

**ROBUST ESTIMATION
OF STRUCTURAL
BREAK POINTS**

Inmaculada Fiteni

98-51



WORKING PAPERS

Working Paper 98-51
Statistics and Econometrics Series 23
July 1998

Departamento de Estadística y Econometría
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (34)91 624-9849

ROBUST ESTIMATION OF STRUCTURAL BREAK POINTS.

Inmaculada Fiteni*

Abstract

This paper is concerned with robust estimation of change points in regression models, possibly with trending regressors. We obtain the rate of convergence and the asymptotic distribution of M-estimators of the regression coefficients and the change point with serially dependent observations. The asymptotic properties of the estimators are developed assuming that the size of the jump is fixed as well as it shrinks to zero as the sample size increases. In the first case, the asymptotic distribution of the change point estimator is difficult to tabulate. The performance of asymptotic inferences in practice is illustrated by means of Monte Carlo simulations.

Keywords:

Structural breaks; Robustness; M-estimators; τ -estimators; Serial dependence.

*Fiteni, Departamento de Estadística y Econometría, Universidad Carlos III de Madrid. C/ Madrid, 126 28903 Madrid. Spain. Ph: 34-1-624.98.90; Fax: 34-1-624.98.49, e-mail: fiteni@est-econ.uc3m.es. Research funded by Dirección General de Enseñanza Superior, reference number: PB95-0292. This paper is part of my doctoral dissertation at Universidad Carlos III de Madrid. I thank Miguel Delgado and Javier Hidalgo for their suggestions and advice.

1. INTRODUCTION

The development of statistical inferences on structural change points has been of major concern in the statistic and econometric literature. Among survey papers on this topic we mention Zacks (1983), Wölfe and Schechtman (1984), Deshayes and Picard (1986), Husková and Sen (1989), Krishnaiah and Miao (1988), Csörgo and Horváth (1988), Perron (1993) and Stock (1997). There are also several monographs, like Broemeling (1982), Broemeling and Tsurumi (1986), Hackl (1989), Hackl and Westlund (1989, 91), and Brodsky and Darkhhansky (1994). The problem is to test whether or not a change in the parameters of the model has occurred, and if so, to estimate when and by how much.

Much has been written about testing for structural breaks, but comparatively little is known about estimation of the location point. Hinkley (1971), Bhattacharya (1987) and Yao (1987) consider the maximum likelihood estimation of the break date for i.i.d. variables with a simple shift, Picard (1985) for a Gaussian autoregressive process, and Feder (1975) for segmented regressions. Bai (1994) estimates the unknown change point by the method of least squares in a linear process. But these classical estimators are sensitive to deviations from the model distribution, to outlying observations, and to model misspecifications, which can produce disastrous effects on the estimates. Departures from the assumed model can be solved, in part, estimating nonparametrically the underlying regression model, as proposed by Carlstein (1988), Dümbgen (1991), Chu and Wu (1994), and Delgado and Hidalgo (1997) among others. Alternatively, robust methods, which are insensitive to small deviations from the assumptions, can also be applied. Bai (1995) proposes the use of LAD estimation, which has good properties in terms of robustness (qualitative robustness, B-robustness and maximum breakdown point). However, this estimator is not robust in the presence of leverage, and is very inefficient under normality. Antoch and Husková (1997) propose an M-estimator for the location of a change in the mean of i.i.d. variables, obtaining the best trade-off between efficiency under the true model and robustness, but with breakdown zero if we consider stochastic regressors.

This paper generalizes the existing literature on the estimation in three aspects. First, we study M-estimation of the break points in the context of regression models with structural change, obtaining the asymptotic behavior for the estimators of the location and size of the break date, such as both pre-break and post-break parameters, which allows per-

forming inferences. The above mentioned estimators can be considered as particular cases of this one. Second, most of the existing estimation procedures are constructed based on the *i.i.d.* assumption which, unfortunately, often does not hold in real world analyses. Thus, we introduce general conditions on serially dependence. At last, we discuss the problem of robust estimation under leverage, studying the extension to τ -estimators, considered as a particular class of globally robust estimators, with maximum breakdown point.

The rest of the paper is organized as follows. Section 2 introduces the model and regularity conditions. The asymptotic properties of the M-estimators are studied in Section 3, under two assumptions about the size of the jump, fixed and converging to zero with the sample size. In Section 4, we discuss the extension to τ -estimators, which are robust in the presence of leverage. Section 5 shows the results of Monte Carlo experiments, which illustrates the performance of the asymptotic approximations in practice.

2. MODEL AND ASSUMPTIONS

Let $\{Z_t = (Y_t, X_t), t = 1, \dots, T\}$ be a sample of a $R \times R^p$ -valued stochastic process defined on the probability space (Ω, F, P) . We are interested in estimating the parameter vector $\xi_0 = (\beta'_{10}, \beta'_{20}, \tau_0)'$ where $\beta_{j0} \in \Theta \subset R^p, j = 1, 2, \beta_{10} \neq \beta_{20}$, and $\tau_0 \in \Pi \subset (0, 1)$, defined by means of the moment condition

$$\xi_0 = \underset{\xi \in \Theta^2 \times \Pi}{\text{Arg min}} \quad \underset{T \rightarrow \infty}{\text{Lim}} S_T(\xi), \quad (1)$$

where

$$S_T(\xi) = S_{1T}(\beta_1, \tau) + S_{2T}(\beta_2, \tau), \quad (2)$$

$$S_{1T}(\beta, \tau) = \frac{1}{T} \sum_{t=1}^{[T\tau]} \rho(Y_t - X_t' \beta), \quad (3)$$

$$S_{2T}(\beta, \tau) = \frac{1}{T} \sum_{t=[T\tau]+1}^T \rho(Y_t - X_t' \beta). \quad (4)$$

$[\cdot]$ means the nearest integer, and $\rho: R \rightarrow R$ is a function which identifies the parameters of the model. That is, each ρ function defines a particular linear predictor of Y_t given X_t , with changing parameters at a given moment of time τ_0 . For instance, $\rho(u) = u^2$ defines the least squares predictor, $\rho(u) = |u|$ the least absolute deviation predictor, and

$\rho(u) = \frac{1}{2}u^2I(|u| \leq c) + |u|I(|u| > c)$ is the Huber predictor, a compromise between the two above, where $I(A)$ is the indicator function of the event A , and c is a suitable chosen constant. Different ρ functions may define different parameter values, except in certain circumstances. For instance, when the conditional distribution of Y_t given X_t is symmetric with respect to its mean, which is a linear combination of the X_t with changing parameters at X_t , least squares, least absolute deviations and Huber predictors are identical. However, the resulting estimators will have very different statistical properties.

Natural estimates of ξ_0 are defined as the sample analogues of (1),

$$\hat{\xi}_T = \underset{\xi \in \Theta^2 \times \Pi}{\text{Arg min}} S_T(\xi), \quad (5)$$

and can be obtained by means of an iterative procedure, using the fact that

$$\hat{\tau}_T = \underset{\tau \in \Pi}{\text{Arg min}} \left(S_{1T}(\hat{\beta}_{1T}(\tau), \tau) + S_{2T}(\hat{\beta}_{2T}(\tau), \tau) \right),$$

where

$$\hat{\beta}_{jT}(\tau) = \underset{\beta \in \Theta}{\text{Arg min}} S_{jT}(\beta, \tau), \quad \forall \tau \in (0, 1), \quad j = 1, 2 \quad (6)$$

which are estimating

$$\beta_j(\tau) = \underset{\beta \in \Theta}{\text{Arg min}} \lim_{T \rightarrow \infty} S_{jT}(\beta, \tau), \quad j = 1, 2$$

Thus, $\hat{\beta}_{jT} = \hat{\beta}_{jT}(\hat{\tau}_T)$, for $j = 1, 2$ and the size of the jump $\lambda_0 = \beta_{10} - \beta_{20}$ is estimated by $\hat{\lambda}_T = \hat{\beta}_{1T} - \hat{\beta}_{2T}$.

Remark 1 *In order to get a scale-invariant estimator, we must consider the objective function (2)-(4), with $S_{jT}(\xi)$ replaced by*

$$\begin{aligned} S_{1T}(\beta, \tau) &= \frac{1}{T} \sum_{t=1}^{[T\tau]} \rho\left(\frac{Y_t - X_t' \beta}{\hat{\sigma}_T}\right), \\ S_{2T}(\beta, \tau) &= \frac{1}{T} \sum_{t=[T\tau]+1}^T \rho\left(\frac{Y_t - X_t' \beta}{\hat{\sigma}_T}\right), \end{aligned}$$

where $\hat{\sigma}_T$ is the scale estimate of the error, which can be obtained either separately or simultaneously with $\hat{\xi}_T$. These estimators are called *M-estimates with general scale*. A consistent and robust estimator of the scale can be obtained from a preliminary estimator $\tilde{\xi}_T$ of ξ_0 . For instance, the median absolute deviations, defined as

$$\hat{\sigma}_T = \underset{\sigma}{\text{Arg min}} \frac{1}{T} \sum_{t=1}^{[T\tilde{\tau}_T]} \left| |Y_t - X_t' \tilde{\beta}_{1T}| - \sigma \right| + \frac{1}{T} \sum_{t=[T\tilde{\tau}_T]+1}^T \left| |Y_t - X_t' \tilde{\beta}_{2T}| - \sigma \right|,$$

assuming that ξ_0 is identified by $\rho(u) = |u|$. In cases where we want to allow different scales for each subsample, it can be defined

$$(\hat{\sigma}_{1T}, \hat{\sigma}_{2T}) = \underset{\sigma_1, \sigma_2}{\text{Arg min}} \frac{1}{T} \sum_{t=1}^{\lfloor T\bar{\tau}_T \rfloor} \left| |Y_t - X_t' \tilde{\beta}_{1T}| - \sigma_1 \right| + \frac{1}{T} \sum_{t=\lfloor T\bar{\tau}_T \rfloor + 1}^T \left| |Y_t - X_t' \tilde{\beta}_{2T}| - \sigma_2 \right|,$$

and $\hat{\sigma}_T = \hat{\sigma}_{1T} I(t \leq \lfloor T\bar{\tau}_T \rfloor) + \hat{\sigma}_{2T} I(t > \lfloor T\bar{\tau}_T \rfloor)$. By Lemma 5.1 of Yohai (1985), it can be shown that, with a consistent estimator of the scale, the asymptotic behavior of the estimates are equivalent with the scale estimated or assumed known. For simplicity henceforth, we will consider the objective function composed by (3) and (4) instead of the previous one.

The properties of the estimators will be derived assuming 'Near Epoch Dependence'. This class of dependence goes back to Ibramigov (1962), and has been formalized in different ways by Billingsley (1968), McLeisch (1975), Bierens (1981), Wooldridge and White (1988), Andrews (1988), and Pötscher and Prucha (1991), among others. In order to obtain the asymptotic properties of the estimators, it is necessary to make the following assumptions:

A.1 Assumptions on $\rho(\cdot)$

A.1.1 Let $\rho(\cdot)$ be a convex real function twice continuously differentiable in R , with first derivative ψ , such that

$$\lim_{x \rightarrow \pm\infty} \rho(x) = \infty.$$

A.1.2 $\psi(\cdot)$ is a bounded and strictly negative (positive) function for large negative (positive) values of its argument; $\dot{\psi}(u) = d\psi(u)/du$.

A.2 Model assumptions.

Define the sequence $\{\eta_t(\theta) = \psi(u_t + \theta' X_t) X_t, \forall t \leq T\} \forall \theta \in \Theta$, and $\eta_t = \eta_t(\theta) |_{\theta=0}$, $\forall t \leq T$.

A.2.1 $\Theta \subset R^p$, is a convex set.

A.2.2 $\tau_0 \in \Pi \subset (0, 1)$.

A.2.3 $\{Z_t = (Y_t, X_t)' : t \leq T\}$, is a random vector with domain in Z , $L^0 - NED$ on a strong mixing base $\{w_t : t = \dots, 0, 1, \dots\}$, where Z is a Borel subset of R^{p+1} . Let $F_t^{(z)}$ be the distribution function of Z_t and $F_T^{(z)} = (1/T) \sum_1^T F_t^{(z)}$; then $\{F_T^{(z)}, T \geq 1\}$ is tight on Z .

A.2.4 For some $r > 2$, $\{\eta_t : t \leq T\}$ is a random vector sequence of mean zero, $L^2 - NED$ of size $-1/2$ on a strong mixing base $\{w_t : t = \dots, 0, 1, \dots\}$ of size $-\tau / (r - 2)$ and $Sup_{t \leq T} E |\eta_t|^r < \infty$.

A.2.5 $\eta_t(\theta)$ is Borel measurable in $Z_t \forall \theta \in \Theta$, and $\partial \eta_t(\theta) / \partial \theta'$, that is continuous in $(Z_t, \theta) \in Z \times \Theta$ by A.1.1, satisfy that,

$$Sup_{t \leq T} E \left[Sup_{\theta \in \Theta} \left\| \frac{\partial \eta_t(\theta)}{\partial \theta'} \right\|^{1+\varepsilon} \right] < \infty,$$

for some $\varepsilon > 0$:

A.2.6 The $Lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^{[T\tau]} E [\partial \eta_t(\theta) / \partial \theta']$ exists uniformly in $(\theta, \tau) \in \Theta \times \Pi$ and equals $\tau M(\theta) \forall (\theta, \tau) \in \Theta \times \Pi$, where

$$M(\theta) = Lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T E \left[\frac{\partial \eta_t(\theta)}{\partial \theta'} \right]. \quad (7)$$

a positive definite matrix. For notational convenience, define $M = M(\theta) |_{\theta=0}$.

A.2.7 $\forall \tau \in (0, 1)$ we assume that:

$$Lim_{T \rightarrow \infty} Var \left[\frac{1}{\sqrt{T}} \sum_{t=1}^{[T\tau]} \eta_t \right] = \tau S,$$

where

$$S = Lim_{T \rightarrow \infty} Var \left[\frac{1}{\sqrt{T}} \sum_{t=1}^T \eta_t \right] \quad (8)$$

is a finite and positive $p \times p$ matrix.

Assumptions A.1 are standard in robust estimation. The differentiability of ρ allow us to express (6) by the first order conditions,

$$\begin{aligned} \{\hat{\beta}_{1T}(\tau)\} &= \left\{ \beta \in \Theta \ / \sum_{t=1}^{[T\tau]} \eta_t(\beta_0 - \beta) = 0 \right\}, \\ \{\hat{\beta}_{2T}(\tau)\} &= \left\{ \beta \in \Theta \ / \sum_{t=[T\tau]+1}^T \eta_t(\beta_0 - \beta) = 0 \right\}, \end{aligned}$$

estimators of

$$\begin{aligned}\{\beta_1(\tau)\} &= \left\{ \beta \in \Theta \ / \ Lim_{T \rightarrow \infty} \sum_{t=1}^{[T\tau]} \eta_t(\beta_0 - \beta) = 0 \right\}, \\ \{\beta_2(\tau)\} &= \left\{ \beta \in \Theta \ / \ Lim_{T \rightarrow \infty} \sum_{t=[T\tau]+1}^T \eta_t(\beta_0 - \beta) = 0 \right\},\end{aligned}$$

Thus, $\{\hat{\beta}_{1T}(\tau)\}$ and $\{\hat{\beta}_{2T}(\tau)\}$ define the subsets such that, for fixed τ , the objective function is minimized. The convexity of $\rho(\cdot)$ implies the convexity of these subsets. Under A.1, the subsets of Θ , $\{\beta_1(\tau)\}$ and $\{\beta_2(\tau)\}$, are non empty, convex and compact. If $\rho(\cdot)$ is strictly convex, these sets shrinks to an unique point (Huber, 1964). But, this property, although would simplify the problem, rules out estimators like Hubers, which are of interest. The same occurs with the continuous assumption of the second derivative of $\rho(\cdot)$, that we make in A.1.1. However, in order to solve this, we obtain, in Proposition 1 below, an uniformly convergent smoothed version of the Huber score function.

In order to derive the asymptotic properties of estimator parameters, we suppose convexity of the parametric space, by assumptions A.2.1 and A.2.2, that also consider the shift location far away of the interval extremes. A.2.3 and A.2.4 are standard of weak dependence for this case of robust regression, and A.2.5, A.2.6 and A.2.7 are covariance stationary conditions.

Proposition 1 *Given the Huber ψ -function,*

$$\psi(t) = \begin{cases} c, & t > c \\ t, & |t| < c \\ -c, & t < -c \end{cases},$$

we obtain the following sequence of twice differentiable functions,

$$h_T(t) = \left. \begin{cases} c & t > c + \frac{1}{2T^p} \\ -\frac{T^p}{2}t^2 + T^p\left(c + \frac{1}{2T^p}\right)t - \left(\frac{T^p}{2}c^2 - \frac{c}{2} + \frac{1}{8T^p}\right), & c - \frac{1}{2T^p} \leq t \leq c + \frac{1}{2T^p} \\ t, & |t| < c - \frac{1}{2T^p} \\ \frac{T^p}{2}t^2 + T^p\left(c + \frac{1}{2T^p}\right)t + \left(\frac{T^p}{2}c^2 - \frac{c}{2} + \frac{1}{8T^p}\right), & -c \leq t \leq -c + \frac{1}{2T^p} \\ -c & t < -c - \frac{1}{2T^p} \end{cases} \right\},$$

such that $h_T(t)$ converges to $\psi(t)$ uniformly in t , for $p > 0$ and fixed c .

Thus, the asymptotic results derived for smooth ψ functions can apply as in Theorem 2 of Bloomfield and Steiger (1983) for LAD estimation. This is an alternative way to solve the problem of non-differentiability of the objective function, which have been already analyzed by others, such as Jureckova (1996), who considers this type of functions as a sum of three functions with different degrees of smoothing, and studies the asymptotic behavior of each one. The proposition can be extended to another non differentiable functionals.

3. ASYMPTOTIC PROPERTIES

In order to establish the asymptotic distribution of the estimators, firstly we need to derive their rate of consistency. This is obtained in Theorem 1 below.

Theorem 1 *Assume A.1 and A.2, then*

$$\begin{aligned} (\hat{\beta}_{jT} - \beta_{j0}) &= Op\left(\frac{1}{\sqrt{T}}\right), \quad j = 1, 2, \\ (\hat{\tau}_T - \tau_0) &= Op\left(\frac{1}{T\|\lambda_0\|^2}\right), \end{aligned} \quad (9)$$

where $\|x\|$ represents the Euclidean norm of the vector x .

To obtain the previous results, we analyze the global behavior of the objective function $S_T(\xi)$ over the whole parameter space. To this end observe that the parameter estimator (5) can be also defined as

$$\hat{\xi}_T = \underset{\xi \in \Theta^2 \times \Pi}{Arg \min} (S_T(\xi) - S_T(\xi_0)), \quad (10)$$

and we want to prove that $\forall \varepsilon > 0, \exists C > 0$ such that

$$\Pr \left\{ \left(\|\hat{\beta}_{1T} - \beta_{10}\| > \frac{C}{\sqrt{T}} \right) \cup \left(\|\hat{\beta}_{2T} - \beta_{20}\| > \frac{C}{\sqrt{T}} \right) \cup \left(\|\hat{\tau}_T - \tau_0\| > \frac{C}{T\|\lambda_0\|^2} \right) \right\} < 3\varepsilon. \quad (11)$$

The upper bound of 3ε is only for notational convenience and, without loss of generality, it corresponds to one ε for each of the three sets. By definition of $\hat{\xi}_T$ in (10), $S_T(\hat{\xi}_T) - S_T(\xi_0) \leq 0$, so the left side of (11) is upper bounded by

$$\Pr \left\{ \underset{A \cup B \cup D}{Inf} (S_T(\xi) - S_T(\xi_0)) < 0 \right\} = \Pr \left\{ \underset{A \cup B \cup D}{Sup} (S_T(\xi_0) - S_T(\xi)) < 0 \right\},$$

where the sets A , B and D are defined as follows

$$A = \left\{ \beta_1 \in \Theta / \left\| \hat{\beta}_{1T} - \beta_{10} \right\| > \frac{C}{\sqrt{T}} \right\}, \quad (12)$$

$$B = \left\{ \beta_2 \in \Theta / \left\| \hat{\beta}_{2T} - \beta_{20} \right\| > \frac{C}{\sqrt{T}} \right\}, \quad (13)$$

$$D = \left\{ \tau \in \Pi / \left\| \hat{\tau}_T - \tau_0 \right\| > \frac{C}{T \|\lambda_0\|^2} \right\}. \quad (14)$$

Thus, Theorem 1 is a consequence of the following result.

Theorem 2 *Under A.1 and A.2, $\forall \varepsilon > 0$, $\exists C > 0$ such that*

$$\Pr \left\{ \sup_{A \cup B \cup D} (S_T(\xi_0) - S_T(\xi)) < 0 \right\} < 3\varepsilon,$$

where A , B and D are defined in (12), (13) and (14) respectively.

The estimators of the regression coefficients are, as usual, \sqrt{T} -consistent. For the rate of convergence of the structural break point estimator, we can consider two interesting cases of $\lambda_0 = (\beta_{10} - \beta_{20})$, when λ_0 is constant and when it depends on the sample size, $\lambda_0 = \lambda_T$ with $\|\lambda_T\| \rightarrow 0$, but $T\|\lambda_T\|^2 \rightarrow \infty$, which is satisfied trivially for the constant λ_0 case. From (9), we observe that the rate of convergence of $\hat{\tau}_T$ is $Op(T^{-1})$ when λ_0 is fixed, and $Op(T^{-1}\|\lambda_T\|^{-2})$ otherwise. Finally, we can obtain as a direct result, the \sqrt{T} -consistency of the jump size estimator $\hat{\lambda}_T = \hat{\beta}_{1T} - \hat{\beta}_{2T}$.

As viewed before, Theorem 2 describes the global behavior of $S_T(\beta_1, \beta_2, \tau)$ in the whole set $\Theta^2 \times \Pi$. Next, the limiting distribution is obtained by studying the local behavior of this objective function in a compact set determined by Theorem 2. Thus, β_j , $j = 1, 2$, is constrained to be in a $T^{-1/2}$ neighborhood of the true parameters values, β_{j0} , $j = 1, 2$ respectively, and a similar comment applies to the estimated shift point, in a T^{-1} neighborhood of τ_0 . In order to derive the asymptotic distribution, it is necessary to obtain the rate of convergence previously, because the argmin functional, used to obtain the location estimate, is not continuous when the minimized function is defined in an unbounded set.

Therefore, it is convenient to reparametrize the objective function in the following way:

$$\Lambda_T(v) = S_T \left(\xi_0 + \left(\frac{v'_1}{\sqrt{T}}, \frac{v'_2}{\sqrt{T}}, \frac{v_3 P_\lambda}{T} \right)' \right) - S_T(\xi_0),$$

for $v = (v'_1, v'_2, v_3)' \in V_N$, the argument of this new function, such that

$$V_N = \{v : |v_j| < N, j = 1, 2, 3\},$$

is a compact set defined for every $N > 0$. Thus,

$$\begin{aligned} \beta_j &= \beta_{j0} + \frac{v_j}{\sqrt{T}}, \quad j = 1, 2, \\ \tau &= \tau_0 + v_3 \frac{P_\lambda}{T}, \quad \text{where } P_\lambda = O(\|\lambda_0\|^{-2}); \end{aligned} \quad (15)$$

so that considering (15), we have $P_{\lambda_T} = O(\|\lambda_T\|^{-2})$ for $\lambda_0 = \lambda_T$ decreasing case, and $P_\lambda = 1$ for constant λ_0 .

The weak convergence results for the estimators follows taking into account that:

$$\begin{aligned} \sqrt{T}(\hat{\beta}_{jT} - \beta_{j0}) &= \hat{v}_j, \quad j = 1, 2, \\ T(\hat{\tau}_T - \tau_0) &= \hat{v}_3 P_\lambda, \end{aligned}$$

and

$$(\hat{v}'_1, \hat{v}'_2, \hat{v}_3)' = \underset{v \in V_N}{\text{Arg min}} \Lambda_T(v),$$

defined in a compact set. This is obtained in Theorem 3 bellow.

Theorem 3 Assume A1 and A2, then,

(i)

$$\begin{bmatrix} \hat{v}_1 \\ \hat{v}_2 \end{bmatrix} \xrightarrow{d} \begin{bmatrix} M^{-1} S^{1/2} \frac{B(\tau_0)}{\tau_0} & 0 \\ 0 & M^{-1} S^{1/2} \frac{B(1-\tau_0)}{1-\tau_0} \end{bmatrix}, \quad (16)$$

where " \xrightarrow{d} " represents convergence in distribution, M and S are defined in (7) and (8) respectively, and $B(\cdot)$ is a p -vector of independent Brownian motions defined in $[0, 1]$.

(ii) Assuming λ_0 constant,

$$\hat{v}_3 \Rightarrow \underset{w}{\text{Arg max}} \left\{ \lambda'_0 S^{1/2} W^*(w) - \frac{1}{2} \lambda'_0 M(\lambda_0) \lambda_0 |w| \right\}, \quad (17)$$

where $M(\theta)$ is defined by (7) and $W^*(\cdot)$ represents a process on Z , the integer set, such that:

$$W^*(w) = \begin{cases} 0 & w = 0 \\ \sum_{t=v}^{-1} \eta_t & w = -1, -2, \dots \\ \sum_{t=1}^v \eta_t & w = 1, 2, \dots \end{cases}.$$

(iii) Assuming $\lambda_0 = \lambda_T \rightarrow 0$ with $T \|\lambda_T\|^2 \rightarrow \infty$,

$$\frac{(\lambda_T' M \lambda_T)^2}{(\lambda_T' S \lambda_T)} \hat{v}_3 \Rightarrow \underset{w}{\text{Arg max}} \left\{ W(w) - \frac{1}{2} |w| \right\}, \quad (18)$$

where $W(\cdot)$ represents an independent two-sided standard Brownian motion defined in R .

(iv) Finally, we obtain that $(\hat{v}'_1, \hat{v}'_2)'$ and \hat{v}_3 are asymptotically independent both, when λ_0 is fixed or converging to zero.

The limiting distribution for the estimated regression parameters is standard, and is the same as if the true change point were known. For the shift estimator, we only obtain a free-distribution under the assumption of λ_0 decreasing with T . A two sided Brownian motion is represented by

$$W(w) = \begin{cases} W_1(-w), & w < 0 \\ W_2(w), & w \geq 0 \end{cases},$$

with $\{W_1(t) : t \in [0, \infty)\}$ and $\{W_2(t) : t \in [0, \infty)\}$ being independent standard Brownian processes. The explicit form of distribution (18) is given by:

$$F(t) = \frac{1}{2} + \frac{1}{\sqrt{2\pi}} \sqrt{t} e^{-\frac{1}{2}t} + \frac{3}{2} e^t \Phi\left(-\frac{3}{2}\sqrt{t}\right) - \left(\frac{1}{2}t + \frac{5}{2}\right) \Phi\left(-\frac{1}{2}\sqrt{t}\right), \quad (19)$$

for $t > 0$. See, e.g., Bai (1994) or Antoch and Husková (1997), and references in there.

It can be easily seen, in the local change case, that the asymptotic variance of the shift estimators depends on the ratio

$$\frac{(\lambda_T' M \lambda_T)^2}{(\lambda_T' S \lambda_T)} = \frac{(\lambda_T' M(F, \psi) \lambda_T)^2}{(\lambda_T' S(F, \psi) \lambda_T)},$$

i.e., depends on the distribution F and the score function ψ . The larger is the ratio, the smaller is the asymptotic variance. Thus, the smallest variance corresponds to $\psi(x; \beta) = \frac{\partial f(x; \beta) / \partial \theta}{f(x; \beta)}$, $x \in R^p$, $\beta \in \Theta$, and $f(\cdot; \beta)$ the density function of $F(\cdot; \beta)$. If this density were unknown, an estimator of the optimal score function could be developed and this estimator can be used as the proper score function.

Corollary 1 *Considering part (i) of Theorem 2, it is possible to derive the asymptotic distribution of the jump size M -estimator:*

$$S^{-1/2} M \sqrt{T} (\hat{\lambda}_T - \lambda_0) \xrightarrow{d} \left(\frac{B(\tau_0) - \tau_0 B(1)}{\tau_0(1 - \tau_0)} \right), \quad (20)$$

defined as the generalized Bessel tied-down distribution of order p .

Corollary 2 *The following particular cases are worth mentioning:*

(i) *The regressors $\{X_t : t \leq T\}$ are i.i.d. Then we obtain that:*

$$\begin{aligned} M &= \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T E [\dot{\psi}_t X_t X_t'] = E [\dot{\psi}(u_t)] E [X_t X_t'], \\ S &= \lim_{T \rightarrow \infty} \text{Var} \left[\frac{1}{\sqrt{T}} \sum_{t=1}^T \psi_t X_t \right] = E [\psi^2(u_t)] E [X_t X_t'], \end{aligned}$$

and therefore,

$$M^{-1} S^{1/2} = \frac{E [\psi^2(u_t)]^{1/2}}{E [\dot{\psi}(u_t)]} D^{-1/2}, \quad (21)$$

where $D = E [X_t X_t']$.

(ii) *The regressors $\{X_t : t \leq T\}$ are functions of time trends. Let be $X_t = g(t/T)$, where g is a bounded vector valued-function defined on $(0, 1)$ and is continuously differentiable. Then*

$$\begin{aligned} M &= \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T E [\dot{\psi}_t X_t X_t'] = E [\dot{\psi}(u_t)] \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T (g(t/T) g(t/T)'), \\ S &= \lim_{T \rightarrow \infty} \text{Var} \left[\frac{1}{\sqrt{T}} \sum_{t=1}^T \psi_t X_t \right] = E [\psi^2(u_t)] \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T (g(t/T) g(t/T)'), \end{aligned}$$

and therefore,

$$M^{-1} S^{1/2} = \frac{E [\psi^2(u_t)]^{1/2}}{E [\dot{\psi}(u_t)]} D^{-1/2}, \quad (22)$$

where $D = \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T (g(t/T) g(t/T)').$

Using the expressions (21) and (22) in (16) and (20) we can derive the limit distribution of the regressor coefficients and the jump size, for both cases (i) and (ii) respectively.

Corollary 3 *We consider the limit distribution of the shift estimator for the two cases of $\{X_t : t \leq T\}$, i.i.d. and function trends. Assuming that the jump size decreases with T , we obtain that,*

(i) *For the i.i.d. case we have,*

$$\frac{(\lambda_T' M \lambda_T)^2}{(\lambda_T' S \lambda_T)} = \frac{(\lambda_T' D \lambda_T) E [\dot{\psi}(u_t)]^2}{E [\psi^2(u_t)]},$$

where $D = E [X_t X_t']$. Therefore,

$$(\lambda_T' D \lambda_T) T (\hat{\tau}_T - \tau_0) \Rightarrow \frac{E [\psi^2(u_t)]}{E [\dot{\psi}(u_t)]^2} \underset{v}{\text{Arg max}} \left\{ W(v) - \frac{1}{2} |v| \right\}.$$

(ii) For the trending case with the function g as in Corollary 2, we obtain that

$$(\lambda_T' g(\tau_0) g(\tau_0)' \lambda_T) T (\hat{\tau}_T - \tau_0) \Rightarrow \frac{E[\psi^2(u_t)]}{E[\psi(u_t)]^2} \underset{v}{\text{Argmax}} \left\{ W(v) - \frac{1}{2} |v| \right\},$$

a result that cannot be derived directly by Theorem 2. The proof is given in Appendix 2.

In order to make inferences about the parameters of the model, we need to estimate the asymptotic variance. In particular, M , defined in (7), can be consistently estimated by,

$$\hat{M} = \hat{M}_1 + \hat{M}_2, \quad (23)$$

where,

$$\hat{M}_1 = \frac{1}{T\hat{\tau}} \sum_{t=1}^{[T\hat{\tau}]} \psi \left(Y_t - \hat{\beta}_{1T}(\hat{\tau}_T)' X_t \right) X_t X_t', \quad (24)$$

$$\hat{M}_2 = \frac{1}{T - T\hat{\tau}} \sum_{t=[T\hat{\tau}]+1}^T \psi \left(Y_t - \hat{\beta}_{2T}(\hat{\tau}_T)' X_t \right) X_t X_t', \quad (25)$$

and S , defined in (8), is consistently estimated by,

$$\hat{S} = \hat{S}_1 + \hat{S}_2, \quad (26)$$

such that, the estimation will depend on the serial dependence of the process. Thus, if $\{\eta_t, t \geq 1\}$ is a sequence of zero mean and uncorrelated random vectors, then:

$$S = \underset{T \rightarrow \infty}{\text{Lim}} \frac{1}{T} \sum_{t=1}^T E(\eta_t \eta_t'),$$

and, therefore, its estimator will be defined by

$$\begin{aligned} \hat{S}_1 &= \frac{1}{T\hat{\tau}} \sum_{t=1}^{[T\hat{\tau}]} \left(\psi \left(Y_t - \hat{\beta}_{1T}(\hat{\tau}_T)' X_t \right) X_t - \bar{m}_{1T}(\hat{\tau}_T) \right) \\ &\quad \times \left(\psi \left(Y_t - \hat{\beta}_{1T}(\hat{\tau}_T)' X_t \right) X_t - \bar{m}_{1T}(\hat{\tau}_T) \right)', \end{aligned} \quad (27)$$

$$\begin{aligned} \hat{S}_2 &= \frac{1}{T - T\hat{\tau}} \sum_{t=[T\hat{\tau}]+1}^T \left(\psi \left(Y_t - \hat{\beta}_{2T}(\hat{\tau}_T)' X_t \right) X_t - \bar{m}_{2T}(\hat{\tau}_T) \right) \\ &\quad \times \left(\psi \left(Y_t - \hat{\beta}_{2T}(\hat{\tau}_T)' X_t \right) X_t - \bar{m}_{2T}(\hat{\tau}_T) \right)', \end{aligned} \quad (28)$$

where:

$$\bar{m}_{1T}(\hat{\tau}_T) = \frac{1}{T\hat{\tau}} \sum_{t=1}^{[T\hat{\tau}]} \psi \left(Y_t - \hat{\beta}_{1T}(\hat{\tau}_T)' X_t \right) X_t,$$

$$\bar{m}_{2T}(\hat{\tau}_T) = \frac{1}{T - T\hat{\tau}} \sum_{t=[T\hat{\tau}_T]+1}^T \psi(Y_t - \hat{\beta}_{2T}(\hat{\tau}_T)' X_t) X_t.$$

On the other hand, if $\{\eta_t, t \geq 1\}$ is a zero mean and serially dependent random vectors, then

$$S = \sum_{v=0}^{\infty} \Gamma_v + \sum_{v=1}^{\infty} \Gamma'_v,$$

where:

$$\Gamma_v = \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T E(\eta_t \eta'_{t-v}),$$

and the estimator will be defined by

$$\begin{aligned} \hat{S}_1 &= \sum_{v=0}^{[T\hat{\tau}_T]-1} k\left(\frac{v}{l(T)}\right) \frac{1}{T\hat{\tau}} \sum_{t=v+1}^{[T\hat{\tau}_T]} \left(\psi(Y_t - \hat{\beta}_{1T}(\hat{\tau}_T)' X_t) X_t - \bar{m}_{1T}(\hat{\tau}_T) \right) \\ &\quad \times \left(\psi(Y_t - \hat{\beta}_{1T}(\hat{\tau}_T)' X_t) X_t - \bar{m}_{1T}(\hat{\tau}_T) \right)' \\ &\quad + \sum_{v=1}^{[T\hat{\tau}_T]-1} k\left(\frac{v}{l(T)}\right) \frac{1}{T\hat{\tau}} \sum_{t=v+1}^{[T\hat{\tau}_T]} \left(\psi(Y_{t-v} - \hat{\beta}_{1T}(\hat{\tau}_T)' X_{t-v}) X_{t-v} - \bar{m}_{1T}(\hat{\tau}_T) \right) \\ &\quad \times \left(\psi(Y_{t-v} - \hat{\beta}_{1T}(\hat{\tau}_T)' X_{t-v}) X_{t-v} - \bar{m}_{1T}(\hat{\tau}_T) \right)', \\ \hat{S}_2 &= \sum_{v=[T\hat{\tau}_T]}^{[T\hat{\tau}_T]-1} k\left(\frac{v}{l(T)}\right) \frac{1}{T - T\hat{\tau}} \sum_{t=v+[T\hat{\tau}_T]+1}^T \left(\psi(Y_{t-v} - \hat{\beta}_{2T}(\hat{\tau}_T)' X_{t-v}) X_{t-v} - \bar{m}_{2T}(\hat{\tau}_T) \right) \\ &\quad \times \left(\psi(Y_{t-v} - \hat{\beta}_{2T}(\hat{\tau}_T)' X_{t-v}) X_{t-v} - \bar{m}_{2T}(\hat{\tau}_T) \right)' \\ &\quad + \sum_{v=[T\hat{\tau}_T]+1}^{T-1} k\left(\frac{v}{l(T)}\right) \frac{1}{T - T\hat{\tau}} \sum_{t=v+[T\hat{\tau}_T]+1}^T \left(\psi(Y_t - \hat{\beta}_{2T}(\hat{\tau}_T)' X_t) X_t - \bar{m}_{2T}(\hat{\tau}_T) \right) \\ &\quad \times \left(\psi(Y_t - \hat{\beta}_{2T}(\hat{\tau}_T)' X_t) X_t - \bar{m}_{2T}(\hat{\tau}_T) \right)', \end{aligned} \tag{29}$$

where $k(\cdot)$ is a real-valued kernel and $l(T)$ is a bandwidth parameter (possibly data-dependent). This class of estimators corresponds to Parzen's (1957) class of kernel estimators of the spectral density matrix at frequency zero of the random vectors $\{\psi(Y_t - \beta'_{10} X_t) X_t, t \leq [T\tau_0]\}$ and $\{\psi(Y_t - \beta'_{20} X_t) X_t, t > [T\tau_0]\}$. It be developed by Andrews (1990), who finds an asymptotically optimal choice of both bandwidth and kernel parameters, and applied by Andrews (1993) in tests for parameter instability under the same weak dependence conditions we are considering here. The consistency property is formalized in the following theorem.

Theorem 4 Assume A.1 and A.2, then

$$\left\| \hat{M} - M \right\| \xrightarrow{P} 0 \quad \text{and} \quad \left\| \hat{S} - S \right\| \xrightarrow{P} 0,$$

where \hat{M} is defined in (23), (24) and (25), and \hat{S} , in (26), (27) and (28), or in (26), (29) and (30), depending on the temporal structure of data generating process.

4.- ON ESTIMATORS WITH BOUNDED INFLUENCE FUNCTION AND HIGH BREAKDOWN POINT

The M-estimators discussed before protects against outlying Y_t , but cannot cope with leverage points, namely outliers in the factor space, which may have large influence on the fit. This could be a problem when we consider stochastic regressors, with a possible contaminated distribution. Therefore, is advisable to consider robust estimators under possible leverage points.

The τ -estimators, defined by Yohai and Zamar (1988), have the following properties: (a) they are qualitatively robust, (b) their breakdown-point is maximum in the presence of contamination in both axes, and (c) they are highly efficient for regression models with normal errors. Asymptotically, a τ -estimator is equivalent to an M-estimator with a score function ψ given by a weighted average of two ψ functions, one corresponding to a very robust estimate and the other to a highly efficient estimate. The weights are adaptative and depend on the underlying error distribution.

The τ -estimators are defined by the minimization of a particular residual scale estimate. Thus, considering the objective function $S_T(\xi)$ in (2), define

$$S_{1T}(\beta_1, \tau) = s_T^2(\beta, \tau) \frac{1}{T} \sum_{t=1}^{[T\tau]} \rho_2 \left(\frac{Y_t - X_t' \beta_1}{s_T(\beta, \tau)} \right), \quad (31)$$

$$S_{2T}(\beta_2, \tau) = s_T^2(\beta, \tau) \frac{1}{T} \sum_{t=\{T\tau\}+1}^T \rho_2 \left(\frac{Y_t - X_t' \beta_2}{s_T(\beta, \tau)} \right), \quad (32)$$

where for $j = 1, 2$, $S_{jT}(\beta, \tau)$ is the τ -scale estimator pre- $[T\tau]$ and post- $[T\tau]$ respectively, and $s_T(\beta, \tau)$ is the scale M-estimator defined by Huber (1981) such that,

$$s_T^2(\beta, \tau) \mid \frac{1}{T} \sum_{t=1}^{[T\tau]} \rho_1 \left(\frac{Y_t - X_t' \beta_1}{s_T(\beta, \tau)} \right) + \frac{1}{T} \sum_{t=\{T\tau\}+1}^T \rho_1 \left(\frac{Y_t - X_t' \beta_2}{s_T(\beta, \tau)} \right) = b, \quad (33)$$

with b , a constant properly defined by:

$$E_{\Phi} [\rho_1 (u_t)] = b,$$

Φ represents the standard normal distribution. For the particular case $\rho_2(u) = u^2$ we obtain the sample standard deviation (least squares), and if $\rho_1 = \rho_2$ we have $S_{nj} = \sqrt{b} s_{nj}$, the corresponding to the S -estimator (Davies, 1990).

The parameter regression estimators $\hat{\beta}_T(\tau) = (\hat{\beta}'_{1T}(\tau), \hat{\beta}'_{2T}(\tau))'$ work out from the first order conditions. Then, for each τ we obtain that,

$$\begin{aligned} \frac{1}{T} \sum_{t=1}^{[T\tau]} \left(W_T(\hat{\beta}_T(\tau)) \psi_1 \left(\frac{Y_t - X_t' \hat{\beta}_{1T}(\tau)}{s_T(\hat{\beta}_T(\tau), \tau)} \right) + \psi_2 \left(\frac{Y_t - X_t' \hat{\beta}_{1T}(\tau)}{s_T(\hat{\beta}_T(\tau), \tau)} \right) \right) X_t &= 0, \\ \frac{1}{T} \sum_{t=[T\tau]+1}^T \left(W_T(\hat{\beta}_T(\tau)) \psi_1 \left(\frac{Y_t - X_t' \hat{\beta}_{2T}(\tau)}{s_T(\hat{\beta}_T(\tau), \tau)} \right) + \psi_2 \left(\frac{Y_t - X_t' \hat{\beta}_{2T}(\tau)}{s_T(\hat{\beta}_T(\tau), \tau)} \right) \right) X_t &= 0, \end{aligned}$$

where the first derivatives are given by Proposition 17 in Appendix 2, considering $\psi_i = \rho'_i$, $i = 1, 2$, and

$$W_T(\hat{\beta}_T(\tau)) = \frac{\frac{1}{T} \sum_{t=1}^{[T\tau]} \tilde{\rho}_2 \left(\frac{Y_t - X_t' \hat{\beta}_{1T}(\tau)}{s_T(\hat{\beta}_T(\tau), \tau)} \right) + \frac{1}{T} \sum_{t=[T\tau]+1}^T \tilde{\rho}_2 \left(\frac{Y_t - X_t' \hat{\beta}_{2T}(\tau)}{s_T(\hat{\beta}_T(\tau), \tau)} \right)}{\frac{1}{T} \sum_{t=1}^{[T\tau]} \tilde{\psi}_1 \left(\frac{Y_t - X_t' \hat{\beta}_{1T}(\tau)}{s_T(\hat{\beta}_T(\tau), \tau)} \right) + \frac{1}{T} \sum_{t=[T\tau]+1}^T \tilde{\psi}_1 \left(\frac{Y_t - X_t' \hat{\beta}_{2T}(\tau)}{s_T(\hat{\beta}_T(\tau), \tau)} \right)}.$$

with $\tilde{\psi}_i(z) = z\psi_i(z)$ and $\tilde{\rho}_i(z) = 2\rho_i(z) - \tilde{\psi}_i(z)$, for $i = 1, 2$. In order to obtain the asymptotic properties for this new estimators, we need to replace assumption A.1 and A.2 by B.1 and B.2 below:

B.1 Let ρ_1 and ρ_2 be two real functions satisfying the same properties of ρ in A.1 and, additionally assume that,

- (i) $\rho_i(0) = 0$.
- (ii) $\rho_i(-u) = \rho_i(u)$.
- (iii) $0 \leq u \leq v$ implies $\rho_i(u) \leq \rho_i(v)$.
- (iv) ρ is even and twice continuously differentiable.
- (v) Let $a = \sup \rho_i(u)$, then $0 < a < \infty$. There exist a constant m such that $\rho_i(u)$ is constant for $|u| > m$.

(vi) If $\rho(u) < a$ and $0 \leq u < v$, then $\rho_i(u) < \rho_i(v)$.

for $i = 1, 2$, and ρ_2 also satisfies that,

(vii) $2\rho_2(u) - \psi_2(u)u \geq 0$.

B.2 Assume that $Z_t = (Y_t, X_t)$ is an *i.i.d.* random vector distributed as $Z = (Y, X) \in \mathbb{R}^{p+1}$, such that,

$$F(Y, X) = H(X)G(Y - \beta'_0 X),$$

and it holds the following conditions:

$$E \left[\|X_t X_t'\|^{1+\epsilon} \right] < \infty, \quad (34)$$

$$E \left[\sup_{\theta \in \Theta} \left| \dot{\psi}_j \left(\frac{u_t + \theta' X_t}{s(\theta)} \right) \left(\frac{u_t + \theta' X_t}{s(\theta)} \right)^2 \right|^{1+\epsilon} \right] < \infty. \quad (35)$$

Assumption A.1 is standard in the context of τ -estimators (see Yohai and Zamar (1988), Assumption 1), and, for the sake of presentation, we consider the *i.i.d.* case, without loss of generality. In the next theorem we obtain the asymptotic behavior for this type of estimators.

Theorem 5 Under B.1 and B.2, we obtain that Theorems 2 and 3 holds for the τ -estimators, and the asymptotic distribution is given by (16), (17) and (18) such that

$$S = E \left[\frac{X_t X_t'}{\sigma_0^2} \right]^{-1} E \left[\psi_0^2 \left(\frac{u_t}{\sigma_0} \right) \right],$$

$$M(\lambda_0) = E \left[\dot{\psi}_0 \left(\frac{u_t + \lambda'_0 X_t}{\sigma_0} \right) \right] E[X_t X_t'] + (1 - \tau_0) \frac{E \left[2\dot{\psi}_2 \left(\frac{u_t}{\sigma_0} \right) \left(\frac{u_t}{\sigma_0} \right) \right]}{E \left[\psi_1 \left(\frac{u_t}{\sigma_0} \right) \left(\frac{u_t}{\sigma_0} \right) \right]} E \left[\psi_1 \left(\frac{u_t + \lambda'_0 X_t}{\sigma_0} \right) X_t \right],$$

with

$$\begin{aligned} \psi_0 \left(\frac{u}{\sigma_0} \right) &= W_0 \psi_1 \left(\frac{u}{\sigma_0} \right) + \psi_2 \left(\frac{u}{\sigma_0} \right), \\ W_0 &= \frac{2E_{F_0} \left[\rho_2 \left(\frac{u}{\sigma_0} \right) \right] - E_{F_0} \left[\psi_2 \left(\frac{u}{\sigma_0} \right) \left(\frac{u}{\sigma_0} \right) \right]}{E_{F_0} \left[\psi_1 \left(\frac{u}{\sigma_0} \right) \left(\frac{u}{\sigma_0} \right) \right]}, \end{aligned} \quad (36)$$

and σ_0 is defined by

$$E_{F_0} \left[\rho_1 \left(\frac{u}{\sigma_0} \right) \right] = b.$$

Finally the matrix M is given by $M(\lambda_0)$ for $\lambda_0 = 0_p$, and therefore,

$$M = E \left[\dot{\psi}_0 \left(\frac{u_t}{\sigma_0} \right) \right] E [X_t X_t'] .$$

The previous theorem allows us to consider the τ -estimator as an M-estimator type with a particular score function $\psi = \psi_0$, given by (36).

6.- MONTE CARLO EXPERIMENTS

In this section, simulations are performed in order to verify some theoretical properties of the change point estimator for a finite sample situation. First, we compare, in terms of bias and mean square error (MSE), the performance of the Least Squares (LS), Least Absolute Deviations (LAD) and Huber estimators under different distribution scenarios for the error term. We also compare the sample and asymptotic distribution of the location estimators, obtaining the proportion of times that the estimate is outside the asymptotic confidence intervals at different confidence levels. Computational costs have prevented us for including the τ -estimators in this simulation study. Nevertheless, we perform a small exercise to illustrate Section 4, generating a model with structural change, and analyzing the sensitivity of the different location estimators in the presence of a solely outlier in both, OX and OY axes.

6.1. Simulation results

Data are generated according to the following model:

$$Y_t = 1 + X_t + \lambda_T I(\tau_t > \tau_0) + u_t, \quad t = 1, \dots, T$$

where $\tau_0 = 0.5$, $x_t \sim i.i.d.N(0, 1)$ and $u_t \sim i.i.d.F(u)$, with F generated as an standard normal, double exponential, t_3 , t_5 , $0.9N(0, 1) + 0.1N(0, 9)$ and $0.75N(0, 1) + 0.25N(0, 9)$, which will be denoted by $N90$ and $N75$ respectively. The error term is standardized in order to get a variance equal to one in all the distribution cases. The size of the jump λ_T is considered decreasing with the sample size, at a rate such that $T \|\lambda_T\|^2 \rightarrow \infty$. Thus, for $T = 100, 200$ and 500 , the value of $\lambda_T = 2.1892, 2.0$ and 1.79527 respectively (we fixed the value of $\lambda_T = 2$ for $T = 200$, and the other ones were obtained considering the dependence with T). For each type of distribution $F(u)$, 5000 replications are performed

and the LS, LAD and Huber estimators are considered, in order to compare them under these different scenarios. The computed Huber estimator is scale-invariant, considering the MAD as the scale estimator, and the constant c , that defines the estimator, is equal to 1.345, corresponding to the minimax version. The programs are written in FORTRAN90 Double Precision, and the IMSL routines were used for the random number generation.

TABLES 1 AND 2 ABOUT HERE

Respect to the point estimation performance, the results are as expected in terms of MSE. For the standard normal case, LS is the best estimator in all the cases of T . The same occurs for the LAD estimator under the double exponential distribution. Under the mixed and t -distributions, the LS estimator performs comparatively rather badly, and the other ones are very similar. Huber estimator is better than LAD estimator for the distributions t_5 and $N90$ in all cases of T , and the difference between them for the rest of the cases are not meaningful.

We also report the proportion of times that the estimator is outside the asymptotic confidence intervals constructed from the asymptotic distribution (18), where the standard errors are estimated as indicated in Theorem 4. The asymptotic critical values c_α , are equal to 7.69, 11.035 and 19.78 for $\alpha = 0.1, 0.05$ and 0.01 respectively. Inspecting the tables 1 to 3, we can observe that LAD estimator underestimates the tail probabilities in all the distribution scenarios, and this problem is not be solved increasing the sample size. Otherwise, LS and Huber estimators approximate rather well the probabilities, obtaining a good result for the biggest sample size. In order to gain a better understanding about the sample distribution of the break estimator in the whole range, not only in the tails, we estimated nonparametrically its density, considering a bandwidth $a \sim T^{-1/5}$. Although not reported here, we obtained that the fit of the sample densities to the theoretical ones is quite good for all the estimators, and near the sample mean, it looks that the LAD estimators is better for the smaller sample sizes. However for $T = 500$ the behavior of all the estimators is very similar.

Finally, it is worth mentioning that, in this simulation study, we have supposed that the jump size is decreasing with T , and we have compared the estimated probabilities with respect to the asymptotic distribution of the break estimator only for this case, which, as we know, is free-distribution. However, it would be of interest to compare them with the

asymptotic distribution under the assumption of a jump size constant, as we can see in Bai (1994) for LAD estimator. In fact, he concludes that the approximation in this case is better than in the first one.

Although not reported here, we also considered another possible designs, with the possibility of a change in the slope too. The results does not change in a significative way.

6.2. Outlier sensitivity analysis

This exercise has been performed in order to illustrate the sensitivity that particular break estimators has in presence of outliers in both, OY or OX axes. To this end, we generated a sample of the following model:

$$\begin{aligned} Y_t &= 1 + X_t + u_t, & t = 1, \dots, T\tau_0 \\ Y_t &= 2 + X_t + u_t, & t = T\tau_0 + 1, \dots, T \end{aligned}$$

where $x_t \sim i.i.d.N(0, 0.1)$, $u_t \sim i.i.d.N(0, 0.1)$, $\tau_0 = 0.5$ and $T = 30$. Table 3 shows the sample data, which is drawn in Figure 1.

TABLE 3 AND FIGURE 1 ABOUT HERE

For illustrative purposes, as in Hampel (1986), we change the point corresponding to $t = 22$, e.g., $P = (2.84, 22, 0.06)$, to $P_1 = (-0.80, 22, 0.08)$, an outlier in OY axis. Next, the same point was changed to $P_2 = (0, 22, 2.0)$, an outlier in OX axis or leverage point. For this three sample date, we estimate, with LS, LAD, Huber and τ -estimators the break point $\hat{\tau}_T$ and the regression parameters pre- $[T\hat{\tau}_T]$ and post- $[T\hat{\tau}_T]$. This is shown in Table 4.

TABLE 4 ABOUT HERE

The results are as expected. In absence of outliers, all the estimators behaves identically, and the estimates of the break are situated in the true break point. In the presence of an outlying observation in the error term, P_1 , the LS estimator results very sensitive, in such a way that the break estimator corresponds to $t = 22$, the outlying observation. The same occurs in presence of a leverage point, P_2 , which affects in the same way to all the estimators except the τ -estimators, the only one which prevents against observations of

this nature. This can be observe clearly in Figures 2, 3 and 4, which draw the dependent variable and the fit respect the variable time in the three data cases.

FIGURES 2, 3 AND 4 ABOUT HERE

REFERENCES:

- Andrews, D.W.K. (1988) Laws of large numbers for dependent non-identically distributed random variables. *Econometric Theory*, 4, 458-467.
- Antoch, J. and M. Husková. (1997) Estimators of changes. *Manuscript*, Charles University (Praha).
- Bai, J. (1995) Least absolute deviation estimation of a shift. *Econometric Theory*, 11, 403-436.
- Ibramigov, I.A.. (1962) Some limit theorems for stationary processes. *Theory of Probability and its Applications*, 7, 349-382.
- Bai, J. (1994) Least squares estimation of a shift in linear processes. *Journal of Time Series Analysis*, 15, 453-472.
- Bhattacharya, P.K. (1987) Maximum likelihood estimation of a change-point in the distribution of independent random variables, general multiparameter case. *Journal of Multivariate Analysis*, 23, 183-208.
- Bierens, H. J. (1981) Robust methods and asymptotic theory. *Lecture Notes in Economics and Mathematical System*, 192. Berlin: Springer-Verlag.
- Billingsley, P. (1968) *Convergence of Probability Measures*. New York: Wiley.
- Bloomfield, P. and W.L. Steiger. (1983) *Least Absolute deviations: Theory, applications and algorithms*, Progress in Probability and Statistics, Huber, P. and M. Rosenblatt eds.
- Brodsky, B.S. and B.E. Darkhovsky. (1993) *Non parametric methods in change-point problems*. Kluwer Academic Publ., Dordrecht.
- Broemeling, L.D. (1982) Introduction. *Journal of Econometrics*, 19, 1-5. (special issue on structural change in econometrics).
- Broemeling, L.D. and H. Tsurumi. (1986) *Econometric and Structural Change*. Marcel Dekker, New York.
- Carlstein, E. (1993) Non parametric estimation of a change-point. *Annals of Statistics*, 16, 188-197.

Carlstein, E. (1993) Non parametric estimation of a change-point. *Annals of Statistics*, 16, 188-197.

Chu, J.S. and C.K. Wu. (1993) Kernel-type estimators of jump points and values of a regression function. *Annals of Statistics*, 21, 1545-1566.

Csörgö M. and L. Horváth. (1986) Non parametric methods for the change-point problem. *Handbook of Statistics*, vol 7, Krishnaiah P.R. and C.R. Rao eds., North Holland, 403-425.

Davies, P.J. (1990) the asymptotics of S-estimates in the linear regression model. *Annals of Statistics*, 18, 1651-1675.

Delgado, M.A. and J. Hidalgo. (1997) Nonparametric inference on structural breaks. *Manuscript*, Department of Econometrics, Universidad Carlos III.

Deshayes, J. and D. Picard. (1986) Off-line statistical analysis of change-point models using non parametric and likelihood methods. *Detection of Abrupt Change in Signals and Dynamical Systems*. M. Basseville and A. Benveniste, eds. Springer-Verlag Lecture Notes in Control and Information Sciences, no. 77, 103-168.

Dümbgen, L. (1991) The asymptotic behavior of some nonparametric change-point estimators. *Annals of Statistics*, 19, 1471-1495.

Feder, P. (1975) The log likelihood ratio in segmented regression. *Annals of Statistics*, 3, 84-97.

Hackl, P. (1989) ed. *Statistical Analysis and Forecasting of Economic Structural Change*. Berlin: Springer-Verlag.

Hackl, P. and A. Westlund (1989) Statistical analysis of 'structural change': an annotated bibliography. *Empirical Economics*, 143, 167-192.

Hackl, P. and A. Westlund (1991) eds. *Economic Structural Change: Analysis and Forecasting*. Springer-Verlag: Berlin, 95-119.

Hampel, F.R. (1986) *Robust statistics: the approach based on influence*. New York: Wiley.

Hinkley, D.V.. (1970) Inference about the change point in a sequence of random variables. *Biometrika*, 57, 1-17.

- Huber, P.J. (1964) Robust estimation of a location parameter. *Annals of Mathematical Statistics*, 35, 73-101.
- Huber, P.J. (1981) *Robust Statistics*, New York: Wiley.
- Husková, M. and P.K. Sen. (1989) Nonparametric tests for shift and change in regression at an unknown time point. *Statistical Analysis and Forecasting of Economic Structural Change*. Hackl ed., Berlin: Springer-Verlag, 73-87.
- Jureková, J. (1993) *Robust statistical procedures: asymptotic and interrelations*. New York: Wiley.
- Krishnaiah, P.R. and B.Q. Miao. (1988) Review about estimation of change points. *Handbook of Statistics*, vol. 7, New York: Elsevier.
- McLeish, D.L. (1975) Invariance principles for dependent variables. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 32, 165-178.
- Parzen, E. (1957) On consistent estimates of the spectrum of a stationary time series. *Annals of Mathematical Statistics*, 28, 329-348.
- Perron, P. (1993) Trend, unit root and structural change in macroeconomic time series. *Manuscript*, Department of Economics, University of Montreal.
- Picard, D. (1985) Testing and estimating change-points in time series. *Advances in Applied Probability*, 17, 841-867.
- Pötscher, P.M. and I.R. Prucha. (1991) Basic structure of the asymptotic theory in dynamic nonlinear econometric models, part I: Consistency and approximation concepts. *Econometric Reviews*, 10, 125-216.
- Stock, J.H. (1986) Unit roots, structural breaks and trends. *Handbook of Econometrics*, vol. 4, chapter 46. Engle R.F. and D.L. McFadden eds., New York: Elsevier.
- Wolfe, D.A. and E. Schechtman. (1984) Non parametric procedures for the change-point problem. *Journal of Statistical Planning and Inference*, vol 9, 389-396.
- Wooldridge, J.M., and H. White. (1988) Some invariance principles and central limit theorems for dependent heterogeneous processes. *Econometric Theory*, 4, 210-230.
- Yao, Y.-C. (1993) Approximating the distribution of the ML estimate of the change-point in a sequence in independent r.v.'s. *Annals of Statistics*, 15,

1321-1328.

Yohai, V.J. (1985) High Breakdown-point and high efficiency robust estimates for regression. *Technical Report*, Department of Statistics, University of Washington, 66.

Yohai, V.J. and R. Zamar (1988) High breakdown-point estimates of regression by means of the minimization of an efficient scale. *Journal of the American Statistical Association*, 402, 406-413.

Zacks, S. (1983) Survey of classical and bayesian approaches to the change-point problem: fixed sample and sequential procedures of testing and estimation. *Recent Advances in Statistics*. M.H. Rizvi et al., eds., New York: Academic Press, 245-269.

APPENDIX 1: PROOF OF THEOREMS.

In order to get more clarity in the proofs the subscript T of the estimators is omitted.

Proof of Theorem 2.

The proof of this theorem is based on Propositions 1 to 16, at the end of this proof, and will be made simultaneously for both cases of λ (fixed and decreasing); we will distinguish only when needed.

We want to prove that

$$\Pr \left\{ \underset{A \cup B \cup C}{Sup} (S_T(\xi_0) - S_T(\xi)) > 0 \right\} < 3\varepsilon. \quad (1)$$

From the definition of $S_T(\xi)$, after some straightforward calculations,

$$S_T(\xi_0) - S_T(\xi) = \frac{1}{T} \sum_{t=1}^{[T\tau_0]} \rho(Y_t - \beta'_{10} X_t) + \frac{1}{T} \sum_{t=[T\tau_0]+1}^T \rho(Y_t - \beta'_{20} X_t) \quad (2)$$

$$- \frac{1}{T} \sum_{t=1}^{[T\tau_0]} \rho(Y_t - \beta'_1 X_t) - \frac{1}{T} \sum_{t=[T\tau_0]+1}^T \rho(Y_t - \beta'_2 X_t) \quad (3)$$

$$+ \frac{1}{T} \sum_{t=[T\tau]+1}^{[T\tau_0]} \rho(Y_t - \beta'_1 X_t) - \frac{1}{T} \sum_{t=[T\tau]+1}^{[T\tau_0]} \rho(Y_t - \beta'_2 X_t), \quad (4)$$

where we have assumed, with no loss of generality, that $\tau \leq \tau_0$ (the $\tau \geq \tau_0$ case is similar and can be proved by identical methods).

Because by A.1.1, ρ is a twice continuously differentiable function, applying the Mean Value Theorem (MVT) up to the second term, we obtain that,

$$\begin{aligned} (2) + (3) &= -(\beta_{10} - \beta_1)' \frac{1}{T} \sum_{t=1}^{[T\tau_0]} \eta_t \\ &\quad - \frac{1}{2} (\beta_{10} - \beta_1)' \frac{1}{T} \sum_{t=1}^{[T\tau_0]} \dot{\eta}_t (\delta_1 (\beta_{10} - \beta_1)) (\beta_{10} - \beta_1) \\ &\quad - (\beta_{20} - \beta_2)' \frac{1}{T} \sum_{t=[T\tau_0]+1}^T \eta_t \\ &\quad - \frac{1}{2} (\beta_{20} - \beta_2)' \frac{1}{T} \sum_{t=[T\tau_0]+1}^T \dot{\eta}_t (\delta_2 (\beta_{20} - \beta_2)) (\beta_{20} - \beta_2) \end{aligned}$$

where $0 < \delta_1, \delta_2 < 1$, $\dot{\eta}_t(\theta) = \partial \eta_t(\theta) / \partial \theta'$, and we have used that $[T\tau] = k$ and $[T\tau_0] = k_0$, for notational convenience.

Next, adding and subtracting $\frac{1}{T} \sum_{t=k+1}^{k_0} \rho(u_t)$ in (4) we obtain that

$$(4) = (\beta_{10} - \beta_1)' \frac{1}{T} \sum_{t=k+1}^{k_0} \eta_t \quad (5)$$

$$+ \frac{1}{2} (\beta_{10} - \beta_1)' \frac{1}{T} \sum_{t=k+1}^{k_0} \dot{\eta}_t (\delta_3 (\beta_{10} - \beta_1)) (\beta_{10} - \beta_1) \quad (6)$$

$$- (\beta_{10} - \beta_2)' \frac{1}{T} \sum_{t=k+1}^{k_0} \eta_t$$

$$- \frac{1}{2} (\beta_{10} - \beta_2)' \frac{1}{T} \sum_{t=k+1}^{k_0} \dot{\eta}_t (\delta_4 (\beta_{10} - \beta_2)) (\beta_{10} - \beta_2),$$

by the MVT up the second term, where $0 < \delta_3, \delta_4 < 1$. Observe that $(\beta_{10} - \beta_2) = \lambda + (\beta_{20} - \beta_2)$, so the previous expression becomes:

$$(4) = (5) + (6)$$

$$- \lambda' \frac{1}{T} \sum_{t=k+1}^{k_0} \eta_t - (\beta_{20} - \beta_2)' \frac{1}{T} \sum_{t=k+1}^{k_0} \eta_t$$

$$- \frac{1}{2} \lambda' \frac{1}{T} \sum_{t=k+1}^{k_0} \dot{\eta}_t (\delta_4 (\beta_{10} - \beta_2)) \lambda$$

$$- \lambda' \frac{1}{T} \sum_{t=k+1}^{k_0} \dot{\eta}_t (\delta_4 (\beta_{10} - \beta_2)) (\beta_{20} - \beta_2)$$

$$- \frac{1}{2} (\beta_{20} - \beta_2)' \frac{1}{T} \sum_{t=k+1}^{k_0} \dot{\eta}_t (\delta_4 (\beta_{10} - \beta_2)) (\beta_{20} - \beta_2).$$

Gathering (2), (3) and (4) and arranging the terms, we obtain that,

$$S_T(\xi_0) - S_T(\xi) = -\frac{1}{2} (\beta_{10} - \beta_1)' M_{1t} (\beta_{10} - \beta_1) - (\beta_{10} - \beta_1)' N_{1t} \quad (7)$$

$$- \frac{1}{2} (\beta_{20} - \beta_2)' M_{2t} (\beta_{20} - \beta_2) - (\beta_{20} - \beta_2)' N_{2t} \quad (8)$$

$$- 2 \left(\frac{1}{2} \lambda' + \frac{1}{2} (\beta_{20} - \beta_2)' \right) M_{4t} \left(\frac{1}{2} \lambda + \frac{1}{2} (\beta_{20} - \beta_2) \right) \quad (9)$$

$$- (\lambda' + (\beta_{20} - \beta_2)') N_{4t}. \quad (10)$$

where

$$M_{1t} = \frac{1}{T} \sum_{t=1}^k \dot{\eta}_t (\theta_1), \quad N_{1t} = \frac{1}{T} \sum_{t=1}^k \eta_t,$$

$$M_{2t} = \frac{1}{T} \sum_{t=k_0+1}^T \dot{\eta}_t (\theta_2), \quad N_{2t} = \frac{1}{T} \sum_{t=k_0+1}^T \eta_t, \quad (11)$$

$$M_{4t} = \frac{1}{T} \sum_{t=k+1}^{k_0} \dot{\eta}_t (\theta_4), \quad N_{4t} = \frac{1}{T} \sum_{t=k+1}^{k_0} \eta_t,$$

and

$$\theta_1 = \delta_1 (\beta_{10} - \beta_1) = \delta_3 (\beta_{10} - \beta_1), \theta_2 = \delta_2 (\beta_{20} - \beta_2) \text{ and } \theta_4 = \delta_4 (\beta_{10} - \beta_2). \quad (12)$$

Finally, observe that the parameter space,

$$\begin{aligned} A \cup B \cup D &= [(A \cup B) \cap D] \cup [(A \cup B) \cap \bar{D}] \cup [(\overline{A \cup B}) \cap D] \\ &= G_1 \cup G_2 \cup G_3, \end{aligned} \quad (13)$$

a union of disjoint sets, where \bar{E} denotes the complementary set of E . The left side of (1) is upper bounded by

$$\Pr \left\{ \underset{G_1}{Sup} (S_T(\xi_0) - S_T(\xi)) > 0 \right\} + \Pr \left\{ \underset{G_2}{Sup} (S_T(\xi_0) - S_T(\xi)) > 0 \right\} + \Pr \left\{ \underset{G_3}{Sup} (S_T(\xi_0) - S_T(\xi)) > 0 \right\},$$

with $S_T(\xi_0) - S_T(\xi)$ defined in (7) – (10). Thus, it suffices to show that each of the terms of the above expression is bounded by ε for C large enough. This is obtained below.

1st Case: $G_1 = [(A \cup B) \cap D]$

In order to bound the probability (1) restricted in this parametric set, we need to consider the following possible cases of β_2 :

- $(\beta_{20} - \beta_2) \neq -\lambda$
 - b.1** $F_1 = \left\{ (\beta_{20} - \beta_2) \neq -\lambda, \|\beta_{20} - \beta_2\| < \frac{C}{\sqrt{T}} \right\}$.
 - b.2** $F_2 = \left\{ (\beta_{20} - \beta_2) \neq -\lambda, \|\beta_{20} - \beta_2\| > \frac{C}{\sqrt{T}} \right\}$.
- $(\beta_{20} - \beta_2) = -\lambda$
 - b.3** $F_3 = \left\{ (\beta_{20} - \beta_2) = -\lambda, \|\beta_{20} - \beta_2\| < \frac{C}{\sqrt{T}} \right\}$.
 - b.4** $F_4 = \left\{ (\beta_{20} - \beta_2) = -\lambda, \|\beta_{20} - \beta_2\| > \frac{C}{\sqrt{T}} \right\}$.

Observe that $(A \cup B)$ implies $\|\beta_1 - \beta_{10}\| > \frac{C}{\sqrt{T}}$ and/or $\|\beta_2 - \beta_{20}\| > \frac{C}{\sqrt{T}}$. First, consider $\|\beta_1 - \beta_{10}\| < \frac{C}{\sqrt{T}}$; in this case we have $B \cap D$, because of G_1 . In this parametric subset, the probability (1) will be upper bounded by

$$\Pr \left\{ \underset{A}{Sup} (7) > 0 \right\} + \Pr \left\{ \underset{B \cap D}{Sup} ((8) + (9) + (10)) > 0 \right\},$$

where the first term is asymptotically negligible by Proposition 4. Because $B = F_2 \cup F_4$, an union of disjoint sets, an upper bound for the second term is given by:

$$\Pr \left\{ \underset{B}{Sup} (8) > 0 \right\} + \Pr \left\{ \underset{F_2 \cap D}{Sup} (9) + (10) > 0 \right\} + \Pr \left\{ \underset{F_4 \cap D}{Sup} (9) + (10) > 0 \right\}, \quad (14)$$

which tends to zero for large enough C , because of each of its terms are asymptotically negligible by Propositions 7, 9 and 11 respectively..

Next, we consider $\|\beta_1 - \beta_{10}\| > \frac{C}{\sqrt{T}}$, so that may happen $B \cap D$ or $\bar{B} \cap D$. First, let be $A \cap B \cap D$, a subset of G_1 for which the probability (1) will be less or equal than:

$$\Pr \left\{ Sup_A (7) > 0 \right\} + \Pr \left\{ Sup_{B \cap D} (8) + (9) + (10) > 0 \right\}.$$

By Proposition 5, the first term converges to zero for large enough C , similar as the second one, bounded by the expression (14) as in the previous case. .

Finally, it remains the subset $A \cap \bar{B} \cap D$. Observe that $\bar{B} = F_1 \cup F_3$, where $F_1 \cap F_3 = \emptyset$, and then (1) is bounded by

$$\Pr \left\{ Sup_A (7) > 0 \right\} + \Pr \left\{ Sup_{\bar{B}} (8) > 0 \right\} + \Pr \left\{ Sup_{F_1 \cap D} (9) + (10) > 0 \right\} + \Pr \left\{ Sup_{F_3 \cap D} (9) + (10) > 0 \right\},$$

which converges to zero by Propositions 5, 6, 8 and 13 respectively.

$$2^{nd} \text{ Case: } G_2 = [(A \cup B) \cap \bar{D}]$$

Proceeding as above we have that if \bar{A} , then $B \cap \bar{D}$. In this case we obtain that, in this subset, (1) is bounded by

$$\Pr \left\{ Sup_{\bar{A}} (7) > 0 \right\} + \Pr \left\{ Sup_{B \cap \bar{D}} (8) + (9) + (10) > 0 \right\},$$

which converges to zero by Propositions 4 and 15.

Next, if A then we have $B \cap \bar{D}$ or $\bar{B} \cap \bar{D}$. In the first case (1) is bounded by

$$\Pr \left\{ Sup_A (7) > 0 \right\} + \Pr \left\{ Sup_{B \cap \bar{D}} (8) + (9) + (10) > 0 \right\},$$

and each of the terms converge to zero by Propositions 5 and 15 respectively. In the second case, $A \cap \bar{B} \cap \bar{D}$, (1) will be less or equal to,

$$\Pr \left\{ Sup_A (7) > 0 \right\} + \Pr \left\{ Sup_{\bar{B}} (8) > 0 \right\} + \Pr \left\{ Sup_{\bar{B} \cap \bar{D}} (9) + (10) > 0 \right\},$$

which tends to zero by Propositions 5, 6 and 16.

$$3^{rd} \text{ Case: } G_3 = [(\bar{A} \cap \bar{B}) \cap D]$$

As in the first case, observe that $\bar{B} = F_1 \cup F_3$, where $F_1 \cap F_3 = \emptyset$, and then (1) is bounded by

$$\Pr \left\{ Sup_{\bar{A}} (7) > 0 \right\} + \Pr \left\{ Sup_{\bar{B}} (8) > 0 \right\} + \Pr \left\{ Sup_{F_1 \cap D} (9) + (10) > 0 \right\} + \Pr \left\{ Sup_{F_3 \cap D} (9) + (10) > 0 \right\},$$

which converges to zero by Propositions 4, 6, 8 and 13 respectively.

We conclude the proof of the theorem. ■

Propositions 1 to 16 below are used for the proof of Theorem 2, and they are obtained under the Assumptions A.1 and A.2. We assume, with no loss of generality, that $\tau \leq \tau_0$.

Proposition 1

$$\begin{aligned} M_{1t} &= \tau M(\theta_1) + op(1), \\ M_{2t} &= (1 - \tau_0) M(\theta_2) + op(1), \\ \frac{T}{(k_0 - k)} M_{4t} &= M(\theta_4) + op(1). \end{aligned}$$

where $M(\theta)$ is defined by the Assumption A.2.6, and, for $i = 1, 2$ and 4, M_{it} and θ_i given in (11) and (12).

Proof: By Lemma 1.

Proposition 2 $\forall \alpha \in R^p$,

$$\begin{aligned} \sqrt{T} \alpha' N_{1t} &\Rightarrow (\alpha' S \alpha)^{1/2} B(\tau), \\ \sqrt{T} \alpha' N_{2t} &\Rightarrow (\alpha' S \alpha)^{1/2} B(1 - \tau_0), \\ \frac{T}{(k_0 - k)^{1/2}} \alpha' N_{4t} &\Rightarrow (\alpha' S \alpha)^{1/2} B(1). \end{aligned}$$

where, N_{it} is defined in (11) for $i = 1, 2$ and 4, and $B(\cdot)$ is a Brownian motion process on $(0, 1)$.

Proof: By Lemma 3.

Proposition 3 Let $M(\theta)_{p \times p}$, be the finite and positive definite matrices, by assumption A.2.6, with eigenvalues $a_j, j = 1, \dots, p$. Then, $\forall \gamma \in R^p$:

$$\gamma' M(\theta) \gamma \geq q \|\gamma\|^2 > 0,$$

where $q = \min_j \{a_j\}$.

Proof: By the Courant Theorem,

$$\gamma' M(\theta) \gamma \geq \inf_{\gamma} \frac{\gamma' M(\theta) \gamma}{\gamma' \gamma} \gamma' \gamma \geq \inf_{\gamma} \frac{\gamma' M(\theta) \gamma}{\gamma' \gamma} \inf_{\gamma} \gamma' \gamma = \min_j \{a_j\} \|\gamma\| = q \|\gamma\|^2 > 0.$$

Proposition 4 Define the event

$$\{E_1\} = \left\{ \sup_{\bar{A}} \left[-\frac{1}{2} (\beta_{10} - \beta_1)' M_{1t} (\beta_{10} - \beta_1) - (\beta_{10} - \beta_1)' N_{1t} \right] > 0 \right\},$$

then it holds that $\Pr \{E_1\} < \varepsilon$ for C large enough.

Proof: For $\beta_1 \in \bar{A}$,

$$\begin{aligned} \sup_{\bar{A}} [(\beta_{10} - \beta_1)' M_{1t} (\beta_{10} - \beta_1)] &= Op(T^{-1}), \\ \sup_{\bar{A}} [(\beta_{10} - \beta_1)' N_{1t}] &= Op(T^{-1}), \end{aligned}$$

by Propositions 1 and 2. Thus, by definition, if $X_t = Op(T^{-1})$, then for every $\varepsilon > 0$, exists a constant $C > 0$ such that $\Pr\{|X_t| > T^{-1}C\} < \varepsilon$, therefore, the result follows for all C . \blacksquare

Proposition 5 *Define the event*

$$\{E_2\} = \left\{ \mathit{Sup}_A \left[-\frac{1}{2} (\beta_{10} - \beta_1)' M_{1t} (\beta_{10} - \beta_1) - (\beta_{10} - \beta_1)' N_{1t} \right] > 0 \right\},$$

then it holds that $\Pr\{E_2\} < \varepsilon$ for C large enough.

Proof: Observe that $\{E_2\}$ implies the event

$$\left\{ \mathit{Sup}_A \left[\frac{-\frac{1}{2} (\beta_{10} - \beta_1)' M_{1t} (\beta_{10} - \beta_1) - (\beta_{10} - \beta_1)' N_{1t}}{\|\beta_{10} - \beta_1\|} \right] > 0 \right\}, \quad (15)$$

and given that $(\beta_{10} - \beta_1)' M_{1t} (\beta_{10} - \beta_1)$ (15) will imply that,

$$\begin{aligned} & \left\{ \mathit{Sup}_A \frac{|(\beta_{10} - \beta_1)' \sqrt{T} N_{1t}|}{\|\beta_{10} - \beta_1\|} > \mathit{Inf}_A \frac{\sqrt{T} |(\beta_{10} - \beta_1)' M_{1t} (\beta_{10} - \beta_1)|}{\|\beta_{10} - \beta_1\|} \right\} \\ \equiv & \left\{ \mathit{Sup}_A \left| \frac{(\beta_{10} - \beta_1)' \sqrt{T} N_{1t}}{\|\beta_{10} - \beta_1\|} \right| > \mathit{Inf}_A \sqrt{T} \left(\frac{\|(M(\theta_1) + op(1))^{1/2} (\beta_{10} - \beta_1)\|^2}{\|\beta_{10} - \beta_1\|} \right) \right\} \end{aligned} \quad (16)$$

by Proposition 1. The symbol " \equiv " indicates equivalence of events in probability. Using Proposition 3 we obtain that:

$$\frac{\|M(\theta_1)^{1/2} (\beta_{10} - \beta_1)\|^2}{\|\beta_{10} - \beta_1\|} \geq \frac{q_1 \|\beta_{10} - \beta_1\|^2}{\|\beta_{10} - \beta_1\|} \geq q_1 \|\beta_{10} - \beta_1\|,$$

where q_1 is the least eigenvalue of the matrix $M(\theta_1)$. Thus, (16) implies the event:

$$\left\{ \mathit{Sup}_A \left| \frac{(\beta_{10} - \beta_1)' \sqrt{T} N_{1t}}{\|\beta_{10} - \beta_1\|} \right| > \mathit{Inf}_A (q_1 + op(1)) \sqrt{T} \|\beta_{10} - \beta_1\| \right\},$$

and then, we obtain that $\Pr\{E_1\}$ is upper bounded by

$$\Pr \left\{ \mathit{Sup}_A \left| \frac{(\beta_{10} - \beta_1)' \sqrt{T} N_{1t}}{\|\beta_{10} - \beta_1\|} \right| > q_1 C \right\} = O\left(\frac{1}{C^2}\right),$$

by Lemma 4. From here the conclusion is standard. \blacksquare

Proposition 6 *Define the event*

$$\{E_3\} = \left\{ \mathit{Sup}_B \left[-\frac{1}{2} (\beta_{20} - \beta_2)' M_{2t} (\beta_{20} - \beta_2) - (\beta_{20} - \beta_2)' N_{2t} \right] > 0 \right\},$$

the it holds that $\Pr\{E_3\} < \varepsilon$ for large enough C .

Proof: By Propositions 1 and 2, the result follows as in Proposition 4. \blacksquare

Proposition 7 Define the event,

$$\{E_4\} = \left\{ \text{Sup}_B \left[-\frac{1}{2} (\beta_{20} - \beta_2)' M_{2t} (\beta_{20} - \beta_2) - (\beta_{20} - \beta_2)' N_{2t} \right] > 0 \right\},$$

then it holds that $\Pr \{E_4\} < \varepsilon$ for C large enough.

Proof: By Propositions 1 and 2, and Lemma 5, the result follows as in Proposition 5. \blacksquare

Proposition 8 Define the event

$$\{E_5\} = \left\{ \text{Sup}_{F_1 \cap D} \left[-2 \left(\frac{1}{2} \lambda' + \frac{1}{2} (\beta_{20} - \beta_2)' \right) M_{4t} \left(\frac{1}{2} \lambda + \frac{1}{2} (\beta_{20} - \beta_2) \right) - (\lambda' + (\beta_{20} - \beta_2)') N_{4t} \right] > 0 \right\},$$

then it holds that $\Pr \{E_5\} < \varepsilon$ for C large enough.

Proof: Observe that, considering the subset F_1 , the event $\{E_5\}$ implies that

$$\left\{ \text{Sup}_{F_1 \cap D} \frac{T \left[-2 \left(\frac{1}{2} \lambda' + \frac{1}{2} (\beta_{20} - \beta_2)' \right) M_{4t} \left(\frac{1}{2} \lambda + \frac{1}{2} (\beta_{20} - \beta_2) \right) - (\lambda' + (\beta_{20} - \beta_2)') N_{4t} \right]}{\|\lambda\| (k_0 - k)^{1/2}} > 0 \right\},$$

and therefore,

$$\left\{ \text{Sup}_{F_1 \cap D} \left| (\lambda' + (\beta_{20} - \beta_2)') \frac{TN_{4t}}{\|\lambda\| (k_0 - k)^{1/2}} \right| > 2 \left(\frac{1}{2} \lambda' + \frac{1}{2} (\beta_{20} - \beta_2)' \right) \frac{TM_{4t}}{\|\lambda\| (k_0 - k)^{1/2}} \left(\frac{1}{2} \lambda + \frac{1}{2} (\beta_{20} - \beta_2) \right) \right\} \quad (17)$$

because the absolute value argument on the right side is strictly positive. Given F_1 ,

$(\beta_{20} - \beta_2) = Op(T^{-1/2})$, and then, (17) will imply that,

$$\begin{aligned} & \left\{ \text{Sup}_D \left| \left(\lambda' + Op(T^{-1/2}) \right) \frac{TN_{4t}}{\|\lambda\| (k_0 - k)^{1/2}} \right| \right. \\ & \left. > \text{Inf}_D \left| 2 \left(\frac{1}{2} \lambda' + \frac{1}{2} Op(T^{-1/2}) \right) \frac{(k_0 - k)^{1/2}}{\|\lambda\|} \frac{TM_{4t}}{(k_0 - k)} \left(\frac{1}{2} \lambda' + \frac{1}{2} Op(T^{-1/2}) \right) \right| \right\} \\ \Rightarrow & \left\{ \text{Sup}_D \left| \frac{\lambda' + Op(T^{-1/2})}{\|\lambda\|} \frac{T}{(k_0 - k)^{1/2}} N_{4t} \right| > \text{Inf}_D \left| 2 \frac{(k_0 - k)^{1/2}}{\|\lambda\|} \left\| (M(\theta_4) + op(1))^{1/2} \left(\frac{1}{2} \lambda' + Op(T^{-1/2}) \right) \right\| \right| \right\} \end{aligned}$$

by Proposition 1. In this context " \Rightarrow " indicates implication of events. Let q_4 be the least eigenvalue of $M(\theta_4)$, finite and positive definite, therefore,

$$\left\| M(\theta_4)^{1/2} \left(\frac{1}{2} \lambda' + \frac{1}{2} Op(T^{-1/2}) \right) \right\|^2 \geq \frac{q_4 (\|\lambda\|^2 + Op(T^{-1}))}{4} \geq \frac{q_4}{4} (\|\lambda\|^2 + Op(T^{-1})),$$

by Proposition 3, so that (18) implies the event

$$\left\{ \text{Sup}_D \left| \left(\frac{\lambda'}{\|\lambda\|} + Op \left(\frac{1}{T^{1/2} \|\lambda\|} \right) \right) \frac{T}{(k_0 - k)^{1/2}} N_{4t} \right| > \left| \frac{C^{1/2} q_4}{\|\lambda\|^2} \frac{q_4}{2} (\|\lambda\|^2 + Op(T^{-1})) \right| \right\}$$

$$\Rightarrow \left\{ S_{D}^{up} \left| \left(\frac{\lambda'}{\|\lambda\|} + op(1) \right) \frac{T}{(k_0 - k)^{1/2}} N_{4t} \right| > \left| C^{1/2} \frac{q_4}{2} (1 + op(1)) \right| \right\}.$$

From this result, we obtain that $\Pr\{E_5\}$ is upper bounded by

$$\Pr \left\{ S_{D}^{up} \left| \frac{\lambda'}{\|\lambda\|} \frac{T}{(k_0 - k)^{1/2}} N_{4t} \right| > \frac{q_4}{2} C^{1/2} \right\} = O\left(\frac{1}{C}\right),$$

by Lemma 4. Standard arguments implies the result. \blacksquare

Proposition 9 Define the event

$$\{E_6\} = \left\{ S_{F_2 \cap D}^{up} \left[-2 \left(\frac{1}{2} \lambda' + \frac{1}{2} (\beta_{20} - \beta_2)' \right) M_{4t} \left(\frac{1}{2} \lambda + \frac{1}{2} (\beta_{20} - \beta_2) \right) - (\lambda' + (\beta_{20} - \beta_2)') N_{4t} \right] \right\},$$

then it holds that $\Pr\{E_6\} < \varepsilon$ for C large enough.

Proof: Observe that $\{E_6\}$ implies the event

$$\left\{ S_{F_2 \cap D}^{up} \left[\frac{T}{(k_0 - k)^{1/2}} \frac{-2 \left(\frac{1}{2} \lambda' + \frac{1}{2} (\beta_{20} - \beta_2)' \right) M_{4t} \left(\frac{1}{2} \lambda + \frac{1}{2} (\beta_{20} - \beta_2) \right) - (\lambda' + (\beta_{20} - \beta_2)') N_{4t}}{\|\lambda\| + \|\beta_{20} - \beta_2\|} \right] > 0 \right\},$$

and therefore, will imply that,

$$\left\{ S_{F_2 \cap D}^{up} \left| \frac{\lambda' + (\beta_{20} - \beta_2)'}{\|\lambda\| + \|\beta_{20} - \beta_2\|} \frac{T}{(k_0 - k)^{1/2}} N_{4t} \right| > \inf_{F_2 \cap D} \left| \frac{q_4 (k_0 - k)^{1/2} \|\lambda + (\beta_{20} - \beta_2)\|^2}{2 \|\lambda\| + \|\beta_{20} - \beta_2\|} \right| \right\}, \quad (19)$$

by Propositions 1 and 3. Now observe that,

$$\|\lambda + (\beta_{20} - \beta_2)\|^2 = \|\lambda\|^2 + \|\beta_{20} - \beta_2\|^2 + 2\lambda'(\beta_{20} - \beta_2),$$

such that, in F_2 , we must consider two possible cases for β_2 :

a) $(\beta_{20} - \beta_2) > -\lambda$. Given that, $\lambda > 0$, it holds that,

$$2\lambda'(\beta_{20} - \beta_2) > -2\lambda'\lambda = -2\|\lambda\|^2,$$

thus,

$$\|\lambda + (\beta_{20} - \beta_2)\|^2 > \|\beta_{20} - \beta_2\|^2 - \|\lambda\|^2 = (\|\beta_{20} - \beta_2\| + \|\lambda\|)(\|\beta_{20} - \beta_2\| - \|\lambda\|).$$

b) $(\beta_{20} - \beta_2) < -\lambda$. Given that $(\beta_{20} - \beta_2) < 0$, we obtain:

$$2\lambda'(\beta_{20} - \beta_2) > -2(\beta_{20} - \beta_2)'(\beta_{20} - \beta_2) = -2\|\beta_{20} - \beta_2\|^2,$$

therefore,

$$\|\lambda + (\beta_{20} - \beta_2)\|^2 > \|\lambda\|^2 - \|\beta_{20} - \beta_2\|^2 = (\|\lambda\| + \|\beta_{20} - \beta_2\|)(\|\lambda\| - \|\beta_{20} - \beta_2\|).$$

So, considering first the case a), we obtain that (19) implies the event

$$\begin{aligned} & \left\{ \sup_{F_2 \cap D} \left| \frac{\lambda' + (\beta_{20} - \beta_2)'}{\|\lambda\| + \|\beta_{20} - \beta_2\|} \frac{T}{(k_0 - k)^{1/2}} N_{4t} \right| > \inf_{F_2 \cap D} \left| \frac{q_4 (k_0 - k)^{1/2}}{2} (\|\beta_{20} - \beta_2\| - \|\lambda\|) \right| \right\} \\ \Rightarrow & \left\{ \sup_{F_2 \cap D} \left| \frac{\lambda' + (\beta_{20} - \beta_2)'}{\|\lambda\| + \|\beta_{20} - \beta_2\|} \frac{T}{(k_0 - k)^{1/2}} N_{4t} \right| > \inf_{F_2 \cap D} \left| \frac{q_4 C^{1/2}}{2 \|\lambda\|} \left(\frac{C^{1/2}}{\sqrt{T}} - \|\lambda\| \right) \right| \right\}. \end{aligned}$$

With this result, we obtain that $\Pr \{E_6\}$ will be less than or equal to,

$$\begin{aligned} & \Pr \left\{ \sup_{F_2 \cap D} \left| \frac{\lambda' + (\beta_{20} - \beta_2)'}{\|\lambda\| + \|\beta_{20} - \beta_2\|} \frac{T}{(k_0 - k)^{1/2}} N_{4t} \right| > \inf_{F_2 \cap D} \left| \frac{q_4 C^{1/2}}{2 \|\lambda\|} \left(\frac{C^{1/2}}{\sqrt{T}} - \|\lambda\| \right) \right| \right\}, \\ & = O \left(\frac{1}{\left(\frac{C^{1/2}}{\sqrt{T}} - \|\lambda\| \right)^2 \frac{C}{\|\lambda\|^2}} \right) = O \left(\frac{1}{\left(\frac{C^{1/2}}{\sqrt{T} \|\lambda\|} - 1 \right)^2 C} \right), \end{aligned}$$

by Lemma 5, and thus, $o(1)$ for arbitrary large C . Following equivalent arguments for the case b), we can also obtain an upper bound of the probability of interest, let be

$$\begin{aligned} & \Pr \left\{ \sup_{F_2 \cap D} \left| \frac{\lambda' + (\beta_{20} - \beta_2)'}{\|\lambda\| + \|\beta_{20} - \beta_2\|} \frac{T}{(k_0 - k)^{1/2}} N_{4t} \right| > \inf_{F_2 \cap D} \left| \frac{q_4 C^{1/2}}{2 \|\lambda\|} \left(\|\lambda\| - \frac{C^{1/2}}{\sqrt{T}} \right) \right| \right\} \\ & = O \left(\frac{1}{\left(\|\lambda\| - \frac{C^{1/2}}{\sqrt{T}} \right)^2 \frac{C}{\|\lambda\|^2}} \right) = O \left(\frac{1}{\left(1 - \frac{C^{1/2}}{\sqrt{T} \|\lambda\|} \right)^2 C} \right) \end{aligned}$$

by Lemma 5. The result follows by standard arguments. \blacksquare

Proposition 10 Define the event

$$\{E_7\} = \left\{ \sup_D \left[\frac{1}{2} \lambda' M_{4t} \lambda - \lambda' N_{4t} \right] > 0 \right\},$$

then it holds that $\Pr \{E_7\} < \varepsilon$ for C large enough.

Proof: The above event implies that:

$$\left\{ \sup_D \left[\frac{T \left(\frac{1}{2} \lambda' M_{4t} \lambda - \lambda' N_{4t} \right)}{\|\lambda\| (k - k_0)^{1/2}} \right] > 0 \right\},$$

and, since $\lambda' M_{4t} \lambda > 0$, will imply

$$\begin{aligned} & \left\{ \sup_D \left| \frac{\lambda'}{\|\lambda\|} \frac{T}{(k - k_0)^{1/2}} N_{4t} \right| > \inf_D \left| \frac{1}{2} \lambda' \frac{(k_0 - k)^{1/2}}{\|\lambda\|} \frac{T}{(k_0 - k)} M_{4t} \lambda \right| \right\} \\ \Rightarrow & \left\{ \sup_D \left| \frac{\lambda'}{\|\lambda\|} \frac{T}{(k - k_0)^{1/2}} N_{4t} \right| > C^{1/2} (q_4 + op(1)) \right\}, \end{aligned}$$

and thus, we obtain that $\Pr \{E_7\}$ is less than or equal to

$$\Pr \left\{ \sup_D \left| \frac{\lambda'}{\|\lambda\|} \frac{T}{(k - k_0)^{1/2}} N_{4t} \right| > C^{1/2} q_4 \right\} = O \left(\frac{1}{C} \right),$$

which tends to zero for large enough C . Lemma 5 and Propositions 1 and 3 has been used. \blacksquare

Proposition 11 Define the event

$$\{E_8\} = \left\{ \text{Sup}_{F_3 \cap D} \left[-\frac{1}{2} (\beta_{20} - \beta_2)' M_{4t} (\beta_{20} - \beta_2) - (\beta_{20} - \beta_2)' M_{4t} \lambda - (\beta_{20} - \beta_2)' N_{4t} \right] > 0 \right\},$$

then it holds that $\Pr \{E_8\} < \varepsilon$ for C large enough.

Proof: $\{E_8\}$ imply the event

$$\left\{ \text{Sup}_{F_3 \cap D} \left[\frac{T}{(k - k_0)^{1/2}} \frac{-\frac{1}{2} (\beta_{20} - \beta_2)' M_{4t} (\beta_{20} - \beta_2) - (\beta_{20} - \beta_2)' M_{4t} \lambda - (\beta_{20} - \beta_2)' N_{4t}}{\|\beta_{20} - \beta_2\|} \right] > 0 \right\} \quad (20)$$

Observe that $(\beta_{20} - \beta_2) = -\lambda$, and therefore:

$$(\beta_{20} - \beta_2)' M_{4t} \lambda = -(\beta_{20} - \beta_2)' M_{4t} (\beta_{20} - \beta_2),$$

strictly negative. Thus, (20) will imply that:

$$\begin{aligned} & \left\{ \text{Sup}_{F_3 \cap D} \left| \frac{(\beta_{20} - \beta_2)' T N_{4t}}{\|\beta_{20} - \beta_2\| (k - k_0)^{1/2}} \right| \right. \\ & > \text{Inf}_{F_3 \cap D} \left. \left| \frac{1}{2} (\beta_{20} - \beta_2)' \frac{T M_{4t}}{(k - k_0)} (\beta_{20} - \beta_2) - (\beta_{20} - \beta_2)' \frac{T M_{4t}}{(k - k_0)} \lambda \right| \frac{(k - k_0)^{1/2}}{\|\beta_{20} - \beta_2\|} \right\} \\ \Rightarrow & \left\{ \text{Sup}_{F_3 \cap D} \left| \frac{(\beta_{20} - \beta_2)' T N_{4t}}{\|\beta_{20} - \beta_2\| (k - k_0)^{1/2}} \right| > \text{Inf}_{F_3 \cap D} \left| \frac{(q_4 + op(1)) \|\beta_{20} - \beta_2\|^2}{\|\beta_{20} - \beta_2\|} \frac{C^{1/2}}{\|\lambda\|} \right| \right\}, \end{aligned}$$

therefore, the probability of the event $\{E_8\}$ will be upper bounded by

$$\Pr \left\{ \text{Sup}_{F_3 \cap D} \left| \frac{(\beta_{20} - \beta_2)' T N_{4t}}{\|\beta_{20} - \beta_2\| (k - k_0)^{1/2}} \right| > q_4 C^{1/2} \right\} = O \left(\frac{1}{C} \right)$$

by Lemma 5, and thus, tends to zero for arbitrary large C . \blacksquare

Proposition 12 Define the event

$$\{E_9\} = \left\{ \text{Sup}_{F_4 \cap D} \left[-\frac{1}{2} (\beta_{20} - \beta_2)' M_{4t} (\beta_{20} - \beta_2) - (\beta_{20} - \beta_2)' M_{4t} \lambda - (\beta_{20} - \beta_2)' N_{4t} \right] \right\},$$

then it holds that $\Pr \{E_9\} < \varepsilon$ for C large enough .

Proof: The proof is as in Proposition 11, given that the parametric space β_2 belongs to, is not used for the proof. The result is established taking into account only the equality condition respect to λ . \blacksquare

Proposition 13 Define the event $\{E_{10}\}$ as

$$\left\{ \text{Sup}_{F_3 \cap D} \left[-2 \left(\frac{1}{2} \lambda' + \frac{1}{2} (\beta_{20} - \beta_2)' \right) M_{4t} \left(\frac{1}{2} \lambda' + \frac{1}{2} (\beta_{20} - \beta_2)' \right) - (\lambda' + (\beta_{20} - \beta_2)') N_{4t} \right] > 0 \right\},$$

then it holds that $\Pr \{E_{10}\} < \varepsilon$ for large enough C .

Proof: The above probability is less than or equal to:

$$\begin{aligned} & \Pr \left\{ \text{Sup}_D \left[-\frac{1}{2} \lambda' M_{4t} \lambda' - \lambda' N_{4t} \right] > 0 \right\} \\ & + \Pr \left\{ \text{Sup}_{F_3 \cap D} \left[-\frac{1}{2} (\beta_{20} - \beta_2)' M_{4t} (\beta_{20} - \beta_2)' - (\beta_{20} - \beta_2)' M_{4t} \lambda' - (\beta_{20} - \beta_2)' N_{4t} \right] > 0 \right\}, \end{aligned}$$

and the results follows by Propositions 10 and 11. \blacksquare

Proposition 14 Define the event $\{E_{11}\}$ as

$$\left\{ \text{Sup}_{F_4 \cap D} \left[-2 \left(\frac{1}{2} \lambda' + \frac{1}{2} (\beta_{20} - \beta_2)' \right) M_{4t} \left(\frac{1}{2} \lambda' + \frac{1}{2} (\beta_{20} - \beta_2)' \right) - (\lambda' + (\beta_{20} - \beta_2)' N_{4t} \right] > 0 \right\},$$

then it holds that $\Pr \{E_{11}\} < \varepsilon$ for large enough C .

Proof: Similar to Proposition 13, an upper bound of this probability is given by:

$$\begin{aligned} & \Pr \left\{ \text{Sup}_D \left[-\frac{1}{2} \lambda' M_{4t} \lambda' - \lambda' N_{4t} \right] > 0 \right\} \\ & + \Pr \left\{ \text{Sup}_{F_4 \cap D} \left[-\frac{1}{2} (\beta_{20} - \beta_2)' M_{4t} (\beta_{20} - \beta_2)' - (\beta_{20} - \beta_2)' M_{4t} \lambda' - (\beta_{20} - \beta_2)' N_{4t} \right] > 0 \right\} \end{aligned}$$

and the result follows by Propositions 10 and 12. \blacksquare

Proposition 15 Define the event $\{E_{12}\}$ as

$$\begin{aligned} & \left\{ \text{Sup}_{B \cap \bar{D}} \left[-(\beta_{20} - \beta_2)' N_{2t} - \frac{1}{2} (\beta_{20} - \beta_2)' M_{2t} (\beta_{20} - \beta_2) \right. \right. \\ & \left. \left. - 2 \left(\frac{1}{2} \lambda' + \frac{1}{2} (\beta_{20} - \beta_2)' \right) M_{4t} \left(\frac{1}{2} \lambda' + \frac{1}{2} (\beta_{20} - \beta_2)' \right) - (\lambda' + (\beta_{20} - \beta_2)' N_{4t} \right] > 0 \right\}, \end{aligned}$$

then it holds that $\Pr \{E_{12}\} < \varepsilon$ for large enough C .

Proof: Observe that $\{E_{12}\}$ implies the event

$$\begin{aligned} & \left\{ \text{Sup}_{B \cap \bar{D}} \left| \frac{(\beta_{20} - \beta_2)' \sqrt{T} N_{2t} + \frac{\lambda' + (\beta_{20} - \beta_2)' \sqrt{T} N_{4t}}{\|\beta_{20} - \beta_2\|}}{\|\beta_{20} - \beta_2\|} \right| \right. \\ & \left. > \text{Inf}_{B \cap \bar{D}} \left| \frac{\sqrt{T} (\beta_{20} - \beta_2)' M_{2t} (\beta_{20} - \beta_2)}{2 \|\beta_{20} - \beta_2\|} + \frac{2 \sqrt{T} \left(\frac{1}{2} \lambda' + \frac{1}{2} (\beta_{20} - \beta_2)' \right) M_{4t} \left(\frac{1}{2} \lambda' + \frac{1}{2} (\beta_{20} - \beta_2)' \right)}{\|\beta_{20} - \beta_2\|} \right| \right\} \end{aligned} \quad (21)$$

for the set \bar{D} , $|\tau - \tau_0| < \frac{C}{T \|\lambda\|^2}$, and therefore:

$$\begin{aligned} \alpha' M_{4t} \alpha &= \frac{1}{T} \sum_{t=k+1}^{k_0} \alpha' \eta_t(\theta_4) \alpha = \frac{k_0 - k}{T} \alpha' M(\theta_4) \alpha + op(1) = Op \left(\frac{\|\alpha\|^2}{T \|\lambda\|^2} \right), \\ \sqrt{T} \alpha' N_{4t} &= \frac{1}{\sqrt{T}} \alpha' \sum_{t=k+1}^{k_0} \eta_t = \left(\frac{k_0 - k}{T} \right)^{1/2} (\alpha' S \alpha)^{1/2} B(1) = Op \left(\frac{\|\alpha\|}{\sqrt{T} \|\lambda\|} \right), \end{aligned}$$

and thus,

$$\frac{\lambda' + (\beta_{20} - \beta_2)' \sqrt{T} N_{4t}}{\|\beta_{20} - \beta_2\|} = Op \left(\frac{\|\lambda\| + \|\beta_{20} - \beta_2\|}{T \|\lambda\|^2 \|\beta_{20} - \beta_2\|} \right) = Op \left(\frac{1}{\sqrt{T} \|\beta_{20} - \beta_2\|} + \frac{1}{\sqrt{T} \|\lambda\|} \right) = op(1),$$

$$\begin{aligned}
& \frac{2\sqrt{T} \left(\frac{1}{2}\lambda' + \frac{1}{2}(\beta_{20} - \beta_2)' \right) M_{4t} \left(\frac{1}{2}\lambda + \frac{1}{2}(\beta_{20} - \beta_2) \right)}{\|\beta_{20} - \beta_2\|} \\
&= O_p \left(\frac{\sqrt{T} \|\lambda\|^2 + \|\beta_{20} - \beta_2\|^2}{T \|\lambda\|^2 \|\beta_{20} - \beta_2\|} \right) = O_p \left(\frac{1}{\sqrt{T} \|\beta_{20} - \beta_2\|} \right) + O_p \left(\frac{\sqrt{T} \|\beta_{20} - \beta_2\|}{T \|\lambda\|^2} \right) \\
&= op(1) + \sqrt{T} \|\beta_{20} - \beta_2\| O_p \left(\frac{1}{T \|\lambda\|^2} \right)
\end{aligned}$$

Then, by Proposition 3, (21) will imply the event

$$\begin{aligned}
& \left\{ \sup_{B \cap \bar{D}} \left| \frac{(\beta_{20} - \beta_2)'}{\|\beta_{20} - \beta_2\|} \sqrt{T} N_{2t} \right| \right. \\
& \left. > \inf_{B \cap \bar{D}} \left[(q_2 + op(1)) \sqrt{T} \|\beta_{20} - \beta_2\| + \sqrt{T} \|\beta_{20} - \beta_2\| O_p \left(\frac{1}{T \|\lambda\|^2} \right) \right] \right\} + op(1).
\end{aligned}$$

Therefore, $\Pr \{E_{12}\}$ is less than or equal to

$$\Pr \left\{ \sup_{B \cap \bar{D}} \left| \frac{(\beta_{20} - \beta_2)'}{\|\beta_{20} - \beta_2\|} \sqrt{T} N_{2t} \right| > C \left(q_2 + O_p \left(\frac{1}{T \|\lambda\|^2} \right) \right) \right\} = O \left(\frac{1}{C^2} \right),$$

by Lemma 6. Standard arguments conclude the proof. \blacksquare

Proposition 16 Define the event $\{E_{13}\}$ as

$$\left\{ \sup_{B \cap \bar{D}} \left[-2 \left(\frac{1}{2}\lambda' + \frac{1}{2}(\beta_{20} - \beta_2)' \right) M_{4t} \left(\frac{1}{2}\lambda + \frac{1}{2}(\beta_{20} - \beta_2) \right) - (\lambda' + (\beta_{20} - \beta_2)') N_{4t} \right] > 0 \right\}$$

then it holds that $\Pr \{E_{13}\} < \varepsilon$ for large enough C .

Proof: $\{E_{13}\}$ implies that

$$\left\{ \sup_{B \cap \bar{D}} \left[-\sqrt{T} \left(\frac{1}{2}\lambda' M_{4t} \lambda + \frac{1}{2}(\beta_{20} - \beta_2)' M_{4t} (\beta_{20} - \beta_2) + \lambda' M_{4t} (\beta_{20} - \beta_2) \right) - (\lambda' + (\beta_{20} - \beta_2)') \sqrt{T} N_{4t} \right] > 0 \right\}$$

Observe that, in this subset,

$$(\lambda' + (\beta_{20} - \beta_2)') \sqrt{T} N_{4t} = \left(O_p(\|\lambda\|) + O_p \left(\frac{1}{\sqrt{T}} \right) \right) O_p \left(\frac{1}{\sqrt{T} \|\lambda\|} \right) = op(1),$$

by Proposition 2, and

$$\sqrt{T} \lambda' M_{4t} \lambda = O_p \left(\sqrt{T} \frac{\|\lambda\|^2}{T \|\lambda\|^2} \right) = op(1),$$

$$\sqrt{T} (\beta_{20} - \beta_2)' M_{4t} (\beta_{20} - \beta_2) = O_p \left(\frac{1}{\sqrt{T}} \right) O_p \left(\frac{1}{T \|\lambda\|^2} \right) = op(1),$$

$$\sqrt{T} \lambda' M_{4t} (\beta_{20} - \beta_2) = O_p \left(\frac{1}{T \|\lambda\|^2} \right) O_p \left(\frac{\sqrt{T} \|\lambda\|}{\sqrt{T}} \right) = op(1),$$

by Proposition 1. Then by definition, if a random vector X_t is $op(1)$, then it holds that $\forall \varepsilon > 0, \forall C > 0, \Pr \{|X_t| > C\} < \varepsilon$, that proves the result for all C . \blacksquare

Proof of Theorem 3.

Again, we will only consider the case of $v_3 < 0$, without loss of generality because of symmetry, and for notational convenience $[v_3 P_\lambda]$ will be denoted by $v_3 P_\lambda$.

By (7)-(10), we obtain that,

$$T\Lambda_T(v) = -\frac{1}{2}v'_1 \frac{1}{T} \sum_{t=1}^{k_0} \dot{\eta}_t \left(\delta_1 \frac{v_1}{\sqrt{T}} \right) v_1 - \frac{1}{\sqrt{T}} v'_1 \sum_{t=1}^{k_0} \eta_t \quad (22)$$

$$+ \frac{1}{2}v'_1 \frac{1}{T} \sum_{t=k_0+v_3 P_\lambda+1}^{k_0} \dot{\eta}_t \left(\delta_3 \frac{v_1}{\sqrt{T}} \right) v_1 + \frac{1}{\sqrt{T}} v'_1 \sum_{t=k_0+v_3 P_\lambda+1}^{k_0} \eta_t \quad (23)$$

$$- \frac{1}{2}v'_2 \frac{1}{T} \sum_{t=k_0+1}^T \dot{\eta}_t \left(\delta_2 \frac{v_2}{\sqrt{T}} \right) v_2 - \frac{1}{\sqrt{T}} v'_2 \sum_{t=k_0+1}^T \eta_t$$

$$- \frac{1}{2}\lambda' \sum_{t=k_0+v_3 P_\lambda+1}^{k_0} \dot{\eta}_t \left(\delta_4 \left(\lambda + \frac{v_2}{\sqrt{T}} \right) \right) \lambda - \frac{1}{2}v'_2 \frac{1}{T} \sum_{t=k_0+v_3 P_\lambda+1}^{k_0} \dot{\eta}_t \left(\delta_4 \left(\lambda + \frac{v_2}{\sqrt{T}} \right) \right) v_2$$

$$- \lambda' \frac{1}{\sqrt{T}} \sum_{t=k_0+v_3 P_\lambda+1}^{k_0} \dot{\eta}_t \left(\delta_4 \left(\lambda + \frac{v_2}{\sqrt{T}} \right) \right) v_2 - \lambda' \sum_{t=k_0+v_3 P_\lambda+1}^{k_0} \eta_t - \frac{1}{\sqrt{T}} v'_2 \sum_{t=k_0+v_3 P_\lambda+1}^{k_0} \eta_t,$$

where for the terms (22) and (23) we have employed that $\sum_{t=1}^k = \sum_{t=1}^{k_0} - \sum_{t=k+1}^{k_0}$. Next, by Lemma 2 and Assumption A.2.1,

$$-\frac{1}{2}v'_1 (\tau_0 M + op(1)) v_1 - \frac{1}{2}v'_2 ((1 - \tau_0) M + op(1)) v_2 \quad (24)$$

$$+ \frac{1}{2}v'_1 \left(\frac{v_3 P_\lambda}{T} M + op(1) \right) v_1 - \frac{1}{2}\lambda' (v_3 P_\lambda M(\lambda) + op(1)) \lambda \quad (25)$$

$$- \frac{1}{2}v'_2 \left(\frac{v_3 P_\lambda}{T} M(\lambda) + op(1) \right) v_2 - \lambda' \frac{1}{\sqrt{T}} (v_3 P_\lambda M(\lambda) + op(1)) v_2 \quad (26)$$

$$- v'_1 \frac{1}{\sqrt{T}} \sum_{t=1}^{k_0} \eta_t - v'_2 \frac{1}{\sqrt{T}} \sum_{t=k_0+1}^T \eta_t + v'_1 \frac{1}{\sqrt{T}} \sum_{t=k_0+v_3 P_\lambda+1}^{k_0} \eta_t \quad (27)$$

$$- \lambda' \sum_{t=k_0+v_3 P_\lambda+1}^{k_0} \eta_t - v'_2 \frac{1}{\sqrt{T}} \sum_{t=k_0+v_3 P_\lambda+1}^{k_0} \eta_t, \quad (28)$$

Hence, as $T \rightarrow \infty$:

$$\begin{aligned} \frac{1}{2}v'_1 \left(M \frac{v_3 P_\lambda}{T} + op(1) \right) v_1 &= op(1), \\ \frac{1}{2}v'_2 \left(M(\lambda) \frac{v_3 P_\lambda}{T} + op(1) \right) v_2 &= op(1), \\ \frac{1}{2}\lambda' \left(M(\lambda) \frac{v_3 P_\lambda}{\sqrt{T}} + op(1) \right) v_2 &= op(1), \end{aligned}$$

in both cases, where λ is fixed or decreasing to zero with T , because

$$\text{when } \lambda \text{ decreasing: } \frac{P_{\lambda_T}}{T} = O\left(\frac{1}{T\|\lambda_T\|^2}\right) = o(1),$$

and

$$\text{when } \lambda \text{ fixed: } \frac{P_\lambda}{T} = O\left(\frac{1}{T}\right) = o(1).$$

Thus, we obtain that,

$$\begin{aligned} & (24) + (25) + (26) \\ &= -\frac{1}{2}v'_1\tau_0 Mv_1 - \frac{1}{2}v'_2(1-\tau_0) Mv_2 - \frac{1}{2}\lambda' M(\lambda)\lambda v_3 + op(1). \end{aligned}$$

Next, we analyze the terms, (27) and (28). By Lemma 3,

$$\begin{aligned} v'_1 \frac{1}{\sqrt{T}} \sum_{t=1}^{k_0} \eta_t &\Rightarrow v'_1 S^{1/2} B(\tau_0), \\ v'_2 \frac{1}{\sqrt{T}} \sum_{t=k_0+1}^T \eta_t &\Rightarrow v'_2 S^{1/2} B(1-\tau_0), \\ v'_1 \frac{1}{\sqrt{T}} \sum_{t=k_0+v_3 P_\lambda+1}^T \eta_t &\Rightarrow v'_1 S^{1/2} B\left(\frac{v_3 P_\lambda}{T}\right) = op(1), \end{aligned}$$

and

$$v'_2 \frac{1}{\sqrt{T}} \sum_{t=k_0+v_3 P_\lambda+1}^T \eta_t \Rightarrow v'_2 S^{1/2} B\left(\frac{v_3 P_\lambda}{T}\right) = op(1).$$

And, for the remaining terms is convenient to study the cases when λ decreases and is constant, separately.

(a) When $\lambda = \lambda_T$ such as $T\|\lambda_T\|^2 \rightarrow \infty$. For this case:

$$-\lambda'_T \sum_{t=k_0+v_3 P_{\lambda_T}+1}^{k_0} \eta_t \Rightarrow -\sqrt{P_{\lambda_T}} \lambda'_T S^{1/2} W(-v_3) = -W(-P_{\lambda_T} \lambda'_T S \lambda_T v_3) = -W(-v_3),$$

by the invariance principle, Lemma 3, for $v_3 < 0$ with $P_{\lambda_T} = (\lambda'_T S \lambda_T)^{-1}$; we have that $W(\cdot)$ is a Brownian motion process in R^- . When $v_3 > 0$, the above expression converges to $W(v_3)$, which is a Brownian process defined on the positive half of the real line.

Next, by Lemma 2, $M(\lambda_T) = M$ for this case of λ decreasing, and therefore,

$$-\frac{1}{2}\lambda'_T v_3 P_{\lambda_T} M(\lambda_T) \lambda_T = -\frac{1}{2}\lambda'_T M \lambda_T v_3 (\lambda'_T S \lambda_T)^{-1} = -\frac{1}{2} \frac{\lambda'_T M \lambda_T}{\lambda'_T S \lambda_T} v_3.$$

Thus:

$$(24) + (25) + (26) = -\frac{1}{2}v'_1\tau_0 Mv_1 - \frac{1}{2}v'_2(1-\tau_0) Mv_2 - \frac{1}{2} \frac{\lambda'_T M \lambda_T}{\lambda'_T S \lambda_T} v_3 + op(1), \quad (29)$$

$$(27) + (28) \Rightarrow -v'_1 S^{1/2} B(\tau_0) - v'_2 S^{1/2} B(1-\tau_0) - W(-v_3). \quad (30)$$

(b) When λ is constant. Consider the process

$$W^*(v) = \begin{cases} 0 & v = 0 \\ \sum_{t=v}^{-1} \eta_t & v = -1, -2, \dots \\ \sum_{t=1}^v \eta_t & v = 1, 2, \dots \end{cases}. \quad (31)$$

and let be $W^\#(k) = \sum_{t=k+1}^{k_0} \eta_t$, for $k \leq k_0$ (taking $W^\#(k_0) = 0$) and $W^\#(k) = \sum_{t=k_0+1}^k \eta_t$, for $k > k_0$. Thus, $W^\#(k)$ has the same distribution as $W^*(k - k_0)$, where $W^*(\cdot)$ is defined by (31). For $k - k_0 = v_3$, $v_3 < 0$,

$$-\lambda' \sum_{t=k_0+v_3+1}^{k_0} \eta_t \Rightarrow -\lambda' W^*(v_3).$$

For $v_3 > 0$, the limit distribution is also defined in (31). Therefore:

$$(24) + (25) + (26) = -\frac{1}{2} v_1' \tau_0 M v_1 - \frac{1}{2} v_2' (1 - \tau_0) M v_2 - \frac{1}{2} \lambda' M(\lambda) \lambda v_3 + op(1) \quad (32)$$

$$(27) + (28) \Rightarrow -v_1' S^{1/2} B(\tau_0) - v_2' S^{1/2} B(1 - \tau_0) - \lambda' W^*(-v_3). \quad (33)$$

And the separate treatment for the cases a) and b) is concluded.

After realizing the distribution limit of the objective function above, defined on the compact set $\{|v_i| < M, i = 1, 2, 3\}$, we obtain the weak convergence of the estimators applying the mapping continuous theorem. Thus, \hat{v}_1 and \hat{v}_2 are derived by the first order conditions:

$$-\tau_0 M \hat{v}_1 - S^{1/2} B(\tau_0) = 0 \Rightarrow \hat{v}_1 = \frac{1}{\tau_0} M^{-1} S^{1/2} B(\tau_0),$$

$$-(1 - \tau_0) M \hat{v}_2 - S^{1/2} B(1 - \tau_0) = 0 \Rightarrow \hat{v}_2 = \frac{1}{1 - \tau_0} M^{-1} S^{1/2} B(1 - \tau_0),$$

while \hat{v}_3 , under the assumption $\lambda = \lambda_T$ decreasing, is obtained by considering (29) and (30), thus:

$$\begin{aligned} \hat{v}_3 &= \underset{v}{\text{Arg min}} -\lambda' W(|v|) + \frac{1}{2} \frac{\lambda_T' M \lambda_T}{\lambda_T' S \lambda_T} |v| \\ &= \underset{v}{\text{Arg max}} \lambda' W(|v|) - \frac{1}{2} \frac{\lambda_T' M \lambda_T}{\lambda_T' S \lambda_T} |v|. \end{aligned}$$

Using the change of variable $w = \left(\frac{\lambda_T' M \lambda_T}{\lambda_T' S \lambda_T} \right)^2 v$,

$$\begin{aligned} \frac{\hat{k} - k_0}{P_{\lambda_T}} &= \left(\frac{\lambda_T' S \lambda_T}{\lambda_T' M \lambda_T} \right)^2 \underset{w}{\text{Arg max}} \left\{ \left| \frac{\lambda_T' S \lambda_T}{\lambda_T' M \lambda_T} \right| \left[W(|w|) - \frac{1}{2} |w| \right] \right\} \\ &= \left(\frac{\lambda_T' S \lambda_T}{\lambda_T' M \lambda_T} \right)^2 \underset{w}{\text{Arg max}} \left\{ W(|w|) - \frac{1}{2} |w| \right\}, \end{aligned}$$

and by definition of P_{λ_T} ,

$$\frac{(\lambda'_T M \lambda_T)^2}{\lambda'_T S \lambda_T} (\hat{k} - k_0) \Rightarrow \underset{w}{\text{Arg max}} \left\{ W(|w|) - \frac{1}{2} |w| \right\}.$$

Finally, for λ constant, we consider (32) and (33), therefore,

$$\begin{aligned} \hat{v}_3 &= \underset{v}{\text{Arg min}} -\lambda' W^*(v) + \frac{1}{2} \lambda' M(\lambda) \lambda |v| \\ &= \underset{v}{\text{Arg max}} \lambda' W^*(v) - \frac{1}{2} \lambda' M(\lambda) \lambda |v|. \end{aligned}$$

This concludes the proof of the theorem. ■

Proof of Corollary 1.

By Theorem 3 we have obtained that

$$\sqrt{T} \begin{bmatrix} \hat{\beta}_1 - \beta_{10} \\ \hat{\beta}_2 - \beta_{20} \end{bmatrix} \xrightarrow{d} \begin{bmatrix} M^{-1} S^{1/2} \frac{B(\tau_0)}{\tau_0} & 0 \\ 0 & M^{-1} S^{1/2} \frac{B(1-\tau_0)}{1-\tau_0} \end{bmatrix}.$$

Define the vector $H = \begin{bmatrix} I_p \\ -I_p \end{bmatrix} \in R^{p \times 2p}$. Then,

$$\sqrt{T} (\hat{\lambda} - \lambda_0) = \sqrt{T} H \begin{bmatrix} \hat{\beta}_1 - \beta_{10} \\ \hat{\beta}_2 - \beta_{20} \end{bmatrix},$$

which converges in distribution to

$$H \begin{bmatrix} M^{-1} S^{1/2} \frac{B(\tau_0)}{\tau_0} & 0 \\ 0 & M^{-1} S^{1/2} \frac{B(1-\tau_0)}{1-\tau_0} \end{bmatrix} = M^{-1} S^{1/2} \left(\frac{B(\tau_0) - \tau_0 B(1)}{\tau_0 (1-\tau_0)} \right)$$

a Bessel tied down process of order p . ■

Proof of Corollary 3.

We only proof the part (ii) of Corollary 3. The rest of this corollary is obtained in a trivial way as a direct result of Theorem 3. Trending regressors $X_t = g(t/T)$ satisfy all necessary assumptions of Theorem 2. Then, to obtain the asymptotic distribution of the break date estimator, we only need to consider, from Theorem 3, the limiting process of

$$-\lambda'_T \sum_{t=k_0+v_3 P_{\lambda_T}+1}^{k_0} \eta_t - \frac{1}{2} \lambda'_T \sum_{t=k_0+v_3 P_{\lambda}+1}^{k_0} \dot{\eta}_t \left(\delta_4 \left(\lambda_T + \frac{v_2}{\sqrt{T}} \right) \right) \lambda_T + op(1), \quad (34)$$

considering again $v_3 < 0$. For the opposite case, the proof is similar.

Then, we have in this case that the first term of (34) is equal to

$$\lambda'_T \sum_{t=k_0+v_3 P_{\lambda_T}+1}^{k_0} \psi_t g(t/T) = \lambda'_T \sum_{t=k_0+v_3 P_{\lambda_T}+1}^{k_0} \psi_t g(k_0/T) \quad (35)$$

$$+ \lambda'_T \sum_{t=k_0+v_3 P_{\lambda_T}+1}^{k_0} \psi_t (g(t/T) - g(k_0/T)). \quad (36)$$

By Lemma 3,

$$(35) \Rightarrow (\lambda'_T g(\tau_0) S g(\tau_0)' \lambda_T)^{1/2} W(-v_3 P_{\lambda_T}) = W(-v_3),$$

for $v_3 < 0$ with $P_{\lambda_T} = (\lambda'_T g(\tau_0) S g(\tau_0)' \lambda_T)^{-1}$ and

$$S = \lim_{T \rightarrow \infty} \text{Var} \left[\frac{1}{\sqrt{T}} \sum_{t=1}^T \psi_t \right]. \quad (37)$$

$W(\cdot)$ is a Brownian motion process in R .

The term (36) is uniformly negligible because its variance will be an $op(1)$. To obtain this, observe that, an upper bound of this variance is given by

$$\begin{aligned} & \text{Sup}_x \left\| \frac{dg(x)}{dx} \right\|^2 \text{Var} \left(\lambda'_T \sum_{t=k_0+v_3 P_{\lambda_T}+1}^{k_0} \psi_t \left(\frac{t}{T} - \frac{k_0}{T} \right) \right) \\ & \leq \text{Sup}_x \left\| \frac{dg(x)}{dx} \right\|^2 \left(\frac{k_0}{T} - \frac{k_0 + v_3 P_{\lambda_T}}{T} \right)^2 \text{Var} \left(\lambda'_T \sum_{t=k_0+v_3 P_{\lambda_T}+1}^{k_0} \psi_t \right) \\ & \leq \text{Sup}_x \left\| \frac{dg(x)}{dx} \right\|^2 \left(\frac{N}{T \|\lambda_T\|^2} \right)^2 S \|\lambda_T\|^2 \frac{N}{\|\lambda_T\|^2}, \end{aligned}$$

that is an $op(1)$ because $T \|\lambda_T\|^2 \rightarrow \infty$ and $g(\cdot)$ bounded. The first equality is obtained by (37) and $P_{\lambda_T} = O(\|\lambda_T\|^{-2})$, and the last inequality holds for a finite constant N .

Now, we analyze the second term of (34). Observe that,

$$\begin{aligned} & \frac{1}{2} \lambda'_T \sum_{t=k_0+v_3 P_{\lambda}+1}^{k_0} \dot{\eta}_t \left(\delta_4 \left(\lambda_T + \frac{v_2}{\sqrt{T}} \right) \right) \lambda_T \\ & = \frac{1}{2} \lambda'_T \sum_{t=k_0+v_3 P_{\lambda}+1}^{k_0} \dot{\psi} \left(u_t + g(t/T)' \theta_T \right) g(t/T) g(t/T)' \lambda_T \end{aligned} \quad (38)$$

where $\theta_T = \delta_4 \left(\lambda_T + v_2/\sqrt{T} \right)$, which converges to zero because $\|\lambda_T\| \rightarrow 0$ and $\|v_2\| < N$ by definition. Thus, we have that (38) is equal to

$$\frac{1}{2} \lambda'_T \sum_{t=k_0+v_3 P_{\lambda_T}+1}^{k_0} \dot{\psi} \left(u_t + g(t/T)' \theta_T \right) (g(t/T) - g(k_0/T)) (g(t/T) - g(k_0/T))' \lambda_T \quad (39)$$

$$+\frac{1}{2}\lambda'_T \sum_{t=k_0+v_3P_{\lambda_T}+1}^{k_0} \dot{\psi}(u_t + g(t/T)' \theta_T) g(k_0/T) g(k_0/T)' \lambda_T \quad (40)$$

$$+\lambda'_T \sum_{t=k_0+v_3P_{\lambda_T}+1}^{k_0} \dot{\psi}(u_t + g(t/T)' \theta_T) (g(t/T) - g(k_0/T)) g(k_0/T)' \lambda_T. \quad (41)$$

and we study each of above terms.

$$(39) = \frac{1}{2}\lambda'_T \sum_{t=k_0+v_3P_{\lambda_T}+1}^{k_0} \dot{\psi}(u_t + g(t/T)' \theta_T) \frac{dg(x^*)}{x} \frac{dg(x^*)'}{x} \left(\frac{t}{T} - \frac{k_0}{T}\right)^2 \lambda_T,$$

applying the mean value theorem, with $x^* \in (t/T - k_0/T)$. An upper bound of this term is given by

$$\begin{aligned} & \left(\text{Sup}_x \left\| \frac{dg(x)}{x} \right\| \right)^2 \left(\frac{k_0}{T} - \frac{k_0 + v_3 P_{\lambda_T}}{T} \right)^2 v_3 P_{\lambda_T} \text{Sup}_t |\dot{\psi}(u_t)| \lambda'_T \lambda_T \\ & \leq K \frac{N^2}{T^2 \|\lambda_T\|^4} \|\lambda_T\|^2 \frac{N}{\|\lambda_T\|^2} = op(1), \end{aligned}$$

with K less than infinity because of boundness of $g(\cdot)$ and $\dot{\psi}(\cdot)$. The last results holds by $T \|\lambda_T\|^2 \rightarrow \infty$ and N , a finite constant.

$$(40) = \frac{1}{2}\lambda'_T v_3 P_{\lambda_T} (M + op(1)) g(\tau_0) g(\tau_0)' \lambda_T,$$

by Lemma 2, given that $g(t/T)' \theta_T \leq \|g(t/T)\| \|\theta_T\| \rightarrow 0$, because $g(\cdot)$ is bounded and $\theta_T \rightarrow 0$ and $M = \lim_{T \rightarrow \infty} \sum_{t=1}^T E[\dot{\psi}_t]$. And, finally,

$$\begin{aligned} (41) & \leq \text{Sup}_x \left\| \frac{dg(x)}{x} \right\| \text{Sup}_t |\dot{\psi}(u_t)| \|\lambda_T\| g(\tau_0) \sum_{t=k_0+v_3P_{\lambda_T}+1}^{k_0} \left(\frac{t}{T} - \frac{k_0}{T} \right) \\ & \leq K \|\lambda_T\|^2 \frac{N}{\|\lambda_T\|^2} \frac{C}{T \|\lambda_T\|^2} = op(1). \end{aligned}$$

Then, we have obtained that (34) converges to

$$-\frac{1}{2}\lambda'_T v_3 P_{\lambda_T} M g(\tau_0) g(\tau_0)' \lambda_T - W(-v_3) = -\frac{M}{S} v_3 - W(-v_3)$$

by definition of $P_{\lambda_T} = (\lambda'_T g(\tau_0) S g(\tau_0)' \lambda_T)^{-1}$. Then, applying a similar proof for $v_3 > 0$, we obtain that,

$$\hat{v}_3 \Rightarrow \underset{v}{\text{Arg min}} \left\{ -W(v) + \frac{1}{2} \frac{M}{S} |v| \right\} = \underset{v}{\text{Arg max}} \left\{ W(v) - \frac{1}{2} \frac{M}{S} |v| \right\}.$$

Let be $w = (M^2/S^2) v$ and, by definition, $\hat{v}_3 = (\hat{k} - k_0) / P_{\lambda_T}$. Therefore,

$$\begin{aligned} \frac{(\hat{k} - k_0)}{P_{\lambda_T}} & \Rightarrow \frac{S^2}{M^2} \underset{w}{\text{Arg max}} \left\{ \left| \frac{S}{M} \right| W(w) + \frac{1}{2} \frac{M}{S} \frac{S^2}{M^2} |w| \right\} \\ & = \frac{S^2}{M^2} \underset{w}{\text{Arg max}} \left\{ W(w) + \frac{1}{2} |w| \right\}, \end{aligned}$$

and then,

$$(\lambda_T' g(\tau_0) g(\tau_0)' \lambda_T) (\hat{k} - k_0) \Rightarrow \frac{S}{M^2} \text{Arg max}_w \left\{ W(w) + \frac{1}{2} |w| \right\},$$

the result of the Corollary. ■

Proof of Theorem 4.

Suppose, with no loss of generality, that $\hat{\tau} > \tau_0$. However, by Theorem 2, we know that $\hat{\tau} = \tau_0 + O_p(T^{-1})$. Then, we have,

$$\hat{M}_1 = \frac{1}{T\hat{\tau}} \sum_{t=1}^{[T\hat{\tau}]} \dot{\psi} \left(Y_t - \hat{\beta}_1(\hat{\tau})' X_t \right) X_t X_t' = \frac{1}{T\hat{\tau}} \sum_{t=1}^{[T\tau_0]} \dot{\eta}_t \left(\beta_{10} - \hat{\beta}_1(\hat{\tau}) \right) + \frac{1}{T\hat{\tau}} \sum_{t=[T\tau_0]+1}^{[T\hat{\tau}]} \dot{\eta}_t \left(\beta_{20} - \hat{\beta}_1(\hat{\tau}) \right).$$

By Theorem 2 we obtain that $\left(\hat{\beta}_1(\hat{\tau}) - \beta_{10} \right) = O_p(T^{-1/2})$ and $\left(\hat{\beta}_1(\hat{\tau}) - \beta_{20} \right) = \lambda_0 + O_p(T^{-1/2})$, and by Lemma 2,

$$\hat{M}_1 = \frac{\tau_0}{\hat{\tau}} M + \frac{\hat{\tau} - \tau_0}{\hat{\tau}} M + o_p(1) = \frac{\tau_0}{\tau_0 + O_p(T^{-1})} M + \frac{O_p(T^{-1})}{\tau_0 + O_p(T^{-1})} M + o_p(1) = M + o_p(1).$$

For \hat{M}_2 the result follows in a similar way. This proves the first part of the Theorem.

Now, we consider \hat{S}_1 for the *i.i.d.* case. In particular, we have that, supposing again that $\hat{\tau} > \tau_0$, and using the same arguments as before,

$$\begin{aligned} \hat{S}_1 &= \frac{1}{T\hat{\tau}} \sum_{t=1}^{[T\tau_0]} \left(\eta_t \left(\beta_{10} - \hat{\beta}_1(\hat{\tau}) \right) - \bar{m}_{1T}(\hat{\tau}) \right) \times \left(\eta_t \left(\beta_{10} - \hat{\beta}_1(\hat{\tau}) \right) - \bar{m}_{1T}(\hat{\tau}) \right)' \\ &\quad + \frac{1}{T\hat{\tau}} \sum_{t=[T\tau_0]+1}^{[T\hat{\tau}]} \left(\eta_t \left(\beta_{20} - \hat{\beta}_1(\hat{\tau}) \right) - \bar{m}_{1T}(\hat{\tau}) \right) \times \left(\eta_t \left(\beta_{20} - \hat{\beta}_1(\hat{\tau}) \right) - \bar{m}_{1T}(\hat{\tau}) \right)' \\ &= \frac{1}{T\hat{\tau}} \sum_{t=1}^{[T\tau_0]} \eta_t \left(\beta_{10} - \hat{\beta}_1(\hat{\tau}) \right) \eta_t \left(\beta_{10} - \hat{\beta}_1(\hat{\tau}) \right)' + \frac{1}{T\hat{\tau}} \sum_{t=[T\tau_0]+1}^{[T\hat{\tau}]} \eta_t \left(\beta_{20} - \hat{\beta}_1(\hat{\tau}) \right) \eta_t \left(\beta_{20} - \hat{\beta}_1(\hat{\tau}) \right)' \\ &\quad - \bar{m}_{1T}(\hat{\tau}) \bar{m}_{1T}'(\hat{\tau}), \end{aligned}$$

where

$$\bar{m}_{1T}(\hat{\tau}) = \frac{1}{T\hat{\tau}} \sum_{t=1}^{[T\tau_0]} \eta_t \left(\beta_{10} - \hat{\beta}_1(\hat{\tau}) \right) + \frac{1}{T\hat{\tau}} \sum_{t=[T\tau_0]+1}^{[T\hat{\tau}]} \eta_t \left(\beta_{20} - \hat{\beta}_1(\hat{\tau}) \right).$$

Applying the MVT in both terms of the above expression we obtain that:

$$\begin{aligned} \bar{m}_{1T}(\hat{\tau}) &= \frac{1}{T\hat{\tau}} \sum_{t=1}^{[T\hat{\tau}]} \eta_t + \left(\beta_{10} - \hat{\beta}_1(\hat{\tau}) \right)' \frac{1}{T\hat{\tau}} \sum_{t=1}^{[T\hat{\tau}]} \dot{\eta}_t \left(\delta_1 \left(\beta_{10} - \hat{\beta}_1(\hat{\tau}) \right) \right) \\ &\quad + \frac{1}{T\hat{\tau}} \sum_{t=[T\tau_0]+1}^{[T\hat{\tau}]} \eta_t + \left(\beta_{20} - \hat{\beta}_1(\hat{\tau}) \right)' \frac{1}{T\hat{\tau}} \sum_{t=[T\tau_0]+1}^{[T\hat{\tau}]} \dot{\eta}_t \left(\delta_2 \left(\beta_{20} - \hat{\beta}_1(\hat{\tau}) \right) \right), \end{aligned}$$

with $0 < \delta_1, \delta_2 < 1$. By Lemma 3 we have that the first term plus the third term is $Op(T^{-1/2})$. By Theorem 2 and Lemma 2 the second plus the fourth terms is an $Op(T^{-1/2})$ term. Therefore, $\bar{m}_{1T}(\hat{\tau}) = Op(T^{-1/2})$.

Next, observe that,

$$\begin{aligned} & \left\| \hat{S}_1 - S \right\| \\ & \leq \left\| \frac{1}{T\hat{\tau}} \sum_{t=1}^{[T\tau_0]} \eta_t (\beta_{10} - \hat{\beta}_1(\hat{\tau})) \eta_t (\beta_{10} - \hat{\beta}_1(\hat{\tau}))' + \frac{1}{T\hat{\tau}} \sum_{t=[T\tau_0]+1}^{[T\hat{\tau}]} \eta_t (\beta_{20} - \hat{\beta}_1(\hat{\tau})) \eta_t (\beta_{20} - \hat{\beta}_1(\hat{\tau}))' - S \right\| \\ & \quad + Op(T^{-1}) \\ & \leq \left\| \frac{1}{T\hat{\tau}} \sum_{t=1}^{[T\hat{\tau}]} \eta_t \eta_t' - S \right\| \end{aligned} \tag{42}$$

$$\begin{aligned} & + \left\| \left(\beta_{10} - \hat{\beta}_1(\hat{\tau}) \right)' \frac{1}{T\hat{\tau}} \sum_{t=1}^{[T\tau_0]} 2\eta_t \left(\delta(\beta_{10} - \hat{\beta}_1(\hat{\tau})) \right) \dot{\eta}_t \left(\delta(\beta_{10} - \hat{\beta}_1(\hat{\tau})) \right) \right. \\ & \quad \left. + \left(\beta_{20} - \hat{\beta}_1(\hat{\tau}) \right)' \frac{1}{T\hat{\tau}} \sum_{t=[T\tau_0]+1}^{[T\hat{\tau}]} 2\eta_t \left(\delta(\beta_{20} - \hat{\beta}_1(\hat{\tau})) \right) \dot{\eta}_t \left(\delta(\beta_{20} - \hat{\beta}_1(\hat{\tau})) \right) \right\|. \end{aligned} \tag{43}$$

By Assumption A.2.7 we obtain that (42) \xrightarrow{p} 0, with $S = E[\eta_t \eta_t']$. Also we have that (43) \xrightarrow{p} 0, given that, by Theorem 2, $(\hat{\beta}_1(\hat{\tau}) - \beta_{10}) = Op(T^{-1/2})$ and $(\hat{\tau} - \tau_0) = Op(T^{-1})$. Then, by Assumptions A.2.5 we have

$$\begin{aligned} & \left\| \left(\beta_{10} - \hat{\beta}_1(\hat{\tau}) \right)' \frac{1}{T\hat{\tau}} \sum_{t=1}^{[T\tau_0]} 2\eta_t \left(\delta(\beta_{10} - \hat{\beta}_1(\hat{\tau})) \right) \dot{\eta}_t \left(\delta(\beta_{10} - \hat{\beta}_1(\hat{\tau})) \right) \right\| \\ & = \left\| Op(T^{-1/2}) 2 \frac{\tau_0}{\hat{\tau}} E \left[\eta_t \left(Op(T^{-1/2}) \right) \dot{\eta}_t \left(Op(T^{-1/2}) \right) \right] + op(1) \right\| = op(1), \end{aligned}$$

and

$$\begin{aligned} & \left\| \left(\beta_{20} - \hat{\beta}_1(\hat{\tau}) \right)' \frac{1}{T\hat{\tau}} \sum_{t=[T\tau_0]+1}^{[T\hat{\tau}]} 2\eta_t \left(\delta(\beta_{20} - \hat{\beta}_1(\hat{\tau})) \right) \dot{\eta}_t \left(\delta(\beta_{20} - \hat{\beta}_1(\hat{\tau})) \right) \right\| \\ & = \left\| \left(\lambda_0 + Op(T^{-1/2}) \right) 2 \frac{\hat{\tau} - \tau_0}{\hat{\tau}} E \left[\eta_t \left(\lambda_0 + Op(T^{-1/2}) \right) \dot{\eta}_t \left(\lambda_0 + Op(T^{-1/2}) \right) \right] \right\| = op(1). \end{aligned}$$

Then, we have obtained that $\hat{S}_1 \xrightarrow{p} E[\eta_t \eta_t'] = S$, that proves the result of this Theorem. For \hat{S}_2 , the proof is similar. To consider the case of weak dependence, see Andrews (1991), who gives the necessary conditions for consistency estimation.

Proof of Theorem 5.

To prove this theorem we are going to analyze the objective function $S_T(\xi)$ for the τ -estimators in order to obtain an equivalent form of a particular M-estimator. We use

Propositions 17 and 18, at the end of this proof.

The objective function is defined as:

$$S_T(\xi) = s_T^2(\beta, \tau) \frac{1}{T} \sum_{t=1}^{[T\tau]} \rho_2 \left(\frac{Y_t - X_t' \beta_1}{s_T(\beta, \tau)} \right) + s_T^2(\beta, \tau) \frac{1}{T} \sum_{t=[T\tau]+1}^T \rho_2 \left(\frac{Y_t - X_t' \beta_2}{s_T(\beta, \tau)} \right) \quad (44)$$

where $\xi = (\beta_1', \beta_2', \tau)' \in \Theta^2 \times \Pi$ and $s_T(\beta, \tau)$ is such that

$$s_T^2(\beta, \tau) / \frac{1}{T} \sum_{t=1}^{[T\tau]} \rho_1 \left(\frac{Y_t - X_t' \beta_1}{s_T(\beta, \tau)} \right) + \frac{1}{T} \sum_{t=[T\tau]+1}^T \rho_1 \left(\frac{Y_t - X_t' \beta_2}{s_T(\beta, \tau)} \right) = b. \quad (45)$$

For notational simplicity we express in next $s_T(\beta, \tau) = s_T(\beta)$. The estimator of ξ is given by:

$$\hat{\xi} = \underset{\xi \in \Theta^2 \times \Pi}{\text{Arg min}} S_T(\xi) = \underset{\xi \in \Theta^2 \times \Pi}{\text{Arg min}} (S_T(\xi) - S_T(\xi_0)) \quad (46)$$

where $\xi_0 = (\beta_{1t}', \beta_{2t}', \tau_0)'$, $\beta_{1t} = \beta_{10} I(t \leq T\tau)$ and $\beta_{2t} = \beta_{10} I(T\tau \leq t \leq T\tau_0) + \beta_{20} I(T\tau_0 \leq t \leq T)$.

Thus, the objective function evaluated in ξ_0 is as follows,

$$\begin{aligned} S_T(\xi_0) &= s_T^2(\beta_t) \frac{1}{T} \sum_{t=1}^{[T\tau]} \rho_2 \left(\frac{Y_t - X_t' \beta_{1t}}{s_T(\beta_t)} \right) + s_T^2(\beta_t) \frac{1}{T} \sum_{t=[T\tau]+1}^T \rho_2 \left(\frac{Y_t - X_t' \beta_{2t}}{s_T(\beta_t)} \right) \\ &= s_T^2(\beta_t) \frac{1}{T} \sum_{t=1}^{[T\tau]} \rho_2 \left(\frac{Y_t - X_t' \beta_{10} I(t \leq [T\tau])}{s_T(\beta_t)} \right) \\ &\quad + s_T^2(\beta_t) \frac{1}{T} \sum_{t=[T\tau]+1}^T \rho_2 \left(\frac{Y_t - X_t' \beta_{10} I([T\tau] + 1 \leq t < [T\tau_0]) - X_t' \beta_{20} I([T\tau_0] + 1 \leq t < T)}{s_T(\beta_t)} \right) \\ &= s_T^2(\beta_t) \frac{1}{T} \sum_{t=1}^T \rho_2 \left(\frac{u_t}{s_T(\beta_t)} \right) \end{aligned}$$

with

$$s_T^2(\beta_t) / \frac{1}{T} \sum_{t=1}^T \rho_1 \left(\frac{u_t}{s_T(\beta_t)} \right) = b.$$

Now, we approximate the objective function in (46) by the MVT up to the second term.

By Propositions 17 and 18 we obtain that,

$$\begin{aligned} &S_T(\xi) - S_T(\xi_0) \quad (47) \\ &= -\frac{1}{T} \sum_{t=1}^{[T\tau]} (\beta_1 - \beta_{1t})' s_T(\beta_t) \left(W_T(\beta_t) \psi_1 \left(\frac{r_t(\beta_{1t})}{s_T(\beta_t)} \right) + \psi_2 \left(\frac{r_t(\beta_{1t})}{s_T(\beta_t)} \right) \right) X_t \\ &\quad - \frac{1}{T} \sum_{t=[T\tau]+1}^T (\beta_2 - \beta_{2t})' s_T(\beta_t) \left(W_T(\beta_t) \psi_1 \left(\frac{r_t(\beta_{2t})}{s_T(\beta_t)} \right) + \psi_2 \left(\frac{r_t(\beta_{2t})}{s_T(\beta_t)} \right) \right) X_t \\ &\quad + \frac{1}{2} \frac{1}{T} \sum_{t=1}^{[T\tau]} (\beta_1 - \beta_{1t})' \left(\left(W_T(\beta_t^*) \dot{\psi}_1 \left(\frac{r_t(\beta_{1t}^*)}{s_T(\beta_t^*)} \right) + \dot{\psi}_2 \left(\frac{r_t(\beta_{1t}^*)}{s_T(\beta_t^*)} \right) \right) X_t X_t' \right) \end{aligned}$$

$$\begin{aligned}
& + \psi_1 \left(\frac{r_t(\beta_{1t}^*)}{s_T(\beta_t^*)} X_t W_{1T}(\beta_t^*) \right) (\beta_1 - \beta_{1t}) \\
& + \frac{1}{2} \frac{1}{T} \sum_{t=[T\tau]+1}^T (\beta_2 - \beta_{2t})' \left(\left(W_T(\beta_t^*) \dot{\psi}_1 \left(\frac{r_t(\beta_{2t}^*)}{s_T(\beta_t^*)} \right) + \dot{\psi}_1 \left(\frac{r_t(\beta_{2t}^*)}{s_T(\beta_t^*)} \right) \right) X_t X_t' \right. \\
& \left. + \psi_1 \left(\frac{r_t(\beta_{2t}^*)}{s_T(\beta)} \right) X_t W_{2T}(\beta_t^*) \right) (\beta_2 - \beta_{2t})
\end{aligned}$$

where $\beta_t = (\beta'_{1t}, \beta'_{2t})'$, $\beta_t^* = (\beta'^*_{1t}, \beta'^*_{2t})'$, with $\beta^*_{1t} = \beta_{1t} + \delta_1(\beta_1 - \beta_{1t})$ and $\beta^*_{2t} = \beta_{2t} + \delta_2(\beta_2 - \beta_{2t})$, $0 < \delta_1, \delta_2 < 1$, and $W_{jT} = (\gamma_{jT}(\beta) + \gamma_T(\beta) \dot{s}_{jT}(\beta))' / D_T(\beta)$ for $j = 1, 2$. The cross product are zero because the indicator functions define disjoint sets.

In order to obtain a equivalent form of an M-estimator, a notational change is needed.

From now on, consider that, for $\theta = (\theta'_1, \theta'_2, \theta'_3)' \in \Theta^3 \subset R^{3p}$:

$$W_T(\theta) = \frac{\frac{1}{T} \sum_{t=1}^{[T\tau]} \tilde{\rho}_2 \left(\frac{u_t + \theta'_1 X_t}{s_T(\theta)} \right) + \frac{1}{T} \sum_{t=[T\tau]+1}^{[T\tau_0]} \tilde{\rho}_2 \left(\frac{u_t + \theta'_2 X_t}{s_T(\theta)} \right) + \frac{1}{T} \sum_{t=[T\tau]+1}^T \tilde{\rho}_2 \left(\frac{u_t + \theta'_3 X_t}{s_T(\theta)} \right)}{\frac{1}{T} \sum_{t=1}^{[T\tau]} \tilde{\psi}_1 \left(\frac{u_t + \theta'_1 X_t}{s_T(\theta)} \right) + \frac{1}{T} \sum_{t=[T\tau]+1}^{[T\tau_0]} \tilde{\psi}_1 \left(\frac{u_t + \theta'_2 X_t}{s_T(\theta)} \right) + \frac{1}{T} \sum_{t=[T\tau]+1}^T \tilde{\psi}_1 \left(\frac{u_t + \theta'_3 X_t}{s_T(\theta)} \right)}.$$

$$s_T^2(\theta) / \frac{1}{T} \sum_{t=1}^{[T\tau]} \rho_1 \left(\frac{u_t + \theta'_1 X_t}{s_T(\theta)} \right) + \frac{1}{T} \sum_{t=[T\tau]+1}^{[T\tau_0]} \rho_1 \left(\frac{u_t + \theta'_2 X_t}{s_T(\theta)} \right) + \frac{1}{T} \sum_{t=[T\tau]+1}^T \rho_1 \left(\frac{u_t + \theta'_3 X_t}{s_T(\theta)} \right) = b.$$

The same for $W_{jT}(\theta)$, that will be given by $W_{jT}(\theta) = \gamma_{jT}(\theta) + \gamma_T(\theta) \dot{s}_{jT}(\theta) / D_T(\theta)$, with

$$\begin{aligned}
\gamma_{1T}(\theta) &= \frac{1}{T} \sum_{t=1}^{[T\tau]} 2\dot{\psi}_2 \left(\frac{u_t + \theta'_1 X_t}{s_T(\theta)} \right) \left(\frac{u_t + \theta'_1 X_t}{s_T(\theta)} \right) \frac{X_t}{D_T(\theta)}, \\
\gamma_{2T}(\theta) &= \frac{1}{T} \sum_{t=[T\tau]+1}^{[T\tau_0]} 2\dot{\psi}_2 \left(\frac{u_t + \theta'_2 X_t}{s_T(\theta)} \right) \left(\frac{u_t + \theta'_2 X_t}{s_T(\theta)} \right) \frac{X_t}{D_T(\theta)} \\
&\quad + \frac{1}{T} \sum_{t=[T\tau]+1}^T 2\dot{\psi}_2 \left(\frac{u_t + \theta'_3 X_t}{s_T(\theta)} \right) \left(\frac{u_t + \theta'_3 X_t}{s_T(\theta)} \right) \frac{X_t}{D_T(\theta)},
\end{aligned}$$

$$\begin{aligned}
\gamma_T(\theta) &= \frac{1}{T} \sum_{t=1}^{[T\tau]} \tilde{\Upsilon}_T \left(\frac{u_t + \theta'_1 X_t}{s_T(\theta)} \right) \left(\frac{u_t + \theta'_1 X_t}{s_T(\theta)} \right) + \frac{1}{T} \sum_{t=[T\tau]+1}^{[T\tau_0]} \tilde{\Upsilon}_T \left(\frac{u_t + \theta'_2 X_t}{s_T(\theta)} \right) \left(\frac{u_t + \theta'_2 X_t}{s_T(\theta)} \right) \\
&\quad + \frac{1}{T} \sum_{t=[T\tau]+1}^T \tilde{\Upsilon}_T \left(\frac{u_t + \theta'_3 X_t}{s_T(\theta)} \right) \left(\frac{u_t + \theta'_3 X_t}{s_T(\theta)} \right),
\end{aligned}$$

and, finally, $\dot{s}_{jT}(\theta) = -A_{jT}(\theta) / D_T(\theta)$ for $j = 1, 2$, where,

$$\begin{aligned}
A_{1T}(\theta) &= \frac{1}{T} \sum_{t=1}^{[T\tau]} \psi_1 \left(\frac{u_t + \theta'_1 X_t}{s_T(\theta)} \right) X_t, \\
A_{2T}(\theta) &= \frac{1}{T} \sum_{t=[T\tau]+1}^{[T\tau_0]} \psi_1 \left(\frac{u_t + \theta'_2 X_t}{s_T(\theta)} \right) X_t + \frac{1}{T} \sum_{t=[T\tau]+1}^T \psi_1 \left(\frac{u_t + \theta'_3 X_t}{s_T(\theta)} \right) X_t,
\end{aligned}$$

$$\begin{aligned}
D_T(\theta) &= \sum_{t=1}^{[T\tau]} \psi_1 \left(\frac{u_t + \theta'_1 X_t}{s_T(\theta)} \right) \left(\frac{u_t + \theta'_1 X_t}{s_T(\theta)} \right) + \frac{1}{T} \sum_{t=[T\tau]+1}^{[T\tau_0]} \psi_1 \left(\frac{u_t + \theta'_2 X_t}{s_T(\theta)} \right) \left(\frac{u_t + \theta'_2 X_t}{s_T(\theta)} \right) \\
&\quad + \sum_{t=[T\tau_0]+1}^T \psi_1 \left(\frac{u_t + \theta'_3 X_t}{s_T(\theta)} \right) \left(\frac{u_t + \theta'_3 X_t}{s_T(\theta)} \right).
\end{aligned}$$

Then, the Taylor approximation of the objective function (47) up to the second term results as follows:

$$\begin{aligned}
&S_T(\xi) - S_T(\xi_0) \\
&= -(\beta_1 - \beta_{10})' s_T(0_{3p}) \frac{1}{T} \sum_{t=1}^{[T\tau]} \left(W_T(0_{3p}) \psi_1 \left(\frac{u_t}{s_T(0_{3p})} \right) + \psi_2 \left(\frac{u_t}{s_T(0_{3p})} \right) \right) X_t \\
&\quad - (\beta_2 - \beta_{10})' s_T(0_{3p}) \frac{1}{T} \sum_{t=[T\tau]+1}^{[T\tau_0]} \left(W_T(0_{3p}) \psi_1 \left(\frac{u_t}{s_T(0_{3p})} \right) + \psi_2 \left(\frac{u_t}{s_T(0_{3p})} \right) \right) X_t \\
&\quad - (\beta_2 - \beta_{20})' s_T(0_{3p}) \frac{1}{T} \sum_{t=[T\tau_0]+1}^T \left(W_T(0_{3p}) \psi_1 \left(\frac{u_t}{s_T(0_{3p})} \right) + \psi_2 \left(\frac{u_t}{s_T(0_{3p})} \right) \right) X_t \\
&\quad + \frac{1}{2} (\beta_1 - \beta_{10})' \left(\frac{1}{T} \sum_{t=1}^{[T\tau]} \left(W_T(\theta) \dot{\psi}_1 \left(\frac{u_t + \theta'_1 X_t}{s_T(\theta)} \right) + \dot{\psi}_1 \left(\frac{u_t + \theta'_1 X_t}{s_T(\theta)} \right) \right) X_t X_t' \right. \\
&\quad \left. + \frac{1}{T} \sum_{t=1}^{[T\tau]} W_{1T}(\theta) \psi_1 \left(\frac{u_t + \theta'_1 X_t}{s_T(\theta)} \right) X_t \right) (\beta_1 - \beta_{10}) \\
&\quad + \frac{1}{2} (\beta_2 - \beta_{10})' \left(\frac{1}{T} \sum_{t=[T\tau]+1}^{[T\tau_0]} \left(W_T(\theta) \dot{\psi}_1 \left(\frac{u_t + \theta'_2 X_t}{s_T(\theta)} \right) + \dot{\psi}_2 \left(\frac{u_t + \theta'_2 X_t}{s_T(\theta)} \right) \right) X_t X_t' \right. \\
&\quad \left. + \frac{1}{T} \sum_{t=[T\tau]+1}^{[T\tau_0]} W_{2T}(\theta) \psi_1 \left(\frac{u_t + \theta'_2 X_t}{s_T(\theta)} \right) X_t \right) (\beta_2 - \beta_{10}) \\
&\quad + \frac{1}{2} (\beta_2 - \beta_{20})' \left(\frac{1}{T} \sum_{t=[T\tau_0]+1}^T \left(W_T(\theta) \dot{\psi}_1 \left(\frac{u_t + \theta'_3 X_t}{s_T(\theta)} \right) + \dot{\psi}_2 \left(\frac{u_t + \theta'_3 X_t}{s_T(\theta)} \right) \right) X_t X_t' \right. \\
&\quad \left. + \frac{1}{T} \sum_{t=[T\tau_0]+1}^T W_{2T}(\theta) \psi_1 \left(\frac{u_t + \theta'_3 X_t}{s_T(\theta)} \right) X_t \right) (\beta_2 - \beta_{20})
\end{aligned}$$

where $\theta = (\theta'_1, \theta'_2, \theta'_3)'$ with $\theta_1 = \delta_1 (\beta_{10} - \beta_1)$, $\theta_2 = \delta_2 (\beta_{10} - \beta_2)$ and $\theta_3 = \delta_2 (\beta_{20} - \beta_2)$. Besides, we have defined $0_{3p} = (0'_p, 0'_p, 0'_p)'$, where $0'_p$ is a p -vector with all its elements equal to zero, and therefore,

$$W_T(0_{3p}) = \frac{\frac{1}{T} \sum_{t=1}^T \tilde{\rho}_2 \left(\frac{u_t}{s_T(0_{3p})} \right)}{\frac{1}{T} \sum_{t=1}^T \tilde{\psi}_1 \left(\frac{u_t}{s_T(0_{3p})} \right)} \text{ and } s_T^2(0_{3p}) / \frac{1}{T} \sum_{t=1}^T \rho_1 \left(\frac{u_t}{s_T(0_{3p})} \right) = b.$$

With this, we have obtained the expression of the objective function of a particular M-

estimator:

$$\begin{aligned}
& -(\beta_1 - \beta_{10})' \frac{1}{T} \sum_{t=1}^{[T\tau]} \eta_t + \frac{1}{2} (\beta_1 - \beta_{10})' \frac{1}{T} \sum_{t=1}^{[T\tau]} (\dot{\eta}_t(\theta_1) + \dot{\eta}_{1t}(\theta_1)) (\beta_1 - \beta_{10}) \\
& -(\beta_2 - \beta_{10})' \frac{1}{T} \sum_{t=[T\tau]+1}^{[T\tau_0]} \eta_t + \frac{1}{2} (\beta_2 - \beta_{10})' \frac{1}{T} \sum_{t=[T\tau]+1}^{[T\tau_0]} (\dot{\eta}_t(\theta_2) + \dot{\eta}_{2t}(\theta_2)) (\beta_2 - \beta_{10}) \\
& -(\beta_2 - \beta_{20})' \frac{1}{T} \sum_{t=[T\tau_0]+1}^T \eta_t + \frac{1}{2} (\beta_2 - \beta_{20})' \frac{1}{T} \sum_{t=[T\tau_0]+1}^T (\dot{\eta}_t(\theta_3) + \dot{\eta}_{2t}(\theta_3)) (\beta_2 - \beta_{20})
\end{aligned}$$

where, for $i = 1, 2, 3$ and $j = 1, 2$,

$$\begin{aligned}
\eta_t(\theta_i) &= s_T(\theta) \left(W_T(\theta) \psi_1 \left(\frac{u_t + \theta_i' X_t}{s_T(\theta)} \right) + \psi_2 \left(\frac{u_t + \theta_i' X_t}{s_T(\theta)} \right) \right) X_t, \\
\dot{\eta}_t(\theta_i) &= \left(W_T(\theta) \dot{\psi}_1 \left(\frac{u_t + \theta_i' X_t}{s_T(\theta)} \right) + \dot{\psi}_2 \left(\frac{u_t + \theta_i' X_t}{s_T(\theta)} \right) \right) X_t X_t', \\
\dot{\eta}_{jt}(\theta_i) &= W_{jT}(\theta) \psi_1 \left(\frac{u_t + \theta_i' X_t}{s_T(\theta)} \right) X_t,
\end{aligned}$$

and, for $\theta = 0_{3p}$, we have that $\eta_t(\theta_i) = \eta_t$.

Now, observe that to prove Theorem 2 we studied the Taylor expansion of $S_T(\xi_0) - S_T(\xi)$ instead of $S_T(\xi) - S_T(\xi_0)$. Then to obtain a similar form in this case we express the objective function as

$$\begin{aligned}
& S_T(\xi_0) - S_T(\xi) \\
&= -(\beta_{10} - \beta_1)' N_{1t} - \frac{1}{2} (\beta_{10} - \beta_1)' M_{1t} (\beta_{10} - \beta_1) \\
&\quad - (\beta_{10} - \beta_2)' N_{4t} - \frac{1}{2} (\beta_{10} - \beta_2)' M_{4t} (\beta_{10} - \beta_2) \\
&\quad - (\beta_{20} - \beta_2)' N_{2t} - \frac{1}{2} (\beta_{20} - \beta_2)' M_{2t} (\beta_{20} - \beta_2),
\end{aligned} \tag{49}$$

taking into account that,

$$\begin{aligned}
N_{1t} &= \frac{1}{T} \sum_{t=1}^{[T\tau]} \eta_t, & M_{1t} &= \frac{1}{T} \sum_{t=1}^{[T\tau]} (\dot{\eta}_t(\theta_1) + \dot{\eta}_{1t}(\theta_1)), \\
N_{2t} &= \frac{1}{T} \sum_{t=[T\tau_0]+1}^T \eta_t, & M_{2t} &= \frac{1}{T} \sum_{t=[T\tau_0]+1}^T (\dot{\eta}_t(\theta_3) + \dot{\eta}_{2t}(\theta_3)), \\
N_{4t} &= \frac{1}{T} \sum_{t=[T\tau]+1}^{[T\tau_0]} \eta_t, & M_{4t} &= \frac{1}{T} \sum_{t=[T\tau]+1}^{[T\tau_0]} (\dot{\eta}_t(\theta_2) + \dot{\eta}_{2t}(\theta_2)).
\end{aligned}$$

Noting that $(\beta_{10} - \beta_2) = \lambda + (\beta_{20} - \beta_2)$, we can express the term (49) by,

$$\begin{aligned}
& -\lambda' N_{4t} - (\beta_{10} - \beta_2)' N_{4t} - \frac{1}{2} \lambda' M_{4t} \lambda - \frac{1}{2} (\beta_{10} - \beta_2)' M_{4t} (\beta_{10} - \beta_2) - \lambda' M_{4t} (\beta_{10} - \beta_2) \\
&= -(\lambda + (\beta_{10} - \beta_2))' N_{4t} - 2 \left(\frac{1}{2} \lambda' + \frac{1}{2} (\beta_{10} - \beta_2)' \right) M_{4t} \left(\frac{1}{2} \lambda + \frac{1}{2} (\beta_{10} - \beta_2) \right),
\end{aligned}$$

and finally, we obtain that

$$\begin{aligned}
S_T(\xi_0) - S_T(\xi) &= -\frac{1}{2}(\beta_{10} - \beta_1)' M_{1t}(\beta_{10} - \beta_1) - (\beta_{10} - \beta_1)' N_{1t} \\
&\quad -\frac{1}{2}(\beta_{20} - \beta_2)' M_{2t}(\beta_{20} - \beta_2) - (\beta_{20} - \beta_2)' N_{2t} \\
&\quad -2\left(\frac{1}{2}\lambda' + \frac{1}{2}(\beta_{10} - \beta_2)'\right) M_{4t}\left(\frac{1}{2}\lambda' + \frac{1}{2}(\beta_{10} - \beta_2)'\right) \\
&\quad -(\lambda + (\beta_{10} - \beta_2))' N_{4t},
\end{aligned}$$

an expression similar to (7)-(10) in the proof of Theorem 2.

Thus, to obtain the results of Theorems 2 and 3 for this estimators, we only need to verify the Lemmas 1 to 6 for this case, which constitute the necessary and sufficient conditions of the mentioned Theorems. The equivalent results of Lemmas 1, 2 and 3 are obtained in Lemmas 12, 13 and 14 respectively. The rest of the Lemmas follows in identical form.

By Theorem 3, we can obtain the asymptotic distribution for the τ -estimator as a particular case, considering that, by Lemma 14,

$$S = \sigma_0^2 E \left[\psi_0^2 \left(\frac{u_t}{\sigma_0} \right) \right] E [X_t X_t'],$$

with $\psi_0(u) = W_0 \psi_1(u) + \psi_2(u)$. Following the proof of Theorem 3 for this case, we also obtain that $M(\lambda)$ is given by $M(\theta_\lambda) + M_2(\theta_\lambda)$, where $\theta_\lambda = (0'_p, \lambda', 0'_p)'$, and then,

$$M(\theta_\lambda) = E \left[W(\theta_\lambda) \dot{\psi}_1 \left(\frac{u_t + \lambda' X_t}{s(\theta_\lambda)} \right) + \dot{\psi}_2 \left(\frac{u_t + \lambda' X_t}{s(\theta_\lambda)} \right) \right] E [X_t X_t'], \quad (50)$$

$$M_2(\theta_\lambda) = E \left[W_2(\theta_\lambda) \psi_1 \left(\frac{u_t + \lambda' X_t}{s(\theta_\lambda)} \right) X_t \right]. \quad (51)$$

By the same way $s(\theta_\lambda)$ is defined by

$$\tau E \left[\rho_1 \left(\frac{u_t}{s(\theta_\lambda)} \right) \right] + (\tau_0 - \tau) E \left[\rho_1 \left(\frac{u_t + \lambda' X_t}{s(\theta_\lambda)} \right) \right] + (1 - \tau_0) E \left[\rho_1 \left(\frac{u_t}{s(\theta_\lambda)} \right) \right] = b, \quad (52)$$

and, given that $(\tau_0 - \tau) = v_3 P_\lambda / T$, which tends to zero in both cases of λ (fixed or decreasing), we obtain (52) equals to

$$E \left[\rho_1 \left(\frac{u_t}{s(\theta_\lambda)} \right) \right] = b,$$

and then, $s(\theta_\lambda) = \sigma_0$ by definition.

Now, with similar arguments,

$$\begin{aligned}
W(\theta_\lambda) &= \frac{\tau E \left[\tilde{\rho}_2 \left(\frac{u_t}{s(\theta_\lambda)} \right) \right] + (\tau_0 - \tau) E \left[\tilde{\rho}_2 \left(\frac{u_t + \lambda' X_t}{s(\theta_\lambda)} \right) \right] + (1 - \tau_0) E \left[\tilde{\rho}_2 \left(\frac{u_t}{s(\theta_\lambda)} \right) \right]}{\tau E \left[\tilde{\psi}_1 \left(\frac{u_t}{s(\theta_\lambda)} \right) \right] + (\tau_0 - \tau) E \left[\tilde{\psi}_1 \left(\frac{u_t + \lambda' X_t}{s(\theta_\lambda)} \right) \right] + (1 - \tau_0) E \left[\tilde{\psi}_1 \left(\frac{u_t}{s(\theta_\lambda)} \right) \right]} \\
&= \frac{E \left[\tilde{\rho}_2 \left(\frac{u_t}{\sigma_0} \right) \right]}{E \left[\tilde{\psi}_1 \left(\frac{u_t}{\sigma_0} \right) \right]} + op(1) = W_0 + op(1),
\end{aligned}$$

and, to derive $W_2(\theta_\lambda)$, observe firstly that,

$$\begin{aligned} A_2(\theta_\lambda) &= (\tau_0 - \tau) E \left[\psi_1 \left(\frac{u_t + \lambda' X_t}{s(\theta_\lambda)} \right) X_t \right] + (1 - \tau_0) E \left[\psi_1 \left(\frac{u_t}{s(\theta_\lambda)} \right) X_t \right] \\ &= \frac{v_3 P_\lambda}{T} E \left[\psi_1 \left(\frac{u_t + \lambda' X_t}{\sigma_0} \right) X_t \right] + (1 - \tau_0) E \left[\psi_1 \left(\frac{u_t}{\sigma_0} \right) X_t \right], \end{aligned}$$

an $op(1)$ term, given that $v_3 P_\lambda / T \rightarrow 0$ and the second term is equal to zero. Secondly,

$$\begin{aligned} D(\theta_\lambda) &= \tau E \left[\psi_1 \left(\frac{u_t}{s(\theta_\lambda)} \right) \left(\frac{u_t}{s(\theta_\lambda)} \right) \right] + (\tau_0 - \tau) E \left[\psi_1 \left(\frac{u_t + \lambda' X_t}{s(\theta_\lambda)} \right) \left(\frac{u_t + \lambda' X_t}{s(\theta_\lambda)} \right) \right] \\ &\quad + (1 - \tau_0) E \left[\psi_1 \left(\frac{u_t}{s(\theta_\lambda)} \right) \left(\frac{u_t}{s(\theta_\lambda)} \right) \right] \\ &= E \left[\psi_1 \left(\frac{u_t}{\sigma_0} \right) \left(\frac{u_t}{\sigma_0} \right) \right] + op(1), \\ \dot{s}_2(\theta_\lambda) &= -\frac{A_2(\theta_\lambda)}{D(\theta_\lambda)} = op(1), \end{aligned} \tag{53}$$

and finally,

$$\begin{aligned} \gamma_2(\theta_\lambda) &= (\tau_0 - \tau) E \left[2\dot{\psi}_2 \left(\frac{u_t + \lambda' X_t}{s(\theta_\lambda)} \right) \left(\frac{u_t + \lambda' X_t}{s(\theta_\lambda)} \right) \frac{X_t}{D(\theta_\lambda)} \right] \\ &\quad + (1 - \tau_0) E \left[2\dot{\psi}_2 \left(\frac{u_t}{s(\theta_\lambda)} \right) \left(\frac{u_t}{s(\theta_\lambda)} \right) \frac{X_t}{D(\theta_\lambda)} \right] \\ &= (1 - \tau_0) \frac{E \left[2\dot{\psi}_2 \left(\frac{u_t}{\sigma_0} \right) \left(\frac{u_t}{\sigma_0} \right) X_t \right]}{E \left[\psi_1 \left(\frac{u_t}{\sigma_0} \right) \left(\frac{u_t}{\sigma_0} \right) \right]}. \end{aligned} \tag{54}$$

By the results (53) and (54), we obtain that,

$$W_2(\theta_\lambda) = (1 - \tau_0) \frac{E \left[2\dot{\psi}_2 \left(\frac{u_t}{\sigma_0} \right) \left(\frac{u_t}{\sigma_0} \right) X_t \right]}{E \left[\psi_1 \left(\frac{u_t}{\sigma_0} \right) \left(\frac{u_t}{\sigma_0} \right) \right]} + op(1),$$

because $\gamma(\theta)$ is a finite for all θ .

Therefore, $M(\lambda)$ will be defined by

$$E \left[W_0 \dot{\psi}_1 \left(\frac{u_t + \lambda' X_t}{\sigma_0} \right) + \dot{\psi}_2 \left(\frac{u_t + \lambda' X_t}{\sigma_0} \right) \right] E [X_t X_t'] + (1 - \tau_0) \frac{E \left[2\dot{\psi}_2 \left(\frac{u_t}{\sigma_0} \right) \left(\frac{u_t}{\sigma_0} \right) X_t \right]}{E \left[\psi_1 \left(\frac{u_t}{\sigma_0} \right) \left(\frac{u_t}{\sigma_0} \right) \right]} E \left[\psi_1 \left(\frac{u_t + \lambda' X_t}{\sigma_0} \right) X_t \right]$$

and for λ converging to zero, the second term of the above expression is an $op(1)$, and then $M = M(0)$ will be equal to

$$E \left[W_0 \dot{\psi}_1 \left(\frac{u_t}{\sigma_0} \right) + \dot{\psi}_2 \left(\frac{u_t}{\sigma_0} \right) \right] E [X_t X_t'].$$

This concludes the proof of the theorem. ■

The following Propositions 17 and 18 correspond to the proof of Theorem 5.

Proposition 17

$$\frac{\partial S_T(\xi)}{\partial \beta'_1} = -s_T(\beta) \frac{1}{T} \sum_{t=1}^{[T\tau]} \left(W_T(\beta) \psi_1 \left(\frac{r_t(\beta_1)}{s_T(\beta)} \right) + \psi_2 \left(\frac{r_t(\beta_1)}{s_T(\beta)} \right) \right) X_t, \quad (55)$$

$$\frac{\partial S_T(\xi)}{\partial \beta'_2} = -s_T(\beta) \frac{1}{T} \sum_{t=[T\tau]+1}^T \left(W_T(\beta) \psi_1 \left(\frac{r_t(\beta_2)}{s_T(\beta)} \right) + \psi_2 \left(\frac{r_t(\beta_2)}{s_T(\beta)} \right) \right) X_t, \quad (56)$$

where

$$W_T(\beta) = \frac{\frac{1}{T} \sum_{t=1}^{[T\tau]} \tilde{\rho}_2 \left(\frac{r_t(\beta_1)}{s_T(\beta)} \right) + \frac{1}{T} \sum_{t=[T\tau]+1}^T \tilde{\rho}_2 \left(\frac{r_t(\beta_2)}{s_T(\beta)} \right)}{\frac{1}{T} \sum_{t=1}^{[T\tau]} \tilde{\psi}_1 \left(\frac{r_t(\beta_2)}{s_T(\beta)} \right) + \frac{1}{T} \sum_{t=[T\tau]+1}^T \tilde{\psi}_1 \left(\frac{r_t(\beta_2)}{s_T(\beta)} \right)}. \quad (57)$$

with $\tilde{\psi}_i(z) = z\psi_i(z)$ and $\tilde{\rho}_i(z) = 2\rho_i(z) - \tilde{\psi}_i(z)$, for $i = 1, 2$.

Proof: Taking into account the objective function $S_T(\xi)$ defined by (44) we have that,

$$\begin{aligned} \frac{\partial S_T(\xi)}{\partial \beta'_1} &= 2s_T(\beta) \frac{\partial s_T(\beta)}{\partial \beta'_1} \left(\frac{1}{T} \sum_{t=1}^{[T\tau]} \rho_2 \left(\frac{r_t(\beta_1)}{s_T(\beta)} \right) + \frac{1}{T} \sum_{t=[T\tau]+1}^T \rho_2 \left(\frac{r_t(\beta_2)}{s_T(\beta)} \right) \right) \\ &\quad + s_T^2(\beta) \frac{1}{T} \sum_{t=1}^{[T\tau]} \psi_2 \left(\frac{r_t(\beta_1)}{s_T(\beta)} \right) \frac{-X_t s_T(\beta) - r_t(\beta_1) \partial s_T(\beta) / \partial \beta'_1}{s_T^2(\beta)} \\ &\quad + s_T^2(\beta) \frac{1}{T} \sum_{t=[T\tau]+1}^T \psi_2 \left(\frac{r_t(\beta_2)}{s_T(\beta)} \right) \frac{-r_t(\beta_2) \partial s_T(\beta) / \partial \beta'_1}{s_T^2(\beta)}, \end{aligned} \quad (58)$$

where $r_t(\theta) = Y_t - X_t'\theta$ is the residual function. Now, the partial derivatives of the scale M-estimator $s_T(\beta)$, defined by (45), is given by

$$\begin{aligned} &\frac{1}{T} \sum_{t=1}^{[T\tau]} \psi_1 \left(\frac{r_t(\beta_1)}{s_T(\beta)} \right) \frac{-X_t s_T(\beta) - r_t(\beta_1) \partial s_T(\beta) / \partial \beta'_1}{s_T^2(\beta)} \\ &\quad + \frac{1}{T} \sum_{t=[T\tau]+1}^T \psi_1 \left(\frac{r_t(\beta_2)}{s_T(\beta)} \right) \frac{-r_t(\beta_2) \partial s_T(\beta) / \partial \beta'_1}{s_T^2(\beta)} = 0, \end{aligned}$$

and therefore

$$\frac{\partial s_T(\beta)}{\partial \beta'_1} = - \frac{\frac{1}{T} \sum_{t=1}^{[T\tau]} \psi_1 \left(\frac{r_t(\beta_1)}{s_T(\beta)} \right) X_t}{\frac{1}{T} \sum_{t=1}^{[T\tau]} \psi_1 \left(\frac{r_t(\beta_1)}{s_T(\beta)} \right) \left(\frac{r_t(\beta_1)}{s_T(\beta)} \right) + \frac{1}{T} \sum_{t=[T\tau]+1}^T \psi_1 \left(\frac{r_t(\beta_2)}{s_T(\beta)} \right) \left(\frac{r_t(\beta_2)}{s_T(\beta)} \right)} = - \frac{A_{1T}(\beta)}{D_T(\beta)}. \quad (59)$$

Thus, inserting (59) in (58), we obtain the result (55). For (56) the proof is similar. ■

Proposition 18

$$\frac{\partial^2 S_T(\xi)}{\partial \beta_1 \partial \beta'_1} = \frac{1}{T} \sum_{t=1}^{[T\tau]} \left(W_T(\beta) \dot{\psi}_1 \left(\frac{r_t(\beta_1)}{s_T(\beta)} \right) + \dot{\psi}_2 \left(\frac{r_t(\beta_1)}{s_T(\beta)} \right) \right) X_t X_t' + \psi_1 \left(\frac{r_t(\beta_1)}{s_T(\beta)} \right) X_t W_{1T}(\beta)$$

$$\frac{\partial^2 S_T(\xi)}{\partial \beta_2 \partial \beta_2'} = \frac{1}{T} \sum_{t=[T\tau]+1}^T \left(W_T(\beta) \dot{\psi}_1 \left(\frac{r_t(\beta_2)}{s_T(\beta)} \right) + \dot{\psi}_2 \left(\frac{r_t(\beta_2)}{s_T(\beta)} \right) \right) X_t X_t' + \psi_1 \left(\frac{r_t(\beta_2)}{s_T(\beta)} \right) X_t W_{2T}(\beta)$$

where

$$W_{jT}(\beta) = \gamma_{jT}(\beta) + \gamma_T(\beta) \frac{\dot{s}_{jT}(\beta)'}{D_T(\beta)}, \quad (60)$$

for $j = 1, 2$,

$$\gamma_{1T}(\beta) = \frac{1}{T} \sum_{t=1}^{[T\tau]} 2\dot{\psi}_2 \left(\frac{r_t(\beta_1)}{s_T(\beta)} \right) \left(\frac{r_t(\beta_1)}{s_T(\beta)} \right) \frac{X_t}{D_T(\beta)}, \quad (61)$$

$$\gamma_{2T}(\beta) = \frac{1}{T} \sum_{t=[T\tau]+1}^T 2\dot{\psi}_2 \left(\frac{r_t(\beta_2)}{s_T(\beta)} \right) \left(\frac{r_t(\beta_2)}{s_T(\beta)} \right) \frac{X_t}{D_T(\beta)}, \quad (62)$$

$$\gamma_T(\beta) = \left(\frac{1}{T} \sum_{t=1}^{[T\tau]} \tilde{\Upsilon}_T \left(\frac{r_t(\beta_1)}{s_T(\beta)} \right) \frac{r_t(\beta_1)}{s_T(\beta)} + \frac{1}{T} \sum_{t=[T\tau]+1}^T \tilde{\Upsilon}_T \left(\frac{r_t(\beta_2)}{s_T(\beta)} \right) \frac{r_t(\beta_2)}{s_T(\beta)} \right). \quad (63)$$

and, finally,

$$\dot{s}_{jT}(\beta) = -\frac{A_{jT}(\beta)}{D_T(\beta)} \quad (64)$$

for $j = 1, 2$, defined by (59) for $j = 1$, and for $j = 2$ obtained in a equivalent way.

Proof: Taking into account Proposition 17 we obtain that,

$$\frac{\partial^2 S_T(\xi)}{\partial \beta_1 \partial \beta_1'} = -\frac{1}{T} \sum_{t=1}^{[T\tau]} \left(W_T(\beta) \Psi_1 \left(\frac{r_t(\beta_1)}{s_T(\beta)} \right) + \Psi_2 \left(\frac{r_t(\beta_1)}{s_T(\beta)} \right) \right) X_t \dot{s}_{1T}(\beta)' \quad (65)$$

$$-s_T(\beta) \frac{1}{T} \sum_{t=1}^{[T\tau]} \psi_1 \left(\frac{r_t(\beta_1)}{s_T(\beta)} \right) X_t \dot{W}_{1T}(\beta)' \quad (66)$$

$$+ \frac{1}{T} \sum_{t=1}^{[T\tau]} \left(W_T(\beta) \dot{\psi}_1 \left(\frac{r_t(\beta_1)}{s_T(\beta)} \right) + \dot{\psi}_2 \left(\frac{r_t(\beta_1)}{s_T(\beta)} \right) \right) X_t X_t,$$

where $\dot{s}_{1T}(\beta) = \partial s_T(\beta) / \partial \beta_1'$, $\dot{W}_{1T}(\beta) = \partial W_T(\beta) / \partial \beta_1'$ and $\Psi_i(z) = \psi_i(z) + z \dot{\psi}_i(z)$ for $i = 1, 2$ with $\dot{\psi}_i(z) = \partial \psi_i(z) / \partial z$.

But observe that, inserting the partial derivatives of the M-scale, given by (59), we obtain that (65) is equal to

$$\begin{aligned} & \frac{1}{T} \sum_{t=1}^{[T\tau]} \left(W_T(\beta) \Psi_1 \left(\frac{r_t(\beta_1)}{s_T(\beta)} \right) + \Psi_2 \left(\frac{r_t(\beta_1)}{s_T(\beta)} \right) \right) X_t \frac{A_{1T}(\beta)'}{D_T(\beta)} \\ &= \frac{1}{T} \sum_{t=1}^{[T\tau]} \Upsilon_T \left(\frac{r_t(\beta_1)}{s_T(\beta)} \right) \frac{X_t}{D_T(\beta)} \frac{1}{T} \sum_{t=1}^{[T\tau]} \psi_1 \left(\frac{r_t(\beta_1)}{s_T(\beta)} \right) X_t' \end{aligned} \quad (67)$$

where $\Upsilon_T(z) = W_T(\beta) \Psi_1(z) + \Psi_2(z)$.

For the term (66), after some algebra we obtain that,

$$\begin{aligned}
& s_T(\beta) \dot{W}_{1T}(\beta)' \\
&= \frac{1}{T} \sum_{t=1}^{[T\tau]} \tilde{\Upsilon}_T \left(\frac{r_t(\beta_1)}{s_T(\beta)} \right) \frac{X_t'}{D_T(\beta)} \\
&+ \left(\frac{1}{T} \sum_{t=1}^{[T\tau]} \tilde{\Upsilon}_T \left(\frac{r_t(\beta_1)}{s_T(\beta)} \right) \frac{r_t(\beta_1)}{s_T(\beta)} + \frac{1}{T} \sum_{t=[T\tau]+1}^T \tilde{\Upsilon}_T \left(\frac{r_t(\beta_2)}{s_T(\beta)} \right) \frac{r_t(\beta_2)}{s_T(\beta)} \right) \frac{\dot{s}_{1T}(\beta)'}{D_T(\beta)}
\end{aligned} \tag{68}$$

with $\tilde{\Psi}_2(x) = \psi_2(z) - z\dot{\psi}_2(z)$, and $\tilde{\Upsilon}_T(z) = W_T(\beta)\Psi_1(z) - \tilde{\Psi}_2(z)$. Then, considering (67) and (68),

$$(65) + (66) = \frac{1}{T} \sum_{t=1}^{[T\tau]} \psi_1 \left(\frac{r_t(\beta_1)}{s_T(\beta)} \right) X_t \left(\gamma_{1T}(\beta) + \gamma_T(\beta) \frac{\dot{s}_{1T}(\beta)'}{D_T(\beta)} \right),$$

with $\gamma_{1T}(\beta)$ and $\gamma_T(\beta)$ defined by (61) and (63), given that $\Upsilon_T(z) - \tilde{\Upsilon}_T(z) = 2z\dot{\psi}_2(z)$. The result for the another derivative is obtained in a similar way. \blacksquare

APPENDIX 2. LEMMATA.

Part A:

The part A of this Lemmata corresponds only to the proof of Theorems 2 and 3, for M-estimators. We assume, with no loss of generality, that $\tau \leq \tau_0$.

Lemma 1 *Assume A.1 and A.2 (except A.2.4),*

$$\text{Sup}_{\tau \in \Pi \subset (0,1)} \text{Sup}_{\theta \in \Theta} \left\| \frac{1}{T} \sum_{t=1}^{T\tau} \dot{\eta}_t(\theta) - \tau M(\theta) \right\| \xrightarrow{P} 0, \tag{69}$$

where $M(\theta)$ is defined by the Assumption A.2.6.

Proof: By the triangular inequality,

$$\begin{aligned}
\text{Sup}_{\tau \in \Pi} \text{Sup}_{\theta \in \Theta} \left\| \frac{1}{T} \sum_{t=1}^{T\tau} \dot{\eta}_t(\theta) - \tau M(\theta) \right\| &\leq \text{Sup}_{\tau \in \Pi} \text{Sup}_{\theta \in \Theta} \left\| \frac{1}{T} \sum_{t=1}^{T\tau} \dot{\eta}_t(\theta) - E[\dot{\eta}_t(\theta)] \right\| \tag{70} \\
&+ \text{Sup}_{\tau \in \Pi} \text{Sup}_{\theta \in \Theta} \left\| \frac{1}{T} \sum_{t=1}^{T\tau} E[\dot{\eta}_t(\theta)] - \tau M(\theta) \right\|. \tag{71}
\end{aligned}$$

By assumption A.2.6, (71) $\rightarrow 0$, while (70) $\xrightarrow{P} 0$ by A.1.1, A.2.3 and A.2.5 applying Lemma A.3 of Andrews (1993). \blacksquare

Lemma 2 Let $\Theta_0 \subset \Theta$ be a compact subset of R^p , containing neighborhoods of θ_0 . Consider a sequence $\{\theta_T, T \geq 1\} \in \Theta_0$ such that $\theta_T \xrightarrow{T \uparrow \infty} \theta_0$. Then, under A.1 and A.2 (except A.2.4),

$$\text{Sup}_{\tau \in \Pi} \left\| \frac{1}{T} \sum_{t=1}^{T\tau} \dot{\eta}_t(\theta_T) - \tau M(\theta_0) \right\| \xrightarrow{P} 0.$$

Proof: By the triangle inequality, the left side of the above expression is bounded by

$$\text{Sup}_{\tau \in \Pi} \left\| \frac{1}{T} \sum_{t=1}^{T\tau} \dot{\eta}_t(\theta_T) - E[\dot{\eta}_t(\theta_T)] \right\| \quad (72)$$

$$+ \text{Sup}_{\tau \in \Pi} \left\| \frac{1}{T} \sum_{t=1}^{T\tau} E[\dot{\eta}_t(\theta_T)] - E[\dot{\eta}_t(\theta_0)] \right\| \quad (73)$$

$$+ \text{Sup}_{\tau \in \Pi} \left\| \frac{1}{T} \sum_{t=1}^{T\tau} E[\dot{\eta}_t(\theta_0)] - \tau M(\theta_0) \right\|. \quad (74)$$

By A.2.6, (74) \xrightarrow{P} 0, while (72) \xrightarrow{P} 0, by A.1.1, A.2.3 and A.2.5, because

$$(72) \leq \text{Sup}_{\tau \in \Pi} \text{Sup}_{\theta \in \Theta_0} \left\| \frac{1}{T} \sum_{t=1}^{T\tau} \dot{\eta}_t(\theta) - E[\dot{\eta}_t(\theta)] \right\| \xrightarrow{P} 0,$$

by Lemma A.3 of Andrews (1993). It remains to examine (73): (i) by tightness condition of $\{F_T, T \geq 1\}$ we obtain that $\frac{1}{T} \sum_{t=1}^T P(z_t \notin C_j) \rightarrow 0$ as $j \rightarrow \infty$, for some sequence of compact sets $\{C_j, j \geq 1\}$ in Z , and (ii) $\forall j \geq 1$:

$$\begin{aligned} & \text{Sup}_{\tau \in \Pi} \left\| \frac{1}{T} \sum_{t=1}^{T\tau} E[\dot{\eta}_t(\theta_T) - \dot{\eta}_t(\theta_0)] I(z_t \in C_j) \right\| \\ & \leq \text{Sup}_{z_t \in C_j} \|\dot{\eta}_t(\theta_T) - \dot{\eta}_t(\theta_0)\| \rightarrow 0 \quad \text{for } \theta_T \rightarrow \theta_0, \end{aligned} \quad (75)$$

for a function $\dot{\eta}_t(\cdot)$ defined in $(z, \theta) \in Z \times \Theta$, continuous by assumption A.1.1, and thus, uniformly continuous in the compact set C_j . So, the difference of the function, evaluated in θ_T and θ_0 will be asymptotically negligible under the supreme metric, as we obtain in (75). (iii) By results (i) and (ii):

$$\text{Sup}_{\tau \in \Pi} \left\| \frac{1}{T} \sum_{t=1}^{T\tau} E[\dot{\eta}_t(\theta_T) - \dot{\eta}_t(\theta_0)] \right\| \rightarrow 0,$$

as $\theta_T \rightarrow \theta_0$. The lemma result follows. ■

Lemma 3 Consider $v_T(\tau) = \frac{1}{\sqrt{T}} \sum_{t=1}^{T\tau} \eta_t$ such as $\{v_T(\tau), T \geq 1\}$ belongs to the bounded cadlag function space in R^p and is defined on $\Pi \subset [0, 1]$. Under A.1 and A.2 (except A.2.5),

$$v_T(\tau) \Rightarrow S^{1/2} B(\tau),$$

where " \Rightarrow " denotes weak convergence, in the $D([0, 1])$ space under the Skorokhod metric. $B(\tau)$ is an p -vector independent motion Brownian process. The process can be defined in $C([0, 1])$, for which exists an equivalence between the Skorokhod and Uniform metrics.

Proof: To prove this lemma we use the results of Lemma A.4 of Andrews (1993), with the difference that the triangular arrays he used, are constant across the subscript T in our case. Then, we have to prove that the process $\{v_T(\tau), T \geq 1\}$ is such that:

- (i) $\forall \alpha \in R^p, \alpha' v_T(\tau) \Rightarrow \alpha' S^{1/2} B(\tau)$,
- (ii) $\{v_T(\tau), T \geq 1\}$ has asymptotically independent increments.

To obtain (i), under A.2.1, A.2.2 and A.2.7, we apply Cor. 3.2 of Wooldridge and White (1988), which utilizes the results of McLeish (1977). Noting that Cor. 3.2. yields weak convergence of the standard partial sum process in $D([0, 1])$ with the Skorohod metric and the σ -algebra generated by it, this can be converted into weak convergence in $D([0, 1])$ with the uniform metric and the σ -algebra generated by the closed balls under the uniform metric. This is treated by Andrews (1993).

To obtain (ii) is enough to prove:

$$\begin{pmatrix} v_T(\tau_2) - v_T(\tau_1) \\ v_T(\tau_0) \end{pmatrix} \xrightarrow{d} N \left(0, \begin{pmatrix} (\tau_2 - \tau_1) S & 0 \\ 0 & \tau_0 S \end{pmatrix} \right), \quad \forall 0 \leq \tau_0 < \tau_1 \leq \tau_2 \leq 1.$$

By Cramér-Wold device, leads to

$$\alpha'_1 v_T(\tau_2) - \alpha'_1 v_T(\tau_1) + \alpha'_2 v_T(\tau_0) \xrightarrow{d} N(0, (\tau_2 - \tau_1) \alpha'_1 S \alpha_1 + \tau_0 \alpha'_2 S \alpha_2), \quad \forall \alpha_1, \alpha_2 \in R^p.$$

To obtain above result, we use again Cor. 3.2 of Wooldridge and White (1988),

$$\begin{aligned} \alpha'_1 v_T(\tau_2) - \alpha'_1 v_T(\tau_1) + \alpha'_2 v_T(\tau_0) &= \alpha'_1 \frac{1}{\sqrt{T}} \sum_{t=T\tau_1+1}^{T\tau_2} \eta_t + \alpha'_2 \frac{1}{\sqrt{T}} \sum_{t=1}^{T\tau_0} \eta_t \\ &\Rightarrow \alpha'_1 S^{1/2} B(\tau_2 - \tau_1) + \alpha'_2 S^{1/2} B(\tau_0) \\ &\stackrel{d}{=} N(0, (\tau_2 - \tau_1) \alpha'_1 S \alpha_1 + \tau_0 \alpha'_2 S \alpha_2). \end{aligned}$$

And we get the desired result. ■

Lemma 4 Let $v_T(\tau)$ be defined as in Lemma 3. Under assumptions A.1 and A.2 (except A.2.5), we obtain that $\forall \alpha \in R^p$ and for any positive constant k :

$$\Pr \left\{ \sup_{\tau \leq \tau_0 - \frac{M}{T}} \alpha' |v_T(\tau)| > k \right\} = O \left(\frac{\|\alpha\|}{k^2} \right),$$

where M is a positive arbitrary large constant.

Proof: Firstly, observe that,

$$\Pr \left\{ \sup_{\tau \leq \tau_0 - \frac{M}{k}} \alpha' |v_T(\tau)| > k \right\} \leq \Pr \left\{ \sup_{\tau \leq 1} \alpha' |v_T(\tau)| > k \right\},$$

and by Lemma 3 we obtain,

$$\alpha' v_T(\tau) \Rightarrow (\alpha' S \alpha)^{1/2} B(\tau),$$

by Cramer-Wold device. Let be $Y = \sup_{\tau} |B(\tau)|$, a process with finite second moment, because $P \{ \sup_{\tau} B(\tau) \leq d \} = 2\Phi(d)$, $d \geq 0$ and $B(\cdot)$ is symmetric around zero. Thus, by Markov inequality,

$$\Pr \left\{ \sup_{\tau \leq 1} \alpha' |v_T(\tau)| > k \right\} \rightarrow \Pr \left\{ Y > k (\alpha' S \alpha)^{-1/2} \right\} \leq \frac{E(Y^2)}{k^2} (\alpha' S \alpha).$$

Considering the Cauchy-Swartz inequality, $(\alpha' S \alpha) = \|S^{1/2} \alpha\|^2 \leq \|S^{1/2}\|^2 \|\alpha\|^2$ and S finite matrix, the result follows. \blacksquare

Lemma 5 *Let $v_T(\tau)$ be as in Lemma 3. Under the assumptions A.1 and A.2 (except A.2.5), we obtain that $\forall \alpha \in R^p$ and for any positive constant k :*

$$\Pr \left\{ \sup_{\tau \leq \tau_0 - \frac{M}{k}} \alpha' \frac{|v_T(\tau_0) - v_T(\tau)|}{\sqrt{\tau_0 - \tau}} > k \right\} = O \left(\frac{\|\alpha\|}{k^2} \right),$$

where M is a positive arbitrary large constant.

Proof: The proof follows as in Lemma 4, defining Y by the absolute value of a standard normal variable, and therefore, with a finite second moment. \blacksquare

Lemma 6 *Let $v_T(\tau)$ be as in Lemma 3. Under assumptions A.1 and A.2 (except A.2.5), we obtain that $\forall \alpha \in R^p$ and for any positive constant k :*

$$\Pr \{ \alpha' |v_T(1) - v_T(\tau_0)| > k \} = O \left(\frac{\|\alpha\|}{k^2} \right).$$

Proof: The proof follows as in Lemma 5. \blacksquare

Part B:

The part B of this Lemmata corresponds only to the proof of Theorem 5, of τ -estimators.

Lemma 7 *Let Θ a compact set in R^p and $[h_1, h_2]$ a closed interval with $h_1 > 0$. Then, under the assumptions B.1 and B.2 we obtain that*

$$\sup_{\tau \in \bar{\Pi}} \sup_{\substack{\theta_i \in \Theta \\ s \in [h_1, h_2]}} \left| \frac{1}{T\tau} \sum_{t=1}^{[T\tau]} \rho_1 \left(\frac{u_t + \theta'_i X_t}{s} \right) - E \left[\rho_1 \left(\frac{u_t + \theta'_i X_t}{s} \right) \right] \right| \xrightarrow{a.s.} 0$$

Proof: As in Lemma 4.3 of Yohai (1986), the proof follows directly from Lemma 1 of Yohai (1974). ■

Lemma 8 By the assumptions B.1 and B.2 we obtain that

$$\sup_{\theta \in \Theta^3} |s_T(\theta) - s(\theta)| \xrightarrow{a.s.} 0,$$

where, given $\theta = (\theta'_1, \theta'_2, \theta'_3)' \in \Theta^3$, $s_T(\theta)$ and $s(\theta)$ are such as,

$$B_T(\theta, \tau, s_T(\theta)) = b \text{ and } B(\theta, \tau, s(\theta)) = b,$$

with

$$B_T(b, \tau, s) = \frac{1}{T} \sum_{t=1}^{[T\tau]} \rho_1 \left(\frac{u_t + b'_1 X_t}{s} \right) + \frac{1}{T} \sum_{t=[T\tau]+1}^{[T\tau_0]} \rho_1 \left(\frac{u_t + b'_2 X_t}{s} \right) + \frac{1}{T} \sum_{t=[T\tau_0]+1}^T \rho_1 \left(\frac{u_t + b'_3 X_t}{s} \right),$$

$$B(b, \tau, s) = \tau E \left[\rho_1 \left(\frac{u_t + b'_1 X_t}{s} \right) \right] + (\tau_0 - \tau) E \left[\rho_1 \left(\frac{u_t + b'_2 X_t}{s} \right) \right] + (1 - \tau_0) E \left[\rho_1 \left(\frac{u_t + b'_3 X_t}{s} \right) \right],$$

for $b = (b'_1, b'_2, b'_3)' \in R^{3p}$, $s > 0$ and $\tau \in (0, 1)$.

Proof: Define $h_1 = \inf_{\theta \in \Theta^3} s(\theta)$ and $h_2 = \sup_{\theta \in \Theta^3} s(\theta)$, such that $h_1 > 0$ and $h_2 < \infty$. By Lemma 7 we obtain that

$$\lim_{T \rightarrow \infty} \sup_{\substack{\theta \in \Theta^3 \\ s \in [h_1, h_2]}} |B_T(\theta, \tau, s) - B(\theta, \tau, s)| = 0. \quad (76)$$

Let ε be such that $0 \leq \varepsilon \leq h_1/2$ and define $g_1(\theta) = B(\theta, \tau, s(\theta) + \varepsilon)$ and $g_2(\theta) = B(\theta, \tau, s(\theta) - \varepsilon)$, then it holds that

$$g_1(\theta) < b < g_2(\theta).$$

Given that $g_1(\cdot)$ and $g_2(\cdot)$ are continuous function, we have that

$$\gamma_1 = \inf_{\theta \in \Theta^3} g_1(\theta) < b < \sup_{\theta \in \Theta^3} g_2(\theta),$$

and define $\delta = \min\{b - \gamma_1, \gamma_2 - b\}$. If (76) holds, then $\exists T_0$ such that $\forall T \geq T_0$,

$$\sup_{\substack{\theta \in \Theta^3 \\ s \in [h_1/2, 2h_2]}} |B_T(\theta, \tau, s) - B(\theta, \tau, s)| < \frac{\delta}{2}. \quad (77)$$

Now, observe that:

a) Given that $s(\theta) - \varepsilon \geq h_1 - h_1/2 = h_1/2$ then $s(\theta) - \varepsilon \in [h_1/2, 2h_2]$, and therefore, by (77),

$$\sup_{\theta \in \Theta^3} |B_T(\theta, \tau, s(\theta) - \varepsilon) - B(\theta, \tau, s(\theta) - \varepsilon)| \leq \frac{\delta}{2},$$

inequality which also holds for the *Inf* in θ . Then, because $\text{Inf}|A - B| > |\text{Inf}(A) - \text{Sup}(B)|$,

$$S_{\text{sup}1} \leq \text{Inf}_{\theta \in \Theta^3} B_T(\theta, \tau, s(\theta) - \varepsilon) \leq S_{\text{sup}2}, \quad (78)$$

where

$$S_{\text{sup}1} = \sup_{\theta \in \Theta^3} B(\theta, \tau, s(\theta) - \varepsilon) - \frac{\delta}{2} \text{ and } S_{\text{sup}2} = \sup_{\theta \in \Theta^3} B(\theta, \tau, s(\theta) - \varepsilon) + \frac{\delta}{2}.$$

Therefore (78) takes to

$$S_{\text{inf}1} \leq \text{Inf}_{\theta \in \Theta^3} B_T(\theta, \tau, s(\theta) - \varepsilon), \quad (79)$$

where $S_{\text{inf}1}$ is defined as $S_{\text{sup}1}$, considering the *Inf* instead of the *Sup*. However, because $S_{\text{inf}1} = \gamma_2 - \delta/2 \geq b + \delta - \delta/2 = b + \delta/2$, by (79),

$$\text{Inf}_{\theta \in \Theta^3} B_T(\theta, \tau, s(\theta) - \varepsilon) \geq b + \frac{\delta}{2}.$$

b) Given that $s(\theta) + \varepsilon \leq h_2 + h_1/2 \leq 2h_2$ then $s(\theta) + \varepsilon \in [h_1/2, 2h_2]$ and therefore, by (77),

$$\sup_{\theta \in \Theta^3} |B_T(\theta, \tau, s(\theta) + \varepsilon) - B(\theta, \tau, s(\theta) + \varepsilon)| \leq \frac{\delta}{2},$$

and because of $\text{Sup}|A - B| \geq |\text{Sup}(A - B)| \geq |\text{Sup}(A) - \text{Sup}(B)|$,

$$S'_{\text{sup}1} \leq \sup_{\theta \in \Theta^3} B_T(\theta, \tau, s(\theta) + \varepsilon) \leq S'_{\text{sup}2},$$

with

$$S'_{\text{sup}1} = \sup_{\theta \in \Theta^3} B(\theta, \tau, s(\theta) + \varepsilon) - \frac{\delta}{2} \text{ and } S'_{\text{sup}2} = \sup_{\theta \in \Theta^3} B(\theta, \tau, s(\theta) + \varepsilon) + \frac{\delta}{2}.$$

Next observe that $S'_{\text{sup}2} = \gamma_1 + \delta/2 \leq b - \delta + \delta/2 = b - \delta/2$.

With the results in a) and b) we have obtained that, given $\varepsilon > 0$, $\exists T_0$ such that $\forall T \geq T_0$

$$\sup_{\theta \in \Theta^3} B_T(\theta, \tau, s(\theta) + \varepsilon) \leq b - \frac{\delta}{2} \leq b \leq b + \frac{\delta}{2} \leq \text{Inf}_{\theta \in \Theta^3} B_T(\theta, \tau, s(\theta) - \varepsilon),$$

and taking into account that $b = B_T(\theta, \tau, s_T(\theta))$, the proof concludes because we have obtained that $\forall \varepsilon > 0$, $\exists T_0$ such that $\forall T \geq T_0$

$$s(\theta) - \varepsilon \leq s_T(\theta) \leq s(\theta) + \varepsilon,$$

uniformly in $\theta \in \Theta^3$. ■

Lemma 9 Under the assumptions B.1 and B.2 we obtain that, for $\theta_i \in \Theta$, $i = 1, 2, 3$, $\theta = (\theta'_1, \theta'_2, \theta'_3)' \in \Theta^3$ and $j = 1, 2$.

$$\text{Sup}_{\tau \in \Pi} \text{Sup}_{\theta \in \Theta^3} \left\| \frac{1}{T} \sum_{t=1}^{[T\tau]} \dot{\psi}_j \left(\frac{u_t + \theta'_i X_t}{s_T(\theta)} \right) X_t X'_t - \tau E \left[\dot{\psi}_j \left(\frac{u_t + \theta'_i X_t}{s(\theta)} \right) X_t X'_t \right] \right\| \xrightarrow{P} 0, \quad (80)$$

$$\text{Sup}_{\tau \in \Pi} \text{Sup}_{\theta \in \Theta^3} \left\| \frac{1}{T} \sum_{t=1}^{[T\tau]} \rho_2 \left(\frac{u_t + \theta'_i X_t}{s_T(\theta)} \right) - \tau E \left[\rho_2 \left(\frac{u_t + \theta'_i X_t}{s(\theta)} \right) \right] \right\| \xrightarrow{P} 0, \quad (81)$$

$$\text{Sup}_{\tau \in \Pi} \text{Sup}_{\theta \in \Theta^3} \left\| \frac{1}{T} \sum_{t=1}^{[T\tau]} \dot{\psi}_j \left(\frac{u_t + \theta'_i X_t}{s_T(\theta)} \right) \left(\frac{u_t + \theta'_i X_t}{s_T(\theta)} \right) - \tau E \left[\dot{\psi}_j \left(\frac{u_t + \theta'_i X_t}{s(\theta)} \right) \left(\frac{u_t + \theta'_i X_t}{s(\theta)} \right) \right] \right\| \xrightarrow{P} 0, \quad (82)$$

$$\text{Sup}_{\tau \in \Pi} \text{Sup}_{\theta \in \Theta^3} \left\| \frac{1}{T} \sum_{t=1}^{[T\tau]} \psi_j \left(\frac{u_t + \theta'_i X_t}{s_T(\theta)} \right) X_t - \tau E \left[\psi_j \left(\frac{u_t + \theta'_i X_t}{s(\theta)} \right) X_t \right] \right\| \xrightarrow{P} 0, \quad (83)$$

$$\text{Sup}_{\tau \in \Pi} \text{Sup}_{\theta \in \Theta^3} \left\| \frac{1}{T} \sum_{t=1}^{[T\tau]} \dot{\psi}_j \left(\frac{u_t + \theta'_i X_t}{s_T(\theta)} \right) \left(\frac{u_t + \theta'_i X_t}{s_T(\theta)} \right) X_t - \tau E \left[\dot{\psi}_j \left(\frac{u_t + \theta'_i X_t}{s(\theta)} \right) \left(\frac{u_t + \theta'_i X_t}{s(\theta)} \right) X_t \right] \right\| \xrightarrow{P} 0, \quad (84)$$

$$\text{Sup}_{\tau \in \Pi} \text{Sup}_{\theta \in \Theta^3} \left\| \frac{1}{T} \sum_{t=1}^{[T\tau]} \psi_j \left(\frac{u_t + \theta'_i X_t}{s_T(\theta)} \right) \left(\frac{u_t + \theta'_i X_t}{s_T(\theta)} \right) - \tau E \left[\psi_j \left(\frac{u_t + \theta'_i X_t}{s(\theta)} \right) \left(\frac{u_t + \theta'_i X_t}{s(\theta)} \right) \right] \right\| \xrightarrow{P} 0, \quad (85)$$

$$\text{Sup}_{\tau \in \Pi} \text{Sup}_{\theta \in \Theta^3} \left\| \frac{1}{T} \sum_{t=1}^{[T\tau]} \dot{\psi}_j \left(\frac{u_t + \theta'_i X_t}{s_T(\theta)} \right) \left(\frac{u_t + \theta'_i X_t}{s_T(\theta)} \right)^2 - \tau E \left[\dot{\psi}_j \left(\frac{u_t + \theta'_i X_t}{s(\theta)} \right) \left(\frac{u_t + \theta'_i X_t}{s(\theta)} \right)^2 \right] \right\| \xrightarrow{P} 0, \quad (86)$$

Proof: We only consider the result (80) because the other ones are obtained in a similar way. Observe that the left side of (80) is upper bounded by

$$\text{Sup}_{\tau \in \Pi} \text{Sup}_{\theta_i \in \Theta} \left\| \frac{1}{T} \sum_{t=1}^{[T\tau]} \dot{\psi}_j \left(\frac{u_t + \theta'_i X_t}{s} \right) X_t X'_t - \tau E \left[\dot{\psi}_j \left(\frac{u_t + \theta'_i X_t}{s} \right) X_t X'_t \right] \right\|, \quad (87)$$

$s \in [s(\theta) - \varepsilon, s(\theta) + \varepsilon]$

because of Lemma 8. By Lemma A.3 of Andrews (1993) we obtain that (87) $\xrightarrow{P} 0$, given that for $j = 1, 2$

$$E \left[\text{Sup}_{\theta \in \Theta^3} \left\| \frac{1}{T} \sum_{t=1}^{[T\tau]} \psi_j \left(\frac{u_t + \theta'_i X_t}{s(\theta)} \right) X_t X'_t \right\|^{1+\varepsilon} \right] < \infty,$$

by assumptions B.1 ($\psi_j(\cdot)$ are bounded) and B.2 ($E[X_t X'_t] < \infty$). ■

Lemma 10 By assumptions B.1 and B.2 we obtain that

$$\text{Sup}_{\theta \in \Theta^3} \|W_T(\theta) - W(\theta)\| \xrightarrow{P} 0, \quad (88)$$

and for $j = 1, 2$,

$$\text{Sup} \|W_{jT}(\theta) - W_j(\theta)\| \xrightarrow{P} 0, \quad (89)$$

where

$$W(\theta) = \frac{\tau E \left[\tilde{\rho}_2 \left(\frac{u_t + \theta'_1 X_t}{s(\theta)} \right) \right] + (\tau_0 - \tau) E \left[\tilde{\rho}_2 \left(\frac{u_t + \theta'_2 X_t}{s(\theta)} \right) \right] + (1 - \tau_0) E \left[\tilde{\rho}_2 \left(\frac{u_t + \theta'_3 X_t}{s(\theta)} \right) \right]}{\tau E \left[\tilde{\psi}_1 \left(\frac{u_t + \theta'_1 X_t}{s(\theta)} \right) \right] + (\tau_0 - \tau) E \left[\tilde{\psi}_1 \left(\frac{u_t + \theta'_2 X_t}{s(\theta)} \right) \right] + (1 - \tau_0) E \left[\tilde{\psi}_1 \left(\frac{u_t + \theta'_3 X_t}{s(\theta)} \right) \right]}$$

and

$$W_j(\theta) = \gamma_j(\theta) + \gamma(\theta) \frac{\dot{s}_j(\theta)}{D(\theta)}.$$

Proof: The result (88) follows by (80) and (81) of Lemma 9. To show (89) we need several previous results. By (83) and (84) in Lemma 9 we obtain that

$$\text{Sup}_{\tau \in \Pi} \text{Sup}_{\theta \in \Theta^3} \left\| \frac{1}{T} \sum_{t=1}^{[T\tau]} \Psi_j \left(\frac{u_t + \theta'_i X_t}{s_T(\theta)} \right) X_t - \tau E \left[\Psi_j \left(\frac{u_t + \theta'_i X_t}{s(\theta)} \right) X_t \right] \right\| \xrightarrow{P} 0, \quad (90)$$

$$\text{Sup}_{\tau \in \Pi} \text{Sup}_{\theta \in \Theta^3} \left\| \frac{1}{T} \sum_{t=1}^{[T\tau]} \tilde{\Psi}_j \left(\frac{u_t + \theta'_i X_t}{s_T(\theta)} \right) X_t - \tau E \left[\tilde{\Psi}_j \left(\frac{u_t + \theta'_i X_t}{s(\theta)} \right) X_t \right] \right\| \xrightarrow{P} 0, \quad (91)$$

by (85) and (86) of Lemma 9,

$$\text{Sup}_{\tau \in \Pi} \text{Sup}_{\theta \in \Theta^3} \left\| \frac{1}{T} \sum_{t=1}^{[T\tau]} \Psi_j \left(\frac{u_t + \theta'_i X_t}{s_T(\theta)} \right) \left(\frac{u_t + \theta'_i X_t}{s_T(\theta)} \right) - \tau E \left[\Psi_j \left(\frac{u_t + \theta'_i X_t}{s(\theta)} \right) \left(\frac{u_t + \theta'_i X_t}{s(\theta)} \right) \right] \right\| \xrightarrow{P} 0,$$

$$\text{Sup}_{\tau \in \Pi} \text{Sup}_{\theta \in \Theta^3} \left\| \frac{1}{T} \sum_{t=1}^{[T\tau]} \tilde{\Psi}_j \left(\frac{u_t + \theta'_i X_t}{s_T(\theta)} \right) \left(\frac{u_t + \theta'_i X_t}{s_T(\theta)} \right) - \tau E \left[\tilde{\Psi}_j \left(\frac{u_t + \theta'_i X_t}{s(\theta)} \right) \left(\frac{u_t + \theta'_i X_t}{s(\theta)} \right) \right] \right\| \xrightarrow{P} 0,$$

such that, these last two results together with (88) imply that,

$$\text{Sup}_{\tau \in \Pi} \text{Sup}_{\theta \in \Theta^3} \left\| \frac{1}{T} \sum_{t=1}^{[T\tau]} \hat{\Upsilon}_T \left(\frac{u_t + \theta'_i X_t}{s_T(\theta)} \right) \left(\frac{u_t + \theta'_i X_t}{s_T(\theta)} \right) - \tau E \left[\hat{\Upsilon} \left(\frac{u_t + \theta'_i X_t}{s(\theta)} \right) \left(\frac{u_t + \theta'_i X_t}{s(\theta)} \right) \right] \right\| \xrightarrow{P} 0,$$

for $\hat{\Upsilon}(z) = W(\theta) \Psi_1(z) - \tilde{\Psi}_2(z)$, and then,

$$\text{Sup}_{\theta \in \Theta^3} \|\gamma_T(\theta) - \gamma(\theta)\| \xrightarrow{P} 0, \quad (92)$$

where

$$\begin{aligned} \gamma(\theta) &= \tau E \left[\hat{\Upsilon} \left(\frac{u_t + \theta'_1 X_t}{s(\theta)} \right) \left(\frac{u_t + \theta'_1 X_t}{s(\theta)} \right) \right] + (\tau_0 - \tau) E \left[\hat{\Upsilon} \left(\frac{u_t + \theta'_2 X_t}{s(\theta)} \right) \left(\frac{u_t + \theta'_2 X_t}{s(\theta)} \right) \right] \\ &\quad + (1 - \tau_0) E \left[\hat{\Upsilon} \left(\frac{u_t + \theta'_3 X_t}{s(\theta)} \right) \left(\frac{u_t + \theta'_3 X_t}{s(\theta)} \right) \right]. \end{aligned}$$

Next, by (83) and (84) we have that,

$$\text{Sup}_{\theta \in \Theta^3} \|\dot{s}_{jT}(\theta) - \dot{s}_j(\theta)\| \xrightarrow{P} 0, \quad (93)$$

for $j = 1, 2$ where

$$\dot{s}_1(\theta) = -\frac{A_1(\theta)}{D(\theta)} \text{ and } \dot{s}_2(\theta) = -\frac{A_2(\theta)}{D(\theta)}$$

with

$$\begin{aligned} A_1(\theta) &= \tau E \left[\psi_1 \left(\frac{u_t + \theta'_1 X_t}{s(\theta)} \right) X_t \right], \\ A_2(\theta) &= (\tau_0 - \tau) E \left[\psi_1 \left(\frac{u_t + \theta'_2 X_t}{s(\theta)} \right) X_t \right] + (1 - \tau_0) E \left[\psi_1 \left(\frac{u_t + \theta'_3 X_t}{s(\theta)} \right) X_t \right], \end{aligned}$$

and

$$D(\theta) = \tau E \left[\tilde{\psi}_1 \left(\frac{u_t + \theta'_1 X_t}{s(\theta)} \right) \right] + (\tau_0 - \tau) E \left[\tilde{\psi}_1 \left(\frac{u_t + \theta'_2 X_t}{s(\theta)} \right) \right] + (1 - \tau_0) E \left[\tilde{\psi}_1 \left(\frac{u_t + \theta'_3 X_t}{s(\theta)} \right) \right]. \quad (94)$$

And, finally we obtain that, by (84)

$$\mathop{Sup}_{\theta \in \Theta^3} \|\gamma_{jT}(\theta) - \gamma_j(\theta)\| \xrightarrow{p} 0, \quad (95)$$

for $j = 1, 2$ where

$$\begin{aligned} \gamma_1(\theta) &= \frac{2\tau E \left[\dot{\psi}_2 \left(\frac{u_t + \theta'_1 X_t}{s(\theta)} \right) \left(\frac{u_t + \theta'_1 X_t}{s(\theta)} \right) X_t \right]}{D(\theta)}, \\ \gamma_2(\theta) &= \frac{2(\tau_0 - \tau) E \left[\dot{\psi}_2 \left(\frac{u_t + \theta'_2 X_t}{s(\theta)} \right) \left(\frac{u_t + \theta'_2 X_t}{s(\theta)} \right) X_t \right] + 2(1 - \tau_0) E \left[\dot{\psi}_2 \left(\frac{u_t + \theta'_3 X_t}{s(\theta)} \right) \left(\frac{u_t + \theta'_3 X_t}{s(\theta)} \right) X_t \right]}{D(\theta)}. \end{aligned}$$

Then, taking into account (92), (93), (94) and (95) the proof is concluded. \blacksquare

Lemma 11 *By assumptions B.1 and B.2 we obtain that*

$$\left\| \frac{1}{\sqrt{T}} \sum_{t=1}^{T\tau} \psi_j \left(\frac{u_t}{s_T(\theta)} \right) X_t - \frac{1}{\sqrt{T}} \sum_{t=1}^{T\tau} \psi_j \left(\frac{u_t}{s(\theta)} \right) X_t \right\| \xrightarrow{p} 0.$$

Proof: By Lemma 8, the result follows as in Lemma 5.1 of Yohai (1985), taking into account that, for this case, the part (i) holds by invariance principle of Donsker. \blacksquare

Lemma 12 *Under B.1 and B.2*

$$\mathop{Sup}_{\tau \in \Pi} \mathop{Sup}_{\theta \in \Theta^3} \left\| \frac{1}{T} \sum_{t=1}^{[T\tau]} (\hat{\eta}_t(\theta_i) + \hat{\eta}_{jt}(\theta_i)) - \tau(M(\theta) + M_j(\theta)) \right\| \xrightarrow{p} 0, \quad (96)$$

where

$$\begin{aligned} M(\theta) &= E \left[\left(W(\theta) \dot{\psi}_1 \left(\frac{u_t + \theta'_i X_t}{s(\theta)} \right) + \dot{\psi}_2 \left(\frac{u_t + \theta'_i X_t}{s(\theta)} \right) \right) X_t X_t' \right], \\ M_j(\theta) &= E \left[W_j(\theta) \psi_1 \left(\frac{u_t + \theta'_i X_t}{s(\theta)} \right) X_t \right]. \end{aligned}$$

Proof: By triangular inequality (96) upper bounded by

$$\begin{aligned} & \text{Sup}_{\tau \in \Pi} \text{Sup}_{\theta \in \Theta^3} \left\| \frac{1}{T} \sum_{t=1}^{[T\tau]} (\dot{\eta}_t(\theta_i) + \dot{\eta}_{jt}(\theta_i)) - \tau(M(\theta) + M_j(\theta)) \right\| \\ & \leq \text{Sup}_{\tau \in \Pi} \text{Sup}_{\theta \in \Theta^3} \left\| \frac{1}{T} \sum_{t=1}^{[T\tau]} \dot{\eta}_t(\theta_i) - \tau M(\theta) \right\| \end{aligned} \quad (97)$$

$$+ \text{Sup}_{\tau \in \Pi} \text{Sup}_{\theta \in \Theta^3} \left\| \frac{1}{T} \sum_{t=1}^{[T\tau]} \dot{\eta}_{jt}(\theta_i) - \tau M_j(\theta) \right\|. \quad (98)$$

(97) $\xrightarrow{P} 0$ by (80) and (88), in Lemmas 9 and 10 respectively, while (98) $\xrightarrow{P} 0$ by (83), in Lemma 9 and (89), in Lemma 10. \blacksquare

Lemma 13 Under B.1, B.2, and the conditions of Lemma 2, given that $\theta_T \xrightarrow[T \uparrow \infty]{} \theta_0$, $\theta_T, \theta_0 \in \Theta^3$,

$$\text{Sup}_{\tau \in \Pi} \left\| \frac{1}{T} \sum_{t=1}^{[T\tau]} (\dot{\eta}_t(\theta_{Ti}) + \dot{\eta}_{jt}(\theta_{Ti})) - \tau(M(\theta_0) + M_j(\theta_0)) \right\| \xrightarrow{P} 0,$$

for $i = 1, 2, 3$, and $j = 1, 2$, where

$$\begin{aligned} M(\theta_0) &= E \left[\left(W(\theta_0) \psi_1 \left(\frac{u_t + \theta'_{0i} X_t}{s(\theta_0)} \right) + \psi_2 \left(\frac{u_t + \theta'_{0i} X_t}{s(\theta_0)} \right) \right) X_t X_t' \right], \\ M_j(\theta_0) &= E \left[W_j(\theta_0) \psi_1 \left(\frac{u_t + \theta'_{0i} X_t}{s(\theta_0)} \right) X_t \right]. \end{aligned}$$

Proof: By Lemma 12, the proof is as in Lemma 2. \blacksquare

Lemma 14 Consider $v_T(\tau) = \frac{1}{\sqrt{T}} \sum_{t=1}^{[T\tau]} \eta_t$, such that $\{v_T(\tau), T \geq 1\}$ belongs to the bounded cadlag functions space in R^p and is defined in $\Pi \subset [0, 1]$. Under B.1 and B.2, it holds the result of Lemma 3.

Proof: By Lemma 8, $\text{Sup}_{\theta} |s_T(\theta) - s(\theta)| \xrightarrow{P} 0$, and for $\theta = (0'_p, 0'_p, 0'_p)' = 0_{3p} \in R^{3p}$, we obtain that $s_T(0) \xrightarrow{P} \sigma_0$, with σ_0 such that

$$E \left[\rho_1 \left(\frac{u_t}{\sigma_0} \right) \right] = b.$$

By Lemma 11, we have that

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{T\tau} \psi_1 \left(\frac{u_t}{s_T(\theta)} \right) X_t \text{ and } \frac{1}{\sqrt{T}} \sum_{t=1}^{T\tau} \psi_2 \left(\frac{u_t}{s_T(\theta)} \right) X_t$$

are asymptotically equivalent to

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{T\tau} \psi_1 \left(\frac{u_t}{\sigma_0} \right) X_t \text{ and } \frac{1}{\sqrt{T}} \sum_{t=1}^{T\tau} \psi_2 \left(\frac{u_t}{\sigma_0} \right) X_t,$$

respectively, and therefore

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{[T\tau]} \eta_t \approx \sigma_0 \frac{1}{\sqrt{T}} \sum_{t=1}^{T\tau} \psi_0 \left(\frac{u_t}{\sigma_0} \right) X_t,$$

where " \approx " denotes asymptotic equivalence, with

$$\psi_0 \left(\frac{u_t}{\sigma_0} \right) = W_0 \psi_1 \left(\frac{u_t}{\sigma_0} \right) + \psi_2 \left(\frac{u_t}{\sigma_0} \right). \quad (99)$$

Then by the invarianza principle of Donsker and Slutsky theorems we obtain that

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{T\tau} \psi_0 \left(\frac{u_t}{\sigma_0} \right) X_t \Rightarrow S^{1/2} B(\tau),$$

with $S = \sigma_0^2 E [\psi_0^2(u_t)] E [X_t X_t']$. The rest of the proof is similar to Lemma 3. ■

APPENDIX 3: TABLES AND FIGURES

TABLE 1

Bias and mean squared for LS, LAD and Huber estimators of structural break point, based on 5000 replications of the model $Y_t = 1 + X_t + \lambda_T I(\tau_t > \tau_0) + u_t$, with $\tau_t = t/T$, $t = 1, \dots, T$, $T = 100, 200$ and 500 , with $\lambda_T = 2.189, 2.0$ and 1.795 respectively. Several distribution for u_t are considered: standard normal, double exponential, t_3 , t_5 and two mixed standard normal distribution, N_{90} and N_{75} , with a 10% and 25% of a normal distribution with variance equal to 9 respectively. The values corresponding to Bias and MSE must be divided by 100.

POINT ESTIMATION		T=100		T=200		T=300	
MODEL	ESTIMATOR	Bias	MSE	Bias	MSE	Bias	MSE
$N(0, 1)$	LS	0.292	2.252	0.418	0.426	0.656	0.312
	LAD	0.390	3.819	0.379	0.738	0.544	0.593
	Huber	0.180	2.395	0.418	0.453	0.583	0.356
$\frac{1}{2} \exp(u)$	LS	-0.364	2.166	-0.062	0.900	-0.037	0.318
	LAD	-0.018	0.987	-0.005	0.351	-0.173	0.138
	Huber	-0.026	2.166	-0.058	0.534	-0.273	0.198
t_3	LS	-1.544	2.972	-0.144	1.527	0.042	0.443
	LAD	-0.306	0.590	-0.084	0.263	0.229	0.089
	Huber	-0.474	0.653	0.084	0.307	0.192	0.104
t_5	LS	-1.000	2.119	-0.189	0.954	0.214	0.311
	LAD	-0.246	1.610	-0.065	0.669	0.359	0.254
	Huber	-0.136	1.283	0.114	0.527	0.296	0.237
N_{90}	LS	-0.156	1.943	-0.525	0.958	-0.028	0.326
	LAD	-0.070	1.339	-0.085	0.621	-0.014	0.217
	Huber	-0.184	1.099	0.027	0.535	-0.025	0.160
N_{75}	LS	0.820	2.039	-0.267	0.949	-0.005	0.032
	LAD	-0.384	0.777	-0.144	0.362	-0.067	0.012
	Huber	-0.122	1.013	-0.228	0.447	-0.012	0.014

TABLE 2

Interval estimation for LS, LAD and Huber estimators of structural break point, based on 5000 replications of the model $Y_t = 1 + X_t + \lambda_T I(\tau_t > \tau_0) + u_t$, with $\tau_t = t/T$, $t = 1, \dots, T$, $T = 100, 200$ and 500 , with $\lambda_T = 2.189, 2.0$ and 1.795 respectively. Several distribution for u_t are considered: standard normal, double exponential, t_3 , t_5 and two mixed standard normal distribution, $N90$ and $N75$, with a 10% and 25% of a normal distribution with variance equal to 9 respectively. The values corresponding to Bias and MSE must be divided by 100.

INTERVAL ESTIMATION		T=100			T=200			T=500		
MODEL	ESTIMATOR	$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.01$
$N(0, 1)$	LS	0.158	0.091	0.036	0.130	0.074	0.018	0.121	0.065	0.016
	LAD	0.120	0.073	0.026	0.088	0.044	0.009	0.088	0.046	0.009
	Huber	0.157	0.096	0.037	0.127	0.073	0.018	0.122	0.067	0.016
$\frac{1}{2} \exp(u)$	LS	0.145	0.093	0.031	0.125	0.070	0.022	0.114	0.062	0.018
	LAD	0.068	0.034	0.007	0.056	0.028	0.004	0.067	0.030	0.006
	Huber	0.148	0.086	0.026	0.118	0.067	0.018	0.114	0.061	0.016
t_3	LS	0.149	0.087	0.032	0.125	0.075	0.025	0.111	0.064	0.016
	LAD	0.052	0.021	0.004	0.058	0.022	0.003	0.057	0.024	0.004
	Huber	0.140	0.078	0.02	0.119	0.067	0.016	0.115	0.062	0.014
t_5	LS	0.144	0.084	0.027	0.127	0.071	0.018	0.116	0.063	0.016
	LAD	0.079	0.038	0.012	0.066	0.034	0.008	0.071	0.032	0.006
	Huber	0.138	0.080	0.023	0.126	0.067	0.017	0.116	0.064	0.018
$N90$	LS	0.140	0.081	0.029	0.136	0.080	0.021	0.114	0.066	0.015
	LAD	0.076	0.039	0.009	0.078	0.034	0.006	0.070	0.034	0.005
	Huber	0.144	0.084	0.026	0.135	0.078	0.023	0.113	0.061	0.016
$N75$	LS	0.147	0.085	0.032	0.133	0.077	0.021	0.111	0.065	0.018
	LAD	0.061	0.094	0.064	0.064	0.030	0.005	0.061	0.028	0.003
	Huber	0.151	0.084	0.028	0.133	0.075	0.021	0.113	0.062	0.015

TABLE 3 AND FIGURE 1

Sample data corresponding to the model $Y_t = 1 + X_t + I(t > T\tau_0) + u_t$ where $x_t \sim i.i.d.N(0, 0.1)$, $u_t \sim i.i.d.N(0, 0.1)$, $\tau_0 = 0.5$ and $T = 30$.

Time	Y	X	Time	Y	X	Time	Y	X
1	1.03	-0.30	11	0.83	-0.14	21	2.08	-0.27
2	0.67	-0.21	12	1.45	-0.00	22	2.84	0.08
3	0.17	-0.16	13	0.87	0.43	23	3.17	0.54
4	1.27	0.06	14	0.98	0.17	24	1.24	-0.41
5	0.87	-0.06	15	1.47	0.08	25	2.50	0.45
6	1.59	0.33	16	1.57	-0.14	26	2.41	0.30
7	1.43	-0.04	17	1.39	-0.28	27	1.34	-0.43
8	0.98	-0.08	18	2.31	0.26	28	1.07	-0.48
9	0.57	-0.04	19	1.89	0.20	29	2.30	0.32
10	1.06	0.07	20	2.18	0.17	30	2.47	-0.37

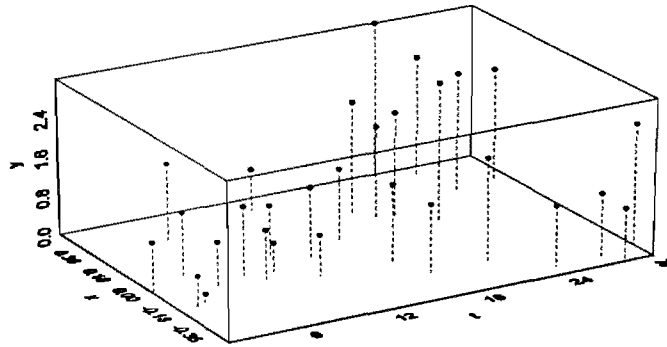


TABLE 4

LS, LAD, Huber and τ - estimator of the break, $[T\hat{\tau}_T]$, and the regression parameters pre-break and post-break for the model $Y_t = 1 + X_t + I(t > T\tau_0) + u_t$ where $x_t \sim i.i.d.N(0, 0.1)$, $u_t \sim i.i.d.N(0, 0.1)$, $\tau_0 = 0.5$ and $T = 30$.

SAMPLE DATA WITH P				
PARAMETER	LS	LAD	HUBER	τ -EST.
Break	15	15	15	15
Constant pre-break	0.99	1.05	1.02	1.03
Constant post-break	2.05	1.89	1.90	1.90
Slope pre-break	0.97	1.62	1.17	1.44
Slope post-break	1.34	1.59	1.51	1.56
SAMPLE DATA WITH P ₁				
PARAMETER	LS	LAD	HUBER	τ -EST
Break	22	15	15	15
Constant pre-break	1.16	1.06	1.02	1.03
Constant post-break	2.07	1.89	1.87	1.89
Slope pre-break	0.60	1.62	1.18	1.40
Slope post-break	1.38	1.59	1.50	1.55
SAMPLE DATA WITH P ₂				
PARAMETER	LS	LAD	HUBER	τ -EST
Break	22	22	22	15
Constant pre-break	1.26	1.25	1.22	1.03
Constant post-break	2.07	1.92	1.91	1.89
Slope pre-break	-0.36	-0.50	-0.42	1.40
Slope post-break	1.38	1.64	1.51	1.55

FIGURE 2

Dependent variable and fit respect to the time for the LS, LAD, Huber and τ - estimators. The model is $Y_t = 1 + X_t + I(t > T\tau_0) + u_t$ where $x_t \sim i.i.d.N(0, 0.1)$, $u_t \sim i.i.d.N(0, 0.1)$, $\tau_0 = 0.5$ and $T = 30$, and the sample data corresponds to \mathcal{P} . The dotted vertical line corresponds to the location break estimate.

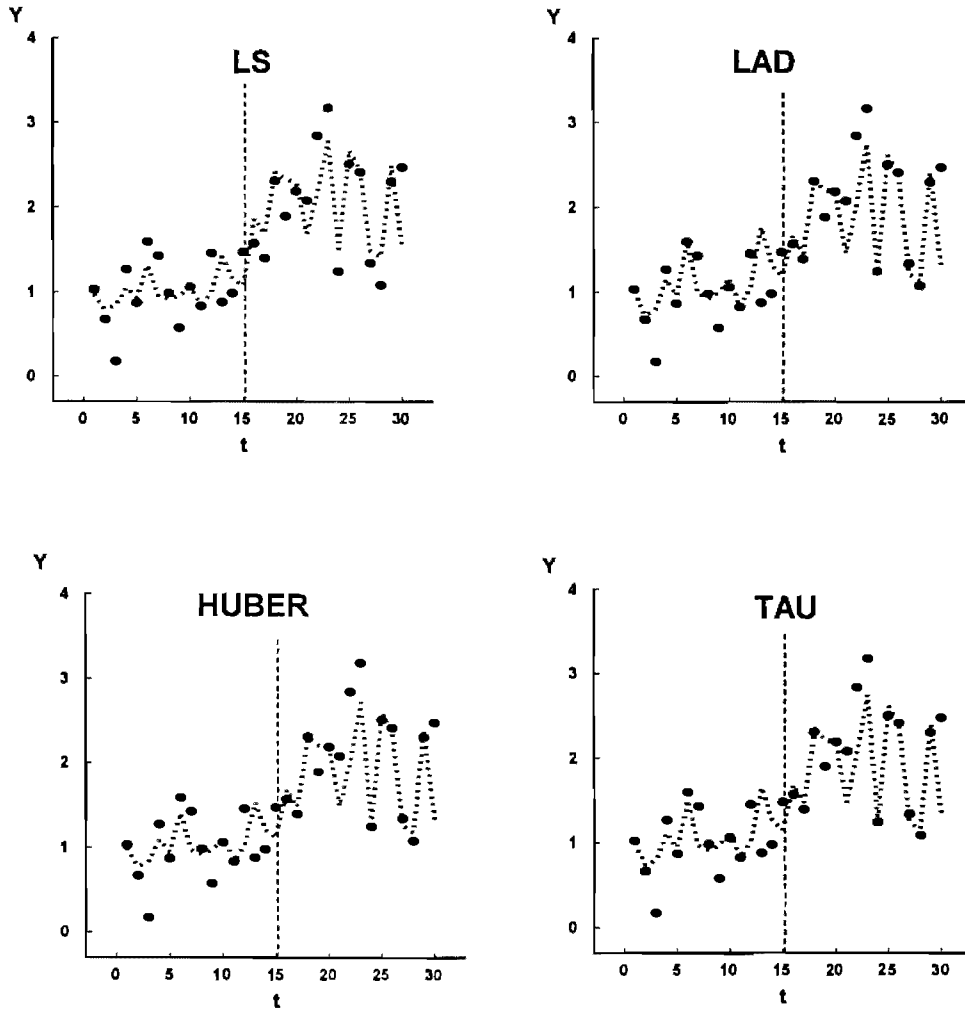


FIGURE 3

Dependent variable and fit respect to the time for the LS, LAD, Huber and τ - estimators. The model is $Y_t = 1 + X_t + I(t > T\tau_0) + u_t$ where $x_t \sim i.i.d.N(0, 0.1)$, $u_t \sim i.i.d.N(0, 0.1)$, $\tau_0 = 0.5$ and $T = 30$, and the sample data corresponds to P_1 . The dotted vertical line corresponds to the location break estimate.

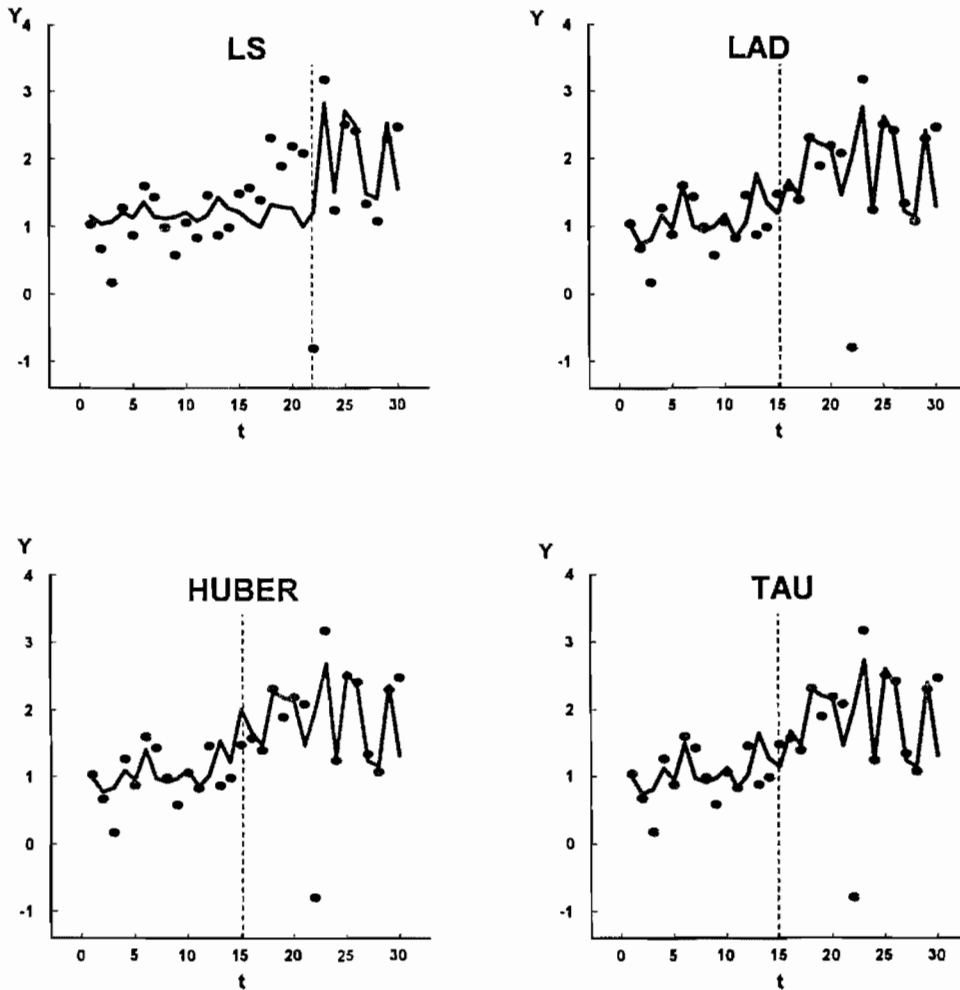


FIGURE 4

Dependent variable and fit respect to the time for the LS, LAD, Huber and τ - estimators. The model is $Y_t = 1 + X_t + I(t > T\tau_0) + u_t$ where $x_t \sim i.i.d.N(0,0.1)$, $u_t \sim i.i.d.N(0,0.1)$, $\tau_0 = 0.5$ and $T = 30$, and the sample data corresponds to P_2 . The dotted vertical line corresponds to the location break estimate.

