_{Ta}\right]^{-1/2} \delta'_{Ta} \sum_{i=1}^{2} \mathbf{K}_i(\pi, u^0) \tag{37}$$

Where,

$$\mathbf{G}\left(\pi, u^0\right) = \left[\left(\mathbf{I}_p - \mathbf{Q}_1(\pi)\mathbf{Q}_1(1)^{-1}\right)\mathbf{Q}_1(\pi) + \left(\mathbf{I}_p - \mathbf{Q}_2(\pi)\mathbf{Q}_2(1)^{-1}\right)\mathbf{Q}_2(\pi)\right] \tag{38}$$

$$\mathbf{K}_i(\pi, u^0) = \left(\mathbf{B}_i(\pi, u^0) - \mathbf{Q}_i(\pi)\mathbf{Q}_i(1)^{-1}\mathbf{B}_i(1, u^0)\right) \tag{39}$$

where $\mathbf{K}_i(\pi, u^0)$ is a Kiefer process on the unit cube $[0,1]^2$ with dimension p and covariance kernel

$$\mathbb{E}\left[\mathbf{K}_i(s, w)\mathbf{K}_i(r, v)'\right] = \sigma_\varepsilon^2 \left(\mathbf{Q}_i(s \wedge r, w \wedge v) - \mathbf{Q}_i(s, w \wedge v)\mathbf{Q}_i(1, w \wedge v)^{-1}\mathbf{Q}_i(r, w \wedge v)\right) \tag{40}$$

Given the symmetry of the result in Proposition 4.15, the analysis for the results in 4.17 follows from Proposition 4.15. A similar version to Corollary 4.16 cannot be obtained of this case since the T-statistic is undetermined. As shown in (Gourieroux and Monfort [1995]) the regressors $\mathbf{Z}_i$ are linear combinations of the regressors $\mathbf{X}_{i\gamma}$ for $q = 0$.

4.3.2 Convergence of the Cox test

Our problem here is to derive the asymptotic limit of the test given in (26). In which we focus in the centered difference between the LS criterion functions of the competing models. We set the problem for testing the model H_{SB} : $\mathbf{Y} = \overline{Z}^0 \delta^0 + \mathbf{u}$ against the alternative belonging to H_{TR}: $\mathbf{Y} = \mathbf{X}^*_{\gamma^0} \alpha^0 + \epsilon$. Hence, the weak convergence result of the sequential test in (26) is analyzed in the next proposition such that,

5 Conclusions

The present paper consider two classes of separate families of hypothesis applied to nonlinear time series models. The proposed non-nested tests compare two rival model in which under correct specification show non-standard asymptotic limits. As a consequence, contrary to the standard results in non-nested hypothesis testing, our test weakly converges to functionals of Gaussian processes.

Although, the complexity of the problem and its results, our tests provides a statistical framework in which researches may infer between economic or financial variables that are subject to structural breaks. And, process with complex dynamics mainly due to nonlinear structures.

MATHEMATICAL APPENDIX

Proof of Proposition 4.6. As Chan [1990] and Hansen [2000], through concentration of the LS estimator of (5), we have,

$$\widetilde{\alpha}(\widetilde{\gamma}) = \left[\mathbf{X}^{**\prime}_{\widetilde{\gamma}} \mathbf{X}^{**}_{\widetilde{\gamma}}\right]^{-1} \mathbf{X}^{**\prime}_{\widetilde{\gamma}} \mathbf{Y} \tag{A-1}$$

Similar to (Gourieroux and Monfort [1994]), under the true model, $\mathbf{Y} = \overline{Z}^0 \delta^0 + \mathbf{u}$, we get,

$$\widetilde{\alpha}(\widetilde{\gamma}) = \left[T^{-1}\mathbf{X}^{**\prime}_{\widetilde{\gamma}} \mathbf{X}^{**}_{\widetilde{\gamma}}\right]^{-1} T^{-1}\mathbf{X}^{**\prime}_{\widetilde{\gamma}} \overline{Z}^0 \delta^0 + o_p(1) \tag{A-2}$$

By Proposition (4.1) and (4.4) $\left[\mathbf{X}_{\widetilde{\gamma}}^{**\prime}\mathbf{X}_{\widetilde{\gamma}}^{**}\right]^{-1}\mathbf{X}_{\widetilde{\gamma}}^{**\prime}\mathbf{u} = O_p\left(T^{-1/2}\right)$. Since, $\mathbf{X}_{\gamma}^{**} = \left[\mathbf{X}_{1\gamma}, \mathbf{X}_{2\gamma}\right]$,

$$\left[\begin{array}{c} \widetilde{\alpha}_1(\widetilde{\gamma}) \\ \widetilde{\alpha}_2(\widetilde{\gamma}) \end{array}\right] = \left[\begin{array}{cc} T^{-1}\mathbf{X}_{1\widetilde{\gamma}}'\mathbf{X}_{1\widetilde{\gamma}} & \mathbf{0} \\ \mathbf{0} & T^{-1}\mathbf{X}_{2\widetilde{\gamma}}'\mathbf{X}_{2\widetilde{\gamma}} \end{array}\right]^{-1} \times \left[\begin{array}{cc} T^{-1}\mathbf{X}_{1\widetilde{\gamma}}'\overline{Z}_1^0 & T^{-1}\mathbf{X}_{1\widetilde{\gamma}}'\overline{Z}_2^0 \\ T^{-1}\mathbf{X}_{2\widetilde{\gamma}}'\overline{Z}_1^0 & T^{-1}\mathbf{X}_{2\widetilde{\gamma}}'\overline{Z}_2^0 \end{array}\right] + o_p(1)$$

Proposition (4.1) imply that,

$$\left[\begin{array}{c} \widetilde{\alpha}_1(\widetilde{\gamma}) \\ \widetilde{\alpha}_2(\widetilde{\gamma}) \end{array}\right] \xrightarrow{p} \left[\begin{array}{c} \alpha_{1a}(\gamma) \\ \alpha_{2a}(\gamma) \end{array}\right] = \left[\begin{array}{cc} \mathbf{C}^*(u) & \mathbf{0} \\ \mathbf{0} & \mathbf{C}^*(1) - \mathbf{C}^*(u) \end{array}\right]^{-1} \times \left[\begin{array}{cc} \pi_0\mathbf{C}_1(u_1) & (1-\pi_0)\mathbf{C}_2(u_2) \\ \pi_0\left[\mathbf{C}_1(1) - \mathbf{C}_1(u_1)\right] & (1-\pi_0)\left[\mathbf{C}_2(1) - \mathbf{C}_2(u_2)\right] \end{array}\right]\delta^0$$

uniformly on $(\pi_0, u) \in [0,1]^2$. Here $\mathbf{C}^*(1) = \pi_0\mathbf{C}_1(1) + (1-\pi_0)\mathbf{C}_2(1)$ and $\mathbf{C}^*(u) = \pi_0\mathbf{C}_1(u_1) + (1-\pi_0)\mathbf{C}_2(u_2)$.. Simple algebra gives the desired result in (4.6). □

Proof of Corollary 4.7. The result is directly implied by Corollary 4.3 and Proposition 4.6, when $q = 0$. That is, for the model with only intercepts, we have that $\pi_0\mathbf{C}_1(1)$ and $(1-\pi_0)\mathbf{C}_2(1)$ equals to π_o and $(1-\pi_0)$, respectively.

Moreover, $\pi_0\mathbf{C}_1(u_1)$ and $(1-\pi_0)\mathbf{C}_2(u_2)$ equals to $\pi_0 p_1(\gamma)$ and $(1-\pi_0)p_2(\gamma)$, respectively. Thus, $p^*(\gamma) = \pi_0 p_1(\gamma) + (1-\pi_0)p_2(\gamma)$. □

Proof of Proposition 4.15. Let first analyze (15), based on the the estimator $\widehat{\omega}(\widehat{\pi}, \widetilde{\gamma}) = \left(\widetilde{\alpha}^*(\widetilde{\gamma})'\mathbf{X}_{\widetilde{\gamma}}^{*\prime}\mathbf{M}_{\overline{Z}}\mathbf{X}_{\widetilde{\gamma}}^*\widetilde{\alpha}^*(\widetilde{\gamma})\right)^{-1}\widetilde{\alpha}^*(\widetilde{\gamma})'\mathbf{X}_{\widetilde{\gamma}}^{*\prime}\mathbf{M}_{\overline{Z}}\mathbf{u}^*$ given in (14). We shall focus on the first component of $\widehat{\omega}(\widehat{\pi}, \widetilde{\gamma})$. we have,

$$\begin{aligned} T^{-1}\widetilde{\alpha}^*(\widetilde{\gamma})'\mathbf{X}_{\widetilde{\gamma}}^{*\prime}\mathbf{M}_{\overline{Z}}\mathbf{X}_{\widetilde{\gamma}}^*\widetilde{\alpha}^*(\widetilde{\gamma}) &= \widetilde{\alpha}^*(\widetilde{\gamma})'\left[T^{-1}\mathbf{X}_{\widetilde{\gamma}}^{*\prime}\mathbf{X}_{\widetilde{\gamma}}^* - T^{-1}\mathbf{X}_{\widetilde{\gamma}}^{*\prime}\overline{Z}\left(T^{-1}\overline{Z}'\overline{Z}\right)^{-1}T^{-1}\overline{Z}'\mathbf{X}_{\widetilde{\gamma}}^*\right]\widetilde{\alpha}^*(\widetilde{\gamma}) \\ &= \widetilde{\alpha}^*(\widetilde{\gamma})'\left[\begin{array}{cc} T^{-1}\mathbf{X}'\mathbf{X} & T^{-1}\mathbf{X}'\mathbf{X}_{\widetilde{\gamma}} \\ T^{-1}\mathbf{X}_{\widetilde{\gamma}}'\mathbf{X} & T^{-1}\mathbf{X}_{\widetilde{\gamma}}'\mathbf{X}_{\widetilde{\gamma}} \end{array}\right]\widetilde{\alpha}^*(\widetilde{\gamma}) - \\ &\quad \widetilde{\alpha}^*(\widetilde{\gamma})'\left[\begin{array}{cc} T^{-1}\mathbf{X}'\overline{Z}_1 & T^{-1}\mathbf{X}'\overline{Z}_2 \\ T^{-1}\mathbf{X}_{\widetilde{\gamma}}'\overline{Z}_1 & T^{-1}\mathbf{X}_{\widetilde{\gamma}}'\overline{Z}_2 \end{array}\right]\left[\begin{array}{cc} T^{-1}\overline{Z}_1'\overline{Z}_1 & \mathbf{0} \\ \mathbf{0} & T^{-1}\overline{Z}_2'\overline{Z}_2 \end{array}\right]^{-1} \\ &\quad \times\left[\begin{array}{cc} T^{-1}\overline{Z}_1'\mathbf{X} & T^{-1}\overline{Z}_1'\mathbf{X}_{\widetilde{\gamma}} \\ T^{-1}\overline{Z}_2'\mathbf{X} & T^{-1}\overline{Z}_2'\mathbf{X}_{\widetilde{\gamma}} \end{array}\right]\widetilde{\alpha}^*(\widetilde{\gamma}) \end{aligned}$$

Proposition (??), (4.1) and (4.6) give,

$$\widetilde{\alpha}^{*}(\widetilde{\gamma})' \mathbf{X}_{\widetilde{\gamma}}^{*\prime}\mathbf{M}_{\overline{Z}}\mathbf{X}_{\widetilde{\gamma}}^{*}\widetilde{\alpha}^{*}(\widetilde{\gamma}) \xrightarrow{p} \alpha_a^{*}(\gamma)' \begin{bmatrix} \mathbf{C}^{*}(1) & \mathbf{C}^{*}(u) \\ \mathbf{C}^{*}(u) & \mathbf{C}^{*}(u) \end{bmatrix} \alpha_a^{*}(\gamma)$$
$$-\alpha_a^{*}(\gamma)' \begin{bmatrix} \pi\mathbf{C}_1(1) & (1-\pi)\mathbf{C}_2(1) \\ \pi\mathbf{C}_1(u_1) & (1-\pi)\mathbf{C}_2(u_2) \end{bmatrix} \begin{bmatrix} \pi\mathbf{C}_1(1) & \mathbf{0} \\ \mathbf{0} & (1-\pi)\mathbf{C}_2(1) \end{bmatrix}^{-1}$$
$$\times \begin{bmatrix} \pi\mathbf{C}_1(1) & \pi\mathbf{C}_1(u_1) \\ (1-\pi)\mathbf{C}_2(1) & (1-\pi)\mathbf{C}_2(u_2) \end{bmatrix} \alpha_a^{*}(\gamma)$$

Standard algebra shows that,

$$\widetilde{\alpha}^{*}(\widetilde{\gamma})' \mathbf{X}_{\widetilde{\gamma}}^{*\prime}\mathbf{M}_{\overline{Z}}\mathbf{X}_{\widetilde{\gamma}}^{*}\widetilde{\alpha}^{*}(\widetilde{\gamma}) \xrightarrow{p} \alpha_a^{*}(\gamma)' \begin{bmatrix} \mathbf{C}^{*}(1) & \mathbf{C}^{*}(u) \\ \mathbf{C}^{*}(u) & \mathbf{C}^{*}(u) \end{bmatrix} \alpha_a^{*}(\gamma)$$
$$-\alpha_a^{*}(\gamma)' \begin{bmatrix} \mathbf{C}^{*}(1) & \mathbf{C}^{*}(u) \\ \mathbf{C}^{*}(u) & \pi\mathbf{C}_1(u_1)\,\mathbf{C}_1(1)^{-1}\,\mathbf{C}_1(u_1) + (1-\pi)\mathbf{C}_2(u_2)\,\mathbf{C}_2(1)^{-1}\,\mathbf{C}_2(u_2) \end{bmatrix} \alpha_a^{*}(\gamma)$$

As $\mathbf{C}^{*}(u) = \pi\mathbf{C}_1(u_1) + (1-\pi)\mathbf{C}_2(u_2)$, $\mathbf{A}(\pi, u)$ in equation (30) is established,

$$T^{-1}\widetilde{\alpha}^{*}(\widetilde{\gamma})' \mathbf{X}_{\widetilde{\gamma}}^{*\prime}\mathbf{M}_{\overline{Z}}\mathbf{X}_{\widetilde{\gamma}}^{*}\widetilde{\alpha}^{*}(\widetilde{\gamma}) \xrightarrow{p} \alpha_a^{*}(\gamma)' \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}(\pi, u) \end{bmatrix} \alpha_a^{*}(\gamma) = \alpha_{Ta}'\mathbf{A}(\pi, u)\,\alpha_{Ta} \tag{A-3}$$

Next, we focus on $\widetilde{\alpha}^{*}(\widetilde{\gamma})' \mathbf{X}_{\widetilde{\gamma}}^{*\prime}\mathbf{M}_{\overline{Z}}\mathbf{u}^{*}$, the second component with $\mathbf{u}^{*} = \mathbf{u} + (\overline{Z}^{0} - \overline{Z})\delta^{0}$,

$$\widetilde{\alpha}^{*}(\widetilde{\gamma})' \mathbf{X}_{\widetilde{\gamma}}^{*\prime}\mathbf{M}_{\overline{Z}}\mathbf{u}^{*} = \widetilde{\alpha}^{*}(\widetilde{\gamma})' \mathbf{X}_{\widetilde{\gamma}}^{*\prime}\mathbf{M}_{\overline{Z}}\mathbf{u} + \widetilde{\alpha}^{*}(\widetilde{\gamma})' \mathbf{X}_{\widetilde{\gamma}}^{*\prime}\mathbf{M}_{\overline{Z}}(\overline{Z}^{0} - \overline{Z})\delta^{0}$$

Similar to **?** and Bai and Perron [1998], we know that $\widetilde{\alpha}^{*}(\widetilde{\gamma})' \mathbf{X}_{\widetilde{\gamma}}^{*\prime}\mathbf{M}_{\overline{Z}}(\overline{Z}^{0} - \overline{Z})\delta^{0} = o_p(1)$ by Assumption (??), and,

$$\begin{aligned} T^{-1/2}\widetilde{\alpha}^{*}(\widetilde{\gamma})' \mathbf{X}_{\widetilde{\gamma}}^{*\prime}\mathbf{M}_{\overline{Z}}\mathbf{u}^{*} &= T^{-1/2}\widetilde{\alpha}^{*}(\widetilde{\gamma})' \mathbf{X}_{\widetilde{\gamma}}^{*\prime}\mathbf{u} - T^{-1/2}\widetilde{\alpha}^{*}(\widetilde{\gamma})' \mathbf{X}_{\widetilde{\gamma}}^{*\prime}\mathbf{P}_{\overline{\mathbf{Z}}}\mathbf{u} + o_p(1) \\ &= \widetilde{\alpha}^{*}(\widetilde{\gamma})' \begin{bmatrix} T^{-1/2}\mathbf{X}'\mathbf{u} \\ T^{-1/2}\mathbf{X}_{\widetilde{\gamma}}'\mathbf{u} \end{bmatrix} - \widetilde{\alpha}^{*}(\widetilde{\gamma})' \begin{bmatrix} T^{-1}\mathbf{X}'\overline{Z}_1 & T^{-1}\mathbf{X}'\overline{Z}_2 \\ T^{-1}\mathbf{X}_{\widetilde{\gamma}}'\overline{Z}_1 & T^{-1}\mathbf{X}_{\widetilde{\gamma}}'\overline{Z}_2 \end{bmatrix} \\ &\quad \times \begin{bmatrix} T^{-1}\overline{Z}_1'\overline{Z}_1 & \mathbf{0} \\ \mathbf{0} & T^{-1}\overline{Z}_2'\overline{Z}_2 \end{bmatrix}^{\mathbf{-1}} \begin{bmatrix} T^{-1/2}\overline{Z}_1'\mathbf{u} \\ T^{-1/2}\overline{Z}_2'\mathbf{u} \end{bmatrix} + o_p(1) \end{aligned}$$

Similarly, by directly using the results of Proposition (4.1), (4.4) and 4.6, yields to,

$$\begin{aligned} T^{-1/2}\widetilde{\alpha}^{*}(\widetilde{\gamma})' \mathbf{X}_{\widetilde{\gamma}}^{*\prime}\mathbf{M}_{\overline{Z}}\mathbf{u}^{*} &\Rightarrow \alpha_a^{*}(\gamma)' \begin{bmatrix} \sigma_u\mathbf{B}_1(\pi,1) + \sigma_u\mathbf{B}_1(\pi,1) \\ \sigma_u\mathbf{B}_2(\pi,u_1) + \sigma_u\mathbf{B}_1(\pi,u_2) \end{bmatrix} \\ &\quad -\alpha_a^{*}(\gamma)' \begin{bmatrix} \mathbf{I}_p & \mathbf{I}_p \\ \mathbf{C}_1(u_1)\,\mathbf{C}_1(1)^{-1} & \mathbf{C}_2(u_2)\,\mathbf{C}_2(1)^{-1} \end{bmatrix} \begin{bmatrix} \sigma_u\mathbf{B}_1(\pi,1) \\ \sigma_u\mathbf{B}_2(\pi,1) \end{bmatrix} \end{aligned}$$

Therefore,

$$T^{-1/2}\widetilde{\alpha}^{*}\left(\widetilde{\gamma}\right)'\mathbf{X}^{*\prime}_{\widetilde{\gamma}}\mathbf{M}_{\overline{Z}}\mathbf{u}^{*}\Rightarrow\alpha^{*}_{a}\left(\gamma\right)'\left[\begin{array}{c}\mathbf{0}\\ \sigma_{u}\sum_{i=1}^{2}\left\{\mathbf{B}_{i}(\pi,u_{i})-\mathbf{C}_{i}\left(u_{i}\right)\mathbf{C}_{i}\left(1\right)^{-1}\mathbf{B}_{i}(\pi,1\right\}\end{array}\right] \tag{A-4}$$

(??) and (??) with continuous mapping theorem (CMT) leads to,

$$T^{1/2}\widehat{\omega}\left(\widehat{\pi},\widetilde{\gamma}\right)\Rightarrow\sigma_{u}\left[\alpha'_{Ta}\mathbf{A}\left(\pi,u\right)\alpha_{Ta}\right]^{-1}\alpha'_{Ta}\sum_{i=1}^{2}\left\{\mathbf{B}_{i}(\pi,u_{i})-\mathbf{C}_{i}\left(u_{i}\right)\mathbf{C}_{i}\left(1\right)^{-1}\mathbf{B}_{i}(\pi,1)\right\} \tag{A-5}$$

on $C\left[0,1\right]^{2}$. By Proposition (??), under the null hypothesis $\widehat{\sigma}_{\varsigma^{*}}\xrightarrow{p}\sigma_{u}$. Therefore, combining this result with equation (??), the sequential t-tests in (15) converges to,

$$T_{1\omega}\Rightarrow\left[\alpha'_{Ta}\mathbf{A}\left(\pi,u\right)\alpha_{Ta}\right]^{-1/2}\alpha'_{Ta}\sum_{i=1}^{2}\mathbf{K}_{i}(\pi,u) \tag{A-6}$$

as established in equation (28) Since, $g\left(\cdot\right)$ in (18) is continuous function using the uniform metric, monotonic on the space of bounded cadlag functions on Γ, (29) holds by (CMT). The sum of two independent $p-$dimensional Kiefer processes on $C\left[0,1\right]^{2}$, the space of all continous functions in the unit cube.Similar to Bai [1996] and Caner and Hansen [2001], the covariance kernel in (32) can be obtained as follows,

$$\begin{aligned}\mathbb{E}\left[\mathbf{K}_{i}(s,w)\mathbf{K}_{i}(r,v)'\right] &= \mathbb{E}[(\mathbf{B}_{i}(s,w)-\mathbf{C}_{i}\left(w\right)\mathbf{C}_{i}\left(1\right)^{-1}\mathbf{B}_{i}(s,1))\\ &\quad\times(\mathbf{B}_{i}(r,v)-\mathbf{C}_{i}\left(v\right)\mathbf{C}_{i}\left(1\right)^{-1}\mathbf{B}_{i}(r,1))']\end{aligned}$$

From Proposition (4.4) we know that, $\mathbb{E}\left[\mathbf{B}_{i}(s,w)\mathbf{B}_{i}(r,v)'\right]=\sigma^{2}_{u}\left(s\wedge r\right)\mathbf{C}_{i}\left(w\wedge v\right)$, thus,

$$\mathbb{E}\left[\mathbf{K}_{i}(s,w)\mathbf{K}_{i}(r,v)'\right]=\sigma^{2}_{u}\left(s\wedge r\right)\left[\mathbf{C}_{i}\left(w\wedge v\right)-\mathbf{C}_{i}\left(w\right)\mathbf{C}_{i}\left(1\right)^{-1}\mathbf{C}_{i}(v)\right] \tag{A-7}$$

which gives the desired result given in (32). □

Proof of Corollary 4.16. Corollaries 4.3, 4.5 and 4.7 applied to Proposition 4.15 establishes the results in the Corollary. □

Proof of Proposition 4.13 Corollary and 4.14. By symmetry of our problem, these results are determined as in Proposition 4.6 and Corollary 4.7, respectively. Similar to Bai [1995] and Bai and Perron [1998], the concentrated LS estimator in the change-point model in (8) is given by,

$$\widetilde{\delta}\left(\widetilde{\pi}\right)=\left[\overline{Z}'\overline{Z}\right]^{-1}\overline{Z}'\mathbf{Y} \tag{A-8}$$

If Proposition (4.8) and (4.11) holds under the true model, $\mathbf{Y} = \mathbf{X}^{**}_{\gamma^0}\alpha^0 + \varepsilon$, the sequence of estimators converge uniformly on (π, u^0) an element on the unit cube set $[0,1]^2$, such that,

$$\begin{aligned}\widetilde{\delta}\left(\widetilde{\pi}\right) &= \begin{bmatrix} \widetilde{\delta}_1\left(\widetilde{\pi}\right) \\ \widetilde{\delta}_2\left(\widetilde{\pi}\right) \end{bmatrix} \xrightarrow{p} \begin{bmatrix} \delta_{1a}\left(\pi, \gamma^0\right) \\ \delta_{2a}\left(\pi, \gamma^0\right) \end{bmatrix} \\ &\xrightarrow{p} \begin{bmatrix} \mathbf{Q}^*\left(\pi\right)^{-1}\mathbf{Q}_1\left(\pi\right) & \mathbf{Q}^*\left(\pi\right)^{-1}\mathbf{Q}_2\left(\pi\right) \\ \left[\mathbf{Q}^*(1) - \mathbf{Q}^*\left(\pi\right)\right]^{-1}\left[\mathbf{Q}_1\left(1\right) - \mathbf{Q}_1\left(\pi\right)\right] & \left[\mathbf{Q}^*(1) - \mathbf{Q}^*\left(\pi\right)\right]^{-1}\left[\mathbf{Q}_2\left(1\right) - \mathbf{Q}_2\left(\pi\right)\right] \end{bmatrix}\alpha^0\end{aligned}$$

The Proposition is proved as $\mathbf{Q}^*\left(1\right) = \mathbf{Q}_1\left(1, u^0\right) + \mathbf{Q}_2\left(1, u^0\right)$ and $\mathbf{Q}^*\left(\pi\right) = \mathbf{Q}_1\left(\pi, u^0\right) + \mathbf{Q}_2\left(\pi, u^0\right)$. Under $q = 0$, combining these results together with Corollary 4.10, we obtain the desired Corollary. □

Proof of Proposition 4.17. Once more, the proof follows similarly to Proposition 4.15. Under the null hypothesis $H_0 : \omega = 1$ (i.e. H_f: $\mathbf{Y} = \mathbf{X}^{**}_{\gamma^0}\alpha^0 + \epsilon$), we shall analyze the estimator of the convex combination for the $P-$test, such that

$$\left(1 - \widehat{\omega}\left(\widetilde{\pi}, \widehat{\gamma}\right)\right) = \left(\widetilde{\delta}^{*\prime}\widetilde{Z}^{*\prime}\mathbf{M}_{\mathbf{X}^{**}_{\widehat{\gamma}}}\widetilde{Z}^*\widetilde{\delta}^*\right)^{-1}\widetilde{\delta}^{*\prime}\widetilde{Z}^{*\prime}\mathbf{M}_{\mathbf{X}^{**}_{\widehat{\gamma}}}\varepsilon^*$$

Where $\varepsilon^* = \varepsilon + (\mathbf{X}^{**}_{\gamma^0} - \mathbf{X}^{**}_{\widehat{\gamma}})\alpha^0$. Consider first the component,

$$\begin{aligned}T^{-1}\widetilde{\delta}^*\left(\widetilde{\pi}\right)'\widetilde{Z}^{*\prime}\mathbf{M}_{\mathbf{X}^{**}_{\widehat{\gamma}}}\widetilde{Z}^*\widetilde{\delta}^*\left(\widetilde{\pi}\right) &= \widetilde{\delta}^*\left(\widetilde{\pi}\right)'\begin{bmatrix} T^{-1}\widetilde{Z}'\widetilde{Z} & T^{-1}\widetilde{Z}'\widetilde{Z}_1 \\ T^{-1}\widetilde{Z}_1'\widetilde{Z} & T^{-1}\widetilde{Z}_1'\widetilde{Z}_1 \end{bmatrix}\widetilde{\delta}^*\left(\widetilde{\pi}\right) - \\ &\quad \widetilde{\delta}^*\left(\widetilde{\pi}\right)'\begin{bmatrix} T^{-1}\widetilde{Z}'\mathbf{X}_{1\widehat{\gamma}} & T^{-1}\widetilde{Z}'\mathbf{X}_{2\widehat{\gamma}} \\ T^{-1}\widetilde{Z}_1'\mathbf{X}_{1\widehat{\gamma}} & T^{-1}\widetilde{Z}_1'\mathbf{X}_{2\widehat{\gamma}} \end{bmatrix}\begin{bmatrix} T^{-1}\mathbf{X}_{1\widehat{\gamma}}'\mathbf{X}_{1\widehat{\gamma}} & \mathbf{0} \\ \mathbf{0} & T^{-1}\mathbf{X}_{2\widehat{\gamma}}'\mathbf{X}_{2\widehat{\gamma}} \end{bmatrix}^{-1} \\ &\quad \times \begin{bmatrix} T^{-1}\mathbf{X}_{1\widehat{\gamma}}'\widetilde{Z} & T^{-1}\mathbf{X}_{1\widehat{\gamma}}'\widetilde{Z}_1 \\ T^{-1}\mathbf{X}_{2\widehat{\gamma}}'\widetilde{Z} & T^{-1}\mathbf{X}_{2\widehat{\gamma}}'\widetilde{Z}_1 \end{bmatrix}\widetilde{\delta}^*\left(\widetilde{\pi}\right)'\end{aligned}$$

Note that here parametrizations in equations (10) and (5) are used. By Proposition (??), (4.8) and (4.13),

$$T^{-1}\widetilde{\delta}^*\left(\widetilde{\pi}\right)'\widetilde{Z}^{*\prime}\mathbf{M}_{\mathbf{X}^{**}_{\widehat{\gamma}}}\widetilde{Z}^*\widetilde{\delta}^*\left(\widetilde{\pi}\right) \xrightarrow{p} \delta_a^*\left(\pi\right)'\begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{G}\left(\pi, u^0\right) \end{bmatrix}\delta_a^*\left(\pi\right) = \delta_{Ta}'\mathbf{G}\left(\pi, u^0\right)\delta_{Ta} \tag{A-9}$$

Where $\mathbf{G}\left(\pi, u^0\right)$ is defined in (38). Now lets consider $\widetilde{\delta}^*\left(\widetilde{\pi}\right)'\widetilde{Z}^{*\prime}\mathbf{M}_{\mathbf{X}^{**}_{\widehat{\gamma}}}\varepsilon^*$. Similar to Hansen [2000], we can prove that $\widetilde{\delta}^*\left(\widetilde{\pi}\right)'\widetilde{Z}^{*\prime}\mathbf{M}_{\mathbf{X}^{**}_{\widehat{\gamma}}}(\mathbf{X}^{**}_{\gamma^0} - \mathbf{X}^{**}_{\widehat{\gamma}})\alpha^0 = o_p(1)$ such that,

$$\begin{aligned}T^{-1/2}\widetilde{\delta}^*\left(\widetilde{\pi}\right)'\widetilde{Z}^{*\prime}\mathbf{M}_{\mathbf{X}^{**}_{\widehat{\gamma}}}\varepsilon^* &= \widetilde{\delta}^*\left(\widetilde{\pi}\right)'\begin{bmatrix} T^{-1/2}\widetilde{Z}'\varepsilon \\ T^{-1/2}\widetilde{Z}_1'\varepsilon \end{bmatrix} - \widetilde{\delta}^*\left(\widetilde{\pi}\right)'\begin{bmatrix} T^{-1}\widetilde{Z}'\mathbf{X}_{1\widehat{\gamma}} & T^{-1}\widetilde{Z}'\mathbf{X}_{2\widehat{\gamma}} \\ T^{-1}\widetilde{Z}_1'\mathbf{X}_{1\widehat{\gamma}} & T^{-1}\widetilde{Z}_1'\mathbf{X}_{2\widehat{\gamma}} \end{bmatrix} \\ &\quad \times \begin{bmatrix} T^{-1}\mathbf{X}_{1\widehat{\gamma}}'\mathbf{X}_{1\widehat{\gamma}} & \mathbf{0} \\ \mathbf{0} & T^{-1}\mathbf{X}_{2\widehat{\gamma}}'\mathbf{X}_{2\widehat{\gamma}} \end{bmatrix}^{-1}\begin{bmatrix} T^{-1/2}\mathbf{X}_{1\widehat{\gamma}}'\varepsilon \\ T^{-1/2}\mathbf{X}_{2\widehat{\gamma}}'\varepsilon \end{bmatrix} + o_p(1)\end{aligned}$$

Proposition (4.8), and 4.13 together with the invariance principle given in Proposition (4.11) leads to the weak convergence,

$$T^{-1/2}\widetilde{\delta}^{*}\left(\widetilde{\pi}\right)'\widetilde{Z}^{*\prime}\mathbf{M}_{\mathbf{X}^{**}_{\widehat{\gamma}}}\varepsilon^{*}\Rightarrow\sigma_{\varepsilon}\delta_{Ta}'\sum_{i=1}^{2}\left\{\mathbf{B}_{i}(\pi,u)-\mathbf{Q}_{i}\left(\pi,u\right)\mathbf{Q}_{\mathbf{i}}\left(1,u\right)^{-1}\mathbf{B}_{i}(1,u)\right\} \tag{A-10}$$

Combining (??) and (??) with (CMT),

$$T^{1/2}(1-\widehat{\omega}\left(\widetilde{\pi},\widehat{\gamma}\right))\Rightarrow\sigma_{\varepsilon}\left[\delta_{Ta}'\mathbf{G}\left(\pi,u^{0}\right)\delta_{Ta}\right]^{-1}\sum_{i=1}^{2}\left\{\mathbf{B}_{i}(\pi,u)-\mathbf{Q}_{i}\left(\pi,u\right)\mathbf{Q}_{i}\left(1,u\right)^{-1}\mathbf{B}_{i}(1,u)\right\} \tag{A-11}$$

Thus,

$$T_{2\omega}\Rightarrow\left[\delta_{Ta}'\mathbf{G}\left(\pi,u^{0}\right)\delta_{Ta}\right]^{-1/2}\delta_{Ta}'\sum_{i=1}^{2}\mathbf{K}_{i}(\pi,u) \tag{A-12}$$

This establishes the desired result in the statistic 36 since under the null hypothesis $\widehat{\sigma}_{\varsigma^{**}}\xrightarrow{p}\sigma_{\varepsilon}$. (37) is implied by CMT . Analogously, the covariance kernel in (40),

$$\begin{aligned}\mathbb{E}\left[\mathbf{K}_{i}(s,w)\mathbf{K}_{i}(r,v)'\right] &= \mathbb{E}[(\mathbf{B}_{i}(s,w)-\mathbf{Q}_{i}\left(s,w\right)\mathbf{Q}_{i}\left(1,w\right)^{-1}\mathbf{B}_{i}(1,w))\\ &\quad\times(\mathbf{B}_{i}(r,v)-\mathbf{Q}_{i}\left(r,v\right)\mathbf{Q}_{i}\left(1,v\right)^{-1}\mathbf{B}_{i}(1,v))']\end{aligned}$$

Taking $\mathbb{E}\left[\mathbf{B}_{i}(s,w)\mathbf{B}_{i}(r,v)'\right]=\sigma_{\varepsilon}^{2}\mathbf{Q}_{i}\left(s\wedge r,w\wedge v\right),$

$$\begin{aligned}\mathbb{E}\left[\mathbf{K}_{i}(s,w)\mathbf{K}_{i}(r,v)'\right] &= \sigma_{\varepsilon}^{2}[\mathbf{Q}_{i}\left(s\wedge r,w\wedge v\right)-\mathbf{Q}_{i}\left(s\wedge 1,w\wedge v\right)\mathbf{Q}_{i}\left(1,v\right)^{-1}\mathbf{Q}_{i}\left(r,v\right)\\ &\quad-\mathbf{Q}_{i}\left(s,w\right)\mathbf{Q}_{i}\left(1,w\right)^{-1}\mathbf{Q}_{i}\left(1\wedge r,w\wedge v\right)\\ &\quad+\mathbf{Q}_{i}\left(s,w\right)\mathbf{Q}_{i}\left(1,w\right)^{-1}\mathbf{Q}_{i}\left(1\wedge 1,w\wedge v\right)\mathbf{Q}_{i}\left(1,v\right)^{-1}\mathbf{Q}_{i}\left(r,v\right)]\end{aligned}$$

Consider $s,r<1$ and $w>v$,

$$\mathbb{E}\left[\mathbf{K}_{i}(s,w)\mathbf{K}_{i}(r,v)'\right]=\sigma_{\varepsilon}^{2}[\mathbf{Q}_{i}\left(s\wedge r,v\right)-\mathbf{Q}_{i}\left(s,v\right)\mathbf{Q}_{i}\left(1,v\right)^{-1}\mathbf{Q}_{i}\left(r,v\right)]$$

Therefore, we conclude,

$$\mathbb{E}\left[\mathbf{K}_{i}(s,w)\mathbf{K}_{i}(r,v)'\right]=\sigma_{\varepsilon}^{2}[\mathbf{Q}_{i}\left(s\wedge r,w\wedge v\right)-\mathbf{Q}_{i}\left(s,w\wedge v\right)\mathbf{Q}_{i}\left(1,w\wedge v\right)^{-1}\mathbf{Q}_{i}\left(r,w\wedge v\right)]$$

which is stated result in (40). □

References

D.W.K. Andrews. Test for parameter instability and structural with unknown change point. *Econometrica*, 61:821–856, 1993.

A.C. Atkinson. A method for discriminating between models. *Journal of the Royal Statistical Society*, B32:211–243, 1970.

J. Bai. Estimation of a change point in multiple regression models. *Review of Economics and Statistics*, 60:551–563, 1995.

J. Bai. Testing for a parameter constancy in linear regressions: An empirical distributions function approach. *Econometrica*, 64:597–622, 1996.

J. Bai and P. Perron. Estimating and testing linear models with multiple structural changes. *Econometrica*, 66(1):47–78, 1998.

P.K. Bhattacharya and P.J. Brockwell. The minimum of an additive process with applications to signal estimation and storage theory. *Z Wahrschein Verw Gebiete*, 37:51–75, 1976.

P.J. Bickel and M.J. Wichura. Convergence criteria for multiparameter stochastic process and some applications. *The Annals of Mathematical Statistics*, 42: 1656–1670, 1971.

P. Billingsley. *Convergence of Probability Measures*. New York: Wiley Series in Probability and Statistics, 1968.

M. Caner and B.E. Hansen. Threshold autoregression with a unit root. *Econometrica*, 69:1555–1592, 2001.

M. Carrasco. Misspecified structural change, threshold, and markov-switching models. *Journal of Econometrics*, 109:239–273, 2002.

K.S. Chan. Testing for threshold autoregression. *The Annals of Statistics*, 4: 1886–1894, 1990.

K.S. Chan. Consistency and limiting distribution of the least squares estimator of a threshold autoregressive model. *The Annals of Statistics*, 21:520–533, 1993.

K.S. Chan and H. Tong. On the use of deterministic lyapunov function for the ergodicity of stochastic difference equations. *Adv. Appl. Probablity*, 17: 666–678, 1985.

D.B.H. Cline and H-M. M. Pu. Geometric ergodicity of nonlinear time series. *Statistica Sinica*, 9:1103–1118, 1999.

D.R. Cox. Tests of separate families of hypothesis. *Proceedings of the Forth Berkeley Symposium on Mathematical Statistics and Probability*, 1:105123, 1961.

D.R. Cox. Further results on tests of separate families of hypothesis. *Journal of the Royal Statistical Society*, pages 406–424, 1962.

M. Csörgo and L. Horváth. *Limit Theorems in Change-Point Analysis*. West Sussex: Wiley Series in Probability and Statistics, 1996.

M. Csörgo and P. Révész. *Strong Approximations in Probability and Statistics*. New York: Academic Press, 1981.

J. Davidson. Establishing conditions for the functional central limit theorem in nonlinear and semiparametric time series process. *Journal of Econometrics*, 106:243–269, 2002.

R. Davidson and J.G. McKinnon. Several tests for model specification in the presence of alternative hypothesis. *Econometrica*, 49:781–793, 1981.

R. Davidson and J.G. McKinnon. *Estimation and Inference in Econometrics*. Oxford: Oxford University Press, 1993.

R.A. Davies, D. Huang, and Y.C. Yao. Testing for change in the parameter values and order of an autoregression. *The Annals of Statistics*, 1:282–304, 1995.

R.B. Davies. Hypothesis testing when a nuisance parameter is present only under the alternative. *Biometrika*, 64:247–254, 1977.

P.H. Franses and D. Van Dijk. *Non-Linear Time Series models in Empirical Finance*. Cambridge University Press, 2000.

J. Gonzalo and R. Montesinos. Threshold stochastic unit root models. 2002.

J. Gonzalo and J.Y. Pitarakis. Estimation and model selection based inference in single and multiple threshold models. *Journal of Econometrics*, 110:319–352, 2002.

C. Gourieroux and A. Monfort. Testing non-nested hypothesis. In R.F. Engle and D.L. McFadden, editors, *Handbook of Econometrics*, volume IV. Holland: Elsevier Science, 1994.

C. Gourieroux and A. Monfort. *Statistics and Econometric Models*. Cambridge University Press, 1995.

B.E. Hansen. Inference when a nuisance parameter is not identified under the null hypothesis. *Econometrica*, 64(2):413–430, 1996.

B.E. Hansen. Sample splitting and threshold estimation. *Econometrica*, 68(3): 575–603, 2000.

T. Hida. *Brownian Motion*. New York: Springer-Verlag, 1980.

G. Kapetanios and M. Weeks. Non-nested models and the likelihood ratio statistic: A comparison of simulation and bootstrap based tests. 2003.

T.H. Lee, H. White, and C.W.J. Granger. Testing for neglected nonlinearity in time series models. *Journal of Econometrics*, 56:269–290, 1993.

G. Mizon and J.F. Richard. The encompassing principle and its applications to testing non-nested hypothesis. *Econometrica*, 3:657–678, 1986.

H.M. Pesaran. On the general problem of model selection. *Review of Economic Studies*, 41:153–171, 1974.

M.H. Pesaran and A.S. Deaton. Testing non-nested nonlinear regression models. *Econometrica*, 46:677–694, 1978.

M.H. Pesaran and B. Pesaran. A simulation approach to the problem od computing cox's statistic for testing nonested models. *Journal of Econometrics*, 57:377–392, 1993.

D. Pollard. *Convergence of Stochastic Processes*. New York: Springer-Verlag, 1985.

G.R. Shorack and J.A. Wellner. *Empirical Processes with Applications to Statistics*. New York: John Wiley and Sons Inc, 1986.

R.J. Smith. Non-nested for competing models estimated by generalized method of moments. *Econometrica*, 4:973–980, 1992.

H. Tong. *Non-Linear Time Series: A Dinamical Approach*. Oxford: Oxford Science Publications, 1990.

A.M. Walker. Some tests of separate families of hypothesis in time series analysis. *Biometrika*, 54:39–68, 1967.