Threshold Integrated Moving Average Models

(Does size Matter? Maybe so)

Preliminary draft

Please do not quote it yet

Jesús Gonzalo *	Oscar Martínez†
Statistics and Econometrics	Statistics and Econometrics
U. Carlos III de Madrid	U. Carlos III de Madrid
jgonzalo@est-econ.uc3m.es	omartin@est-econ.uc3m.es

January 2003

Abstract

The aim of this paper is to identify permanent and transitory shocks. This identification is done according to the size of the shocks or the size of some other important economic variable. In order to be able to carry this identification scheme on, we introduce a new class of threshold models: threshold integrated moving average models (TIMA). These are integrated models with a unit root in the moving average of one regime and an invertible moving average in the other regime. The former regime corresponds to transitory shocks, while the latter corresponds to permanent shocks. The paper analyzes the impulse response function generated by TIMA models and its invertibility. Consistency and asymptotic normality of least squares estimators are established and hypothesis tests for TIMA models are developed. The paper concludes with an application to exchange rates and stock market prices.

Keywords: Asymmetries, Moving Averaged Models, Permanent Shock, Persistence, Threshold Models, Transitory Shock.

JEL: C22, C51

*Corresponding Address: http://www.uc3m.es/uc3m/dpto/DEE/profesorado/personal/jgonzalo.

Financial support DGCYT Grant (SEC01-0890) is gratefully acknowledged.

†This research is part of my doctoral thesis

1 Introduction

None doubts that shocks exist. Where there is a huge controversy is about how to identify them. Clearly their permanent and transitory characteristics can help us in this task. For instance, Blanchard and Quah (1989) identify permanent and transitory shocks with supply and demand shocks respectively. More recently Uhlig (1997) and Faust (1998) use sign restrictions on impulse responses to identify policy shocks. In this paper we identify shocks according to their sizes or the sizes of some extra important economic variable. In order to be able to do that, we introduce a new type of threshold models: Threshold Integrated Moving Average Models (TIMA).

TIMA models have a random walk component in the AR part and different threshold regimes in the moving average side. By placing a unit root in some of the moving average regimes we force the shock in that regime to be transitory (the shocks in the other regimes will be permanent), and by characterizing the regimes according to the shock sizes or the sizes of some other relevant variables we can identify the permanent and transitory shocks from these characteristics.

Our identification approach is not only comparable, but superior in many aspects with respect the two existent single equation decomposition methods: Beveridge-Nelson (1981) decomposition and Unobserved Components (UC) from the signal extraction literature (see Maravall (1995) and Watson (1986)).

In the case of Beveridge-Nelson decomposition the same shock is in both components, transitory and permanent. So they only are able to identify one shock. In the Unobserved Components literature “a priori” information is necessary in order to identify the components. With TIMA models we substitute it for the size criterion, with the advantage that this criterion can be tested. So we can identify more completely the shock and more characteristics in the permanent and transitory components are endogenous.

With the size criterion the TIMA model is completely identified, so estimating the model, we are able to identify the permanent and transitory components. In this paper we propose the Least Squares method to estimate it, proving the consistency of the method. Equally, we provide a way to test the suitability of the size as a criterion for identification. Firstly we test the existence of threshold effect. The null hypothesis will be no threshold effect, then under the null the threshold parameter will be not identified. Therefore we propose a supremum type test. Finally, if the null is rejected, we can test the existence of transitory shocks by testing for a unit root in the threshold MA. The unit root in the MA is a fundamental key to obtain persistent and transitory shocks.

Threshold moving average (TMA) are not new, although until now they have not been used to identify the persistent and transitory components of a time series. They were introduced by Wecker (1981) and generalized by Tong (1990). De Gooijer (1998) studies the properties, consistency and inference of these models although unit root in the MA is not allowed. Engle and Smith (1999) introduce the Stochastic Permanent Breaks models, where the size of the shocks guides the degree of persistence, but again no unit root is allowed. Other related works are Guay and Scaillet (2003) and S.K. Elwood (1998).

The rest of the paper is organized as follows. In section 2 we introduce the model and study the main properties of the process, these related to the Impulse Response Function and invertibility. In section 3 we study the estimation of the TIMA models. Next, in section 4 we discuss the inference about the model. In section 5 we study the finite sample perform. In the last section we present applications to exchange rates and to stock prices. All proofs are gathered in the appendix.

2 TIMA Models

The TIMA model is, in its general form:

$$A(L)\,\Delta y_t = A(L)\,x_t = \mu + \varepsilon_t - \theta_t \varepsilon_{t-1} = \mu + \begin{cases} \varepsilon_t - \theta_1 \varepsilon_{t-1} & if \;\; |z_t| > r \\ \varepsilon_t - \theta_2 \varepsilon_{t-1} & if \;\; |z_t| < r \end{cases} \tag{1}$$

where y_t is the series of interest, x_t is its increment, ε_t is the shock of the model and z_t is the threshold variable. With respect to the parameters, r is the threshold parameter, θ_1 and θ_2 are the parameters of the moving average part. $A(L)$ is the autoregressive part with all the roots outside the unit circle, and μ allows a different mean of x_t.

Along the paper we distinguish two fundamental cases. First, $z_t = \varepsilon_{t-1}$, in this case, it is the size of the shock itself which identifies the permanent and transitory component. We denote this case as Shock-Exciting Threshold Integrated Moving average, STIMA. The second case is when the threshold variable, z_t, is predetermined or observable in t. In this case is the size of another variable, z_t, which identifies the components. Both models may have quite different properties.

For the rest of the paper we require the following assumptions, where $f(.)$ is the density function of ε_t, $\|x\|_\gamma = [E(x^\gamma)]^{1/\gamma}$, is the $L_\gamma - norm$ of x and $F_{t-1} = \sigma(\varepsilon_{t-1}, z_{t-1}, \varepsilon_{t-2}, z_{t-2}, \ldots)$ is the σ-field generated by the random variables ε and z until $t-1$

A.0 $\varepsilon_t \; iid\,(0, \sigma_\varepsilon)$, with $\infty > f(\varepsilon_t) > 0$ for $\forall \varepsilon_t$ and $\|\varepsilon_t\|_{2\gamma} < \infty$ with $\gamma > 2$.

A.1 $E(z_t|\, F_{t-2}) = E(z_t) \Rightarrow E[\varepsilon_t 1(|z_{t+1}| > r)|\, w_{t-1}] = E[\varepsilon_t 1(|z_{t+1}| > r)]$.

A.2 z_t is not mean caused by ε_{t-1}, that is, $E(z_t|\, \varepsilon_{t-1}, F_{t-2}) = E(z_t|\, F_{t-2})$.

A.3 $E\left(\theta^2\left(z_t\right)\middle| F_{t-1}\right) < 1$.

A.4 $\left|\theta_1^0 - \theta_2^0\right| = \partial^0 > 0$.

A.5 z_t is $\alpha - mixing$ of size $-a$.

A.6 $0 < \underline{p} \leq P\left(\left|z_t\right| < r\right| F_{t-1}) \leq \overline{p} < 1$.

A.7 $\max_{r \leq \overline{r}} E\left(\varepsilon_{t-1}^r \middle| \left|z_t\right| = z\right) \leq \sigma_{\varepsilon/z}^{\gamma} < \infty$ for $\gamma \geq 2, \forall z \in [0, \overline{r}]$ and

$mv \leq E\left(1\left(r < \left|z_t\right| < r+v\right)\right| F_{t-2}) \leq Mv$, with $m > 0$ and $M < \infty$.

A.8 $\theta^0 \in \Theta = [-1+\delta, 1-\delta] \times [-1+\delta, 1+\delta'] \times (0, \overline{r}]$ with $\delta > 0$

s.t. $(1-\delta)^4 (1-\overline{p}) + (1+\delta')\,\overline{p} \leq \overline{\lambda} < 1$

A.9 $\lambda_2 = [\partial r M + \lambda_1(\theta_1, \theta_2, r)] < 1$ and $\theta_2 > \theta_1$

$$\lambda_1(\theta_1, \theta_2, r) = E\left[\left|\theta_1\right| 1\left(\left|\varepsilon_{t-1}\right| > r\right) + \left|\theta_2\right| 1\left(\left|\varepsilon_{t-1}\right| < r\right)\right]$$

$$M = \max_e f(-r+e) + f(r+e)$$

$$\partial = \left|\theta_1 - \theta_2\right|$$

A.10 $\theta^0 \in \Theta = [-1+\delta, 1-\delta] \times [-1+\delta, 1] \times (0, \overline{r}]$ with $\delta > 0$ s.t.

$\lambda_2^* = [\partial r M^* + \lambda_1^*(\theta_1, \theta_2, r)] < 1$ and $\theta_2 > \theta_1$ with $\sup_k P\left(\left|\varepsilon_t + k\right| < r\right) \leq \overline{p}(r)$

$$M^* = 4 \max_e f(e)$$

$$\lambda_1^*(\theta_1, \theta_2, r) = \left|\theta_1\right|(1 - \overline{p}(r)) + \left|\theta_2\right| \overline{p}(r)$$

We use the definition of $\alpha-mixing$ used in Davidson (1994). The sequence is $\alpha-mixing$ of size $-a_0$ if $\alpha_m = O(m^{-a})$ for some $a > a_0$. The assumption A.0 and A.4 are usual in threshold models. A.1 and A.2 are the assumptions that we use to prove the existence of persistent and transitory shocks. A.3 is used in the invertibility of the model. In estimation and inference sections, $\theta^0 = \left(\theta_1^0, \theta_2^0, r^0, \sigma_\varepsilon^2\right)$ will be the true parameter vector of the processes. A5-A8 are used in the proof of consistency and asymptotic normality for the case of observable threshold variable. Finally A.9 and A.10 have the same purpose but for STIMA case.

After describing the TIMA models and the necessary assumptions, this section focus on the main properties of TIMA models, i.e. persistence (through the Impulse Response Function) and invertibility. We start with the persistence, where we will see that the possibility of a unit root in the MA is a main aspect of the model since it allows to have two kind of shock, transitory and permanent.

2.1 Impulse Response Function

To see the properties of persistence in the shocks the literature traditionally studies the behavior of the impulse response function (IRF). This function tries to measure the effect of a perturbation in t in the sample path $\{y_{t+k}\}_{k=0}^{\infty}$. Thus, if this effect on y_{t+k} does not vanish when $k \to \infty$ we say that the shock is persistent. With the linear models there is a general consensus about the definition of the IRF. However with nonlinear models three main aspects of a series come up that can determine the definition of the IRF and its relationship with the persistence. These aspects are the history of the series at time $t-1$, the future shocks and the size of the shocks. A deep research about this topic can be found in Potter (2000) and Koop, Pesaran and Potter (1996). They define the Generalized Impulse Response Function (GI) as:

$$GI\left(k, \varepsilon_t, w_{t-1}\right) = E\left[y_{t+k} \middle| \varepsilon_t, w_{t-1}\right] - E\left[y_{t+k} \middle| w_{t-1}\right] \quad for\ k = 0, 1, 2, ..$$

where w_t is the history of the process until t. Obviously this new definition affects and complicates the definition of persistence. But basically it would be that a shock is persistent if the effect of knowing it on the expectation of Y_{t+k} conditional on the past does not vanish when $k \to \infty$. If GI really depends on ε_t and F_{t-1}, there will be shocks with different properties of persistence depending on ε_t and F_{t-1}. Then this new definition includes two of the three main aspects. The problem of future shocks is dealt with by averaging them. In words of Koop et al (1996) the GI is an average of what would happen given the present and the past. As we see later, the average of the future can present problems with the long run properties of the shock's response. Now, in order to understand better the behavior of the GI we consider three examples.

Example 1 In this example we study the GI for a general linear model, $y_t = \Psi(L)\varepsilon_t$, with $\Psi(L) = \sum_{j=0}^{\infty} \theta_j L^j$, where L is the lag operator and ϕ_j is constant. Then as it is easy to prove $y_{t+k} = \sum_{j=0}^{k-1} \theta_j \varepsilon_{t+k-j} + \theta_k \varepsilon_t + \sum_{j=1}^{\infty} \theta_{k+j} \varepsilon_{t-j}$ and assuming that ε_t is a martingale difference sequence (hereafter, mds),

$$E\left[y_{t+k} \middle| w_{t-1}\right] = \sum_{j=1}^{\infty} \theta_{k+j} \varepsilon_{t-j}$$

$$E\left[y_{t+k} \middle| \varepsilon_t, w_{t-1}\right] = \theta_k \varepsilon_t + \sum_{j=1}^{\infty} \theta_{k+j} \varepsilon_{t-j}$$

$$GI\left(k, \varepsilon_t, w_{t-1}\right) = \theta_k \varepsilon_t$$

There are two possibilities: $\lim_{k\to\infty}\theta_k = 0$ or $\lim_{k\to\infty}\theta_k \neq 0$. In the first case, all the shocks ε_t are transitory. In the second one, all of them are persistent. Another important aspect is that the sample path response does not depend on the past and future of the series, as the GI. To see that, if we generate two path of shocks $\{\varepsilon_j\}_{j=1}^{\infty}$ and $\{\varepsilon_j^*\}_{j=1}^{\infty}$ s.t. $\varepsilon_j^* = \varepsilon_j$ for $\forall j \neq t$, then $y_{t+k} - y_{t+k}^* = \theta_k(\varepsilon_t - \varepsilon_t^*)$. The long run properties of the sample path response, $y_{t+k} - y_{t+k}^*$, only depends on ϕ_k, then we can conclude that the GI is a good instrument to define the long run properties of the response to ε_t.

Example 2 Consider now the TAR models, $y_t = \phi_1 y_{t-1} 1\left(z_t \in A\right) + \phi_2 y_{t-1} 1\left(z_t \in A^c\right) + \varepsilon_t$, with $1(.)$ the indicator function. Define $v_t = (\varepsilon_t, z_{t+1})$ and w_{t-1} the history of v_t until $t-1$. Then

$$y_{t+k} = \Psi_{t+k}(L)\,\varepsilon_{t+k} = \sum_{j=0}^{\infty}\theta_{t+k,j}\varepsilon_{t+k-j} = \sum_{j=0}^{k-1}\theta_{t+k,j}\varepsilon_{t+k-j} + \theta_{t+k,k}\varepsilon_t + \theta_{t+k,j-k}\varepsilon_{t+k-j}$$

with $\theta_{t+k,0} = 0$, and $\theta_{t+k,j} = \prod_{i=1}^{j}\left[\phi_1 1\left(z_{t+k-i} \in A\right) + \phi_2 1\left(z_{t+k-i} \in A^c\right)\right]$

$$GI\left(k, v_t, w_{t-1}\right) = \sum_{j=0}^{k-1}\left[E\left(\theta_{t+k,j}\varepsilon_{t+k-j}\middle| w_{t-1}, \varepsilon_t\right) - E\left(\theta_{t+k,j}\varepsilon_{t+k-j}\middle| w_{t-1}\right)\right] + E\left(\theta_{t+k,k}\varepsilon_t\middle| w_{t-1}\right)$$

with $\lambda = E\left[\phi_1 1\left(z_t \in A\right) + \phi_2 1\left(z_t \in A^c\right)\right]$. Now the GI is non-linear and depends on the past of v_{t-1}. For general case the problem is to obtain the expectation of the first summand conditioning on v_t and w_{t-1}. This problem can be solved by simulation and considering the GI as a random variable (see Koop et al. (1996) and Potter (2000)). In any case, it can be proved that if y_t is $\alpha - mixing$, $\lim_{k\to\infty} GI\left(k, v_t, w_{t-1}\right) = 0$ and the probability of a permanent change in the sample path defined as the difference of y_{t+k} and y_{t+k}^* is 0. Then in that case the GI is a good instrument to define the long run properties of the response to ε_t.

In none of the previous examples it is possible to obtain persistent and transitory shocks in the same path of shocks with positive probability. However, as we prove now this behavior can be generated by TIMA models. First we calculate the exact expression of the GI for these models. From the definition of $TIMA$ models, we obtain

$$y_{t+k} = \frac{1}{A(L)(1-L)}\mu + \frac{1-\theta_1 L}{A(L)(1-L)}\varepsilon_{t+k-1} 1\left(|z_{t+k}| > r\right) + \frac{1-\theta_2 L}{A(L)(1-L)}\varepsilon_{t+k-1} 1\left(|z_{t+k}| < r\right)$$

Taking $A(L) = 1 - \phi L$ to simplify the calculations (in the more general case the results still hold), we can write:

$$y_{t+k} = \frac{1}{(1-\phi L)(1-L)}\mu + \frac{1}{(1-\phi L)}\varepsilon_{t+k-1}1\left(|z_{t+k}|>r\right) + \frac{1-\theta_2 L}{(1-\phi L)(1-L)}\varepsilon_{t+k-1}1\left(|z_{t+k}|<r\right) \quad (2)$$

In order to prove the relation between a unit root in some of the MA regimes and the existence of transitory and permanent shocks we use the following lemma

Lemma 1. *Let y_t be a TIMA process as (2); under A.0 and one of the assumptions A.1 or A.2, the GI of y_t is given by*

$$GI\left(k,\varepsilon_t,F_{t-1}\right) = \begin{cases} \phi^{k-1}\varepsilon_t \quad if \;\; \theta_{t+1}=1 \\ \left[(1-\theta)\sum_{j=0}^{k-2}\phi^j + \phi^{k-1}\right]\varepsilon_t \quad if \;\; \theta_{t+1}=\theta\neq 1 \end{cases}$$

Clearly, when $\theta_{t+1}=1$, which implies a unit root in one of the MA part, the shock ε_t will be transitory, since its effect goes to zero when $k\rightarrow\infty$. On the other hand, when $\theta_{t+1}\neq 1$, the shock ε_t will be persistent. It is worth to note that is the size of z_{t+1}, what determines the persistent or transitory effect of ε_t. Now, with the result of the previous lemma we present the example 3.

Example 3 Now we consider the case of $z_t=\varepsilon_{t-1}$, that is, the STIMA model. It is easy to prove that the size of the own shock in the STIMA model can determine its persistence or transitory properties if we allow for a unit root in a regime of the MA. Suppose that $\theta_2=1$, since ε_{t-1} satisfies the assumption $A.0$ and $A.1$ we have

$$GI\left(k,\varepsilon_t,F_{t-1}\right) = \begin{cases} \phi^{k-1}\varepsilon_t \quad if \;\; |\varepsilon_t|<r \\ \left[(1-\theta)\sum_{j=0}^{k-2}\phi^j + \phi^{k-1}\right]\varepsilon_t \quad if \;\; |\varepsilon_t|>r \end{cases}$$

Then, if $|\varepsilon_t|<r$ (that is, the size of the shock is small), ε_t will be transitory. When $|\varepsilon_t|>r$, ε_t is permanent. As we said, note that although the threshold in t depends on ε_{t-1}, it is the size of ε_t, who lays down its transitory or persistent effect. As in the previous examples, when ε_t and ε_t^* are in the same regime, the GI will be a good definition of the long run properties of the sample path response, $y_{t+k}-y_{t+k}^*$.

Another interesting model is when $z_t=x_{t-1}$. This is an special case that does not satisfy any of A.1 and A.2 assumptions, even more, in general, in this model all the shocks are persistent. Attending to the definition of persistent shock that is based on this GI, the shock will be persistent if it affects in a permanent way the expectation of the series. The problem of this definition arises when the expectation changes although the future series, that is, the sample path response, does not change with a positive probability.

As in the previous examples, if we generate two path of shocks $\{\varepsilon_j\}_{j=1}^{\infty}$ and $\{\varepsilon_j^*\}_{j=1}^{\infty}$ s.t. $\varepsilon_j^* = \varepsilon_j$ for $\forall j \neq t$, and $1(|x_{t+i}| < r) = 1(|x_{t+i}^*| < r)$ for $i = 0, 1$, it can be proved that $\{y_j\}_{j=t+1}^{\infty} = \{y_j^*\}_{j=t+1}^{\infty}$. In this case, the shock ε_t^* does not affect the path of y_{t+k}.

In this context, if the probability of generating a persistent effect on y_{t+k} is positive, our expectation about y_{t+k} will change in a persistent way. But in the TIMA model with $z_t = x_{t-1}$ there is also a positive probability that the shock does not change the sample path. In that case, this GI is not a good definition of the long run properties of the shocks. Clearly the introduction of these kind of models requires a wider definition of persistent shock, based on the GI, able to resume better the effect of the shocks in the sample path of the series. A possible solution would be to include future events on the actual information set.

2.2 Invertibility

In this section we study the invertibility of TIMA models. At first, we do not focus on the level of the series, y_t, but in the increment, x_t. Like in the IRF case, the non-linearity of TIMA models disables the use of the classical definition of invertibility. Then we use the general invertibility definition introduced by Granger and Andersen (1978) and later improved by Hallin (1980). He proved that under nonlinearity with constant coefficients both definitions are equivalent. Clearly, the TIMA models are within this class. The definition of Granger and Andersen is the following:

Definition 1. *(Granger and Andersen) The process $x_t = g(x_{t-1}, \varepsilon_{t-1}, ..., x_{t-p}, \varepsilon_{t-p}) + \varepsilon_t$ will be invertible if*

$$\lim_{t\to\infty} E\left(e_t^2\right) = 0$$

with

$$e_t = \varepsilon_t - \widehat{\varepsilon}_t = \varepsilon_t - (x_t - g(x_{t-1}, \widehat{\varepsilon}_{t-1}, ..., x_{t-p}, \widehat{\varepsilon}_{t-p}))$$

The results about the invertibility property for TIMA models are summarized in the following lemmas:

Lemma 2. *Let x_t be the first difference of a STIMA process; under A.0 and A.9 x_t is invertible.*

Lemma 3. *Let x_t be the first difference of a TIMA process, under A.0 and A.3 x_t is invertible.*

The proof of both lemmas can be found in the appendix.

Lemma 2 allows that $\theta_2 = 1$; this is the case of small shock transitory, and big shock persistent. Awfully this assumption excludes the opposite case, big transitory shock.

If we study the assumption A.9 we can distinguish two different parts. In the first one, we have $\lambda_1(\theta_1, \theta_2, r)$, that must be less than 1. This checks the case of overdifferentiability, specially when we allow a unit root in the MA. But the non-invertibility is not only a problem of overdifferentiability, it is a problem of non-linearity too. The second part, ∂rM, checks the non-linearity degree, measured as a product of the *gap*, ∂r, and its probability, M. For the case of $\theta_1 = 1$, (small persistent shock), the conditions for invertibility are quite restrictive preventing the estimation of the model[1]. Therefore, in this paper we focus on identifying the big shocks as persistent in the case of STIMA models.

In the case of TIMA models, there is no restriction about the non-linearity, since the model is linear in the shocks. For the case $z_t = x_{t-1}$, the process will satisfy the assumption A.3 if $f(\varepsilon_t) > 0$ for all $\varepsilon_t \in R$.

3 Estimation

In this section we propose the Least Squares (LS) method to estimate the parameters of TIMA models. We concentrate in the simple case of TIMA:

$$\Delta y_t = x_t = \varepsilon_t - \theta_t \varepsilon_{t-1} = \begin{cases} \varepsilon_t - \theta_1 \varepsilon_{t-1} & if \ |z_t| > r \\ \varepsilon_t - \theta_2 \varepsilon_{t-1} & if \ |z_t| < r \end{cases}$$

The parameters to estimate are $(\theta_1, \theta_2, r, \sigma_\varepsilon^2)$ where σ_ε^2 is the variance of ε_t. We prove that all the parameters are identified and that the estimators are consistent under several assumptions. As the previous section, we distinguish between the case of STIMA and observed threshold variable, since the properties of estimators and the way to prove them are quite different.

In general, for Threshold Autoregressive (TAR) models, the rate of convergence for the estimators of the parameter is $T^{1/2}$ except for r, the threshold parameter, which is T. That result is due to the kind of discontinuity of the model in r via the indicator function. In the case of TIMA models we obtain the same results, since the model is continuous in θ_i, but nor in r through the indicator function. For the STIMA case, the rate of convergence is T for all the estimators, since the threshold variable is not observed and must be estimated. In this way, all the parameters to estimate enter in the indicator function.

For the rest of the paper, $\theta^0 = (\theta_1^0, \theta_2^0, r^0, \sigma_\varepsilon^2)$ will be the true parameter vector of the process. In the proof of consistency r^0 is unknown. In the TIMA case, the assumption of

[1] In fact, the condition for invertibility in that case exists, although it implies that the shocks do not have moments greater than 2.

r^0 known simplifies the proof, since the model is continuous in the parameters to estimate. However in the case of STIMA models, knowing r^0 simplifies in no way the proof, since the model is still discontinuous in the rest of the parameters.

3.1 Observable Threshold Variable

In this section we define the LS estimator for the parameters as the value that minimizes a properly objective function that we will define later. The assumptions for this case are A.0 and A.4-A.8. These six assumptions for consistency are similar to these of Hansen (2000) about threshold estimation, except that we assume *iid* in ${\varepsilon_t}^2$ and we do not assume continuity of the moments in r. Instead of this, we bounds the conditional and unconditional moments. Finally, the consistency proof is limited to the invertible case, since assumption A.8 implies assumption A.3.

Taking $\theta = (\theta_1, \theta_2, r)$, the objective function will be the sum of the square estimated errors, that is

$$Q_T(\omega, \theta) = \sum_{t=1}^{T} e_t^2(\omega, \theta)$$

with

$$e_t(\omega, \theta) = \theta_{t-1} e_{t-1}(\omega, \theta) + x_t \qquad e_0 = 0$$

$$\theta_{t-1} = \begin{cases} \theta_1 & if \ \ |z_{t-1}| > r \\ \theta_2 & if \ \ |z_{t-1}| < r \end{cases}$$

We define the LS estimator of θ as $\widehat{\theta}_T$, that must satisfy:

$$\widehat{\theta}_T = \arg\min_{\theta \in \Theta} Q_T(\omega, \theta)$$

Following Potcher and Prucha (1997) to prove the consistency of the estimator we define the following distance:

$$\rho\left(\theta^1, \theta^2\right) = \max\left\{\max_{i=1,2}\left\{\left|\theta_i^1 - \theta_i^2\right|\right\}, \left|r^1 - r^2\right|\right\}$$

We say that $\widehat{\theta}_T$ is consistent if $\rho\left(\widehat{\theta}_T, \theta^0\right) \xrightarrow{p} 0$. A sufficient condition for $\rho\left(\widehat{\theta}_T, \theta^0\right) \xrightarrow{p} 0$, is that for $\forall \epsilon > 0$

$$\lim\inf_{T\to\infty} P\left(\left[\inf_{\theta|\rho(\theta,\theta^0)>\epsilon} Q_T(\omega,\theta) - Q_T\left(\omega,\theta^0\right)\right] > 0\right) \overset{T\to\infty}{\longrightarrow} 1$$

[2] This assumption can be reduced to a stationary $\alpha - mixing$ martingale difference sequence.

where $Q_T(\omega,\theta^0) = \sum_{t=1}^T \varepsilon_t^2$. To obtain this equality we are assuming that $\varepsilon_0 = 0$. In fact it is an unnecessary assumption since using A.6 it can be proved that $\left|\frac{1}{T}Q_T(\omega,\theta^0) - \frac{1}{T}\sum_{t=1}^T \varepsilon_t^2\right| \overset{p}{\longrightarrow} 0$.

It easy to prove that

$$\begin{aligned}
\inf_{\theta|\rho(\theta,\theta^0)>\epsilon} Q_T(\omega,\theta) - Q_T(\omega,\theta^0) &= \inf_{\theta|\rho(\theta,\theta^0)>\epsilon} \sum_{t=1}^T \left[e_t^2(\omega,\theta) - \varepsilon_t^2\right] \\
&\leq \inf_{\theta|\rho(\theta,\theta^0)>\epsilon} \sum_{t=1}^T \left(\theta_{t-1}e_{t-1}(\omega,\theta) - \theta_{t-1}^0\varepsilon_{t-1}\right)^2 + \\
&+ \inf_{\theta|\rho(\theta,\theta^0)>\epsilon} \sum_{t=1}^T 2\varepsilon_t\left(\theta_{t-1}e_{t-1}(\omega,\theta) - \theta_{t-1}^0\varepsilon_{t-1}\right)
\end{aligned}$$

So the scheme of the proof is based on the proof of the two followings claims,

Claim 1. *For all $\epsilon > 0$ and $\eta > 0$*

$$P\left[\sup_{\theta|\rho(\theta,\theta^0)>\epsilon} \frac{\sum_{t=1}^T 2\varepsilon_t\left(\theta_{t-1}e_{t-1}(\omega,\theta) - \theta_{t-1}^0\varepsilon_{t-1}\right)}{T} < \eta\right] \overset{T\to\infty}{\longrightarrow} 1$$

Claim 2. *For all $\epsilon > 0$ and $\eta > 0$ $\exists a(\epsilon) > 0$ for all $\epsilon > 0$ s.t*

$$P\left[\inf_{\theta|\rho(\theta,\theta^0)>\epsilon} \frac{1}{T}\sum_{t=1}^T \left(\theta_{t-1}e_{t-1}(\omega,\theta) - \theta_{t-1}^0\varepsilon_{t-1}\right)^2 > a(\epsilon) - \eta\right] \overset{T\to\infty}{\longrightarrow} 1$$

Then, taking $\eta < a(\epsilon)/2$ we have the result.

To prove these both claims we use

$$\left(\theta_{t-1}e_{t-1}(\omega,\theta) - \theta_{t-1}^0\varepsilon_{t-1}\right) = \left(\theta_{t-1}e_{t-1}(\omega,\theta) - \theta_{t-1}'e_{t-1}'\right) + \left(\theta_{t-1}'e_{t-1}' - \theta_{t-1}^0\varepsilon_{t-1}\right)$$

where $\theta = (\theta_1, \theta_2, r)$ and $\theta' = \left(\theta_2^0, \theta_2^0, r\right)$ and follow the proof of Hansen (2000) and Chan (1993) for consistency fitted to this more complex case. For a given r, $\theta_{t-1}e_{t-1}(\omega,\theta)$ is differentiable with respect to $\theta = (\theta_1,\theta_2)$, and using a Taylor expansion we obtain

$$\begin{aligned}
(e_t - e_t') &= \left(\theta_{t-1}e_{t-1}(\omega,\theta) - \theta_{t-1}'e_{t-1}'\right) \qquad (3)\\
&= \sum_{i=1}^2 \left(\theta_i - \theta_i^0\right) h_t^i(\theta') + \sum_{i=1}^2\sum_{l=1}^2 \left(\theta_i - \theta_i^0\right) h_t^{il}(\theta^*)\left(\theta_l - \theta_l^0\right) \\
&= \sum_{i=1}^2 \left(\theta_i - \theta_i^0\right) h_t^i(\theta^0) + \sum_{i=1}^2 \left(\theta_i - \theta_i^0\right)\left(h_t^i(\theta^0) - h_t^i(\theta')\right) + \sum_{i=1}^2\sum_{l=1}^2 \left(\theta_i - \theta_i^0\right) h_t^{il}(\theta^*)\left(\theta_l - \theta_l^0\right)
\end{aligned}$$

with

$$h_t^i(\theta') = \left.\frac{\partial e_t}{\partial \theta_i}\right|_{\theta'} \quad h_t^{il}(\theta^*) = \left.\frac{\partial^2 e_t}{\partial \theta_i \partial \theta_l}\right|_{\theta^*} \quad \theta_i^* \in \left(\theta_i, \theta_i^0\right)$$

The proof of both claims are in the appendix.

Finally, as we said in the beginning of the section we prove in this case that the rate of convergence to the true value of the estimators is $T^{1/2}$ in the case of θ_1 and θ_2 and T for r. This will state the inference in the model. This result is in the following theorem:

Theorem 1. *Under* A.0, *and* A.4-A.8, $\widehat{\theta}_{i,T} = \theta_i^0 + O_p\left(T^{-1/2}\right)$ *and* $\widehat{r}_T = r_0 + O_p\left(T^{-1}\right)$.

The proof of the theorem can be found in the appendix.

3.2 Shock-Exciting TIMA

The specific assumptions to obtain the consistency of the LS estimator in the case of STIMA are A.0, A.4 and A.10. In this paper we only prove the consistency for the case of x_t invertible, as it is reflected in assumption $A.10$, which implies $A.9$. With respect to the STIMA process that satisfy these assumptions, at first, the smaller are the kurtosis and r, the bigger is Θ. The main problem of this condition is that it depends on $f(.)$, the density function of ε_t, although we can obtain lower bounds for r given (θ_1, θ_2) for unknown $f(.)$. As in the observable case, the objective function is

$$Q_T(\omega,\theta) = \sum_{t=1}^{T} e_t^2(\omega,\theta)$$

with

$$e_t(\omega,\theta) = \theta_{t-1} e_{t-1}(\omega,\theta) + x_t \qquad e_0 = 0$$

$$\theta_{t-1} = \begin{cases} \theta_1 & if \ \ |e_{t-1}| > r \\ \theta_2 & if \ \ |e_{t-1}| < r \end{cases}$$

We will follow the same steps to prove consistency, then we must prove

$$\lim_{T\to\infty} \inf \left\{ P\left[\inf_{\{\theta\in\Theta:\rho(\theta,\theta^0)\geq\varepsilon\}} Q_T(\omega,\theta) - Q_T\left(\omega,\theta^0\right) \right] > 0 \right\} = 1$$

The main change is in θ_{t-1}. Now z_t is not observable and we must estimate it. As we will see this is a very important issue , although at first glance, the decomposition of $Q_T(\omega,\theta) - Q_T\left(\omega,\theta^0\right)$ does not change, and we have

$$\inf_{\{\theta\in\Theta:\rho(\theta,\theta^0)\geq\varepsilon\}} \frac{1}{T} Q_T(\omega,\theta) - \frac{1}{T} Q_T\left(\omega,\theta^0\right) \geq \inf_{\{\theta\in\Theta:\rho(\theta,\theta^0)\geq\varepsilon\}} \frac{1}{T} \sum_{t=1}^{T} \left(\theta_{t-1} e_{t-1}(\omega,\theta) - \theta_{t-1}^0 \varepsilon_{t-1}\right)^2$$
$$+ \inf_{\{\theta\in\Theta:\rho(\theta,\theta^0)\geq\varepsilon\}} \frac{1}{T} \sum_{t=1}^{T} 2\varepsilon_t \left(\theta_{t-1} e_{t-1}(\omega,\theta) - \theta_{t-1}^0 \varepsilon_{t-1}\right)$$

Again we have the same two claims to prove. As in the observable case, we identify $Q_T\left(\omega, \theta^0\right) = \sum_{t=1}^{T} \varepsilon_t^2$ using the same argument. The main difference in the STIMA case is in the proof of claim 1, where there is no differentiability with respect to any parameter. All the proofs are in the appendix.

Another important difference is the rate of convergence of the estimator. As we said at the beginning of the section the rate of convergence for this case is T. This is established in the following theorem,

Theorem 2. *Under* A.0*,* A.4 *and* A.10 $\widehat{\theta}_{i,T} = \theta_i^0 + O_p\left(T^{-1}\right)$ *and* $\widehat{r}_T = r_0 + O_p\left(T^{-1}\right).$

The proof of this theorem can be found in the appendix. As we said, the reason which produces the change in the rate of convergence of θ_i is the continuity. In STIMA models, all the parameters enter in the indicator function through $e_t(\theta)$, and the objective function is not continuous in any parameter.

4 Inference

In this section we present the way to test the main hypothesis about the model. Clearly, the most important hypothesis is the existence of transitory shocks depending on the size. This implies two kind of tests. Firstly, we must test that both parameters, θ_1 and θ_2, are different. The null hypothesis is $\theta_1 = \theta_2$. Under the null hypothesis, the parameter r is not defined. This issue has been broadly studied in TAR models. This problem was pointed out by Davies (1977,1987); see also Andrews and Ploberger (1994) and Andrews (1994). We will apply the approximation of Hansen (1996). Guay and Scaillet (2003) adapt this procedure to threshold moving average models for indirect inference.

Secondly, we must contrast that $\theta_i = 1$. To do that, we need to know the asymptotic distribution of the estimators. This will be made for the TIMA models when the threshold variable is observable and r^0 is known. When r^0 is unknown we estimate it and using that it converges at T ratio, we can take it as known. In the case of STIMA, we propose to estimate in two steps. In the second one we can use that the rate of all the estimators is T to consider the shocks as known. In that case, the second estimation is in a TIMA model with threshold observable variable.

In the following subsection we prove the asymptotic normality of the estimator of θ_i.

4.1 Asymptotic Normality

In this subsection we prove the asymptotic normality of the estimator when r^0 is known. For a given r, the objective function, as we saw, is differentiable with respect to θ_i, and then

$$Q_T(\omega,\theta)-Q_T(\omega,\theta^0)=\sum_{t=1}^{T}\left(\theta_{t-1}e_{t-1}(\omega,\theta)-\theta_{t-1}^0\varepsilon_{t-1}\right)^2+\sum_{t=1}^{T}2\varepsilon_t\left(\theta_{t-1}e_{t-1}(\omega,\theta)-\theta_{t-1}^0\varepsilon_{t-1}\right)$$

If we define

$$\begin{aligned}H_{t-1}^{i,l}(\theta^*) &= \left.\frac{\partial^2\left(\theta_{t-1}e_{t-1}(\omega,\theta)-\theta_{t-1}^0\varepsilon_{t-1}\right)^2}{\partial\theta_i\partial\theta_l}\right|_{\theta^*}\\ &= 2h_{t-1}^l(\theta^*)\,h_{t-1}^i(\theta^*)+2\left(\theta_{t-1}^*e_{t-1}^*(\omega,\theta^*)-\theta_{t-1}^0\varepsilon_{t-1}\right)h_{t-1}^{i,l}(\theta^*)\end{aligned}$$

and applying a second Taylor expansion and using that $\left.\frac{\partial\left(\theta_{t-1}e_{t-1}(\omega,\theta)-\theta_{t-1}^0\varepsilon_{t-1}\right)^2}{\partial\theta_i}\right|_{\theta^0}=0$

$$\begin{aligned}Q_T\left(\omega,\widehat{\theta}\right)-Q_T(\omega,\theta^0) &= \frac{1}{2}\sum_{i=1}^{2}\sum_{l=1}^{2}\sum_{t=1}^{T}\left(\widehat{\theta}_i-\theta_i^0\right)H_{t-1}^{i,l}(\theta^*)\left(\widehat{\theta}_l-\theta_l^0\right)+\\ &+\sum_{i=1}^{2}\sum_{t=1}^{T}2\left(\widehat{\theta}_i-\theta_i^0\right)\varepsilon_t h_t^i(\theta^0)+\frac{1}{2}\sum_{i=1}^{2}\sum_{l=1}^{2}\sum_{t=1}^{T}2\left(\widehat{\theta}_i-\theta_i^0\right)\varepsilon_t h_t^{il}(\theta^*)\left(\widehat{\theta}_l-\theta_l^0\right)\end{aligned}$$

with $\rho(\theta^*,\theta^0)<\rho\left(\widehat{\theta},\theta^0\right)$ and $\widehat{\theta}=\arg\min_{\theta\in\Theta}Q_T(\omega,\theta)$. Now using the results of the consistence and rate of convergence of $\widehat{\theta}$ of previous sections we have that for all $\eta>0$

$$\lim_{T\to\infty}P\left[\sup_{\theta^*\in\Theta}\left|\frac{1}{2}\frac{1}{T}\sum_{i=1}^{2}\sum_{l=1}^{2}\sum_{t=1}^{T}2\varepsilon_t h_t^{il}(\theta^*)\right|>\eta\right]=0$$

and $\widehat{\theta}$ is $O_p\left(T^{-1/2}\right)$, then $T\left(\widehat{\theta}_i-\theta_i^0\right)\left(\widehat{\theta}_l-\theta_l^0\right)=O_p(1)$, and

$$\frac{1}{2}\sum_{i=1}^{2}\sum_{l=1}^{2}\sum_{t=1}^{T}2\left(\widehat{\theta}_i-\theta_i^0\right)\varepsilon_t h_t^{il}(\theta^*)\left(\widehat{\theta}_l-\theta_l^0\right)=o_p(1)$$

Define

$$H_T(\theta)=\frac{1}{T}\sum_{t=1}^{T}\begin{bmatrix}H_{t-1}^{1,1}(\theta) & H_{t-1}^{1,2}(\theta)\\ H_{t-1}^{2,1}(\theta) & H_{t-1}^{2,2}(\theta)\end{bmatrix}$$

$$\lim_{T\to\infty}E\left(H_T(\theta^0)\right)=H(\theta^0)$$

$$D_T(\theta^0)=\frac{1}{T^{1/2}}\sum_{t=1}^{T}\begin{bmatrix}\varepsilon_t h_t^1(\theta^0)\\ \varepsilon_t h_t^2(\theta^0)\end{bmatrix}$$

$$\Omega=\begin{bmatrix}E\left(\varepsilon_t h_t^1(\theta^0)\right)^2 & E\left(\varepsilon_t^2 h_t^1(\theta^0)\,h_t^2(\theta^0)\right)\\ E\left(\varepsilon_t^2 h_t^1(\theta^0)\,h_t^2(\theta^0)\right) & E\left(\varepsilon_t h_t^2(\theta^0)\right)^2\end{bmatrix}$$

Theorem 3. *Under A.0, and A.4-A.8 with* $H\left(\theta^0\right)$ *a positive matrix,*

$$T^{1/2}\left(\widehat{\theta}-\theta^0\right) \xrightarrow{d} N\left(\begin{pmatrix}0\\0\end{pmatrix}, 4H^{-1}\left(\theta^0\right)\Omega H^{-1}\left(\theta^0\right)\right)$$

The proof of the theorem can be found in the appendix. It can be proved that a consistent estimator of the variance-covariance matrix of $T^{1/2}\left(\widehat{\theta}-\theta^0\right)$ is $H_T\left(\widehat{\theta}_T\right)$.

4.2 The r^0 unknown and STIMA cases

For these cases we propose to estimate $1\left(|z_{t-1}|<r^0\right)$ and $1\left(|z_{t-1}|>r^0\right)$, and use that the rate of convergence is T. Using it the asymptotic distribution will be the same. To prove it we use the approximation of Pötscher and Prucha (1997). Define $\tau=\left\{1\left(|z_{t-1}|<r^0\right)\right\}_{t=1}^T$ as a nuisance parameter. Then the objective function depends on (θ,τ) and as we have proved it is twice differentiable in θ. Define $\widehat{\tau}=\{1\left(|\widehat{e}_{t-1}|<\widehat{r}\right)\}_{t=1}^T$[3], with $\widehat{e}_t$ the LS estimator of ε_t in the STIMA case. First we prove that $\left(\widehat{\theta}\left(\widehat{\tau}\right)-\theta^0\right)=O_p(T^{-1/2})$, with $\widehat{\theta}\left(\widehat{\tau}\right)=\arg\min_\theta Q_T\left(\theta,\widehat{\tau}\right)$.

Lemma 4. *For all* $\varepsilon>0$ $\exists\triangle^*$ *s.t with probability greater than* $1-\varepsilon$*,*

$$\max_i\left|\theta_i-\theta_i^0\right|^2>\frac{\triangle^*}{T}\Rightarrow Q_T\left(\theta,\widehat{\tau}\right)-Q_T\left(\theta^0,\widehat{\tau}\right)>0.$$

Using this lemma and $\widehat{r}=r^0+O_p\left(T^{-1}\right)$, we can prove the following theorem,

Theorem 4. *Under A.0, and A.4-A.8 with* $H\left(\theta^0\right)$ *being a positive matrix,*

$$T^{1/2}\left(\widehat{\theta}\left(\widehat{\tau}\right)-\theta^0\right) \xrightarrow{d} N\left(\begin{pmatrix}0\\0\end{pmatrix}, 4H^{-1}\left(\theta^0\right)\Omega H^{-1}\left(\theta^0\right)\right)$$

With this theorem we prove that the asymptotic distribution for θ_1 and θ_2 is the same if we know $1\left(|z_{t-1}|<r^0\right)$ or if we have to estimate it.

4.3 Threshold Effect Test

In this section we propose a test to contrast the existence of threshold effect. The null hypothesis is no threshold effect, $H_0:\theta_1^0=\theta_2^0$. Under this hypothesis the threshold parameter, r^0, is not defined, and the usual tests do not have the standard properties. We

[3]When r^0 is unknown and z_t observable, we have to follow the same steps, with $\widehat{\tau}=\{1\left(|z_{t-1}|<\widehat{r}\right)\}_{t=1}^T$.

propose a test based on a Wald type statistic of the following form

$$W_T(r) = \left(R\widehat{\theta}_T(r)\right)' \left[RV\left(\widehat{\theta}_T(r)\right)R'\right]^{-1}\left(R\widehat{\theta}_T(r)\right)$$
$$W_T = \sup_r W_T(r)$$

with $R = (1, -1)$. To obtain the asymptotic distribution of W_T we work with the functional $W_T(r)$. First we prove that the finite dimensional distribution converge and then, we prove the tightness. To obtain the asymptotic distribution of W_T we apply the continuous mapping theorem.

First we study the asymptotic distribution of $\widehat{\theta}_T(r)$ for a given r under the null hypothesis of no threshold effect, $x_t = \varepsilon_t - \theta^0\varepsilon_{t-1}$.

Lemma 5. *Under H_0 and for a given r,*

$$T^{1/2}\begin{pmatrix} \widehat{\theta}_{1,T}(r) - \theta^0 \\ \widehat{\theta}_{2,T}(r) - \theta^0 \end{pmatrix} = 2T^{-1/2}H^{-1}(r)\begin{pmatrix} \sum_{t=1}^T \varepsilon_t \sum_{j=1}^t \theta^{0^{j-1}} 1\left(|z_{t-j}| > r\right)\varepsilon_{t-j} \\ \sum_{t=1}^T \varepsilon_t \sum_{j=1}^t \theta^{0^{j-1}} 1\left(|z_{t-j}| < r\right)\varepsilon_{t-j} \end{pmatrix} + o_p(1)$$
$$\xrightarrow{d} N\left(0, H^{-1}(r)\,\Omega^* H^{-1}(r)\right)$$

The proof of this lemma can be found in the appendix. Following this proof it is straightforward to prove that the finite dimensional process converge to a multivariate normal distribution, although the variance-covariance matrix is not standard.

Now we prove the tightness for the process $T^{-1/2}\sum_{t=1}^T \varepsilon_t \sum_{j=1}^t \theta^{0^{j-1}} 1\left(|z_{t-j}| > r\right)\varepsilon_{t-j}$ and $T^{-1/2}\sum_{t=1}^T \varepsilon_t \sum_{j=1}^t \theta^{0^{j-1}} 1\left(|z_{t-j}| < r\right)\varepsilon_{t-j}$.

Lemma 6. *Under H_0, for all $\varepsilon > 0$ $\exists\delta > 0$ s.t*

$$\lim_{T\to\infty} P\left(\sup_{|r_1 - r_2| < \delta}\left|T^{-1/2}\sum_{t=1}^T \varepsilon_t \sum_{j=1}^t \theta^{0^{j-1}}\left(1\left(|z_{t-j}| > r_1\right) - 1\left(|z_{t-j}| > r_2\right)\right)\varepsilon_{t-j}\right| > \varepsilon\right) < \varepsilon$$

The proof can be found in the appendix.

With these two lemmas we have the convergence of $W_T(r)$ to a nonstandard gaussian process, whose distribution depends on nuisance parameters. To avoid this problem we use and adaptation of the Hansen (1996) strategy, although in this paper we do not prove the validity of the wild bootstrap.

Table 1: Simple Size of the test for STIMA models.

	T=100		T=200		T=500	
	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.1$
$\theta = -0.8$	0.055	0.105	0.07	0.12	0.035	0.125
$\theta = -0.5$	0.06	0.10	0.07	0.13	0.03	0.125
$\theta = -0.2$	0.05	0.11	0.065	0.135	0.05	0.12
$\theta = 0.2$	0.05	0.085	0.05	0.095	0.055	0.12
$\theta = 0.5$	0.05	0.11	0.065	0.145	0.05	0.105
$\theta = 0.8$	0.055	0.13	0.045	0.15	0.065	0.095

5 Finite Sample Performance

In this section we perform a simulation study to evaluate the small sample properties of the proposed threshold test. Due to the large computational requirements of the estimation and simulation design, we use an approximation to the estimators proposed in the previous section and only 200 simulated samples.

Instead of $\widehat{\theta}_i$, we use $T^{-1/2}\sum_{t=1}^{T}\varepsilon_t\sum_{j=1}^{t}\theta^{0^{j-1}}1\left(|z_{t-j}|>r\right)\varepsilon_{t-j}$ and $T^{-1/2}\sum_{t=1}^{T}\varepsilon_t\sum_{j=1}^{t}\theta^{0^{j-1}}1\left(|z_{t-j}|<r\right)\varepsilon_{t-j}$, where ε_t and θ^0 are estimated under the null.

Table [1] contains the results about the size of the test. Under the null hypothesis, the process x_t follows a moving average. In this simulation design we generate six kind of $M.A$ models, with $\theta = (-0.8, -0.5, -0.2, 0.2, 0.5, 0.8)$. The shocks of the model are iid $N(0,1)$ and we bootstrap taking $B = 1500$.

To evaluate the power of the test we consider two kind of STIMA models, the first one

$$\triangle y_t = x_t = \begin{cases} \varepsilon_t \quad if \;\; |\varepsilon_{t-1}| > r \\ \varepsilon_t - \varepsilon_{t-1} \quad if \;\; |\varepsilon_{t-1}| < r \end{cases}$$

with $r = 0.5$ and 1. In this model, y_t is the sum of a white noise and a random walk.

$$y_t = \sum_{j=0}^{t}\varepsilon_{t-j}1\left(|\varepsilon_{t-j}|>r\right) + \varepsilon_t 1\left(|\varepsilon_t|<r\right) = \sum_{j=0}^{t}\varepsilon_{t-j}^{+} + \varepsilon_t^{-}$$

The results are in table [2] . As we can see, the power of the test for small sample size is small, maybe due to the use of an approximation to the estimators. Clearly, the smaller is r, the smaller is the power, since the probability of change the regime is smaller and then it is more difficult to detect the threshold.

Table 2: Sample Power of the Test for $\theta^0 = (0, 1)$

	T=100		T=200		T=500	
	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.1$
$r = 0.5$	0.085	0.180	0.195	0.335	0.625	0.735
$r = 1$	0.335	0.455	0.635	0.740	0.959	0.975

Table 3: Sample Power of the Test for uncorrelated x_t.

	T=200	
	$\alpha = 0.05$	$\alpha = 0.1$
$(\theta_1 = -0.27, \theta_2 = 1, r = 1)$	0.529	0.714
$(\theta_1 = -0.09, \theta_2 = 1, r = 0.75)$	0.58	0.66

Finally, to evaluate the power of the test we propose another STIMA model,

$$\triangle y_t = x_t = \begin{cases} \varepsilon_t - \theta_1 \varepsilon_{t-1} & if \;\; |\varepsilon_{t-1}| > r \\ \varepsilon_t - \theta_2 \varepsilon_{t-1} & if \;\; |\varepsilon_{t-1}| < r \end{cases}$$

with $\theta = (\theta_1 = -0.27, \theta_2 = 1, r = 1)$ and $\theta = (\theta_1 = -0.09, \theta_2 = 1, r = 0.75)$. In these case the correlogram of the series x_t is similar to a white noise. That is, $E(x_t x_{t-1}) = 0$. Table [3] contains the results for $T = 200$.

6 Applications to Exchange Rates and Stock Prices

This section is still under constructions.

7 Conclusion

In this paper we propose a new model, where shocks can be persistent or transitory depending on the size of some "key" economic variables, including the size of the models shocks. It will be in this case when we will be able to detect whether large shocks have a permanent effect or a transitory one. The identification scheme is done inside the framework of threshold moving average models. By identifying the different threshold regimes we will be identifying not only the different type of shocks but also which are the relevant causes for a shock to be permanent or transitory.

Extensions of this new models to a VAR framework is under current research by the authors. They will allow us to identify which type of permanent and transitory shocks are the ones behind the supply and demand shocks of Blanchard and Quah (1989), as well as which type of permanent shocks are behind the cointegration sources.

8 Appendix

Proof of Lemma 1. Conditioning on the history of $v_t = (\varepsilon_t, z_{t+1})$ until $t-1$, the expectation of the previous expression is given by:

$$\begin{aligned} E\left[y_{t+k}\right| w_{t-1}] &= \frac{1}{A(L)(1-L)}\mu + E\left[\varepsilon_{t+k-1}1\left(|z_{t+k}|>r\right)\right| w_{t-1}] + \\ &+\sum_{i=0}^{k-1}\left[(1-\theta_1)\sum_{j=0}^{i-1}\phi^j+\phi^i\right]E\left[\varepsilon_{t+k-i-1}1\left(|z_{t+k-i}|>r\right)\right| w_{t-1}] + \\ &+\sum_{i=1}^{\infty}\left[(1-\theta_1)\sum_{j=0}^{i+k}\phi^j+\phi^{i+k+1}\right]\varepsilon_{t-i-1}1\left(|z_{t-i}|>r\right) + E\left[\varepsilon_{t+k-1}1\left(|z_{t+k}|<r\right)\right| w_{t-1}] + \\ &+\sum_{i=0}^{k-1}\left[(1-\theta_2)\sum_{j=0}^{i-1}\phi^j+\phi^i\right]E\left[\varepsilon_{t+k-1-i}1\left(|z_{t+k-i}|<r\right)\right| w_{t-1}] + \\ &+\sum_{i=1}^{\infty}\left[(1-\theta_2)\sum_{j=0}^{i+k}\phi^j+\phi^{i+k+1}\right]\varepsilon_{t-i-1}1\left(|z_{t-i}|<r\right) \end{aligned}$$

Then it is straightforward to obtain,

$$\begin{aligned} GI\left(k,V_t,F_{t-1}\right) &= E\left[\varepsilon_{t+k-1}1\left(|z_{t+k}|>r\right)\right|\varepsilon_t, w_{t-1}] - E\left[\varepsilon_{t+k-1}1\left(|z_{t+k}|>r\right)\right| w_{t-1}] + \\ \sum_{i=0}^{k-2}&\left[(1-\theta_1)\sum_{j=0}^{i-1}\phi^j+\phi^i\right]\left\{E\left[\varepsilon_{t+k-1-i}1\left(|z_{t+k-i}|>r\right)\right|\varepsilon_t, w_{t-1}] - E\left[\varepsilon_{t+k-1-i}1\left(|z_{t+k-i}|>r\right)\right| w_{t-1}]\right\} + \\ &\left[(1-\theta_1)\sum_{j=0}^{k-2}\phi^j+\phi^{k-1}\right]\left\{\varepsilon_t 1\left(|z_{t+1}|>r\right) - E\left[\varepsilon_t 1\left(|z_{t+1}|>r\right)\right| w_{t-1}]\right\} + \\ \sum_{i=0}^{k-2}&\left[(1-\theta_2)\sum_{j=0}^{i-1}\phi^j+\phi^i\right]\left\{E\left[\varepsilon_{t+k-1-i}1\left(|z_{t+k-i}|<r\right)\right|\varepsilon_t, w_{t-1}] - E\left[\varepsilon_{t+k-1-i}1\left(|z_{t+k-i}|<r\right)\right| w_{t-1}]\right\} + \\ &\left[(1-\theta_2)\sum_{j=0}^{k-2}\phi^j+\phi^{k-1}\right]\left\{\varepsilon_t 1\left(|z_{t+1}|>r\right) - E\left[\varepsilon_t 1\left(|z_{t+1}|>r\right)\right| w_{t-1}]\right\} \end{aligned}$$

As a result of this assumption we obtain that $E\left[\varepsilon_t 1\left(|z_{t+1}|>r\right)/\, w_{t-1}\right]=0$. This equality holds by

$$\begin{aligned} E\left[\varepsilon_t 1\left(|z_{t+1}|>r\right)\right| w_{t-1}] &= E\left\{E\left[\varepsilon_t 1\left(|z_{t+1}|>r\right)\right|\varepsilon_t, w_{t-1}]\right| w_{t-1}\} = \\ &= E\left\{\varepsilon_t E\left[1\left(|z_{t+1}|>r\right)\right|\varepsilon_t, w_{t-1}]\right| w_{t-1}\} = \\ &= E\left\{\varepsilon_t E\left[1\left(|z_{t+1}|>r\right)\right| w_{t-1}]\right| w_{t-1}\} = \\ &= E\left[1\left(|z_{t+1}|>r\right)\right| w_{t-1}]\, E\left\{\varepsilon_t\right| w_{t-1}\} = 0 \end{aligned}$$

with $F_t = \sigma_t(\varepsilon, z)$.

Then, with any of these assumptions, we obtain

$$GI\left(k,\varepsilon_t,F_{t-1}\right) = \left[(1-\theta_{t+1})\sum_{j=0}^{k-2}\phi^j+\phi^{k-1}\right]\varepsilon_t = \begin{cases} \phi^{k-1}\varepsilon_t \quad if\ \theta_{t+1}=1 \\ \left[(1-\theta)\sum_{j=0}^{k-2}\phi^j+\phi^{k-1}\right]\varepsilon_t \quad if\ \theta_{t+1}=\theta\neq 1 \end{cases}$$

■

Proof of Lemma 2. By definition, $z_t = \varepsilon_{t-1}$, then replacing in equation [7]

$$x_t = \varepsilon_t - \theta_1 1\left(|\varepsilon_{t-1}| > r\right)\varepsilon_{t-1} - \theta_2 1\left(|\varepsilon_{t-1}| < r\right)\varepsilon_{t-1}$$

then

$$\widehat{\varepsilon}_t = x_t + \theta_1 1\left(|\widehat{\varepsilon}_{t-1}| > r\right)\widehat{\varepsilon}_{t-1} + \theta_2 1\left(|\widehat{\varepsilon}_{t-1}| < r\right)\widehat{\varepsilon}_{t-1}$$

Taking $\partial = |\theta_1 - \theta_2|$ and defining the events

$$\begin{array}{ll} A_{1,t-1} = \{|\varepsilon_{t-1}| > r, |\widehat{\varepsilon}_{t-1}| > r\} & A_{2,t-1} = \{|\varepsilon_{t-1}| < r, |\widehat{\varepsilon}_{t-1}| < r\} \\ A_{3,t-1} = \{|\varepsilon_{t-1}| > r, |\widehat{\varepsilon}_{t-1}| < r\} & A_{4,t-1} = \{|\varepsilon_{t-1}| < r, |\widehat{\varepsilon}_{t-1}| > r\} \end{array}$$

and using that $\theta_1 < \theta_2$, we have

$$\begin{aligned} e_t = \varepsilon_t - \widehat{\varepsilon}_t &= \theta_1 1\left(A_{1,t-1}\right)\left(\varepsilon_{t-1} - \widehat{\varepsilon}_{t-1}\right) + \theta_2 1\left(A_{2,t-1}\right)\left(\varepsilon_{t-1} - \widehat{\varepsilon}_{t-1}\right) + \\ &+ \theta_1 1\left(A_{3,t-1}\right)\left(\varepsilon_{t-1} - \widehat{\varepsilon}_{t-1}\right) - \partial 1\left(A_{3,t-1}\right)\widehat{\varepsilon}_{t-1} + \theta_1 1\left(A_{4,t-1}\right)\left(\varepsilon_{t-1} - \widehat{\varepsilon}_{t-1}\right) + \partial 1\left(A_{4,t-1}\right)\varepsilon_{t-1} \\ &= \partial\left[1\left(A_{4,t-1}\right)\varepsilon_{t-1} - 1\left(A_{3,t-1}\right)\widehat{\varepsilon}_{t-1}\right] + \left[\theta_1 1\left(A_{1,t-1}\right) + \theta_2 1\left(A_{2,t-1}\right) + \theta_1 1\left(A_{3,t-1}\right) + \theta_1 1\left(A_{4,t-1}\right)\right] e_{t-1} \end{aligned} \tag{4}$$

$$|e_t| \leq \partial\left[1\left(A_{4,t-1}\right)|\varepsilon_{t-1}| + 1\left(A_{3,t-1}\right)|\widehat{\varepsilon}_{t-1}|\right] + \left[|\theta_1| 1\left(|\varepsilon_{t-1}| > r\right) + |\theta_2| 1\left(|\varepsilon_{t-1}| < r\right)\right]|e_{t-1}|$$

$$E|e_t| \leq \partial r E\left[1\left(A_{4,t-1}\right) + 1\left(A_{3,t-1}\right)\right] + E\left\{\left[|\theta_1| 1\left(|\varepsilon_{t-1}| > r\right) + |\theta_2| 1\left(|\varepsilon_{t-1}| < r\right)\right]|e_{t-1}|\right\}$$

First we calculate $E\left[1\left(A_{4,t-1}\right) + 1\left(A_{3,t-1}\right)\right]$,

$$\begin{aligned} E\left[1\left(A_{4,t-1}\right)\right] &= E\left[1\left(|\varepsilon_{t-1}| < r\right)1\left(|\widehat{\varepsilon}_{t-1}| > r\right)\right] \\ &= E\left[1\left(-r < \varepsilon_{t-1} < -r - e_{t-1}\right) + 1\left(r - e_{t-1} < \varepsilon_{t-1} < r\right)\right] \\ E\left[1\left(A_{3,t-1}\right)\right] &= E\left[1\left(|\varepsilon_{t-1}| > r\right)1\left(|\widehat{\varepsilon}_{t-1}| < r\right)\right] \\ &= E\left[1\left(-r - e_{t-1} < \varepsilon_{t-1} < -r\right) + 1\left(r < \varepsilon_{t-1} < r - e_{t-1}\right)\right] \end{aligned}$$

Now using that e_{t-1} is F_{t-2} measurable and Taking $M = \max_e f\left(-r+e\right) + f\left(r+e\right)$,

$$\begin{aligned} E\left[1\left(A_{4,t-1}\right) + 1\left(A_{3,t-1}\right)\right] &= E\left\{E\left[1\left(-r < \varepsilon_{t-1} < -r - e_{t-1}\right) + 1\left(r - e_{t-1} < \varepsilon_{t-1} < r\right)\right| F_{t-2}\right]\right\} + \\ & E\left\{E\left[1\left(-r - e_{t-1} < \varepsilon_{t-1} < -r\right) + 1\left(r < \varepsilon_{t-1} < r - e_{t-1}\right)\right| F_{t-2}\right]\right\} \\ = E\left\{E\left[\int_{-r}^{-r-e_{t-1}} f(\varepsilon)\,\partial\varepsilon + \int_{r-e_{t-1}}^{r} f(\varepsilon)\,\partial\varepsilon + \int_{-r-e_{t-1}}^{-r} f(\varepsilon)\,\partial\varepsilon + \int_{r}^{r-e_{t-1}} f(\varepsilon)\,\partial\varepsilon \right| F_{t-2}\right]\right\} & \\ &\leq ME\left(|e_{t-1}|\right) \end{aligned}$$

The last inequality using that

$$\begin{aligned} F\left(e_{t-1}\right) &= \int_{-r}^{-r-e_{t-1}} f(\varepsilon)\,\partial\varepsilon + \int_{r-e_{t-1}}^{r} f(\varepsilon)\,\partial\varepsilon + \int_{-r-e_{t-1}}^{-r} f(\varepsilon)\,\partial\varepsilon + \int_{r}^{r-e_{t-1}} f(\varepsilon)\,\partial\varepsilon \\ F\left(e_{t-1}\right) &= F(0) + \left.\frac{\partial F\left(e_{t-1}\right)}{\partial e_{t-1}}\right|_{e^*}\left(e_{t-1} - 0\right) \\ &\leq M1\left(e_{t-1} > 0\right)e_{t-1} - M1(e_{t-1} < 0)e_{t-1} = M|e_{t-1}| \end{aligned}$$

Using that $\lambda_1\left(\theta_1, \theta_2, r\right) = E\left[|\theta_1| 1\left(|\varepsilon_{t-1}| > r\right) + |\theta_2| 1\left(|\varepsilon_{t-1}| < r\right)\right]$,

$$E|e_t| = \left[\partial r M + \lambda_1\left(\theta_1, \theta_2, r\right)\right] E|e_{t-1}|$$

For invertibility we need $E\left(e_{t-1}^2\right)$, then

$$\begin{aligned} e_t^2 \leq \partial^2 r^2\left[1\left(A_{4,t-1}\right) + 1\left(A_{3,t-1}\right)\right] + \left[\theta_1^2 1\left(|\varepsilon_{t-1}| > r\right) + \theta_2^2 1\left(|\varepsilon_{t-1}| < r\right)\right] e_{t-1}^2 + \\ + 2\partial r\left[1\left(A_{4,t-1}\right) + 1\left(A_{3,t-1}\right)\right]\left[|\theta_1| 1\left(|\varepsilon_{t-1}| > r\right) + |\theta_2| 1\left(|\varepsilon_{t-1}| < r\right)\right]|e_{t-1}| \end{aligned}$$

Taking $\lambda_2 = \left[\partial r M + \lambda_1\left(\theta_1, \theta_2, r\right)\right]$ and $K = \partial^2 r^2 M + 4\partial r$

$$E\left(e_t^2\right) \leq KE|e_{t-1}| + \lambda_1\left(\theta_1, \theta_2, r\right)E\left(e_{t-1}^2\right) \leq K\lambda_2^{t-1}E|e_0| + \lambda_2 E\left(e_{t-1}^2\right) \leq tK\lambda_2^{t-1}E|e_0|$$

Then, if $\lambda_2 < 1$ we take that $\lim_{t\to\infty} E\left(e_t^2\right) = 0$. ■

Proof of lemma 3. As z_t is observable, we have

$$\begin{aligned} x_t &= \varepsilon_t + \theta\left(z_{t-1}\right)\varepsilon_{t-1} \\ \widehat{\varepsilon}_t &= x_t - \theta\left(z_{t-1}\right)\widehat{\varepsilon}_{t-1} \\ e_t &= \widehat{\varepsilon}_t - \varepsilon_t = \theta\left(z_{t-1}\right)\left(\varepsilon_{t-1} - \widehat{\varepsilon}_{t-1}\right) \\ \theta\left(x\right) &= \begin{cases} \theta_1 \text{ if } |x| > r \\ \theta_2 \text{ if } |x| < r \end{cases} \end{aligned}$$

Then it is easy to see that

$$E\left(e_t^2\right) = \prod_{j=1}^{t} \theta^2\left(z_{t-j}\right) E\left(e_0^2\right)$$

A sufficient condition for invertibility of x_t is

$$\lim_{t\to\infty} E\left(\prod_{j=1}^{t} \theta^2\left(z_{t-j}\right)\right) = 0$$

This will be true if

$$E\left(\theta^2\left(z_t\right)\middle| F_{t-1}\right) < 1 \text{ with } F_{t-1} = \sigma\left(z_{t-1}, z_{t-2}, ...\right)$$

■

Before proving the claim 1 we need to study the memory and moment properties of the upper limit of some processes as θ_t, e_t, and $h_t^i\left(\theta'\right)$. Let

$$\begin{aligned} \overline{\theta}_t &= \left(1-\delta\right) 1\left(|z_t| > \overline{r}\right) + \left(1+\delta\right) 1\left(|z_t| < \overline{r}\right) \\ \overline{e}_t &= \left(|\varepsilon_t| + |\varepsilon_{t-1}|\right) + \sum_{j=0}^{t-2} \prod_{i=0}^{j} \left|\overline{\theta}_{t-i}\right| \left(|\varepsilon_{t-j-1}| + |\varepsilon_{t-j-2}|\right) \\ \overline{h}_t^i &= \overline{e}_{t-1} + \overline{\theta}_t \overline{h}_{t-1}^i = \overline{e}_{t-1} + \sum_{j=0}^{t-2} \prod_{k=0}^{j} \overline{\theta}_{t-k} \overline{e}_{t-j-2} \end{aligned}$$

Now we define the difference between variables

$$\begin{aligned} 4 \geq \delta_t\left(\theta, \theta^*\right) = \theta_t - \theta_t^* &= 1\left(|z_t| > \max\left(r, r^*\right)\right)\left(\theta_1 - \theta_1^*\right) + 1\left(|z_t| < \min\left(r, r^*\right)\right)\left(\theta_2 - \theta_2^*\right) + \\ &+ \delta 1\left(\min\left(r, r^*\right) < |z_t| < \max\left(r, r^*\right)\right) \\ v_t\left(\theta, \theta^*\right) &= e_t\left(\theta\right) - e_t\left(\theta^*\right) = \delta_t\left(\theta, \theta^*\right) e_{t-1}\left(\theta^*\right) + \theta_{t-1} v_{t-1}\left(\theta, \theta^*\right) \\ \overline{v}_t &= \delta_t\left(\theta, \theta^*\right)\overline{e}_{t-1} + \overline{\theta}_{t-1}\overline{v}_{t-1} \\ s_t^i\left(\theta, \theta^*\right) &= h_t^i\left(\theta\right) - h_t^i\left(\theta^*\right) \leq \overline{s}_t^i \\ \overline{s}_t^i &= 1\left(r^* \leq |z_t| \leq r\right)\overline{e}_{t-1} + \overline{v}_{t-1} + \delta_t\left(\theta, \theta^*\right)\overline{h}_{t-1}^i + \overline{\theta}_{t-1}\overline{s}_{t-1}^i \end{aligned}$$

Lemma 7. *Under assumption A.3 to A.8 $\overline{e}_t$ is a L_4-bounded sequence and a L_2-Near Epoch Dependence (NED) of size $-1/2$ with constant $\|\overline{e}_t\|_4$.*

Proof of lemma 7. Using the Minkowski inequality, the law of iterated expectation and assumption A.8

$$\|\overline{e}_t\|_4 \leq \left\||\varepsilon_t| + |\varepsilon_{t-1}|\right\|_4 + \sum_{j=0}^{t-2} \left\|\prod_{i=0}^{j} \left|\overline{\theta}_{t-i}\right| \left(|\varepsilon_{t-j-1}| + |\varepsilon_{t-j-2}|\right)\right\|_4 \leq 4\|\varepsilon_t\|_4 + \sum_{j=1}^{t-2} \overline{\lambda}^{j-1} 2\|\varepsilon_t\|_4 \leq \frac{4\|\varepsilon_t\|_4}{1-\overline{\lambda}} < \infty$$

To prove that $\overline{e}_t$ is L_2-NED we have to prove that

$$\left\|\overline{e}_t - E\left(\overline{e}_t\right| F_{t-m}^{t+m}\right)\right\|_2 \leq d_t v_m$$

where $v_m \overset{m\to\infty}{\longrightarrow} 0$ and $F_{t-m}^{t+m} = \sigma\left(V_{t-m},...,V_{t+m}\right)$. In our case $V_t = \left(\varepsilon_t, z_t\right)$, then

$$\left\|\overline{e}_t - E\left(\overline{e}_t\right| F_{t-m}^{t+m}\right)\right\|_2 \leq \left\|\sum_{j=m}^{t-2}\prod_{i=0}^{j}\left|\overline{\theta}_{t-i}\right|\left(\left|\varepsilon_{t-j-1}\right| + \left|\varepsilon_{t-j-2}\right|\right)\right\|_2 \leq \sum_{j=m}^{t-2}\overline{\lambda}^j 2\left\|\varepsilon_{t-j-1}\right\|_2 \leq$$

$$\leq 2\left\|\varepsilon_t\right\|_2 \overline{\lambda}^m \sum_{j=0}^{t-2-m}\overline{\lambda}^j \leq \left\|\overline{e}_t\right\|_4 \overline{\lambda}^m$$

taking $v_m = \overline{\lambda}^m$ and $d_t = \left\|\overline{e}_t\right\|_4$ we have proved. ■

Lemma 8. *Under assumption A.3 to A.8* $\overline{h}_t^i$ *is a* L_4-*bounded sequence and a* L_2-Near *Epoch Dependence* (NED) *of size* $-1/2$ *with constant* $\left\|\overline{h}_t^i\right\|_4$.

Proof o lemma 8. As in proof of lemma 7,

$$\left\|\overline{h}_t^i\right\|_4 \leq \left\|\overline{e}_t\right\|_4 + \left\|\overline{\theta}_t \overline{h}_{t-1}^i\right\|_4 \leq 2\left\|\overline{e}_t\right\|_4 + \sum_{j=1}^{t-2}\overline{\lambda}^{j-1}\left\|\overline{e}_{t-j-1}\right\|_4 \leq \frac{2\left\|\overline{e}_t\right\|_4}{1-\overline{\lambda}} < \infty$$

For $L_2 - NED$ we use the result of lemma 7

$$\begin{aligned}
\left\|\overline{h}_t^i - E\left(\overline{h}_t^i\right| F_{t-m}^{t+m}\right)\right\|_2 &\leq \left\|\overline{e}_t - E\left(\overline{e}_t\right| F_{t-m}^{t+m}\right)\right\|_2 + \left\|\sum_{j=0}^{t-2}\prod_{i=0}^{j}\left|\overline{\theta}_{t-i}\right|\left(\overline{e}_{t-j-1} - E\left(\overline{e}_{t-j-1}\right| F_{t-m}^{t+m}\right)\right)\right\|_2 \\
&\leq \left\|\overline{e}_t\right\|_4 \overline{\lambda}^m + \sum_{j=0}^{m-1}\overline{\lambda}^j\left\|\overline{e}_t\right\|_4 \overline{\lambda}^{m-j-1} + \sum_{j=m}^{t-2}\overline{\lambda}^j\left\|\overline{e}_t\right\|_4 \\
&\leq \overline{\lambda}^{m-1}\left[\left\|\overline{e}_t\right\|_4 \overline{\lambda} + m\left\|\overline{e}_t\right\|_4 + \sum_{j=0}^{\infty}\overline{\lambda}^j\left\|\overline{e}_t\right\|_4\right] \leq m\overline{\lambda}^{m-1}\left\|\overline{h}_t^i\right\|_4
\end{aligned}$$

again taking $d_t = \left\|\overline{h}_t^i\right\|_4$ and $v_m = m\overline{\lambda}^{m-1}$ we are done. ■

Lemma 9. *Under assumption A.5 and A.7* $\delta_t\left(\theta,\theta^*\right)$ *is a* $\alpha-mixing$ *of size* $-a$ *with*

$$E\left(\delta_t^k\left(\theta,\theta^*\right)\right) = \max\left(\delta_1^{k/2}, \delta_2^{k/2}, \delta_3\right)\left(2 + \delta^k M\right) \quad for\ k \geq 1$$

Proof of lemma 9. The proof is straight forward using that z_t is $\alpha - mixing$ ■

Lemma 10. *Under assumption A.3 to A.8* $\overline{v}_t$ *is a* L_4-*bounded sequence and a* L_2-NED *of size* $-1/2$ *with constant* $\left\|\overline{v}_t\right\|_4$. *Besides*

$$E\left(\overline{v}_t\right) \leq \max\left(\delta_1^{1/2}, \delta_2^{1/2}, \delta_3\right)\frac{\left(2+\delta M\right)\left[2\sigma_\varepsilon^2 + E\left|e_t^2\right|\right]}{\left(1-\lambda\right)^2} = \max\left(\delta_1^{1/2}, \delta_2^{1/2}, \delta_3\right)K_v$$

$$E\left(\overline{v}_t^2\right) \leq \max\left(\delta_1, \delta_2, \delta_3\right)\frac{\left(2+\delta^2 M\right)\left[2\sigma_\varepsilon^2 + E\left|e_t^2\right|\right]}{\left(1-\lambda\right)^2} = \max\left(\delta_1, \delta_2, \delta_3\right)K_v$$

$\left[\left(2\sigma_{\varepsilon/z}^4 + E\left(\overline{e}_t^4\right)\right)\right]$

Proof of lemma 10. Using assumption $A.7$ and lemma 9

$$\left\|\overline{e}_{t-1}\delta_t\left(\theta,\theta^*\right)\right\|_4 \leq \max\left(\delta_1^{4/2}, \delta_2^{4/2}, \delta_3\right)^{1/4}\left(2+\delta^4 M\right)^{1/4}\left[\left(2\sigma_{\varepsilon/z}^4 + E\left(\overline{e}_t^4\right)\right)\right]^{1/4}$$

$$\|\overline{v}_t\|_4 \leq \|\delta_{t-1}(\theta,\theta^*)\,\overline{e}_{t-1}\|_4 + \left\|\overline{\theta}_{t-1}\overline{v}_{t-1}\right\|_4 \leq$$
$$\leq \max\left(\delta_1^{4/2},\delta_2^{4/2},\delta_3\right)^{1/4}\frac{\left(2+\delta^4 M\right)^{1/4}\left[\left(2\sigma^4_{\varepsilon/z}+E\left(\overline{e}_t^4\right)\right)\right]^{1/4}}{1-\overline{\lambda}} < \infty$$

The rest of the moments, $E(\overline{v}_t)$ and $E\left(\overline{v}_t^2\right)$ are straight forward. For L_2-NED we define the $F_{t-m}^{t+m}-measurable$

$$\overline{v}_t^m = \delta_t(\theta,\theta^*)\,\overline{e}_{t-1}^m + \overline{\theta}_{t-1}\overline{v}_{t-1}^m$$

we use the result of lemma 7

$$\|\overline{v}_t - \overline{v}_t^m\|_2 \leq 4\left\|\overline{e}_t - \overline{e}_{t-1}^m\right\|_2 + \overline{\theta}_{t-1}\left\|\overline{v}_{t-1} - \overline{v}_{t-1}^m\right\|_2$$
$$\leq \sum_{j=0}^{m-1}\overline{\lambda}^j 4\|\overline{e}_t\|_4\overline{\lambda}^{m-j-1} + \overline{\lambda}^m\|\overline{v}_{t-m}\|_2 \leq m\overline{\lambda}^{m-1}\|\overline{v}_{t-m}\|_2$$

taking $d_t = \|\overline{v}_{t-m}\|_2$ and $v_m = m\overline{\lambda}^{m-1}$ we are done. ■

Lemma 11. *Under assumption A.3 to A.8 $\overline{s}_t^i$ is a L_4-bounded sequence and a L_2-NED of size $-1/2$ with constant $\left\|\overline{s}_t^i\right\|_4$. Besides*

$$E(\overline{s}_t) \leq \max\left(\delta_1^{1/2},\delta_2^{1/2},\delta_3\right)K_s \quad E\left(\overline{s}_t^2\right) \leq \max(\delta_1,\delta_2,\delta_3)\,K_s$$

Proof of lemma 11. Here we use the previous lemmas $1(r^* \leq |z_t| \leq r)\,\overline{e}_{t-1} + \overline{v}_{t-1} + \delta_t(\theta,\theta^*)\,\overline{h}_{t-1}^i + \overline{\theta}_{t-1}\overline{s}_{t-1}^i$

$$\left\|\overline{s}_t^i\right\|_4 \leq \left[\delta_3 M\left(2\sigma^4_{\varepsilon/z}+E\left(\overline{e}_t^4\right)\right)\right]^{1/4} + \|\overline{v}_{t-1}\|_4 + \max\left(\delta_1^{4/2},\delta_2^{4/2},\delta_3\right)^{1/4}\left(2+\delta^4 M\right)^{1/4}\left\|\overline{h}_{t-1}^i\right\| +$$
$$+\overline{\lambda}\left\|\overline{s}_{t-1}^i\right\|_4 \leq$$
$$\leq \max\left(\delta_{i=1,2}^{4/2},\delta_3\right)^{1/4}\frac{\left(\left[\left(2+\delta^4 M\right)^{1/4}+M^{1/4}\right]\left[\left(2\sigma^4_{\varepsilon/z}+E\left(\overline{e}_t^4\right)\right)\right]^{1/4}+\left(2+\delta^4 M\right)^{1/4}\left\|\overline{h}_{t-1}^i\right\|\right)}{1-\overline{\lambda}}$$

The other two moments are straight forward. We define the following $F_{t-m}^{t+m}-measurable$ variable

$$\overline{s}_t^{i,m} = 1(r^* \leq |z_t| \leq r)\,\overline{e}_{t-1}^m + \overline{v}_{t-1}^m + \delta_t(\theta,\theta^*)\,\overline{h}_{t-1}^{i,m} + \overline{\theta}_{t-1}\overline{s}_{t-1}^{i,m}$$
$$\left\|\overline{s}_t^i - \overline{s}_t^{i,m}\right\|_2 \leq m^2\overline{\lambda}^m\left\|\overline{s}_t^i\right\|_2$$

■

Lemma 12. *Let w_t be any of the following variables, $\overline{h}_t^i$, $\overline{v}_t$ and $\overline{s}_t^i$. Then*

$$E\left[\max_{1\leq t\leq T}\left(\sum_{j=1}^{t}\left(|\varepsilon_j|\,w_t - E\left(|\varepsilon_j|\,w_t\right)\right)\right)^2\right] \leq K\sum_{j=1}^{t}E\left(|\varepsilon_j|\,w_t\right)^2 \quad K<\infty$$

Proof of lemma 12. The proof of this lemma comes out from the corollary 16.10 of Davidson (1994). To use it we need to prove that $|\varepsilon_j|\,w_t - E\left(|\varepsilon_j|\,w_t\right)$ is a $L_2-Mixingale$ of size $-1/2$. Using the lemmas 8, 10 and 11 we can prove that $|\varepsilon_j|\,w_t$ is $L_4-bounded$ sequence and a L_2-NED of size $-1/2$. Then we can use the Theorem 17.5 of Davidson and we obtain the result. ■

Proof of Claim 1. As is easy to see, $B_\epsilon = \left\{\theta\in\Theta|\,\rho\left(\theta,\theta^0\right)>\epsilon\right\}$ is a compact set, then exist a partition $B_{\epsilon,j}(\theta) = \left\{\theta\in B_\epsilon|\,\rho\left(\theta,\theta^j\right)<\mu\right\}$ such that $B_\epsilon \subset \cup_{j=1}^J B_{\epsilon,j}(\theta)$ with $J=\frac{4\overline{r}}{\mu^3}$.

We can define $\theta^j = \left(\theta_1^0 + j_1\mu, \theta_2^0 + j_2\mu, r^0 + j_3\mu\right)$. We decompose it in a differentiable a non-differentiable part

$$\sup_{\theta\in B_\epsilon} \frac{\sum_{t=1}^T 2\varepsilon_t \left(\theta_{t-1}e_{t-1}(\omega,\theta) - \theta_{t-1}^0\varepsilon_{t-1}\right)}{T} \le \sup_{\theta\in B_\epsilon} \frac{\sum_{t=1}^T 2\varepsilon_t \left(\theta_{t-1}e_{t-1}(\omega,\theta) - \theta'_{t-1}e'_{t-1}\right)}{T} + \sup_{\theta\in B_\epsilon} \frac{\sum_{t=1}^T 2\varepsilon_t \left(\theta'_{t-1}e'_{t-1} - \theta_{t-1}^0\varepsilon_{t-1}\right)}{T}$$

For the differentiable part, and using equation [3]

$$\sup_{\theta\in B_\epsilon} \frac{\sum_{t=1}^T 2\varepsilon_t \left(\theta_{t-1}e_{t-1}(\omega,\theta) - \theta'_{t-1}e'_{t-1}\right)}{T} \le \sup_{\theta\in B_\epsilon} \frac{\sum_{t=1}^T 2\varepsilon_t \sum_{i=1}^2 \left(\theta_i - \theta_i^0\right) h_t^i\left(\theta^0\right)}{T} + + \sup_{\theta\in B_\epsilon} \frac{\sum_{t=1}^T 2\varepsilon_t \sum_{i=1}^2 \left(\theta_i - \theta_i^0\right)\left(h_t^i\left(\theta^0\right) - h_t^i\left(\theta^*\right)\right)}{T}$$

with $\theta^* \in \left(\theta_i, \theta_i^0\right)$. We work with each part. In the first part we use that it is a martingale difference sequence,

$$\sup_{\theta\in B_\epsilon} \frac{\sum_{t=1}^T 2\varepsilon_t \sum_{i=1}^2 \left(\theta_i - \theta_i^0\right) h_t^i\left(\theta^0\right)}{T} \le \left|\frac{2\sum_{t=1}^T 2\varepsilon_t \sum_{i=1}^2 h_t^i\left(\theta^0\right)}{T}\right|$$

$$P\left[\sup_{\theta\in B_\epsilon} \frac{\sum_{t=1}^T 2\varepsilon_t \sum_{i=1}^2 \left(\theta_i - \theta_i^0\right) h_t^i\left(\theta^0\right)}{T} > \eta\right] \le \frac{16\sigma_\varepsilon^2 \sum_{i=1}^2 \left\|h_t^i\left(\theta^0\right)\right\|_2^2}{T\eta} \overset{T\to\infty}{\to} 0$$

For the second and third part we use the partition $B_{\epsilon,j}(\theta)$, but for θ^*, that is

$$\sup_{\theta\in B_\epsilon} \frac{\sum_{t=1}^T 2\varepsilon_t \sum_{i=1}^2 \left(\theta_i - \theta_i^0\right)\left(h_t^i\left(\theta^0\right) - h_t^i\left(\theta^*\right)\right)}{T} \le$$

$$\le \sup_{\theta\in B_\epsilon} \sum_{i=1}^2 \left|\theta_i - \theta_i^0\right| \sup_{\theta\in B_\epsilon} \left|\frac{\sum_{t=1}^T 2\varepsilon_t \sum_{i=1}^2 \left(h_t^i\left(\theta^0\right) - h_t^i\left(\theta^*\right)\right)}{T}\right|$$

$$\sup_{\theta\in B_\epsilon} \left|\frac{\sum_{t=1}^T 2\varepsilon_t \sum_{i=1}^2 \left(h_t^i\left(\theta^0\right) - h_t^i\left(\theta^*\right)\right)}{T}\right| \le \max_j \left|\frac{\sum_{t=1}^T 2\varepsilon_t \sum_{i=1}^2 \left(s_t^i\left(\theta^0, \theta^{*,,j}\right)\right)}{T}\right| + + \max_j \sup_{\theta\in B_{\epsilon,j}(\theta)} \left|\frac{\sum_{t=1}^T 2\varepsilon_t \sum_{i=1}^2 s_t^i\left(\theta^*, \theta^{*,j}\right)}{T}\right|$$

Then, using the lemma 11, 12 and the theory about the martingale difference sequence,

$$P\left[\max_j \left|\frac{\sum_{t=1}^T 2\varepsilon_t \sum_{i=1}^2 s_t^i\left(\theta^0, \theta^{*,,j}\right)}{T}\right| > \eta\right] \le \sum_{j=1}^J \frac{\mu j 2\sigma_\varepsilon^2 \sum_{i=1}^2 K_s}{T\eta^2} \le \frac{4\overline{r}^2\sigma_\varepsilon^2 K_s}{\mu^6 T\eta^2}$$

$$\sup_{\theta\in B_{\epsilon,j}(\theta)} \left|\frac{\sum_{t=1}^T 2\varepsilon_t \sum_{i=1}^2 s_t^i\left(\theta^*, \theta^{*,j}\right)}{T}\right| \le \frac{\sum_{t=1}^T 4\overline{\varepsilon}_t \sum_{i=1}^2 \overline{s}_t^i\left(\theta^*, \theta^{*,j}\right)}{T}$$

Using the lemmas 7 to 12

$$P\left[\max_j \frac{\sum_{t=1}^T 4\overline{\varepsilon}_t \sum_{i=1}^2 \overline{s}_t^i\left(\theta^*, \theta^{*,j}\right)}{T} > \eta\right] \le \frac{32\sigma_\varepsilon^2 K_s \overline{r}}{\mu^2 T\eta^2}$$

$$\max_j \frac{\sum_{t=1}^T 4\overline{\varepsilon}_t \sum_{i=1}^2 \overline{s}_t^i\left(\theta^*, \theta^{*,j}\right)}{T} \le \left(\mu 32\sigma_\varepsilon^2 K_s\right)^{1/2}$$

Taking μ small enough, s.t $\left(\mu 32\sigma_{\varepsilon}^{2}K_{s}\right)^{1/2} < \eta$

$$P\left[\max_{j}\sup_{\theta\in B_{\epsilon,j}(\theta)}\left|\frac{\sum_{t=1}^{T}2\varepsilon_{t}\sum_{i=1}^{2}s_{t}^{i}\left(\theta^{*},\theta^{*,j}\right)}{T}\right|>\eta\right]\overset{T\to\infty}{\longrightarrow}0$$

And since $\sup_{\theta\in B_{\epsilon}}\sum_{i=1}^{2}\left|\theta_{i}-\theta_{i}^{0}\right|\leq 4$

$$P\left[\sup_{\theta\in B_{\epsilon}}\frac{\sum_{t=1}^{T}2\varepsilon_{t}\sum_{i=1}^{2}\left(\theta_{i}-\theta_{i}^{0}\right)\left(h_{t}^{i}\left(\theta^{0}\right)-h_{t}^{i}\left(\theta^{*}\right)\right)}{T}>\eta\right]\overset{T\to\infty}{\longrightarrow}0$$

In the non-differentiable part

$$\sup_{\theta\in B_{\epsilon}}\frac{\sum_{t=1}^{T}2\varepsilon_{t}\left(\theta_{t-1}^{\prime}e_{t-1}^{\prime}-\theta_{t-1}^{0}\varepsilon_{t-1}\right)}{T}\leq\max_{j}\frac{\sum_{t=1}^{T}2\varepsilon_{t}\left(\theta_{t-1}^{\prime,j}e_{t-1}^{\prime,j}-\theta_{t-1}^{0}\varepsilon_{t-1}\right)}{T}+$$
$$+\max_{j}\sup_{B_{\epsilon,j}(\theta)}\frac{\sum_{t=1}^{T}2\varepsilon_{t}\left(\theta_{t-1}e_{t-1}-\theta_{t-1}^{\prime,j}e_{t-1}^{\prime,j}\left(\omega,\theta^{\prime,j}\right)\right)}{T}$$

Now in a similar way of lemma 8 we obtain that $|\varepsilon_{t}|\,\overline{v}_{t}$ is a $L_{4}-bounded$ sequence and a $L_{2}-Near\ Epoch\ Dependence$ (NED) of size $-1/2$ with constant $\left\||\varepsilon_{t}|\,\overline{v}_{t}\right\|_{4}$. Furthermore, using assumption $A.7$

$$\left\|\overline{e}_{t-1}1\left(r<|z_{t}|<r^{j}\right)\right\|_{2}\leq\left|r-r^{j}\right|\left[M\left(2K^{2}+E\left(\overline{e}_{t}^{2}\right)\right)\right]^{1/2}\leq\left|r-r^{j}\right|K_{ev}$$
$$E\left(|\varepsilon_{t}|^{2}\,\overline{v}_{t}^{2}\right)\leq\left|r-r^{j}\right|^{2}\sigma_{\varepsilon}^{2}K_{ev}^{2}$$

Equally, we can extend the lemma 12 to this case and obtain the result about $L_{2}-Mixingale$ of size $-1/2$.Then

$$\sup_{B_{\epsilon,j}(\theta)}\frac{\sum_{t=1}^{T}2\varepsilon_{t}\left(\theta_{t-1}e_{t-1}-\theta_{t-1}^{\prime,j}e_{t-1}^{\prime,j}\left(\omega,\theta^{\prime,j}\right)\right)}{T}\leq\frac{\sum_{t=1}^{T}|\varepsilon_{t}|\,\overline{v}_{t}\left(\theta,\theta^{\prime,j}\right)}{T}$$
$$P\left[\left|\max_{j}\sup_{B_{\epsilon,j}(\theta)}\frac{\sum_{t=1}^{T}2\varepsilon_{t}\left(\theta_{t-1}e_{t-1}-\theta_{t-1}^{\prime,j}e_{t-1}^{\prime,j}\left(\omega,\theta^{\prime,j}\right)\right)}{T}\right|>\eta\right]\leq$$
$$\leq JP\left[\left|\frac{\sum_{t=1}^{T}\left(|\varepsilon_{t}|\,\overline{v}_{t}\left(\theta,\theta^{\prime,j}\right)-E\left(|\varepsilon_{t}|\,\overline{v}_{t}\left(\theta,\theta^{\prime,j}\right)\right)\right)}{T}\right|>\eta-\mu\left(\sigma_{\varepsilon}^{2}K_{ev}^{2}\right)^{1/2}\right]\leq$$
$$\leq J\frac{\mu^{2}\sigma_{\varepsilon}^{2}K_{ev}^{2}}{T\left(\eta-\mu\left(\sigma_{\varepsilon}^{2}K_{ev}^{2}\right)^{1/2}\right)^{2}}\leq\frac{4\overline{r}\sigma_{\varepsilon}^{2}K_{ev}^{2}}{T\mu\left(\eta-\mu\left(\sigma_{\varepsilon}^{2}K_{ev}^{2}\right)^{1/2}\right)^{2}}$$

And for the first part, we can use again that $\varepsilon_{t}\left(\theta_{t-1}^{\prime,j}e_{t-1}^{\prime,j}-\theta_{t-1}^{0}\varepsilon_{t-1}\right)$ is a martingale difference sequence and

$$P\left[\left|\max_{j}\frac{\sum_{t=1}^{T}2\varepsilon_{t}\left(\theta_{t-1}^{\prime,j}e_{t-1}^{\prime,j}-\theta_{t-1}^{0}\varepsilon_{t-1}\right)}{T}\right|>\eta\right]\leq\frac{4}{\mu^{2}}\frac{\sum_{l=1}^{\overline{r}/\mu}\mu^{2}l^{2}\sigma_{\varepsilon}^{2}K_{ev}^{2}}{T\eta^{2}}\leq\frac{4\sigma_{\varepsilon}^{2}K_{ev}^{2}}{\mu^{3}T\eta^{2}}$$

Taking μ small enough, all the probabilities goes to 0 when $T\to\infty$. Then we have proved the claim 1. ■

Proof of claim 2. We do the prove for $r-r^{0}>0$, the other case is equal. If we define

$v_{t-1}\left(\theta,\theta^0\right)=\left(\theta_{t-1}e_{t-1}\left(\omega,\theta\right)-\theta_{t-1}^0\varepsilon_{t-1}\right)$, then, $e_t\left(\omega,\theta\right)=v_{t-1}+\varepsilon_t$ and

$$\inf_{\theta|\rho(\theta,\theta^0)>\epsilon}\sum_{t=1}^{T}\left(\theta_{t-1}e_{t-1}\left(\omega,\theta\right)-\theta_{t-1}^0\varepsilon_{t-1}\right)^2=\inf_{\theta|\rho(\theta,\theta^0)>\epsilon}\sum_{t=1}^{T}v_{t-1}^2$$

$$=\inf_{\theta|\rho(\theta,\theta^0)>\epsilon}\sum_{l=1}^{\left[\frac{T}{2}\right]}v_{k(l)-1}^2+v_{k(l)}^2=\inf_{\theta|\rho(\theta,\theta^0)>\epsilon}\sum_{l=1}^{\left[\frac{T}{2}\right]}W_{k(l)}\qquad\text{with } k\left(l\right)=2\left(l-1\right)+1$$

Analyzing in more detail $v_{k(l)}$, we obtain

$$v_{k(l)}=\left(\theta_{k(l)}e_{k(l)}\left(\omega,\theta\right)-\theta_{k(l)}^0\varepsilon_{k(l)}\right)=I\left(\left|z_{k(l)}\right|>r\right)\left[\left(\theta_1-\theta_1^0\right)\varepsilon_{k(l)}+\theta_1v_{k(l)-1}\right]+$$
$$I\left(\left|z_{k(l)}\right|<r^0\right)\left[\left(\theta_2-\theta_2^0\right)\varepsilon_{k(l)}+\theta_2v_{k(l)-1}\right]+I(r^0<\left|z_{k(l)}\right|<r)\left[\left(\theta_2-\theta_1^0\right)\varepsilon_{k(l)}+\theta_2v_{k(l)-1}\right]$$

There are three possible cases, first one, when $\left|\theta_1-\theta_1^0\right|>\epsilon$,

$$W_{k(l)}=\left(v_{k(l)-1}^2+v_{k(l)}^2\right)\geq\gamma^2 1\left(\left|v_{k(l)-1}\right|>\gamma\right)+1\left(\left|v_{k(l)-1}\right|<\gamma\right)1\left(\left|z_{k(l)}\right|>r\right)\left[\left|\epsilon\varepsilon_{k(l)}\right|-\gamma\right]^2$$

taking $\gamma=\left|\epsilon\frac{1}{2}\varepsilon_{k(l)}\right|$,

$$W_{k(l)}\geq 1\left(\left|z_{k(l)}\right|>\overline{r}\right)\frac{1}{4}\epsilon^2\varepsilon_{k(l)}^2$$

Now using the assumption $A.7$ and $A.1$

$$\left[\frac{2}{T}\right]\sum_{l=1}^{\left[\frac{T}{2}\right]}W_{k(l)}\overset{a.s}{\longrightarrow}\lim_{T\to\infty}\left[\frac{2}{T}\right]\sum_{l=1}^{\left[\frac{T}{2}\right]}E\left(W_{k(l)}\right)\geq\frac{1}{4}\epsilon^2a\left(1-\overline{p}\right)>0$$

The second case, when $\left|\theta_2-\theta_2^0\right|>\epsilon$,

$$\left[\frac{T}{2}\right]\sum_{l=1}^{\left[\frac{T}{2}\right]}W_{k(l)}\overset{a.s}{\longrightarrow}\lim_{T\to\infty}\left[\frac{2}{T}\right]\sum_{l=1}^{\left[\frac{T}{2}\right]}E\left(W_{k(l)}\right)\geq\frac{1}{4}\epsilon^2a\underline{p}>0$$

Finally, $\left|r-r^0\right|>\epsilon$ and $\left|\theta_1-\theta_1^0\right|<\frac{\delta}{2}$ (if the last condition it is not true we are in the first case):

$$W_{k(l)}\geq\gamma^2 1\left(\left|v_{k(l)-1}\right|>\gamma\right)+1\left(\left|v_{k(l)-1}\right|<\gamma\right)1\left(r^0<\left|z_{k(l)}\right|<r^0+\epsilon\right)\left[\frac{\delta}{2}\varepsilon_{k(l)}-\gamma\right]^2$$

Taking $\gamma=\frac{\delta}{4}\varepsilon_{k(l)}$

$$W_{k(l)}\geq 1\left(r^0<\left|z_{k(l)}\right|<r^0+\epsilon\right)\frac{\delta^2}{16}\varepsilon_{k(l)}^2$$

$$\left[\frac{T}{2}\right]\sum_{l=1}^{\left[\frac{T}{2}\right]}W_{k(l)}\overset{a.s}{\longrightarrow}\lim_{T\to\infty}\left[\frac{2}{T}\right]\sum_{l=1}^{\left[\frac{T}{2}\right]}E\left(W_{k(l)}\right)\geq\epsilon m\frac{\delta^2}{16}a>0$$

Let $a\left(\epsilon\right)=\min\left\{\frac{1}{4}\epsilon^2a\left(1-\overline{p}\right),\frac{1}{4}\epsilon^2a\underline{p},\epsilon m\frac{\delta^2}{16}a\right\}>0$, then for all $\eta>0$

$$\lim_{T\to\infty}P\left[\inf_{\theta|\rho(\theta,\theta^0)>\epsilon}\sum_{t=1}^{T}\left(\theta_{t-1}e_{t-1}\left(\omega,\theta\right)-\theta_{t-1}^0\varepsilon_{t-1}\right)^2>a\left(\epsilon\right)-\eta\right]=1$$

■

Before proving the theorem 1 we need study the memory properties of the second derivate of $e_t\left(\theta\right)$ with respect to θ, since we need a second order Taylor approximation. For that we use the following lemma:

Lemma 13. *Let* $h_t^{il}(\theta^*) = \left.\frac{\partial^2 e_t}{\partial\theta_i \partial\theta_l}\right|_{\theta^*}$*, we have*

$$h_t^{il}(\theta) = 1(\theta_{t-1} = \theta_i) h_{t-1}^l + 1(\theta_{t-1} = \theta_l) h_{t-1}^i + \theta_{t-1} h_{t-1}^{il}(\theta)$$
$$\sup_\theta h_t^{il}(\theta) \leq \overline{h}_t^{il} = \sum_{i=1}^{2} \overline{h}_{t-1}^i + \overline{\theta}_{t-1} \overline{h}_{t-1}^{il}$$

Then, under assumption A.3 to A.8, $\overline{h}_t^{il}$ *is a* L_4-*bounded sequence and a* L_2-NED *of size* $-1/2$ *with constant* $\left\|\overline{h}_t^{il}\right\|_4$.

Proof of lemma 13. The proof follows the same steps than the first derivate.

$$\left\|\overline{h}_t^{il}\right\|_4 \leq \sum_{i=1}^{2} \left\|\overline{h}_{t-1}^i\right\|_4 + \overline{\lambda}\left\|\overline{h}_{t-1}^{il}\right\|_4 \leq \frac{\sum_{i=1}^{2}\left\|\overline{h}_{t-1}^i\right\|_4}{1-\overline{\lambda}} < \infty$$

Now define the variable $F_{t-m}^{t+m} - measurable$

$$\overline{h}_t^{il,m} = 2\overline{h}_{t-1}^{i,m} + \overline{\theta}_{t-1}\overline{h}_{t-1}^{il,m}$$

Then

$$\left\|\overline{h}_t^{il} - \overline{h}_t^{il,m}\right\|_2 \leq \sum_{i=1}^{2}\left\|\overline{h}_{t-1}^i - \overline{h}_{t-1}^{i,m}\right\|_2 + \overline{\lambda}\left\|\overline{h}_{t-1}^{il} - \overline{h}_{t-1}^{il,m}\right\|_2 \leq \sum_{j=0}^{m-1} m\overline{\lambda}^{m-1}\sum_{i=1}^{2}\left\|\overline{h}_t^i\right\|_4 + \overline{\lambda}^m\left\|\overline{h}_t^{il}\right\|_4$$

taking $v_m = m^2\overline{\lambda}^{m-1}$ and $d_t = \left\|\overline{h}_t^{il}\right\|_4$ we are done. ■

Lemma 14. *Let* $d_t^{il}(\theta,\theta^*) = h_t^{il}(\theta) - h_t^{il}(\theta^*)$*, we have*

$$d_t^{il}(\theta,\theta^*) \leq \overline{d}_t^{il}(\theta,\theta^*) = \sum_{i=1}^{2}\left[1(r^* \leq |z_t| \leq r)\overline{h}_{t-1}^i(\theta^*) + \overline{s}_{t-1}^i(\theta,\theta^*)\right] + \delta_t(\theta,\theta^*)\overline{h}_{t-1}^{il}(\theta^*) + \overline{\theta}_{t-1}\overline{d}_{t-1}^{il}(\theta,\theta^*)$$

Then, under assumption A.3 to A.8, $\overline{d}_t^{il}(\theta,\theta^*)$ *is a* L_4-*bounded sequence and a* L_2-NED *of size* $-1/2$ *with constant* $\left\|\overline{d}_t^{il}(\theta,\theta^*)\right\|_4$*. Besides*

$$E\left(\overline{d}_t^{il}(\theta,\theta^*)\right) \leq \max\left(\delta_1^{1/2},\delta_2^{1/2},\delta_3\right)K_d \quad E\left(\overline{d}_t^{il}(\theta,\theta^*)\right)^2 \leq \max(\delta_1,\delta_2,\delta_3)K_d$$

Proof of lemma 14. In a same way as previous lemmas and using them

$$\begin{aligned}\left\|\overline{d}_t^{il}(\theta,\theta^*)\right\|_4 &\leq \delta_3^{1/4}\sum_{i=1}^{2}\left\|\overline{h}_{t-1}^i(\theta^*)\right\|_4 + \max\left(\delta_1^{4/2},\delta_2^{4/2},\delta_3\right)^{1/4}\left[2K^{1/4}s + \left\|\overline{h}_{t-1}^{i,l}(\theta^*)\right\|_4\right] + \\ &+ \overline{\lambda}\left\|\overline{d}_t^{il}(\theta,\theta^*)\right\|_4 \\ &\leq \max\left(\delta_1^{4/2},\delta_2^{4/2},\delta_3\right)^{1/4}\frac{\sum_{i=1}^{2}\left\|\overline{h}_{t-1}^i(\theta^*)\right\|_4 + \left[2K^{1/4}s + \left\|\overline{h}_{t-1}^{i,l}(\theta^*)\right\|_4\right]}{1-\overline{\lambda}} \\ &\leq \max\left(\delta_1^{4/2},\delta_2^{4/2},\delta_3\right)^{1/4}K_d^{1/4}\end{aligned}$$

The rest of the moment are straight forward. Define the variable $F_{t-m}^{t+m} - measurable$

$$\overline{d}_t^{il,m}(\theta,\theta^*) = \sum_{i=1}^{2}\left[1(r^* \leq |z_t| \leq r)\overline{h}_{t-1}^{i,m}(\theta^*) + \overline{s}_{t-1}^{i,m}(\theta,\theta^*)\right] + \delta_t(\theta,\theta^*)\overline{h}_{t-1}^{il,m}(\theta^*) + \overline{\theta}_{t-1}\overline{d}_{t-1}^{il,m}(\theta,\theta^*)$$

Then, using the previous lemmas

$$\left\| \overline{d}_t^{il}(\theta,\theta^*) - \overline{d}_t^{il,m}(\theta,\theta^*) \right\|_2 \le m^3 \overline{\lambda}^{m-1} \left\| \overline{d}_t^{il}(\theta,\theta^*) \right\|_2$$

■

Lemma 15. *The lemma 12 can be applied to* $\overline{h}_t^{il}$ *and* $\overline{d}_t^{il}(\theta,\theta^*)$*.*

Proof of lemma 15. Straight forward. ■

Proof of theorem 1. We need to prove $\widehat{\theta}_{i,T} = \theta_i^0 + O_p\left(T^{-1/2}\right)$ and $\widehat{r}_T = r_0 + O_p\left(T^{-1}\right)$. For that define $\Theta_\triangle = \left\{\theta \in \Theta : \frac{\triangle}{T} \le \left|\theta_i - \theta_i^0\right|^2 = \delta_i \le \epsilon \text{ and } \frac{\triangle}{T} \le \left|r - r^0\right| = \delta_3 \le \epsilon\right\}$, that is bounded because Θ is bounded. The limit on $\delta_i \le \epsilon$ will be satisfied for all $T \ge T^*$ with probability big enough for T^* big enough because of consistence of estimators. Then. to prove that the rate of convergence is the wanted, a sufficient condition is for all $\varepsilon > 0$ and T exist $\triangle < \infty$ such that:

$$P\left[\inf_{\{\theta\in\Theta_\triangle\}} Q_T(\omega,\theta) - Q_T\left(\omega,\theta^0\right) > 0\right] \ge 1-\varepsilon$$

The steps are parallel to these of the consistence in claim 1 and 2. Then we have to prove that for all $\eta, \varepsilon > 0$ $\exists \triangle < \infty$ such that

$$P\left[\sup_{\{\theta\in\Theta_\triangle\}} \frac{\sum_{t=1}^T \varepsilon_t\left(\theta_{t-1}e_{t-1}(\omega,\theta) - \theta_{t-1}^0\varepsilon_{t-1}\right)}{T\sup_i \delta_i} > \eta\right] \le \varepsilon \tag{5}$$

and

$$P\left[\sup_{\{\theta\in\Theta_\triangle\}} \frac{\sum_{t=1}^T \left(\theta_{t-1}e_{t-1} - \theta_{t-1}^0\varepsilon_{t-1}\right)^2}{T\sup_i \delta_i} > \eta\right] \ge 1-\varepsilon \tag{6}$$

We divide in first probability to use the derivative part of $\left(\theta_{t-1}e_{t-1}(\omega,\theta) - \theta_{t-1}^0\varepsilon_{t-1}\right)$ with respect to θ_i.

As we saw, with $\theta' = \left(\theta_1^0, \theta_2^0, r\right)$, $\left(\theta_{t-1}e_{t-1}(\omega,\theta) - \theta_{t-1}^0\varepsilon_{t-1}\right) = \left(\theta_{t-1}e_{t-1}(\omega,\theta) - \theta'_{t-1}e'_{t-1}\right) + \left(\theta'_{t-1}e'_{t-1}(\omega,\theta) - \theta_{t-1}^0\varepsilon_{t-1}\right)$.

The process are the same, and their properties too. As $\Theta_\triangle$ is bounded, we can define the partition for (θ_1,θ_2) with $\theta_i^{j_i} = \theta_i^0 + \left(\frac{\triangle}{T}\left(b^{j_i}-1\right)\right)^{1/2}$, $r^{j_3} = r^0 + \frac{\triangle}{T}\left(b^{j_3}-1\right)$ with $\sup_j \frac{\triangle}{T}\left(b^j - 1\right) \le \epsilon$, and $B_{\triangle,j}(\theta) = \left\{\theta_i \in \left(\theta_i^{j_i}, \theta_i^{j_i+1}\right), r \in \left(r^{j_3}, r^{j_3+1}\right)\right\}$. We use a second Taylor expansion as in claim 1 for $\left(e_t(\theta) - e'_t(\theta')\right)$,

$$\left(e_t(\theta) - e'_t(\theta')\right) = \sum_{i=1}^2 \left(\theta_i - \theta'_i\right) h_t^i\left(\theta^0\right) + \sum_{i=1}^2 \left(\theta_i - \theta'_i\right)\left(h_t^i\left(\theta^0\right) - h_t^i(\theta')\right) + \sum_{i=1}^2\sum_{l=1}^2 \left(\theta_i - \theta'_i\right) h_t^{il}(\theta^*)\left(\theta_l - \theta'_l\right)$$

We work with the same parts of claim 1. Then

$$\begin{aligned}
\sup_{\{\theta\in\Theta_\triangle\}} \frac{\sum_{t=1}^T 2\varepsilon_t\left(\theta_{t-1}e_{t-1}(\omega,\theta) - \theta'_{t-1}e'_{t-1}\right)}{T\sup_i\delta_i} &\le \sup_{\{\theta\in\Theta_\triangle\}} \frac{\sum_{t=1}^T 2\varepsilon_t \sum_{i=1}^2 \left(\theta_i - \theta_i^0\right) h_t^i\left(\theta^0\right)}{T\sup_i\delta_i} + \\
&+ \sup_{\{\theta\in\Theta_\triangle\}} \frac{\sum_{t=1}^T 2\varepsilon_t \sum_{i=1}^2 \left(\theta_i - \theta_i^0\right)\left(h_t^i\left(\theta^0\right) - h_t^i(\theta')\right)}{T\sup_i\delta_i} + \\
&+ \sup_{\{\theta\in\Theta_\triangle\}} \frac{\sum_{t=1}^T 2\varepsilon_t \sum_{i=1}^2\sum_{l=1}^2 \left(\theta_i - \theta_i^0\right) h_t^{il}(\theta^*)\left(\theta_l - \theta_l^0\right)}{T\sup_i\delta_i}
\end{aligned}$$

Working with each supremun,

$$P\left[\sup_{\{\theta\in\Theta_{\triangle}\}}\frac{\sum_{t=1}^{T}2\varepsilon_t\sum_{i=1}^{2}\left(\theta_i-\theta_i^0\right)h_t^i\left(\theta^0\right)}{T\sup_i\delta_i}>\eta\right]\leq$$

$$\leq P\left[\sup_{\{\theta\in\Theta_{\triangle}\}}\left|\frac{\max_{1,2}\left(\theta_i-\theta_i^0\right)}{\sup_i\delta_i^{1/2}}\right|\left|\frac{\sum_{t=1}^{T}2\varepsilon_t\sum_{i=1}^{2}h_t^i\left(\theta^0\right)}{T\sup_i\delta_i^{1/2}}\right|>\eta\right]$$

$$\leq\frac{\sigma_{\varepsilon}^2 2^2\sum_{i=1}^{2}\left\|h_t^i\left(\theta^0\right)\right\|}{\Delta\eta^2}$$

The second supremun,

$$\sup_{\{\theta\in\Theta_{\triangle}\}}\frac{\sum_{t=1}^{T}2\varepsilon_t\sum_{i=1}^{2}\left(\theta_i-\theta_i^0\right)\left(h_t^i\left(\theta^0\right)-h_t^i\left(\theta'\right)\right)}{T\sup_i\delta_i}=\max_j\frac{\sum_{t=1}^{T}2\varepsilon_t\sum_{i=1}^{2}\left(\theta_i^j-\theta_i^0\right)\left(s_t^i\left(\theta^0,\theta^{',j}\right)\right)}{T\sup_i\delta_i}+$$

$$\max_j\sup_{B_{\triangle,j}(\theta)}\frac{\sum_{t=1}^{T}2\varepsilon_t\sum_{i=1}^{2}\left(\theta_i^0-\theta_i\right)s_t\left(\theta',\theta^{',j}\right)+\left(\theta_i-\theta_i^j\right)s_t\left(\theta^0,\theta^{',j}\right)}{T\sup_i\delta_i}$$

$$P\left[\max_j\frac{\sum_{t=1}^{T}2\varepsilon_t\sum_{i=1}^{2}\left(\theta_i^{j_i}-\theta_i^0\right)\left(s_t^i\left(\theta^0,\theta^{',j}\right)\right)}{T\sup_i\delta_i}>\eta\right]\leq\sum_{j=1}^{\infty}\frac{\sum_{l=1}^{2}2^8\sigma_{\varepsilon}^2K_s}{\eta^2T\sup_i\delta_i}$$

$$\leq\frac{1}{\triangle}\sum_{j=1}^{\infty}\frac{\sum_{l=1}^{2}2^8\sigma_{\varepsilon}^2K_s}{\eta^2\left(b^j-1\right)}$$

with $\sum_{j=1}^{\infty}\frac{\sum_{l=1}^{2}2^8\sigma_{\varepsilon}^2K_s}{\eta^2(b^j-1)}<\infty$. For the other part, use $\sup_{B_{\triangle,j}(\theta)}\left(\theta_i-\theta_i^j\right)\leq\left(\frac{\triangle}{T}b^j\left(b-1\right)\right)^{1/2}$,

$$\sup_{B_{\triangle,j}(\theta)}\frac{\sum_{t=1}^{T}2\varepsilon_t\sum_{i=1}^{2}\left(\theta_i^0-\theta_i\right)s_t\left(\theta',\theta^{',j}\right)+\left(\theta_i-\theta_i^j\right)s_t\left(\theta^0,\theta^{',j}\right)}{T\sup_i\delta_i}\leq$$

$$\frac{\sum_{t=1}^{T}2\left|\varepsilon_t\right|\sum_{i=1}^{2}\left|\theta_i^0-\theta_i\right|\overline{s}_t\left(\theta',\theta^{',j}\right)+\left|\theta_i-\theta_i^j\right|\overline{s}_t\left(\theta^0,\theta^{',j}\right)}{T\sup_i\delta_i}$$

Using that $\sup_{B_{\triangle,j}(\theta)}\left|\theta_i-\theta_i^j\right|\leq\sup_j\left(\frac{\triangle}{T}b^j\left(b-1\right)\right)^{1/2}\leq\left(\epsilon\left(b-1\right)\right)^{1/2}$,

$$\max_j\frac{\sum_{t=1}^{T}E\left[2\left|\varepsilon_t\right|\sum_{i=1}^{2}\left|\theta_i^0-\theta_i\right|\overline{s}_t\left(\theta',\theta^{',j}\right)+\left|\theta_i-\theta_i^j\right|\overline{s}_t\left(\theta^0,\theta^{',j}\right)\right]}{T\sup_i\delta_i}\leq$$

$$\leq 2^4\left(\sigma_{\varepsilon}^2K_s\right)^{1/2}\left[\left(b-1\right)+\left(\epsilon\left(b-1\right)\right)^{1/2}\right]$$

taking b close enough to 1, this amount is small enough.

$$P\left[\max_j \frac{\sum_{t=1}^T 2\left|\varepsilon_t\right| \sum_{i=1}^2 \left|\theta_i^0 - \theta_i\right| \overline{s}_t\left(\theta', \theta'^{,j}\right) - E\left[2\left|\varepsilon_t\right| \sum_{i=1}^2 \left|\theta_i^0 - \theta_i\right| \overline{s}_t\left(\theta', \theta'^{,j}\right)\right]}{T \sup_i \delta_i} > \eta\right] \leq$$

$$\leq \frac{1}{\triangle} \sum_{j=1}^{\infty} \frac{(b-1)\,\epsilon 2^3 \sigma_\varepsilon^2 K_s}{\left(b^j - 1\right)\eta}$$

$$P\left[\max_j \frac{\sum_{t=1}^T 2\left|\varepsilon_t\right| \sum_{i=1}^2 \left|\theta_i - \theta_i^j\right| \overline{s}_t\left(\theta^0, \theta'^{,j}\right) - E\left[2\left|\varepsilon_t\right| \sum_{i=1}^2 \left|\theta_i - \theta_i^j\right| \overline{s}_t\left(\theta^0, \theta'^{,j}\right)\right]}{T \sup_i \delta_i} > \eta\right] \leq$$

$$\leq \frac{1}{\triangle} \sum_{j=1}^{\infty} \frac{(b-1)\,\epsilon 2^3 \sigma_\varepsilon^2 K_s}{\left(b^j - 1\right)\eta}$$

Then

$$P\left[\sup_{\{\theta \in \Theta_\triangle\}} \frac{\sum_{t=1}^T 2\varepsilon_t \sum_{i=1}^2 \left(\theta_i - \theta_i^0\right) h_t^i\left(\theta^0\right)}{T \sup_i \delta_i} > \eta'\right] \leq \frac{1}{\triangle} \sum_{j=1}^{\infty} \frac{(b-1)\,\epsilon 2^3 \sigma_\varepsilon^2 K_s}{\left(b^j - 1\right)\eta}$$

Finally the third supremum,

$$\sup_{\{\theta \in \Theta_\triangle\}} \frac{\sum_{t=1}^T 2\varepsilon_t \sum_{i=1}^2 \sum_{l=1}^2 \left(\theta_i - \theta_i^0\right) h_t^{il}\left(\theta^*\right)\left(\theta_l - \theta_l^0\right)}{T \sup_i \delta_i} \leq$$

$$\leq \sup_{\{\theta \in \Theta_\triangle\}} \frac{\sum_{i=1}^2 \sum_{l=1}^2 \left(\theta_i - \theta_i^0\right)\left(\theta_l - \theta_l^0\right)}{\left(\sup_i \delta_i\right)^{1-\alpha}} \sup_{\{\theta \in \Theta_\triangle\}} \frac{\sum_{t=1}^T 2\varepsilon_t \sum_{i=1}^2 \sum_{l=1}^2 h_t^{il}\left(\theta^*\right)}{T\left(\sup_i \delta_i\right)^{\alpha}}$$

$$\leq 4\epsilon^\alpha \sup_{\{\theta \in \Theta_\triangle\}} \frac{\sum_{t=1}^T 2\varepsilon_t \sum_{i=1}^2 \sum_{l=1}^2 h_t^{il}\left(\theta^*\right)}{T^{1-\alpha} \triangle^\alpha}$$

for $\alpha < 1/2$. With lemmas 13 and 14 and following the steps of the consistence proof it can be proved that

$$P\left[\sup_{\{\theta \in \Theta_\triangle\}} \frac{\sum_{t=1}^T 2\varepsilon_t \sum_{i=1}^2 \sum_{l=1}^2 h_t^{il}\left(\theta^*\right)}{T^{1-\alpha} \triangle^\alpha} > \eta\right] \leq \frac{K}{\triangle^{2\alpha}}$$

for some $K < \infty$. With this, the differentiable part satisfies that

$$P\left[\sup_{\{\theta \in \Theta_\triangle\}} \frac{\sum_{t=1}^T 2\varepsilon_t \left(\theta_{t-1} e_{t-1}(\omega, \theta) - \theta'_{t-1} e'_{t-1}\right)}{T \sup_i \delta_i} > \eta\right] \leq \frac{K'}{\triangle^{2\alpha}}$$

for some $K' < \infty$.

Now we work with the non-differentiable part. That only depend on the parameter r, the we concentrate in $r^j = r^0 + \frac{\triangle}{T}\left(b^j - 1\right)$ with $\sup_j \frac{\triangle}{T}\left(b^j - 1\right) \leq \epsilon$, and $B_{\triangle,j}(\theta) = \left\{\theta_i = \theta_i^0, r \in \left(r^{j_3}, r^{j_3+1}\right)\right\}$.

$$\sup_{\{\theta \in \Theta_\triangle\}} \frac{\sum_{t=1}^T 2\varepsilon_t \left(\theta'_{t-1} e'_{t-1} - \theta_{t-1}^0 \varepsilon_{t-1}\right)}{T \sup_i \delta_i} \leq \max_j \frac{\sum_{t=1}^T 2\varepsilon_t \left(\theta_{t-1}'^{,j} e_{t-1}'^{,j} - \theta_{t-1}^0 \varepsilon_{t-1}\right)}{T \sup_i \delta_i} +$$

$$+ \max_j \sup_{B_{\epsilon,j}(\theta)} \frac{\sum_{t=1}^T 2\varepsilon_t \left(\theta_{t-1} e_{t-1} - \theta_{t-1}'^{,j} e_{t-1}'^{,j}\left(\omega, \theta'^{,j}\right)\right)}{T \sup_i \delta_i}$$

Using the lemma 10 and $\sup_i \delta_i \geq \frac{\triangle}{T}\left(b^j - 1\right)$ for each j,

$$P\left[\max_j \frac{\sum_{t=1}^T 2\varepsilon_t \left(\theta_{t-1}^{\prime,j} e_{t-1}^{\prime,j} - \theta_{t-1}^0 \varepsilon_{t-1}\right)}{T \sup_i \delta_i} > \eta\right] \leq \sum_{j=1}^{\infty} \frac{\frac{\triangle}{T}\left(b^j - 1\right) K_v}{T \sup_i \delta_i^2 \eta^2} \leq \frac{1}{\triangle} \sum_{j=1}^{\infty} \frac{K_v}{\left(b^j - 1\right)\eta^2}$$

$$P\left[\max_j \frac{\sum_{t=1}^T 2\varepsilon_t \sum_{i=1}^2 \left(\theta_i^{j_i} - \theta_i^0\right)\left(s_t^i\left(\theta^0, \theta^{\prime,j}\right)\right)}{T \sup_i \delta_i} > \eta\right] \leq \sum_{j=1}^{\infty} \frac{\sum_{l=1}^2 2^8 \sigma_\varepsilon^2 K_s}{\eta^2 T \sup_i \delta_i} \leq \frac{1}{\triangle} \sum_{j=1}^{\infty} \frac{\sum_{l=1}^2 2^8 \sigma_\varepsilon^2 K_s}{\eta^2 \left(b^j - 1\right)}$$

$$P\left[\max_j \frac{\sum_{t=1}^T 2\left|\varepsilon_t\right| \overline{v}_t\left(\theta, \theta^{\prime,j}\right) - E\left(2\left|\varepsilon_t\right| \overline{v}_t\left(\theta, \theta^{\prime,j}\right)\right)}{T \sup_i \delta_i} > \eta\right] \leq \sum_{j=1}^{\infty} \frac{\frac{\triangle}{T} b^j (b-1) K_v}{T \sup_i \delta_i^2 \eta^2}$$

$$\leq \frac{1}{\triangle} \sum_{j=1}^{\infty} \frac{2(b-1) K_v}{\left(b^j - 1\right)\eta^2}$$

$$\max_j \frac{\sum_{t=1}^T E\left(2\left|\varepsilon_t\right| \overline{v}_t\left(\theta, \theta^{\prime,j}\right)\right)}{T \sup_i \delta_i} = \max_j \frac{\frac{\triangle}{T} b^j (b-1) K_v}{\sup_i \delta_i \eta^2} \leq \frac{2(b-1) K_v}{\eta^2}$$

Taking $(b-1)$ small enough

$$P\left[\max_j \sup_{B_{\epsilon,j}(\theta)} \frac{\sum_{t=1}^T 2\varepsilon_t \left(\theta_{t-1} e_{t-1} - \theta_{t-1}^{\prime,j} e_{t-1}^{\prime,j}\left(\omega, \theta^{\prime,j}\right)\right)}{T \sup_i \delta_i} > \eta\right] \leq \frac{1}{\triangle} \sum_{j=1}^{\infty} \frac{2(b-1) K_v}{\left(b^j - 1\right)\eta^2}$$

With this we have proved 5.

To prove 6 we use the result of claim 2. As in that case, we do the prove for $r - r^0 > 0$.

$$\inf_{\{\theta \in \Theta_\triangle\}} \sum_{t=1}^T \left(\theta_{t-1} e_{t-1}(\omega, \theta) - \theta_{t-1}^0 \varepsilon_{t-1}\right)^2 = \inf_{\{\theta \in \Theta_\triangle\}} \sum_{l=1}^{\left[\frac{T}{2}\right]} W_{k(l)}$$

Again we have three possible cases, $\left|\theta_i - \theta_i^0\right| > \left(\frac{\triangle}{T}\right)^{1/2}$ for $i = 1, 2$ or $\left|r - r^0\right| > \frac{\triangle}{T}$, if none is true, the $\theta \notin \Theta_\triangle$. First one, when $\left|\theta_1 - \theta_1^0\right|^2 > \frac{\triangle}{T}$,

$$W_{k(l)} = \left(v_{k(l)-1}^2 + v_{k(l)}^2\right) \geq \gamma^2 1\left(\left|v_{k(l)-1}\right| > \gamma\right) + 1\left(\left|v_{k(l)-1}\right| < \gamma\right) 1\left(\left|z_{k(l)}\right| > r\right) \left[\left|\left(\frac{\triangle}{T}\right)^{1/2} \varepsilon_{k(l)}\right| - \gamma\right]^2$$

taking $\gamma = \left|\left(\frac{\triangle}{T}\right)^{1/2} \frac{1}{2} \varepsilon_{k(l)}\right|$,

$$W_{k(l)} \geq 1\left(\left|z_{k(l)}\right| > \overline{r}\right) \frac{1}{4} \frac{\triangle}{T} \varepsilon_{k(l)}^2$$

Now using the assumption $A.7$ and $A.1$

$$\sum_{l=1}^{\left[\frac{T}{2}\right]} W_{k(l)} \geq \frac{1}{T} \sum_{l=1}^{\left[\frac{T}{2}\right]} 1\left(\left|z_{k(l)}\right| > \overline{r}\right) \frac{1}{4} \triangle \varepsilon_{k(l)}^2 \xrightarrow{a.s} \frac{1}{2} \lim_{T \to \infty} \left[\frac{2}{T}\right] \sum_{l=1}^{\left[\frac{T}{2}\right]} E\left(1\left(\left|z_{k(l)}\right| > \overline{r}\right) \frac{1}{4} \triangle \varepsilon_{k(l)}^2\right) \geq \frac{1}{8} \triangle a \left(1 - \overline{p}\right) > 0$$

The second case, when $\left|\theta_2 - \theta_2^0\right|^2 > \frac{\triangle}{T}$,

$$\sum_{l=1}^{\left[\frac{T}{2}\right]} W_{k(l)} \geq \frac{1}{T} \sum_{l=1}^{\left[\frac{T}{2}\right]} \left(\left|z_{k(l)}\right| < r^0\right) \frac{1}{4} \triangle \varepsilon_{k(l)}^2 \xrightarrow{a.s} \frac{1}{2} \lim_{T \to \infty} \left[\frac{2}{T}\right] \sum_{l=1}^{\left[\frac{T}{2}\right]} E\left(1\left(\left|z_{k(l)}\right| < r\right) \frac{1}{4} \triangle \varepsilon_{k(l)}^2\right) \geq \frac{1}{8} \triangle a \underline{p} > 0$$

Finally, $\left|r-r^0\right| > \frac{\triangle}{T}$ and $\left|\theta_1-\theta_1^0\right| < \epsilon < \delta$ (since $\theta \in \Theta_\triangle$):

$$W_{k(l)} \geq \gamma^2 1\left(\left|v_{k(l)-1}\right| > \gamma\right) + 1\left(\left|v_{k(l)-1}\right| < \gamma\right) 1\left(r^0 < \left|z_{k(l)}\right| < r^0 + \frac{\triangle}{T}\right)\left[\left(\delta-\epsilon\right)\varepsilon_{k(l)} - \gamma\right]^2$$

Taking $\gamma = \frac{(\delta-\epsilon)}{2}\varepsilon_{k(l)}$

$$W_{k(l)} \geq 1\left(r^0 < \left|z_{k(l)}\right| < r^0 + \frac{\triangle}{T}\right)\frac{(\delta-\epsilon)^2}{4}\varepsilon_{k(l)}^2$$

$$\sum_{l=1}^{\left[\frac{T}{2}\right]} W_{k(l)} \geq \frac{1}{T}\sum_{l=1}^{\left[\frac{T}{2}\right]} 1\left(r^0 < \left|z_{k(l)}\right| < r^0 + \frac{\triangle}{T}\right) T\frac{(\delta-\epsilon)^2}{4}\varepsilon_{k(l)}^2 \xrightarrow{a.s} \triangle m\frac{(\delta-\epsilon)^2}{4}a > 0$$

Let $a(\triangle) = \min\left\{\frac{1}{8}\triangle a\left(1-\overline{p}\right), \frac{1}{8}\triangle a\underline{p}, \triangle m\frac{(\delta-\epsilon)^2}{4}a\right\} > 0$, then for all $\eta > 0$

$$P\left[\inf_{\{\theta\in\Theta_\triangle\}}\sum_{t=1}^{T}\left(\theta_{t-1}e_{t-1}(\omega,\theta) - \theta_{t-1}^0\varepsilon_{t-1}\right)^2 > a(\triangle) - \eta\right] \leq 1-\varepsilon$$

■

Before proving the claim 1 for STIMA case we need to study the memory and moment properties of some processes, as e_t and θ_t. For that we use the lemma 16 and 17:

Lemma 16.

$$E\left(\left.\theta_t^4\right| F_{t-2}\right) \leq \overline{\lambda} < 1$$

Proof of lemma 16. It is a direct result of applying A.9 since

$$\begin{aligned} E\left(\left.\theta_t^4\right| F_{t-2}\right) &\leq E\left(\left.\left[(1-\delta)^4 1\left(\left|e_{t-1}\right| > r\right) + (1+\delta')^4 1\left(\left|e_{t-1}\right| < r\right)\right]\right| F_{t-2}\right) \leq \\ &\leq (1-\delta)^4\left[1-\sup_k E\left(1\left(\left|\varepsilon_{t-1}+k\right| < r\right)\right)\right] + (1+\delta')^4\sup_k E\left(1\left(\left|\varepsilon_{t-1}+k\right| < r\right)\right) \\ &\leq \overline{\lambda} \end{aligned}$$

■

Lemma 17. *Under the assumptions A.2S, A.3 and A.4 $e_t(w,\theta)$ is a L_4-bounded sequence and a L_2- Near Epoch Dependence (NED) of size $-1/2$ with constant $\left\|e_t\right\|_4$.*

Proof of lemma 17.

$$\left\|e_t\right\|_4 \leq \left\|x_t + \sum_{j=0}^{t-2}\prod_{i=0}^{j}\theta_{t-i}x_{t-j-1}\right\|_4 \leq 2\left\|x_t\right\|_4 + \sum_{j=1}^{t-2}\overline{\lambda}^{j-1}\left\|x_{t-j-1}\right\|_4 \leq \frac{2\left\|x_t\right\|_4}{1-\overline{\lambda}} < \infty$$

With respect to L_2-NED, taking

$$\widehat{e}_t^m = x_t + \sum_{j=1}^{m-1}x_{t-j}\prod_{k=1}^{j}\widehat{\theta}_{t-k}^m \qquad con\ \widehat{\theta}_{t-k}^m = \theta_{t-k}\left(\widehat{e}_{t-k}^m\right)$$

$$\widehat{e}_{t-j}^m = \begin{cases} \widehat{\theta}_{t-j-1}^m\widehat{e}_{t-j-1}^m + x_{t-j} & if\ 0 \leq j < m \\ 0 & if\ j \geq m \end{cases}$$

we know

$$\left\|\overline{e}_t - E\left(\left.\overline{e}_t\right| F_{t-m}^{t+m}\right)\right\|_2 \leq \left\|e_t - \widehat{e}_t^m\right\|_2$$

As in the invertible section

$$e_t - \widehat{e}_t^m = \theta_{t-1} e_{t-1} - \widehat{\theta}_{t-1}^m \widehat{e}_{t-1}^m = (\theta_1 - \theta_2)\, 1\left(|e_{t-1}| > r, \left|\widehat{e}_{t-1}^m\right| < r\right) \widehat{e}_{t-1}^m +$$
$$+ (\theta_2 - \theta_1)\, 1\left(|e_{t-1}| < r, \left|\widehat{e}_{t-1}^m\right| > r\right) e_{t-1} + \left[\theta_1 1\left(|e_{t-1}| > r\right) + \theta_2 1\left(|e_{t-1}| < r\right)\right]\left(e_{t-1} - \widehat{e}_{t-1}^m\right)$$
$$\left|e_t - \widehat{e}_t^m\right| \leq |\theta_2 - \theta_1|\, r\left[1\left(|e_{t-1}| < r, \left|\widehat{e}_{t-1}^m\right| > r\right) + 1\left(|e_{t-1}| > r, \left|\widehat{e}_{t-1}^m\right| < r\right)\right] +$$
$$+ \left[|\theta_1|\, 1\left(|e_{t-1}| > r\right) + |\theta_2|\, 1\left(|e_{t-1}| < r\right)\right]\left|e_{t-1} - \widehat{e}_{t-1}^m\right|$$

Now using the same arguments than invertibility in equation [3] ,we obtain

$$E\left\{\left[1\left(|e_{t-1}| < r, \left|\widehat{e}_{t-1}^m\right| > r\right) + 1\left(|e_{t-1}| > r, \left|\widehat{e}_{t-1}^m\right| < r\right)\right]^\gamma\right\} \leq ME\left[\left|e_{t-1} - e_{t-1}^m\right|\right] \ for\ \gamma > 0$$

then

$$E\left|e_t - \widehat{e}_t^m\right| \leq \left[|\theta_2 - \theta_1|\, rM + \lambda_1(\theta_1, \theta_2, r)\right] E\left|e_{t-1} - \widehat{e}_{t-1}^m\right| \leq \lambda_2^m E\left|e_{t-m}\right|$$

Besides,

$$\left|e_t - \widehat{e}_t^m\right|^2 \leq |\theta_2 - \theta_1|^2 r^2\left[1\left(|e_{t-1}| < r, \left|\widehat{e}_{t-1}^m\right| > r\right) + 1\left(|e_{t-1}| > r, \left|\widehat{e}_{t-1}^m\right| < r\right)\right] +$$
$$+ \left[\theta_1^2 1\left(|e_{t-1}| > r\right) + \theta_2^2 1\left(|e_{t-1}| < r\right)\right]\left|e_{t-1} - \widehat{e}_{t-1}^m\right|^2 +$$
$$2\,|\theta_2 - \theta_1|\, r\left[1\left(|e_{t-1}| < r, \left|\widehat{e}_{t-1}^m\right| > r\right) + 1\left(|e_{t-1}| > r, \left|\widehat{e}_{t-1}^m\right| < r\right)\right]\left[|\theta_1|\, 1\left(|e_{t-1}| > r\right) + |\theta_2|\, 1\left(|e_{t-1}| < r\right)\right]$$
$$E\left|e_t - \widehat{e}_t^m\right|^2 \leq KE\left[\left|e_{t-1} - e_{t-1}^m\right|\right] + \lambda_1(\theta_1, \theta_2, r)\, E\left|e_{t-1} - \widehat{e}_{t-1}^m\right|^2 = K\lambda_2^{m-1} E\left|e_{t-m}\right| +$$
$$+ \lambda_1(\theta_1, \theta_2, r)\, E\left|e_{t-1} - \widehat{e}_{t-1}^m\right|^2$$
$$\leq \sum_{i=0}^{m} K\lambda_2^{m-i-1}\lambda_1^i E\left|e_{t-m}\right| + \lambda_1^m(\theta_1, \theta_2, r)\, E\left|e_{t-m}\right|^2 \leq \lambda_2^m (m+1)\, \frac{KE\left|e_{t-m}\right|^2}{\lambda_2} \overset{m\to\infty}{\longrightarrow} 0$$

Taking $d_t = \frac{K^{1/2}\|e_{t-m}\|_2}{\lambda_2^{1/2}}$ and $v_m = \left[\lambda_2^m (m+1)\right]^{1/2}$ we have proved the lemma. ∎

Now we generalize the definition of $A_{i,t}$ and δ_i in the following way,

$$A_{1,t-1}(\theta, \theta') = \left\{|e_{t-1}(\theta)| > r, |e_{t-1}(\theta')| > r'\right\} \quad A_{2,t-1}(\theta, \theta') = \left\{|e_{t-1}(\theta)| < r, |e_{t-1}(\theta')| < r'\right\}$$
$$A_{3,t-1}(\theta, \theta') = \left\{|e_{t-1}(\theta)| > r, |e_{t-1}(\theta')| < r'\right\} \quad A_{4,t-1}(\theta, \theta') = \left\{|e_{t-1}(\theta)| < r, |e_{t-1}(\theta')| > r'\right\}$$

$$\delta_i = \delta_i(\theta, \theta') = \begin{cases} (\theta_i - \theta_i') \ for\ i = 1, 2 \\ (r - r') \ for\ i = 3 \end{cases}$$

With this we can define and study the following processes, $v_t(\theta, \theta') = (e_t(\theta) - e_t(\theta'))$ and $\overline{v}_t(\theta', \eta) = \sup_{\theta \text{ s.t } \rho(\theta, \theta') < \eta} v_t(\theta, \theta')$.

Lemma 18. *Under the assumptions A.3, A.4 and A.9 $\overline{v}_t(\theta', \delta_i)$ is a L_4-bounded sequence and a $L_2 -$ Near Epoch Dependence (NED) of size $-1/2$ with constant $\|\overline{v}_t(\theta', \delta_i)\|_4$. Besides, $E\left(\overline{v}_t^2(\theta', \delta_i)\right) \leq \max \delta_i K_2$, with $K_2 < \infty$.*

Proof of lemma 18. Our interest is when $\delta_i \to 0$, then we take $\delta_i \leq 1$. We can decompose $(e_t(\theta) - e_t(\theta'))$ as in equation [4], but taking in to account that θ can be different from θ'. Then

$$v_t(\theta, \theta') = e_{t-1}'\left[\delta_1 1\left(A_{1,t-1}'\right) + \delta_2 1\left(A_{2,t-1}'\right) + (\delta_1 + \delta_2)\, 1\left(A_{3,t-1}'\right)\right] +$$
$$+ \left[\theta_1 1\left(|e_{t-1}(\theta')| > r\right) + \theta_2 1\left(|e_{t-1}(\theta')| < r\right)\right] v_{t-1}(\theta, \theta') +$$
$$+ \left(\theta_2 - \theta_1^l\right)\left[e_{t-1} 1\left(A_{4,t-1}'\right) - e_{t-1}^l 1\left(A_{3,t-1}'\right)\right]$$

and

$$\overline{v}_t(\theta',\delta_i) \le 2\max\delta_i\left(\left|e'_{t-1}\right| + \delta + r' + n\right) + \left[\theta_1 1\left(\left|e_{t-1}(\theta')\right| > r\right) + \theta_2 1\left(\left|e_{t-1}(\theta')\right| < r\right)\right]\overline{v}_{t-1}(\theta',\delta_i)$$
$$+ \delta r' \sup_{\theta \text{ s.t } \rho(\theta,\theta')<\eta}\left[1\left(A'_{4,t-1}\right) + 1\left(A'_{3,t-1}\right)\right]$$
$$E\left|\overline{v}_t(\theta',\delta_i)\right| \le 2\max\delta_i\left(E\left|e'_{t-1}\right| + r' + 2\right) + \lambda_1^*(\theta')\,E\left|\overline{v}_{t-1}(\theta',\eta)\right| +$$
$$+ \delta r' E\left|\sup_{\theta \text{ s.t } \rho(\theta,\theta')<\eta}\left[1\left(A'_{4,t-1}\right) + 1\left(A'_{3,t-1}\right)\right]\right|$$

Similar to [**??**]

$$\sup_{\theta \text{ s.t } \rho(\theta,\theta')<\eta}\left[1\left(A'_{4,t-1}\right) + 1\left(A'_{3,t-1}\right)\right] \le \left[1\left(\overline{A}'_{4,t-1}\right) + 1\left(\overline{A}'_{3,t-1}\right)\right]$$
$$1\left(\overline{A}'_{4,t-1}\right) + 1\left(\overline{A}'_{3,t-1}\right) =$$
$$= 1\left[r' - v_{t-1}\left(\theta',\theta^0\right) - \eta - \overline{v}_{t-1}(\theta',\eta) < \varepsilon_{t-1} < r' - v_{t-1}\left(\theta',\theta^0\right) + \overline{v}_{t-1}(\theta',\eta)\right] +$$
$$+ 1\left[-r' - v_{t-1}\left(\theta',\theta^0\right) - \eta - \overline{v}_{t-1}(\theta',\eta) < \varepsilon_{t-1} < -r' - v_{t-1}\left(\theta',\theta^0\right) + \eta + \overline{v}_{t-1}(\theta',\eta)\right]$$

Using that v_{t-1} and $\overline{v}_{t-1}$ are $F_{t-2} - measurables$, the law of iterated expectations and the assumption $A.9$,

$$E\left|\left[1\left(\overline{A}'_{4,t-1}\right) + 1\left(\overline{A}'_{3,t-1}\right)\right]\right|^k \le E\left|\overline{v}_{t-1}(\theta',\delta_i)\right| M^* + \delta_3 M^* \quad \forall k \ge 1$$
$$E\left|\overline{v}_t(\theta',\delta_i)\right| \le \max\delta_i \frac{2\left(E\left|e'_{t-1}\right| + r' + 2 + M^*\right)}{1-\lambda_2^*} \le \max\delta_i \frac{2\left(\left\|e'_{t-1}\right\|_4 + r' + 2 + M^*\right)}{1-\lambda_2^*} \le \max\delta_i \frac{K_3}{1-\lambda_2^*} < \infty$$
$$\left\|\overline{v}_t(\theta',\delta_i)\right\|_4 \le \max\delta_i K_3 + \lambda_1^*(\theta')\left\|\overline{v}_{t-1}(\theta',\delta_i)\right\|_4 + \max\delta_i^{1/4}\left[\frac{K_3}{1-\lambda_2^*}\right]^{1/4}$$
$$\le \max\delta_i^{1/4} 2\frac{K_3}{1-\lambda_2^*} + \lambda_1^*(\theta')\left\|\overline{v}_{t-1}(\theta',\eta)\right\|_4 \le \max\delta_i^{1/4} 2\frac{K_3}{\left(1-\lambda_2^*\right)\left(1-\lambda_1^*\right)} < \infty$$

With this the property $E\left(\overline{v}_t^2(\theta',\delta_i)\right) \le \max\delta_i K_2$ is straight forward. Now we prove that $\overline{v}_t^2(\theta',\delta_i)$ is $L_2 - NED$. As we have proved, e'_t, θ'_t are $L_2 - NED$. Then $v_t\left(\theta',\theta^0\right) = \left(e'_t - \varepsilon_t\right)$ is $L_2 - NED$. Define the following functions $F_{t-m}^t - measurables$,

$$\overline{v}_t^m(\theta',\eta) = 2\eta\left(\left|e'^{,m}_{t-1}\right| + \delta + r' + n\right) + \left[\theta_1 1\left(\left|e^m_{t-1}(\theta')\right| > r\right) + \theta_2 1\left(\left|e^m_{t-1}(\theta')\right| < r\right)\right]\overline{v}^m_{t-1}(\theta',\eta) +$$
$$\delta r'\left[1\left(\overline{A}'^{,m}_{4,t-1}\right) + 1\left(\overline{A}'^{,m}_{3,t-1}\right)\right]$$
$$1\left(\overline{A}'^{,m}_{4,t-1}\right) + 1\left(\overline{A}'^{,m}_{3,t-1}\right) = 1\left(r' - v^m_{t-1}\left(\theta',\theta^0\right) - \eta - \overline{v}^m_{t-1}(\theta',\eta) < \varepsilon_{t-1} < r' - v^m_{t-1}\left(\theta',\theta^0\right) + \eta + \overline{v}^m_{t-1}(\theta',\eta)\right) +$$
$$1\left(r' - v^m_{t-1}\left(\theta',\theta^0\right) - \eta - \overline{v}^m_{t-1}(\theta',\eta) < \varepsilon_{t-1} < r' - v^m_{t-1}\left(\theta',\theta^0\right) + \eta + \overline{v}^m_{t-1}(\theta',\eta)\right)$$

Defining $\Lambda'^{,m}_{t-1} = v_{t-1}\left(\theta',\theta^0\right) - v^m_{t-1}\left(\theta',\theta^0\right)$ and $\overline{\Lambda}'^{,m}_{t-1} = \overline{v}_{t-1}\left(\theta',\theta^0\right) - \overline{v}^m_{t-1}\left(\theta',\theta^0\right)$, we have

$$1\left(\overline{A}'_{4,t-1}\right)-1\left(\overline{A}'^{,m}_{4,t-1}\right)+1\left(\overline{A}'_{3,t-1}\right)-1\left(\overline{A}'^{,m}_{3,t-1}\right)=$$

$$1\left(\begin{array}{c} r'-v_{t-1}\left(\theta',\theta^0\right)+\eta+\overline{v}_{t-1}\left(\theta',\eta\right)<\varepsilon_{t-1}<r'-v_{t-1}\left(\theta',\theta^0\right)+\eta+\overline{v}_{t-1}\left(\theta',\eta\right)+\Lambda'^{,m}_{t-1}-\overline{\Lambda}'^{,m}_{t-1} \\ \Lambda'^{,m}_{t-1}-\overline{\Lambda}'^{,m}_{t-1}>0 \end{array}\right)+$$
$$1\left(\begin{array}{c} r'-v_{t-1}\left(\theta',\theta^0\right)+\eta+\overline{v}_{t-1}\left(\theta',\eta\right)+\overline{\Lambda}'^{,m}_{t-1}-\Lambda'^{,m}_{t-1}<\varepsilon_{t-1}<r'-v_{t-1}\left(\theta',\theta^0\right)+\eta+\overline{v}_{t-1}\left(\theta',\eta\right) \\ \Lambda'^{,m}_{t-1}-\overline{\Lambda}'^{,m}_{t-1}<0 \end{array}\right)+$$
$$1\left(\begin{array}{c} r'-v_{t-1}\left(\theta',\theta^0\right)+\eta-\overline{v}_{t-1}\left(\theta',\eta\right)<\varepsilon_{t-1}<r'-v_{t-1}\left(\theta',\theta^0\right)+\eta-\overline{v}_{t-1}\left(\theta',\eta\right)+\overline{\Lambda}'^{,m}_{t-1}+\Lambda'^{,m}_{t-1} \\ \Lambda'^{,m}_{t-1}+\overline{\Lambda}'^{,m}_{t-1}>0 \end{array}\right)+$$
$$\vdots$$

As before, applying a Taylor expansion in $\left(\Lambda'^{,m}_{t-1},\overline{\Lambda}'^{,m}_{t-1}\right)$ we obtain,

$$E\left|1\left(\overline{A}'_{4,t-1}\right)-1\left(\overline{A}'^{,m}_{4,t-1}\right)+1\left(\overline{A}'_{3,t-1}\right)-1\left(\overline{A}'^{,m}_{3,t-1}\right)\right|\leq 4M^*\left(E\left|\Lambda'^{,m}_{t-1}\right|+E\left|\overline{\Lambda}'^{,m}_{t-1}\right|\right)$$

Now we study $\theta'_{t-1}\overline{v}_{t-1}\left(\theta',\eta\right)-\theta'^{,m}_{t-1}\overline{v}^m_{t-1}\left(\theta',\eta\right),$

$$E\left|\theta'_{t-1}\overline{v}_{t-1}\left(\theta',\eta\right)-\theta'^{,m}_{t-1}\overline{v}^m_{t-1}\left(\theta',\eta\right)\right|\leq E\left|\left(\theta'_{t-1}-\theta'^{,m}_{t-1}\right)\overline{v}_{t-1}\left(\theta',\eta\right)\right|+E\left|\theta'^{,m}_{t-1}\left(\overline{v}_{t-1}\left(\theta',\eta\right)-\overline{v}^m_{t-1}\left(\theta',\eta\right)\right)\right|$$

Using the law of iterated expectation, $\overline{v}_{t-1}\left(\theta',\eta\right)-\overline{v}^m_{t-1}\left(\theta',\eta\right)$ is $F_{t-2}-measurable$ and $E\left(\theta'^{,m}_{t-1}\middle|F_{t-2}\right)\leq\lambda^*_1\left(\theta'\right)$

$$E\left|\theta'^{,m}_{t-1}\left(\overline{v}_{t-1}\left(\theta',\eta\right)-\overline{v}^m_{t-1}\left(\theta',\eta\right)\right)\right|\leq\lambda^*_1\left(\theta'\right)E\left|\overline{\Lambda}'^{,m}_{t-1}\right|$$

For the first part, we need to use the lemma 17.5 of Davidson (1994) p. 271,

$$\left\|\left(\theta'_{t-1}-\theta'^{,m}_{t-1}\right)\overline{v}_{t-1}\left(\theta',\eta\right)\right\|_2\leq 2\left(\left\|\left(\theta'_{t-1}-\theta'^{,m}_{t-1}\right)\right\|_2^{r-2}\left\|\overline{v}_{t-1}\left(\theta',\eta\right)\right\|_2^{r-2}\left\|\left(\theta'_{t-1}-\theta'^{,m}_{t-1}\right)\overline{v}_{t-1}\left(\theta',\eta\right)\right\|_r^r\right)^{\frac{1}{2(r-1)}}$$

To use it, we need that $\left\|\overline{v}_{t-1}\left(\theta',\eta\right)\right\|_r<\infty$ for $r>2$ which is proved before. Denote

$$\left[2\left\|\overline{v}_{t-1}\left(\theta',\eta\right)\right\|_2^{r-2}\left\|\left(\theta'_{t-1}-\theta'^{,m}_{t-1}\right)\overline{v}_{t-1}\left(\theta',\eta\right)\right\|_r^r\right]^{\frac{1}{2(r-1)}}\leq\left[2\left\|\overline{v}_{t-1}\left(\theta',\eta\right)\right\|_2^{2r-2}\right]^{\frac{1}{2(r-1)}}\leq K_4<\infty$$

and taking $r=4$

$$\left\|\left(\theta'_{t-1}-\theta'^{,m}_{t-1}\right)\overline{v}_{t-1}\left(\theta',\eta\right)\right\|_2\leq\left\|\left(\theta'_{t-1}-\theta'^{,m}_{t-1}\right)\right\|_2K_4$$

Then

$$\begin{aligned} E\left|\overline{\Lambda}'^{,m}_t\right| &\leq 2\eta E\left|e'_{t-1}-e'^{,m}_{t-1}\right|+\lambda^*_1\left(\theta'\right)E\left|\overline{\Lambda}'^{,m}_{t-1}\right|+\left\|\left(\theta'_{t-1}-\theta'^{,m}_{t-1}\right)\right\|_2K_4+\delta r'4M^*\left(E\left|\Lambda'^{,m}_{t-1}\right|+E\left|\overline{\Lambda}'^{,m}_{t-1}\right|\right) \\ &\leq(2\eta+K_4)\left\|e'_{t-1}-e'^{,m}_{t-1}\right\|_2+\delta r'4M^*E\left|\Lambda'^{,m}_{t-1}\right|+\lambda^*_2\left(\theta'\right)E\left|\overline{\Lambda}'^{,m}_{t-1}\right| \end{aligned}$$

If $\lambda^*_2\left(\theta'\right)<1$, following the steps of equation [**??**] we have the result. ■

With these results we can begin the prove of claim 1 for the STIMA case. As we did in the case of observable threshold the space $\left\{\theta\in\Theta:\rho\left(\theta,\theta^0\right)\geq\varepsilon\right\}$ is compact, and we can define $\theta^j=\left(\theta^j_1,\theta^j_2,r^j\right)=\left(\theta^0_1+i_j\mu,\theta^0_2+i'_j\mu,r^0+i''_j\mu\right)$ and $B^j=\left\{\theta\in\Theta:\rho\left(\theta,\theta^j\right)\leq\mu\right\}$

such that $\left\{\theta \in \Theta : \rho\left(\theta, \theta^0\right) \geq \varepsilon\right\} \subset \cup_{j=1}^{J} B^j$ with $J = \frac{4\overline{r}}{\mu^3}$. Then

$$\sup_{\{\theta\in\Theta:\rho(\theta,\theta^0)\geq\varepsilon\}} \frac{1}{T}\sum_{t=1}^{T} 2\varepsilon_t\left(\theta_{t-1}e_{t-1}(\theta) - \theta_{t-1}^0\varepsilon_{t-1}\right) \leq \max_j \frac{1}{T}\sum_{t=1}^{T} 2\varepsilon_t\left(\theta_{t-1}^j e_{t-1}\left(\theta^j\right) - \theta_{t-1}^0\varepsilon_{t-1}\right) + \max_j \sup_{\theta\in B^j} \frac{1}{T}\sum_{t=1}^{T} 2\varepsilon_t\left(\theta_{t-1}e_{t-1}(\theta) - \theta_{t-1}^j e_{t-1}\left(\theta^j\right)\right)$$

For the second part,

$$\sup_{\theta\in B^j} \frac{1}{T}\sum_{t=1}^{T} 2\varepsilon_t\left(\theta_{t-1}e_{t-1}(\theta) - \theta_{t-1}^j e_{t-1}\left(\theta^j\right)\right) \leq \frac{1}{T}\sum_{t=1}^{T} 2\left|\varepsilon_t\right| \overline{v}_{t-1}\left(\theta^j, \mu\right)$$

Again, with the lemma 18 like in lemma 12

$$E\left[\max_{1\leq t\leq T}\left(\sum_{l=1}^{t}\left(\left|\varepsilon_l\right|\overline{v}_{l-1}\left(\theta^j,\mu\right) - E\left(\left|\varepsilon_l\right|\overline{v}_{l-1}\left(\theta^j,\mu\right)\right)\right)\right)^2\right] \leq K\sum_{l=1}^{t} E\left(\left|\varepsilon_l\right|\overline{v}_{l-1}\left(\theta^j,\mu\right)\right) \quad K < \infty$$

Then

$$P\left[\left|\max_j \sup_{\theta\in B^j} \frac{1}{T}\sum_{t=1}^{T} 2\varepsilon_t\left(\theta_{t-1}e_{t-1}(\theta) - \theta_{t-1}^j e_{t-1}\left(\theta^j\right)\right)\right| < \eta\right] \leq \sum_{j=1}^{J} P\left[\frac{1}{T}\sum_{t=1}^{T} 2\left|\varepsilon_t\right|\overline{v}_{t-1}\left(\theta^j,\mu\right) < \eta\right] \leq$$

$$\leq \sum_{j=1}^{J} P\left[\frac{1}{T}\sum_{t=1}^{T}\left|2\left|\varepsilon_t\right|\overline{v}_{t-1}\left(\theta^j,\mu\right) - E\left(2\left|\varepsilon_t\right|\overline{v}_{t-1}\left(\theta^j,\mu\right)\right)\right| < \eta - E\left(2\left|\varepsilon_t\right|\overline{v}_{t-1}\left(\theta^j,\mu\right)\right)\right]$$

Using the previous lemma, $E\left(\left|\overline{v}_{t-1}\left(\theta^j,\mu\right)\right|\right) \leq \mu\frac{K_3}{1-\lambda_2^*}$, then using the iterated law of expectations,

$$P\left[\left|\max_j \sup_{\theta\in B^j} \frac{1}{T}\sum_{t=1}^{T} 2\varepsilon_t\left(\theta_{t-1}e_{t-1}(\theta) - \theta_{t-1}^j e_{t-1}\left(\theta^j\right)\right)\right| < \eta\right] \geq \sum_{j=1}^{J} \frac{K\mu\frac{K_3}{1-\lambda_2^*}}{T\left(\eta - \mu\frac{K_3}{1-\lambda_2^*}\right)^2} \geq \frac{4\overline{r}K\frac{K_3}{1-\lambda_2^*}}{T\mu^2\left(\eta - \mu\frac{K_3}{1-\lambda_2^*}\right)^2}$$

Now taking $\mu \leq \frac{\eta(1-\lambda_2^*)}{2K_3}$,

$$P\left[\left|\max_j \sup_{\theta\in B^j} \frac{1}{T}\sum_{t=1}^{T} 2\varepsilon_t\left(\theta_{t-1}e_{t-1}(\theta) - \theta_{t-1}^j e_{t-1}\left(\theta^j\right)\right)\right| < \eta\right] \geq \frac{4\overline{r}K\frac{K_3}{1-\lambda_2^*}}{T\eta^3\left(\frac{K_3}{1-\lambda_2^*}\right)^3}$$

For the first part, we use that is a martingale difference,

$$P\left[\left|\max_j \frac{1}{T}\sum_{t=1}^{T} 2\varepsilon_t\left(\theta_{t-1}^j e_{t-1}\left(\theta^j\right) - \theta_{t-1}^0\varepsilon_{t-1}\right)\right| > \eta\right] \leq$$

$$\leq \sum_{j=1}^{J} P\left[\left|\frac{1}{T}\sum_{t=1}^{T} 2\varepsilon_t\left(\theta_{t-1}^j e_{t-1}\left(\theta^j\right) - \theta_{t-1}^0\varepsilon_{t-1}\right)\right| > \eta\right] \leq$$

$$\leq \sum_{l=1}^{\frac{4\overline{r}}{\mu}} \frac{l^2\mu\sigma_\varepsilon^2 K\frac{K_3}{1-\lambda_2^*}}{T\eta^2} \leq \frac{\sigma_\varepsilon^2 K\frac{K_3}{1-\lambda_2^*}}{T\eta^2\mu^4}$$

Taking again $\mu \leq \frac{\eta(1-\lambda_2^*)}{2K_3}$

$$P\left[\left|\max_j \frac{1}{T}\sum_{t=1}^{T} 2\varepsilon_t\left(\theta_{t-1}^j e_{t-1}\left(\theta^j\right) - \theta_{t-1}^0\varepsilon_{t-1}\right)\right| > \eta\right] \leq \frac{\sigma_\varepsilon^2 K\left(\frac{K_3}{1-\lambda_2^*}\right)^5}{T\eta^6}$$

With that, the claim 1 for the STIMA case is proved since

$$\lim_{T\to\infty} P\left[\sup_{\{\theta\in\Theta:\rho(\theta,\theta^0)\geq\varepsilon\}} \frac{1}{T}\sum_{t=1}^{T} 2\varepsilon_t\left(\theta_{t-1}e_{t-1}(\theta)-\theta^0_{t-1}\varepsilon_{t-1}\right)>\eta\right]\to 0$$

To prove the claim 2 for STIMA case we follow the same steps than in the case of observable variable,

$$\inf_{\{\theta\in\Theta:\rho(\theta,\theta^0)\geq\varepsilon\}} \frac{1}{T}\sum_{t=1}^{T}\left(\theta_{t-1}e_{t-1}(\omega,\theta)-\theta^0_{t-1}\varepsilon_{t-1}\right)^2 = \inf_{\{\theta\in\Theta:\rho(\theta,\theta^0)\geq\varepsilon\}} \frac{1}{T}\sum_{t=1}^{T} v^2_{t-1}\left(\theta,\theta^0\right)$$

$$= \inf_{\{\theta\in\Theta:\rho(\theta,\theta^0)\geq\varepsilon\}} \frac{1}{T}\sum_{l=1}^{\left[\frac{T}{2}\right]} v^2_{k(l)-1}+v^2_{k(l)} = \inf_{\{\theta\in\Theta:\rho(\theta,\theta^0)\geq\varepsilon\}} \frac{1}{T}\sum_{l=1}^{\left[\frac{T}{2}\right]} W_{k(l)} \qquad \text{with } k(l)=2(l-1)+1$$

We do the prove for $r>r^0$, the other case is equal. We distinguish three cases, first one, when $\left|\theta_2-\theta^0_2\right|>\epsilon$,

$$W_{k(l)}\geq I\left(\left|v_{k(l)-1}\right|>\gamma\right)\gamma^2+I\left(\left|v_{k(l)-1}\right|<\gamma\right)I\left(\left|\varepsilon_{k(l)}\right|<r^0-\gamma\right)\left(\left|\epsilon\varepsilon_{k(l)}\right|-\gamma\right)^2$$

taking

$$\gamma=\left|\frac{\epsilon\varepsilon_{k(l)}}{2}\right|$$

$$W_{k(l)}\geq I\left(\left|\varepsilon_{k(l)}\right|<\frac{2r^0}{2+\epsilon}\right)\left(\frac{\epsilon\varepsilon_{k(l)}}{2}\right)^2$$

and using that $E\left[I\left(\left|\varepsilon_{k(l)}\right|<\frac{2r^0}{2+\epsilon}\right)\varepsilon^2_{k(l)}\right]=2a_1>0$ and it is an independent variable,

$$\frac{1}{T}\sum_{l=1}^{\left[\frac{T}{2}\right]} I\left(\left|\varepsilon_{k(l)}\right|<r^0-\epsilon\right)\left(\frac{\epsilon\varepsilon_{k(l)}}{2}\right)^2 \xrightarrow{a.s} \epsilon^2 a_1>0$$

The second case, $r-r^0>\epsilon$. We know that $\left|\theta_2-\theta^0_2\right|<\epsilon$, otherwise we are in the first case. With this, and $\left|\theta^0_1-\theta^0_2\right|=\delta^0$ we have

$$W_{k(l)}\geq I\left(\left|v_{k(l)-1}\right|>\gamma\right)\gamma^2+I\left(\left|v_{k(l)-1}\right|<\gamma\right)I\left(\left|\varepsilon_{k(l)}\right|<r^0+\epsilon-\gamma,\left|\varepsilon_{k(l)}\right|>r^0\right)\left(\left|\delta^0-\epsilon\right|\left|\varepsilon_{k(l)}\right|-\gamma\right)^2$$

We need to take γ s.t $|\gamma|<\epsilon$ and $\left(\left|\delta^0-\epsilon\right|\left|\varepsilon_{k(l)}\right|-\gamma\right)^2>\gamma^2$, taking $\gamma=\left|\frac{\epsilon\varepsilon_{k(l)}}{r^0+\epsilon}\right|$, both conditions are satisfied, and we obtain

$$W_{k(l)}\geq I\left(r^0<\left|\varepsilon_{k(l)}\right|<\frac{\left(r^0+\epsilon\right)^2}{r^0+2\epsilon}\right)\left(\frac{\left(\delta^0-\frac{\epsilon r^0+\epsilon^2-\epsilon}{r^0+\epsilon}\right)\varepsilon_{k(l)}}{2}\right)^2$$

and using $E\left[I\left(r^0<\left|\varepsilon_{k(l)}\right|<\frac{\left(r^0+\epsilon\right)^2}{r^0+2\epsilon}\right)\varepsilon^2_{k(l)}\right]=\epsilon^2 a'_2>0$ for $\epsilon>0$, we obtain

$$\frac{1}{T}\sum_{l=1}^{\left[\frac{T}{2}\right]} I\left(r^0<\left|\varepsilon_{k(l)}\right|<\frac{\left(r^0+\epsilon\right)^2}{r^0+2\epsilon}\right)\left(\frac{\left(\delta^0-\frac{\epsilon r^0+\epsilon^2-\epsilon}{r^0+\epsilon}\right)\varepsilon_{k(l)}}{2}\right)^2 \xrightarrow{a.s} \epsilon^2 a_2>0$$

Finally, the third case, $\left|\theta_1-\theta^0_1\right|>\epsilon$ and $r-r^0<\epsilon$, otherwise we are in the second case. This case is similar to the first case, we have,

$$W_{k(l)}\geq I\left(\left|v_{k(l)-1}\right|>\gamma\right)\gamma^2+I\left(\left|v_{k(l)-1}\right|<\gamma\right)I\left(\left|\varepsilon_{k(l)}\right|>r^0+\epsilon+\gamma\right)\left(\left|\epsilon\varepsilon_{k(l)}\right|-\gamma\right)^2$$

Taking again $\gamma = \left|\frac{\epsilon\varepsilon_{k(l)}}{2}\right|$ we obtain

$$W_{k(l)} \geq I\left(\left|\varepsilon_{k(l)}\right| > \frac{2\left(r^0+\epsilon\right)}{2-\epsilon}\right)\left(\frac{\epsilon\varepsilon_{k(l)}}{2}\right)^2$$

Using that $E\left[I\left(\left|\varepsilon_{k(l)}\right| > \frac{2\left(r^0+\epsilon\right)}{2-\epsilon}\right)\varepsilon_{k(l)}^2\right] = 2a_3 > 0$

$$\frac{1}{T}\sum_{l=1}^{\left[\frac{T}{2}\right]} I\left(\left|\varepsilon_{k(l)}\right| > \frac{2(r^0+\epsilon)}{2-\epsilon}\right)\left(\frac{\epsilon\varepsilon_{k(l)}}{2}\right)^2 \xrightarrow{a.s} \epsilon^2 a_3 > 0$$

Taking $a(\epsilon) = \epsilon^2 \min\{a_1, a_2, a_3\}$, for all $\eta > 0$,

$$\lim_{T\to\infty} P\left[\inf_{\theta|\rho(\theta,\theta^0)>\epsilon}\sum_{t=1}^{T}\left(\theta_{t-1}e_{t-1}(\omega,\theta) - \theta_{t-1}^0\varepsilon_{t-1}\right)^2 > a(\epsilon) - \eta\right] = 1$$

and we have proved the consistence of the GLS estimator for STIMA case.

Proof of theorem 2. The proof of this theorem is similar to the proof of consistence. As in the case of theorem 2, we define $\Theta_\triangle = \left\{\theta \in \Theta : \left|\theta_i - \theta_i^0\right| = \delta_i \geq \frac{\triangle}{T} \text{ and } \left|r - r^0\right| = \delta_3 \geq \frac{\triangle}{T}\right\}$. To prove that the rate of convergence of the estimators is T, a sufficient condition is for all $\varepsilon > 0$ and T exist $\triangle < \infty$ such that:

$$P\left[\inf_{\{\theta\in\Theta_\triangle\}}\frac{Q_T(\omega,\theta) - Q_T\left(\omega,\theta^0\right)}{T\sup\delta_i} > 0\right] \geq 1-\varepsilon \Longrightarrow P\left[\inf_{\{\theta\in\Theta_\triangle\}} Q_T(\omega,\theta) - Q_T\left(\omega,\theta^0\right) > 0\right] \geq 1-\varepsilon$$

We will use the partition $\theta_j = \left(\theta_1^0 \pm \frac{\triangle}{T}\left(b^{j_1}-1\right), \theta_2^0 \pm \frac{\triangle}{T}\left(b^{j_2}-1\right), r_1^0 \pm \frac{\triangle}{T}\left(b^{j_1}-1\right)\right)$ and $B_{\triangle,j} = \left\{\theta \in \left(\theta^j, \theta^{j+1}\right)\right\}$. We work first with

$$P\left[\sup_{\{\theta\in\Theta_\triangle\}}\frac{\sum_{t=1}^T 2\varepsilon_t\left(\theta_{t-1}e_{t-1}(\omega,\theta) - \theta_{t-1}^0\varepsilon_{t-1}\right)}{T\sup_i\delta_i} > \eta\right] \leq \varepsilon$$

We limit this supremum with

$$\sup_{\{\theta\in\Theta_\triangle\}}\frac{\sum_{t=1}^T 2\varepsilon_t\left(\theta_{t-1}e_{t-1}(\omega,\theta) - \theta_{t-1}^0\varepsilon_{t-1}\right)}{T\sup_i\delta_i} \leq \max_j\frac{\sum_{t=1}^T 2\varepsilon_t\left(\theta_{t-1}^j e_{t-1}\left(\omega,\theta^j\right) - \theta_{t-1}^0\varepsilon_{t-1}\right)}{T\sup_i\delta_i} + \max_j\sup_{B_{\triangle,j}}\frac{\sum_{t=1}^T 2\varepsilon_t\left(\theta_{t-1}e_{t-1}(\omega,\theta) - \theta_{t-1}^j e_{t-1}\left(\omega,\theta^j\right)\right)}{T\sup_i\delta_i}$$

For the first term we use that $\varepsilon_t\left(\theta_{t-1}^j e_{t-1}\left(\omega,\theta^j\right) - \theta_{t-1}^0\varepsilon_{t-1}\right)$ is a martingale difference sequence, the lemma 18 about $E\overline{v}_t^2(\theta',\delta_i)$ and $\sup_i\delta_i \geq \sup_{j_i}\frac{\triangle}{T}\left(b^{j_i}-1\right)$

$$P\left[\max_j\frac{\sum_{t=1}^T 2\varepsilon_t\left(\theta_{t-1}^j e_{t-1}\left(\omega,\theta^j\right) - \theta_{t-1}^0\varepsilon_{t-1}\right)}{T\sup_i\delta_i} > \eta\right] \leq \frac{1}{\triangle}\sum_{j=1}^{\infty}\frac{j^2 4\sigma_\varepsilon^2 K_2}{\left(b^j-1\right)\eta^2}$$

with $\sum_{j=1}^{\infty}\frac{j^2 4\sigma_\varepsilon^2 K_2}{\left(b^j-1\right)\eta^2} < \infty$. For the second term

$$\sup_{B_{\triangle,j}}\frac{\sum_{t=1}^T 2\varepsilon_t\left(\theta_{t-1}e_{t-1}(\omega,\theta) - \theta_{t-1}^j e_{t-1}\left(\omega,\theta^j\right)\right)}{T\sup_i\delta_i} \leq \frac{\sum_{t=1}^T 2\left|\varepsilon_t\right|\left(\overline{v}_t\left(\theta,\theta^j\right)\right)}{T\sup_i\delta_i}$$

Now using the lemma 18 and $\sup_{B_{\triangle,j}} \sup_i \delta_i \left(\theta, \theta^j\right) \leq \sup_{j_i} \frac{\triangle}{T} b^{j_i} (b-1)$

$$P\left[\max_j \frac{\sum_{t=1}^{T} 2\left|\varepsilon_t\right|\left(\overline{v}_t\left(\theta,\theta^j\right)\right) - E\left(2\left|\varepsilon_t\right|\left(\overline{v}_t\left(\theta,\theta^j\right)\right)\right)}{T\sup_i \delta_i} > \eta\right] \leq \frac{1}{\triangle}\sum_{j=1}^{\infty}\frac{j^2\left(b-1\right)4\sigma_\varepsilon^2 K_2}{\left(b^j-1\right)\eta^2}$$

$$\max_j \frac{\sum_{t=1}^{T} E\left(2\left|\varepsilon_t\right|\left(\overline{v}_t\left(\theta,\theta^j\right)\right)\right)}{T\sup_i \delta_i} \leq \left(b-1\right)\frac{4\sigma_\varepsilon^2 K_2}{\eta^2}$$

taking $(b-1)$ small enough we have

$$P\left[\sup_{\{\theta\in\Theta_\triangle\}} \frac{\sum_{t=1}^{T} 2\varepsilon_t\left(\theta_{t-1}e_{t-1}\left(\omega,\theta\right) - \theta_{t-1}^0\varepsilon_{t-1}\right)}{T\sup_i \delta_i} > \eta\right] \leq \frac{1}{\triangle}\sum_{j=1}^{\infty}\frac{j^2\left(b-1\right)8\sigma_\varepsilon^2 K_2}{\left(b^j-1\right)\eta^2}$$

Now we work with

$$P\left[\sup_{\{\theta\in\Theta_\triangle\}} \frac{\sum_{t=1}^{T}\left(\theta_{t-1}e_{t-1}\left(\omega,\theta\right) - \theta_{t-1}^0\varepsilon_{t-1}\right)^2}{T\sup_i \delta_i} > \eta\right] \geq 1-\varepsilon$$

As in consistence proof, we do it for the case of $r > r^0, \theta_1 \neq 0$ and $\theta_2 \neq 0$. We distinguish three cases. In all of them we use the result of lemma 18 for

$$P\left(\sup_{\{\theta\in\Theta_\triangle\}}\left|\theta_{t-1}e_{t-1}\left(\omega,\theta\right) - \theta_{t-1}^0\varepsilon_{t-1}\right| < \max_i \delta_i K^*\right) \geq 1 - \frac{K_2}{K^*}$$

As $K_2 < \infty$, $\exists K^* < \infty$ s.t this probability is strictly positive. First one, when $\delta_2 \geq \delta_3, \delta_1$. By consistence, we can take T big enough for $\max_i \delta_i 2K^* < \epsilon < r^0/2$. Define $n \geq \frac{K^*}{r^0 - k^*}$, $k^* > \epsilon$ and

$$W_{k(l)} = \sum_{j=0}^{n} v_{k(l)-j}^2 \qquad \text{con } k\left(l\right) = n\left(l-1\right)$$

Then, the following event $1\left(\left|v_{k(l)-n}\right| < \epsilon\right)\prod_{j=2}^{n-1} 1\left(r^0 - k^* < \varepsilon_{k(l)-j} < r^0 - \epsilon\right)$ implies that $v_{k(l)-1} > \left|\delta_2 k^*\right|$ or $v_{k(l)-1} < -\left|\delta_2 k^*\right|$ depending on the sign of δ_2 and θ_2. First case, $\delta_2 > 0$ and $\theta_2 > 0$

$$1\left(v_{k(l)-n} > 0\right)\prod_{j=1}^{n-1} 1\left(r^0 - k^* < \varepsilon_{k(l)-j} < r^0 - \epsilon\right) \Longrightarrow \epsilon > v_{k(l)-1} > \left|\delta_2 k^*\right|$$

$$1\left(v_{k(l)-n} < 0\right)\prod_{j=1}^{n-1} 1\left(r^0 - k^* < \varepsilon_{k(l)-j} < r^0 - \epsilon\right) \Longrightarrow v_{k(l)-1} > \left|\delta_2 k^*\right|$$

If $v_{k(l)-1} > \left|\delta_2 k^*\right| \Longrightarrow v_{k(l)}^2 \geq 1\left(-r^0 - \left|\delta_2 k^*\right| < \varepsilon_{k(l)} < -r^0\right)\left(\left(\delta-\epsilon\right)\left(r^0-\epsilon\right)-\epsilon\right)^2$
When $\delta_2 < 0$ and $\theta_2 > 0$

$$1\left(v_{k(l)-n} > 0\right)\prod_{j=1}^{n-1} 1\left(r^0 - k^* < \varepsilon_{k(l)-j} < r^0 - \epsilon\right) \Longrightarrow v_{k(l)-1} < -\left|\delta_2 k^*\right|$$

$$1\left(v_{k(l)-n} < 0\right)\prod_{j=1}^{n-1} 1\left(r^0 - k^* < \varepsilon_{k(l)-j} < r^0 - \epsilon\right) \Longrightarrow v_{k(l)-1} < -\left|\delta_2 k^*\right|$$

If $v_{k(l)-1} < -\left|\delta_2 k^*\right| \Longrightarrow v_{k(l)}^2 \geq 1\left(r^0 < \varepsilon_{k(l)} < r^0 + \left|\delta_2 k^*\right|\right)\left(\left(\delta-\epsilon\right)\left(r^0-\epsilon\right)-\epsilon\right)^2$

If $\theta_2 < 0$, $\delta_2 > 0$ and n even, define $k(j) = 2(j-1)$

$$1\left(v_{k(l)-n} > 0\right) 1\left(r^0 - k^* < \varepsilon_{k(l)-k(j)-1} < r^0 - \epsilon\right) \prod_{j=1}^{\frac{n-2}{2}} 1\left(\begin{array}{c} -r^0 + \epsilon < \varepsilon_{k(l)-k(j)} < -r^0 + k^* \\ r^0 - k^* < \varepsilon_{k(l)-k(j)-1} < r^0 - \epsilon \end{array}\right)$$

$$\Longrightarrow v_{k(l)-1} > |\delta_2 k^*|$$

$$1\left(v_{k(l)-n} < 0\right) 1\left(r^0 - k^* < \varepsilon_{k(l)-k(j)-1} < r^0 - \epsilon\right) \prod_{j=1}^{\frac{n-2}{2}} 1\left(\begin{array}{c} -r^0 + \epsilon < \varepsilon_{k(l)-k(j)} < -r^0 + k^* \\ r^0 - k^* < \varepsilon_{k(l)-k(j)-1} < r^0 - \epsilon \end{array}\right)$$

$$\Longrightarrow \epsilon > v_{k(l)-1} > |\delta_2 k^*|$$

If $v_{k(l)-1} > |\delta_2 k^*|\, v_{k(l)}^2 \geq 1\left(-r^0 - |\delta_2 k^*| < \varepsilon_{k(l)} < -r^0\right)\left((\delta - \epsilon)\left(r^0 - \epsilon\right) - \epsilon\right)^2$.

Similar results can be obtained for the rest of the case. Then if we take the first case

$$W_{k(l)} \geq \epsilon^2 1\left(\left|v_{k(l)-n}\right| > \epsilon\right) + 1\left(\left|v_{k(l)-n}\right| < \epsilon\right) \prod_{j=2}^{n-1} 1\left(\begin{array}{c} r^0 - k^* < \varepsilon_{k(l)-j} < r^0 - \epsilon \\ -r^0 - |\delta_2 k^*| < \varepsilon_{k(l)} < -r^0 \end{array}\right)\left((\delta - \epsilon)\left(r^0 - \epsilon\right) - \epsilon\right)^2$$

$$W_{k(l)} \geq \prod_{j=2}^{n-1} 1\left(r^0 - k^* < \varepsilon_{k(l)-j} < r^0 - \epsilon\right) 1\left(-r^0 - |\delta_2 k^*| < \varepsilon_{k(l)} < -r^0\right) \min\left\{\left((\delta - \epsilon)\left(r^0 - \epsilon\right) - \epsilon\right)^2, \epsilon^2\right\}$$

$W_{k(l)}$ is a martingale difference sequence, then

$$\sup_{\{\theta \in \Theta_\triangle\}} \sum_{t=1}^{T}\left(\theta_{t-1} e_{t-1}(\omega, \theta) - \theta_{t-1}^0 \varepsilon_{t-1}\right)^2 \geq \sum_{l=2}^{\left[\frac{T}{n}\right]} W_{k(l)}$$

$$\sum_{l=2}^{\left[\frac{T}{n}\right]} W_{k(l)} \xrightarrow{p} \left[\frac{T}{n}\right] E\left(W_{k(l)}\right) \geq \triangle k^* \frac{1}{n} p^n m \min\left\{\left((\delta - \epsilon)\left(r^0 - \epsilon\right) - \epsilon\right)^2, \epsilon^2\right\} > 0$$

with $p = E\left(1\left(r^0 - k^* < \varepsilon_{k(l)-j} < r^0 - \epsilon\right)\right) \geq m\,|k^* - \epsilon| > 0$.

If $\delta_1 \geq \delta_2, \delta_3$, the proof is easier. We can choose the sign of $v_{k(l)-1}$ asking $\left|\varepsilon_{k(l)-1}\right| > \theta_1 v_{k(l)-2} < \epsilon$. The probability of that even is strictly positive. Finally, if $\delta_3 \geq \delta_2, \delta_1$ then $\delta_3 \geq \frac{\triangle}{T}$ and using the previous results nothing changes. ■

Before to prove the theorem 3 we establish two lemmas which help us to do the proof.

Lemma 19. *Under the assumptions A.3 to A.8*

$$\sup_{\theta^* \ s.t\ \rho(\theta^*, \theta^0) \leq \rho\left(\widehat{\theta}, \theta^0\right)} \left\|H_T(\theta^*) - H_T\left(\theta^0\right)\right\| \xrightarrow{p} 0$$

Lemma 20. *Under the assumption A.3 to A.8 and* Ω *positive matrix,* $D_T\left(\theta^0\right) \xrightarrow{d} N\left(\binom{0}{0}, \Omega\right)$.

Proof of lemma 19. The matrix converges if each elements of it converges.

$$\frac{1}{T} \sum_{t=1}^{T} H_{t-1}^{i,j}(\theta) = 2 h_{t-1}^l(\theta) h_{t-1}^i(\theta) + 2\left(\theta_{t-1} e_{t-1}(\omega, \theta) - \theta_{t-1}^0 \varepsilon_{t-1}\right) h_{t-1}^{i,l}(\theta)$$

$$\frac{1}{T} \sum_{t=1}^{T} H_{t-1}^{i,j}(\theta) - H_{t-1}^{i,j}\left(\theta^0\right) = \frac{1}{T} \sum_{t=1}^{T} 2 s_{t-1}^i\left(\theta, \theta^0\right) h_{t-1}^l\left(\theta^0\right) + \frac{1}{T} \sum_{t=1}^{T} 2 s_{t-1}^l\left(\theta, \theta^0\right) h_{t-1}^i(\theta)$$
$$+ \frac{1}{T} \sum_{t=1}^{T} 2 v_t\left(\theta, \theta^0\right) h_{t-1}^{i,l}(\theta)$$

Now using the lemmas 7 to 11 and the theorem 17.16 of Davidson we have that - $s_{t-1}^i\left(\theta, \theta^0\right) h_{t-1}^l\left(\theta^0\right)$,

$\overline{s}_{t-1}^{i}\left(\rho\left(\widehat{\theta},\theta^{0}\right)\right)\overline{h}_{t-1}^{l}\left(\theta^{0}\right), v_{t}\left(\theta,\theta^{0}\right)h_{t-1}^{i,l}\left(\theta\right)$ and $\overline{v}_{t}\left(\rho\left(\widehat{\theta},\theta^{0}\right)\right)\overline{h}_{t-1}^{i,l}$ are $L_{2}-NED$.

Then $\sup_{\theta^{*}\text{ s.t }\rho(\theta^{*},\theta^{0})\leq\rho\left(\widehat{\theta},\theta^{0}\right)}\frac{1}{T}\sum_{t=1}^{T}H_{t-1}^{i,j}\left(\theta\right)-H_{t-1}^{i,j}\left(\theta^{0}\right)\xrightarrow{p}0$ if for all $\eta>0$

$$\lim_{T\to\infty}P\left[\sup_{\theta^{*}\text{ s.t }\rho(\theta^{*},\theta^{0})\leq\rho\left(\widehat{\theta},\theta^{0}\right)}\frac{1}{T}\sum_{t=1}^{T}H_{t-1}^{i,j}\left(\theta\right)-H_{t-1}^{i,j}\left(\theta^{0}\right)>\eta\right]=0$$

Fist we proof that

$$\lim_{T\to\infty}P\left[\sup_{\theta^{*}\text{ s.t }\rho(\theta^{*},\theta^{0})\leq\rho\left(\widehat{\theta},\theta^{0}\right)}\frac{1}{T}\sum_{t=1}^{T}2s_{t-1}^{i}\left(\theta,\theta^{0}\right)h_{t-1}^{l}\left(\theta^{0}\right)>\eta\right]=0$$

$$\sup_{\theta^{*}\text{ s.t }\rho\ \theta^{*},\theta^{0}\ \leq\rho\ \widehat{\theta},\theta^{0}}\frac{1}{T}\sum_{t=1}^{T}2s_{t-1}^{i}\ \ \theta,\theta^{0}\ \ h_{t-1}^{l}\ \ \theta^{0}\ \leq\frac{1}{T}\sum_{t=1}^{T}2\overline{s}_{t-1}^{i}\ \ \rho\ \ \widehat{\theta},\theta^{0}\ \ \ \overline{h}_{t-1}^{l}\ \ \theta^{0}$$

Now, using the lemma 17.15 of Davidson it can be proved that

$$\left\|\overline{s}_{t-1}^{i}\left(\rho\left(\widehat{\theta},\theta^{0}\right)\right)\overline{h}_{t-1}^{l}\left(\theta^{0}\right)\right\|_{2}\leq\rho\left(\widehat{\theta},\theta^{0}\right)K\quad with\ K<\infty$$

Then, applying the lemma 12 to this processes,

$$P\left[\sup_{\theta^{*}\text{ s.t }\rho(\theta^{*},\theta^{0})\leq\rho\left(\widehat{\theta},\theta^{0}\right)}\frac{1}{T}\sum_{t=1}^{T}2\overline{s}_{t-1}^{i}\left(\rho\left(\widehat{\theta},\theta^{0}\right)\right)\overline{h}_{t-1}^{l}\left(\theta^{0}\right)-E\left(2\overline{s}_{t-1}^{i}\left(\rho\left(\widehat{\theta},\theta^{0}\right)\right)\overline{h}_{t-1}^{l}\left(\theta^{0}\right)\right)>\eta\right]\leq$$

$$\leq\frac{\rho\left(\widehat{\theta},\theta^{0}\right)K}{T}$$

$$\frac{E\left(2\overline{s}_{t-1}^{i}\left(\rho\left(\widehat{\theta},\theta^{0}\right)\right)\overline{h}_{t-1}^{l}\left(\theta^{0}\right)\right)}{T}\leq\rho^{1/2}\left(\widehat{\theta},\theta^{0}\right)K^{1/2}$$

$$\lim_{T\to\infty}P\left[\sup_{\theta^{*}\text{ s.t }\rho(\theta^{*},\theta^{0})\leq\rho\left(\widehat{\theta},\theta^{0}\right)}\frac{1}{T}\sum_{t=1}^{T}2s_{t-1}^{i}\left(\theta,\theta^{0}\right)h_{t-1}^{l}\left(\theta^{0}\right)>\eta\right]=0$$

The other term $\frac{1}{T}\sum_{t=1}^{T}2v_{t}\left(\theta,\theta^{0}\right)h_{t-1}^{i,l}\left(\theta\right)$ is equal, it can be proved

$$\left\|2\overline{v}_{t}\left(\rho\left(\widehat{\theta},\theta^{0}\right)\right)\overline{h}_{t-1}^{i,l}\right\|_{2}\leq\rho\left(\widehat{\theta},\theta^{0}\right)K\quad with\ K<\infty$$

Then

$$\lim_{T\to\infty}P\left[\sup_{\theta^{*}\text{ s.t }\rho(\theta^{*},\theta^{0})\leq\rho\left(\widehat{\theta},\theta^{0}\right)}\frac{1}{T}\sum_{t=1}^{T}2v_{t}\left(\theta,\theta^{0}\right)h_{t-1}^{i,l}\left(\theta\right)>\eta\right]=0$$

The rest of the proof is equal. ■

Proof of lemma 20. Observing that $\varepsilon_{t}h_{t}^{i}\left(\theta^{0}\right)$ is a martingale difference sequence, we apply a Central Limit Theorem for martingales. We have

$$h_{t}^{i}\left(\theta^{0}\right)=1\left(\theta_{t-1}^{0}=\theta_{i}^{0}\right)\varepsilon_{t-1}+\theta_{t-1}^{0}h_{t-1}^{i}\left(\theta^{0}\right)$$
$$E\left(\varepsilon_{t}^{2}h_{t}^{i2}\left(\theta^{0}\right)\right)=E\left(\varepsilon_{t}^{2}\right)E\left(h_{t}^{i2}\left(\theta^{0}\right)\right)=\sigma_{\varepsilon}^{2}\sigma_{hi}^{2}$$
$$E\left(\varepsilon_{t}h_{t}^{i}\left(\theta^{0}\right)\varepsilon_{t}h_{t}^{j}\left(\theta^{0}\right)\right)=\sigma_{\varepsilon}^{2}\sigma_{hh}^{2}$$

the last equalities using the law of iterated expectation. Then

$$\varepsilon_{t}^{2}h_{t}^{i2}\left(\theta^{0}\right)-E\left(\varepsilon_{t}^{2}h_{t}^{i2}\left(\theta^{0}\right)\right)$$
$$\varepsilon_{t}h_{t}^{i}\left(\theta^{0}\right)\varepsilon_{t}h_{t}^{j}\left(\theta^{0}\right)-E\left(\varepsilon_{t}h_{t}^{i}\left(\theta^{0}\right)\varepsilon_{t}h_{t}^{j}\left(\theta^{0}\right)\right)$$

are martingales difference sequences by the independence of ε_{t}. Using assumption A.3 define

$$\sigma^4_{\varepsilon hi} = E\left(\varepsilon_t^2 h_t^{i2}\left(\theta^0\right) - E\left(\varepsilon_t^2 h_t^{i2}\left(\theta^0\right)\right)\right)^2 < \infty$$
$$\sigma^4_{\varepsilon hh} = E\left(\varepsilon_t^2 h_t^{i}\left(\theta^0\right) h_t^{j}\left(\theta^0\right) - E\left(\varepsilon_t^2 h_t^{i}\left(\theta^0\right) h_t^{j}\left(\theta^0\right)\right)\right)^2 < \infty$$

We can use a Law of Large Number for martingale difference sequence as theorem 20.10 of Davidson, with it, we have

$$\Omega_T = \frac{1}{T}\sum_{t=1}^{T}\begin{pmatrix} \varepsilon_t^2 h_t^{1^2}\left(\theta^0\right) & \varepsilon_t^2 h_t^1\left(\theta^0\right) h_t^2\left(\theta^0\right) \\ \varepsilon_t^2 h_t^1\left(\theta^0\right) h_t^2\left(\theta^0\right) & \varepsilon_t^2 h_t^{2^2}\left(\theta^0\right)\end{pmatrix} \xrightarrow{as} \Omega = \begin{pmatrix} \sigma_\varepsilon^2\sigma_{h1}^2 & \sigma_\varepsilon^2\sigma_{hh}^2 \\ \sigma_\varepsilon^2\sigma_{hh}^2 & \sigma_\varepsilon^2\sigma_{h2}^2\end{pmatrix}$$

Then we can apply the CLT for martingale difference (for example theorem 24.3 of Davidson). Define $X_{t,T} = \sum_{i=1}^{2}\mu_i \frac{\varepsilon_t h_t^i(\theta^0)}{\sqrt{T}}$, with $\|\mu\| = 1$.

$$\sum_{t=1}^{T} X_{t,T}^2 = \sum_{t=1}^{T}\sum_{i=1}^{2}\mu_i^2\left(\frac{\varepsilon_t h_t^i\left(\theta^0\right)}{\sqrt{T}}\right)^2 + 2\mu_1\mu_2\sum_{t=1}^{T}\frac{\varepsilon_t^2 h_t^1\left(\theta^0\right) h_t^2\left(\theta^0\right)}{T} \xrightarrow{as} \sigma^2_{x,\mu} < \infty$$

Redefining $X^*_{t,T} = X_{t,T}/\sigma^2_{x,\mu}$ we have $\sum_{t=1}^{T} X^{*^2}_{t,T} \xrightarrow{as} 1$. Now we only have to prove that $\max_{1\le t\le T}\left|X^*_{t,T}\right| \xrightarrow{pr} 0$. Then it is straightforward to prove that

$$P\left[\max_{1\le t\le T}\left|X^*_{t,T}\right| > \eta\right] \le \sum_{t=1}^{T}\frac{\max_i E\left(\varepsilon_t^4 h_t^{i4}\left(\theta^0\right)\right)}{\eta^4 T^2} \xrightarrow{T\to\infty} 0$$

Then using that theorem $\sum_{t=1}^{T} X^{*^2}_{t,T} \xrightarrow{d} N\left(0, \sigma^2_{x,\mu}\right)$. Since this result is for all μ with $\|\mu\| = 1$, then $D_T\left(\theta^0\right) \xrightarrow{d} N\left(\binom{0}{0}, \Omega\right)$ ■

Proof of theorem 3. We have

$$Q_T\left(\omega, \widehat{\theta}\right) - Q_T\left(\omega, \theta^0\right) = T\left(\widehat{\theta} - \theta^0\right)' \frac{1}{2} H_T\left(\theta^*\right)\left(\widehat{\theta} - \theta^0\right) + 2T^{1/2}\left(\widehat{\theta} - \theta^0\right)' D_T\left(\theta^0\right) + o_p(1)$$

Now, define $T^{1/2}\left(\widetilde{\theta} - \theta^0\right) = -2H_T^{-1}\left(\theta^{**}\right) D_T\left(\theta^0\right)$, such that

$$\sum_{t=1}^{T}\left(\widetilde{\theta}_{t-1}\widetilde{e}_{t-1}\left(\omega, \widetilde{\theta}\right) - \theta^0_{t-1}\varepsilon_{t-1}\right)^2 = \frac{1}{2}\sum_{i=1}^{2}\sum_{l=1}^{2}\sum_{t=1}^{T}\left(\widetilde{\theta}_i - \theta_i^0\right) H_{t-1}^{i,l}\left(\theta^{**}\right)\left(\widetilde{\theta}_l - \theta_l^0\right)$$

then

$$\begin{aligned}
Q_T\left(\omega, \widehat{\theta}\right) - Q_T\left(\omega, \theta^0\right) &= \frac{1}{2} T\left(\widehat{\theta} - \theta^0\right)' H_T\left(\theta^*\right)\left(\widehat{\theta} - \theta^0\right) - T\left(\widehat{\theta} - \theta^0\right)' H_T\left(\theta^{**}\right)\left(\widetilde{\theta} - \theta^0\right) + o_p(1) \\
Q_T\left(\omega, \widetilde{\theta}\right) - Q_T\left(\omega, \theta^0\right) &= \frac{1}{2} T\left(\widetilde{\theta} - \theta^0\right)' H_T\left(\theta^{**}\right)\left(\widetilde{\theta} - \theta^0\right) - 2T^{1/2}\left(\widetilde{\theta} - \theta^0\right)' D_T\left(\theta^0\right) + o_p(1) \\
&= -\frac{1}{2} T\left(\widetilde{\theta} - \theta^0\right)' H_T\left(\theta^{**}\right)\left(\widetilde{\theta} - \theta^0\right) + o_p(1)
\end{aligned}$$

Clearly $-2H_T^{-1}\left(\theta^{**}\right) D_T\left(\theta^0\right)$ is $O_p(1)$, then $\rho\left(\widetilde{\theta}, \theta^0\right) \xrightarrow{p} 0$. Then $\rho\left(\theta^{**}, \theta^0\right) \xrightarrow{p} 0$ and $\rho\left(\theta^*, \theta^0\right) \xrightarrow{p} 0$, with this and lemma 19 we have $H_T\left(\theta^*\right) \xrightarrow{p} H_T\left(\theta^0\right)$ and $H_T\left(\theta^{**}\right) \xrightarrow{p}$

$H_T\left(\theta^0\right)$ and $H_T\left(\theta^0\right) \xrightarrow{p} H\left(\theta^0\right) > 0$ then

$$Q_T\left(\omega,\widehat{\theta}\right) - Q_T\left(\omega,\theta^0\right) = \frac{1}{2}T\left(\widehat{\theta}-\theta^0\right)' H_T\left(\theta^0\right)\left(\widehat{\theta}-\theta^0\right) - T\left(\widehat{\theta}-\theta^0\right)' H_T\left(\theta^0\right)\left(\widetilde{\theta}-\theta^0\right) + o_p(1)$$

$$Q_T\left(\omega,\widetilde{\theta}\right) - Q_T\left(\omega,\theta^0\right) = -\frac{1}{2}T\left(\widetilde{\theta}-\theta^0\right)' H_T\left(\theta^0\right)\left(\widetilde{\theta}-\theta^0\right) + o_p(1)$$

$$Q_T\left(\omega,\widetilde{\theta}\right) - Q_T\left(\omega,\theta^0\right) - \left[Q_T\left(\omega,\widehat{\theta}\right) - Q_T\left(\omega,\theta^0\right)\right] \geq 0$$

$$0 \leq -\frac{1}{2}T\left(\widetilde{\theta}-\theta^0\right)' H_T\left(\theta^0\right)\left(\widetilde{\theta}-\theta^0\right) - \frac{1}{2}T\left(\widehat{\theta}-\theta^0\right)' H_T\left(\theta^0\right)\left(\widehat{\theta}-\theta^0\right)$$

$$+T\left(\widehat{\theta}-\theta^0\right)' H_T\left(\theta^0\right)\left(\widetilde{\theta}-\theta^0\right) + o_p(1)$$

$$T\left(\widehat{\theta}-\theta^0\right)' H_T\left(\theta^0\right)\left(\widehat{\theta}-\widetilde{\theta}\right) + T\left(\widetilde{\theta}-\widehat{\theta}\right)' H_T\left(\theta^0\right)\left(\widetilde{\theta}-\theta^0\right) \leq o_p(1)$$

$$T\left(\widehat{\theta}-\theta^0\right)' H_T\left(\theta^0\right)\left(\widehat{\theta}-\widetilde{\theta}\right) - T\left(\widetilde{\theta}-\theta^0\right)' H_T\left(\theta^0\right)\left(\widehat{\theta}-\widetilde{\theta}\right) \leq o_p(1)$$

$$T\left(\widehat{\theta}-\widetilde{\theta}\right)' H_T\left(\theta^0\right)\left(\widehat{\theta}-\widetilde{\theta}\right) \leq o_p(1) \Rightarrow T\left(\widehat{\theta}-\widetilde{\theta}\right)' H\left(\theta^0\right)\left(\widehat{\theta}-\widetilde{\theta}\right) \leq o_p(1) \Rightarrow$$

$$T\left\|\widehat{\theta}-\widetilde{\theta}\right\|^2 C \leq o_p(1)$$

Now with lemma 20 we have

$$D_T\left(\theta^0\right) \xrightarrow{d} N\left(\begin{pmatrix}0\\0\end{pmatrix},\Omega\right)$$

$$T^{1/2}\left(\widetilde{\theta}-\theta^0\right) = -2H_T^{-1}\left(\theta^{**}\right)D_T\left(\theta^0\right) \xrightarrow{d} N\left(\begin{pmatrix}0\\0\end{pmatrix},4H^{-1}\left(\theta^0\right)\Omega H^{-1}\left(\theta^0\right)\right)$$

Using the Slutzky theorem

$$\left\|T^{1/2}\left(\widehat{\theta}-\theta^0\right) - \left[H_T^{-1}\left(\theta^{**}\right)D_T\left(\theta^0\right)\right]\right\| = T^{1/2}\left\|\widehat{\theta}-\widetilde{\theta}\right\| \xrightarrow{p} 0$$

$$T^{1/2}\left(\widehat{\theta}-\theta^0\right) \xrightarrow{d} N\left(\begin{pmatrix}0\\0\end{pmatrix},4H^{-1}\left(\theta^0\right)\Omega H^{-1}\left(\theta^0\right)\right)$$

■

Proof of lemma 4. As we saw

$$Q_T(\theta,\widehat{\tau}) - Q_T\left(\theta^0,\widehat{\tau}\right) = \sum_{t=1}^{T}\left(\theta_{t-1}e_{t-1}(\theta,\widehat{\tau}) - \theta^0_{t-1}\varepsilon_{t-1}\right)^2 - \sum_{t=1}^{T}\left(\theta^o_{t-1}e_{t-1}\left(\theta^0,\widehat{\tau}\right) - \theta^0_{t-1}\varepsilon_{t-1}\right)^2 +$$

$$\sum_{t=1}^{T}2\varepsilon_t\left(\theta_{t-1}e_{t-1}(\theta,\widehat{\tau}) - \theta^0_{t-1}\varepsilon_{t-1}\right) - \sum_{t=1}^{T}2\varepsilon_t\left(\theta^o_{t-1}e_{t-1}\left(\theta^0,\widehat{\tau}\right) - \theta^0_{t-1}\varepsilon_{t-1}\right)$$

and defining $\Theta_{\triangle_1^*} = \left\{\theta\in\Theta : \rho\left(\theta,\theta^0\right) > \left(\frac{\triangle_1^*}{T}\right)^{1/2}\right\}$, it can be proved following the before proofs of the appendix, that for all $\varepsilon^* > 0\ \exists\triangle_1^*$ s.t

$$P\left(\sup_{\theta\in\Theta_{\triangle_1^*}}\frac{\sum_{t=1}^{T}2\varepsilon_t\left(\theta_{t-1}e_{t-1}(\theta,\tau) - \theta^0_{t-1}\varepsilon_{t-1}\right)}{T\left[\rho\left(\theta,\theta^0\right)\right]^2} > \eta\right) \leq \varepsilon^*$$

Then we need to prove that with probability greater than $1-\varepsilon^*$

$$\frac{\sum_{t=1}^{T}\left(\theta_{t-1}e_{t-1}(\theta,\widehat{\tau}) - \theta^0_{t-1}\varepsilon_{t-1}\right)^2}{T\left[\rho\left(\theta,\theta^0\right)\right]^2} - \frac{\sum_{t=1}^{T}\left(\theta^o_{t-1}e_{t-1}\left(\theta^0,\widehat{\tau}\right) - \theta^0_{t-1}\varepsilon_{t-1}\right)^2}{T\rho\left(\theta,\theta^0\right)^2} > \eta$$

for some $\triangle^*$. We know that with probability greater than $1-\varepsilon$ the

$$E\left(\sup_{r,|r-r^0|<\frac{\triangle}{T}}\left(\theta^o_{t-1}e_{t-1}\left(\theta^0,\widehat{\tau}\right)-\theta^0_{t-1}\varepsilon_{t-1}\right)^2\right)\leq\frac{\triangle}{T}K,$$

hen

$$P\left[\frac{\sum_{t=1}^T\left(\theta^o_{t-1}e_{t-1}\left(\theta^0,\widehat{\tau}\right)-\theta^0_{t-1}\varepsilon_{t-1}\right)^2}{T\left[\rho\left(\theta,\theta^0\right)\right]^2}>\eta_1\right]\leq\frac{\triangle K}{\triangle^*\eta_1}$$

And finally, it is easy to prove, following the same arguments than in claim 2 in the appendix,

$$P\left[\sup_{\theta\in\Theta_{\triangle_1^*}}\left|\frac{\sum_{t=1}^T\left(\theta^o_{t-1}e_{t-1}\left(\theta^0,\widehat{\tau}\right)-\theta^0_{t-1}\varepsilon_{t-1}\right)^2}{T\left[\sup_\theta\rho\left(\theta,\theta^0\right)\right]^2}-1\right|>\eta_2\right]\leq\frac{K}{T\eta_2}$$

Then taking $\triangle^*$ and T big enough it is proved ■

Proof of theorem 4. We can make a Taylor expansion in $Q_T(\theta,\tau)$ with respect to θ only

$$Q_T(\theta,\tau)=\sum_{t=1}^T\left(\theta_{t-1}e_{t-1}(\theta,\tau)-\theta^0_{t-1}\varepsilon_{t-1}\right)^2+\sum_{t=1}^T2\varepsilon_t\left(\theta_{t-1}e_{t-1}(\theta,\tau)-\theta^0_{t-1}\varepsilon_{t-1}\right)+\sum_{t=1}^T\varepsilon_t^2$$

$$\begin{aligned}Q_T\left(\widehat{\theta},\widehat{\tau}\right)-Q_T\left(\theta^0,\tau^0\right)=&\sum_{t=1}^T\sum_{i=1}^22\left(\theta_{t-1}e_{t-1}\left(\theta^0,\widehat{\tau}\right)-\theta^0_{t-1}\varepsilon_{t-1}\right)h^i_{t-1}\left(\theta^0,\widehat{\tau}\right)\left(\widehat{\theta}_i-\theta_i^0\right)+\\&\frac{1}{2}\sum_{i=1}^2\sum_{l=1}^2\sum_{t=1}^T\left(\widehat{\theta}_i-\theta_i^0\right)H^{i,l}_{t-1}\left(\theta^*\right)\left(\widehat{\theta}_l-\theta_l^0\right)+\\&\sum_{i=1}^2\sum_{t=1}^T2\left(\widehat{\theta}_i-\theta_i^0\right)\varepsilon_th^i_t\left(\theta^0,\widehat{\tau}\right)+\\&\frac{1}{2}\sum_{i=1}^2\sum_{l=1}^2\sum_{t=1}^T2\left(\widehat{\theta}_i-\theta_i^0\right)\varepsilon_th^{il}_t\left(\theta^*\right)\left(\widehat{\theta}_l-\theta_l^0\right)\end{aligned}$$

As we have proved $\frac{1}{2}\sum_{i=1}^2\sum_{l=1}^2\sum_{t=1}^T2\left(\widehat{\theta}_i-\theta_i^0\right)\varepsilon_th^{il}_t\left(\theta^*\right)\left(\widehat{\theta}_l-\theta_l^0\right)=o_p(1)$. Then we need to prove

$$\sum_{t=1}^T\sum_{i=1}^22\left(\theta_{t-1}e_{t-1}\left(\theta^0,\widehat{\tau}\right)-\theta^0_{t-1}\varepsilon_{t-1}\right)h^i_{t-1}\left(\theta^0,\widehat{\tau}\right)\left(\widehat{\theta}_i-\theta_i^0\right)=o_p(1)$$

$$\sum_{i=1}^2\sum_{t=1}^T2\left(\widehat{\theta}_i-\theta_i^0\right)\varepsilon_t\left(h^i_t\left(\theta^0,\widehat{\tau}\right)-h^i_t\left(\theta^0,\tau^0\right)\right)=o_p(1)$$

which follows using that $\left|\widehat{r}-r^0\right|=O_p\left(T^{-1}\right)$. ■

Proof of lemma 5. Following the steps of before lemmas we have

$$Q_T(\theta,r)-Q_T\left(\theta^0\right)=\sum_{t=1}^T\left(\theta_{t-1}e_{t-1}-\theta^0\varepsilon_{t-1}\right)^2+2\sum_{t=1}^T\varepsilon_t\left(\theta_{t-1}e_{t-1}-\theta^0\varepsilon_{t-1}\right)$$

We prove directly that $\widehat{\theta}_i(r)-\theta^0=O_p\left(T^{-1/2}\right)$. Then for all $\varepsilon>0$, $\exists\triangle(\varepsilon)$ and $T(\varepsilon)$ big enough s.t

$$P\left(\inf_{\theta,s.t|\theta-\theta^0|>\left(\frac{\triangle(\varepsilon)}{T(\varepsilon)}\right)^{1/2}}Q_T(\theta,r)-Q_T\left(\theta^0\right)>0\right)>1-\varepsilon$$

As we have proved that for all $\varepsilon > 0$, $\exists \triangle(\varepsilon)$ and $T(\varepsilon)$ big enough

$$P\left(\sup_{\theta, s.t |\theta-\theta^0| > \left(\frac{\triangle(\varepsilon)}{T(\varepsilon)}\right)^{1/2}} \frac{\sum_{t=1}^{T} \varepsilon_t \left(\theta_{t-1} e_{t-1} - \theta^0 \varepsilon_{t-1}\right)}{Tp\left(\theta, \theta^0\right)} > \eta\right) < \varepsilon^*$$

For the first term, there is a little change,

$$\left(\theta_{t-1} e_{t-1} - \theta^0 \varepsilon_{t-1}\right)^2 = v_t^2(\theta) = 1\left(|z_{t-1}| > r\right)\left(\left(\theta_1 - \theta^0\right)\varepsilon_{t-1} + \theta_1 v_{t-1}(\theta)\right)^2 + 1\left(|z_{t-1}| < r\right)\left(\left(\theta_2 - \theta^0\right)\varepsilon_{t-1} + \theta_2 v_{t-1}(\theta)\right)^2$$

when z_{t-1} is ε_{t-1}, then in the last equality we substitute z_t by e_t. If we impouse that $v_{t-1}(\theta) < \mu$, $1\left(|\varepsilon_{t-1}| < r - \mu\right) \Rightarrow 1\left(|e_{t-1}| < r\right)$ and $1\left(|\varepsilon_{t-1}| > r + \mu\right) \Rightarrow 1\left(|e_{t-1}| > r\right)$. With this and the before proofs we obtain that for all $\varepsilon > 0$, $\exists \triangle(\varepsilon)$ and $T(\varepsilon)$ big enough

$$P\left(\sup_{\theta, s.t |\theta-\theta^0| > \left(\frac{\triangle(\varepsilon)}{T(\varepsilon)}\right)^{1/2}} \frac{\sum_{t=1}^{T} \left(\theta_{t-1} e_{t-1} - \theta^0 \varepsilon_{t-1}\right)^2}{Tp\left(\theta, \theta^0\right)} > \eta\right) \geq 1 - \varepsilon^*$$

Now for r given we can apply a Taylor expansion around θ^0,

$$Q_T\left(\widehat{\theta}, r\right) - Q_T\left(\theta^0\right) = \frac{1}{2}\sum_{i=1}^{2}\sum_{l=1}^{2}\sum_{t=1}^{T}\left(\widehat{\theta}_i - \theta_i^0\right) H_{t-1}^{i,l}\left(\theta^*, r\right)\left(\widehat{\theta}_l - \theta_l^0\right) + \sum_{i=1}^{2}\sum_{t=1}^{T} 2\left(\widehat{\theta}_i - \theta_i^0\right)\varepsilon_t h_t^i\left(\theta^0, r\right) + \\ + \frac{1}{2}\sum_{i=1}^{2}\sum_{l=1}^{2}\sum_{t=1}^{T} 2\left(\widehat{\theta}_i - \theta_i^0\right)\varepsilon_t h_t^{il}\left(\theta^*, r\right)\left(\widehat{\theta}_l - \theta_l^0\right)$$

with

$$H_{t-1}^{i,l}\left(\theta^*, r\right) = \left.\frac{\partial^2\left(\theta_{t-1} e_{t-1}(\theta) - \theta^0 \varepsilon_{t-1}\right)^2}{\partial\theta_i \partial\theta_l}\right|_{\theta^*} = 2h_{t-1}^l\left(\theta^*, r\right) h_{t-1}^i\left(\theta^*, r\right) + \\ + 2\left(\theta_{t-1}^* e_{t-1}^*\left(\theta^*, r\right) - \theta^0 \varepsilon_{t-1}\right) h_{t-1}^{i,l}\left(\theta^*, r\right)$$

As in the proof of theorem 2 $\sum_{i=1}^{2}\sum_{l=1}^{2}\sum_{t=1}^{T} 2\left(\widehat{\theta}_i - \theta_i^0\right)\varepsilon_t h_t^{il}\left(\theta^*, r\right)\left(\widehat{\theta}_l - \theta_l^0\right) = o_p(1)$, and

$$T^{-1}\sum_{t=1}^{T} H_{t-1}^{i,l}\left(\theta^*, r\right) \xrightarrow{p} H^{i,l}(r) = 2T^{-1}\sum_{t=1}^{T} E\left(h_{t-1}^l(r)\, h_{t-1}^i(r)\right)$$

$$\begin{pmatrix} h_{t-1}^1(r) \\ h_{t-1}^2(r) \end{pmatrix} = \begin{pmatrix} h_{t-1}^1\left(\theta^0, r\right) \\ h_{t-1}^2\left(\theta^0, r\right) \end{pmatrix} = \sum_{j=1}^{t-1}\begin{pmatrix} {\theta^0}^{j-1} 1\left(|z_{t-j}| > r\right)\varepsilon_{t-j} \\ {\theta^0}^{j-1} 1\left(|z_{t-j}| < r\right)\varepsilon_{t-j} \end{pmatrix}$$

following the same steps,

$$T^{1/2}\begin{pmatrix} \widehat{\theta}_{1,T}(r) - \theta^0 \\ \widehat{\theta}_{2,T}(r) - \theta^0 \end{pmatrix} = 2T^{-1/2} H^{-1}(r)\begin{pmatrix} \sum_{t=1}^{T}\varepsilon_t \sum_{j=1}^{t} {\theta^0}^{j-1} 1\left(|z_{t-j}| > r\right)\varepsilon_{t-j} \\ \sum_{t=1}^{T}\varepsilon_t \sum_{j=1}^{t} {\theta^0}^{j-1} 1\left(|z_{t-j}| < r\right)\varepsilon_{t-j} \end{pmatrix} + o_p(1) \\ \xrightarrow{d} N\left(0, H^{-1}(r)\,\Omega^* H^{-1}(r)\right)$$

■

Proof of lemma 6. To prove that the following functional in r, $T^{-1/2}\sum_{t=1}^{T}\varepsilon_t \sum_{j=1}^{t} {\theta^0}^{j-1} 1\left(|z_{t-j}| > r\right)\varepsilon_{t-j}$, is tight we divide it in

$$\varepsilon_t \sum_{j=1}^{t} \theta^{0^{j-1}} 1\left(|z_{t-j}| > r\right) \varepsilon_{t-j} = \varepsilon_t \sum_{j=1}^{T^\alpha} \theta^{0^{j-1}} 1\left(|z_{t-j}| > r\right) \varepsilon_{t-j} + \varepsilon_t \sum_{j=T^\alpha+1}^{t} \theta^{0^{j-1}} 1\left(|z_{t-j}| > r\right) \varepsilon_{t-j}$$

$$= \varepsilon_t b_{t,T^\alpha}(r) + \theta^{0^{T^\alpha}} \varepsilon_t b_{t,\infty}(r)$$

$$T^{-1/2} \sum_{t=1}^{T} \varepsilon_t \sum_{j=1}^{t} \theta^{0^{j-1}} 1\left(|z_{t-j}| > r\right) \varepsilon_{t-j} = T^{-1/2} \sum_{t=1}^{T} U_t(r) + T^{-1/2} \sum_{t=1}^{T} V_t(r)$$

First we prove that $T^{-1/2} \sum_{t=1}^{T} U_t(r)$ is tight, using the Chentsov's Criterion. We need to prove that for $r_1 \leq r \leq r_2$, $\alpha > 1/2$ and $F()$ nondecreasing, continuous function.

$$E\left[\left|T^{-1/2} \sum_{t=1}^{T} \left(U_t(r_1) - U_t(r)\right)\right|^2 \left|T^{-1/2} \sum_{t=1}^{T} \left(U_t(r) - U_t(r_2)\right)\right|^2\right] \leq \left[F(r_2) - F(r_1)\right]^{2\alpha}$$

Then, define

$$U_t(r_1) - U_t(r) = \varepsilon_t \sum_{j=1}^{T^\alpha} \theta^{0^{j-1}} 1\left(r_1 < |z_{t-j}| < r\right) \varepsilon_{t-j} = \varepsilon_t b_t(r_1)$$

$$E\left[\left(\sum_{t=1}^{T} \varepsilon_t b_t(r_1)\right)^2 \left(\sum_{s=1}^{T} \varepsilon_t b_t(r_2)\right)^2\right] = \sum_{s,s',t,t'} E\left[\varepsilon_t b_t(r_1) \varepsilon_{t'} b_{t'}(r_1) \varepsilon_{s'} b_{s'}(r_2) \varepsilon_s b_s(r_2)\right] \leq$$

$$\sum_{s,s'<t} E\left[\varepsilon_t^2 b_t^2(r_1) \varepsilon_{s'} b_{s'}(r_2) \varepsilon_s b_s(r_2)\right] + \sum_{s>t,t'} E\left[\varepsilon_t b_t(r_1) \varepsilon_{t'} b_{t'}(r_1) \varepsilon_s^2 b_s^2(r_2)\right] +$$

$$+4 \sum_{t',s'<t=s} E\left[\varepsilon_t^2 b_t(r_1) b_t(r_2) \varepsilon_{t'} b_{t'}(r_1) \varepsilon_{s'} b_{s'}(r_2)\right] + 2 \sum_{s'<s=t=t'} E\left[\varepsilon_t^3 b_t^2(r_1) b_t(r_2) \varepsilon_{s'} b_{s'}(r_2)\right] +$$

$$+2 \sum_{t'<s=t=s'} E\left[\varepsilon_t^3 b_t^2(r_2) b_t(r_1) \varepsilon_{t'} b_{t'}(r_2)\right] + \sum_{t'=s=t=s'} E\left[\varepsilon_t^4 b_t^2(r_2) b_s^2(r_2)\right]$$

We bound the first term,

$$\sum_{s,s'<t} E\left[\varepsilon_t^2 b_t^2(r_1) \varepsilon_{s'} b_{s'}(r_2) \varepsilon_s b_s(r_2)\right] = \sum_{t=1}^{t} E\left[\varepsilon_t^2 b_t^2(r_1) \left(\sum_{s=1}^{t} \varepsilon_s b_s(r_2)\right)^2\right]$$

$$E\left[\varepsilon_t^2 b_t^2(r_1) \left(\sum_{s=1}^{t} \varepsilon_s b_s(r_2)\right)^2\right] \leq 2E\left[\varepsilon_t^2 b_t^2(r_1) \left(\sum_{s=1}^{t-T^\alpha-1} \varepsilon_s b_s(r_2)\right)^2\right] + 2E\left[\varepsilon_t^2 b_t^2(r_1) \left(\sum_{s=t-T^\alpha-2}^{t-1} \varepsilon_s b_s(r_2)\right)^2\right]$$

The first summand, using the ILE

$$E\left[\varepsilon_t^2 b_t^2(r_1) \left(\sum_{s=1}^{t-T^\alpha-1} \varepsilon_s b_s(r_2)\right)^2\right] = \sigma_\varepsilon^2 E\left[\left(\sum_{j=1}^{T^\alpha} \theta^{0^{j-1}} 1\left(r_1 < |z_{t-j}| < r\right) \varepsilon_{t-j}\right)^2 \left(\sum_{s=1}^{t-T^\alpha-1} \varepsilon_s b_s(r_2)\right)^2\right] =$$

$$= \sigma_\varepsilon^2 E\left\{E\left[\left(\sum_{j=1}^{T^\alpha} \theta^{0^{j-1}} 1\left(r_1 < |z_{t-j}| < r\right) \varepsilon_{t-j}\right)^2 \middle| F_{t-T^\alpha-1}\right] \left(\sum_{s=1}^{t-T^\alpha-1} \varepsilon_s b_s(r_2)\right)^2\right\}$$

Using the assumption A.7

$$E\left[\left(\sum_{j=1}^{T^{\alpha}}\theta^{0^{j-1}}1\left(r_1<|z_{t-j}|<r\right)\varepsilon_{t-j}\right)^2\middle|F_{t-T^{\alpha}-1}\right]=$$

$$E\left[\left(\begin{array}{c}\sum_{j=1}^{T^{\alpha}}\theta^{0^{2(j-1)}}1\left(r_1<|z_{t-j}|<r\right)\varepsilon_{t-j}^2+\\ +2\sum_{j=1}^{T^{\alpha}-1}\sum_{l=j+1}^{T^{\alpha}}\theta^{0^{j-1}}\theta^{0^{l-1}}1\left(r_1<|z_{t-j}|<r\right)\varepsilon_{t-j}1\left(r_1<|z_{t-l}|<r\right)\varepsilon_{t-l}\end{array}\right)\middle|F_{t-T^{\alpha}-1}\right]\leq$$

$$\frac{\sigma_{\varepsilon/z}^2|r-r_1|}{1-\theta^{0^2}}+2\frac{\sigma_{\varepsilon/z}^2|r-r_1|^2}{\left(1-\theta^0\right)^2}$$

Now we bound

$$E\left(\sum_{s=1}^{t-T^{\alpha}-1}\varepsilon_s b_s\left(r_2\right)\right)^2\leq\left(t-T^{\alpha}-1\right)\frac{\sigma_{\varepsilon}^2\sigma_{\varepsilon/z}^2|r_2-r|}{1-\theta^0}$$

Then

$$E\left[\varepsilon_t^2b_t^2\left(r_1\right)\left(\sum_{s=1}^{t-T^{\alpha}-1}\varepsilon_s b_s\left(r_2\right)\right)^2\right]\leq\left(t-T^{\alpha}-1\right)\left[\frac{\sigma_{\varepsilon}^2\left(\sigma_{\varepsilon/z}^2\right)^2|r_2-r||r-r_1|}{\left(1-\theta^{0^2}\right)\left(1-\theta^0\right)}+2\frac{\sigma_{\varepsilon}^2\left(\sigma_{\varepsilon/z}^2\right)^2|r_2-r||r-r_1|^2}{\left(1-\theta^0\right)^3}\right]$$

Now we bound the second summand

$$E\left[\varepsilon_t^2b_t^2\left(r_1\right)\left(\sum_{s=t-K-2}^{t-1}\varepsilon_s b_s\left(r_2\right)\right)^2\right]\leq T^{\alpha}\sum_{h=1}^{T^{\alpha}+2}E\left[\varepsilon_t^2b_t^2\left(r_1\right)\varepsilon_{t-h}^2b_{t-h}^2\left(r_2\right)\right]$$

Define $v_t=1\left(r_1<|z_t|<r\right)\varepsilon_t$ and $u_t=1\left(r<|z_t|<r_2\right)\varepsilon_t$, and using the LIE

$$E\left[\varepsilon_t^2b_t^2\left(r_1\right)\varepsilon_{t-h}^2b_{t-h}^2\left(r_2\right)\right]\leq E\left[\varepsilon_t^2\varepsilon_{t-h}^2\sum_{l,l',j,j'}\theta^{0^{j-1}}\theta^{0^{l-1}}\theta^{0^{j'-1}}\theta^{0^{l'-1}}v_{t-j}v_{t-j'}u_{t-h-l}u_{t-h-l'}\right]\leq \tag{7}$$

$$\leq\sigma_{\varepsilon}^2\sigma_{\varepsilon/z}^4\frac{|r_1-r||r-r_2|}{\left(1-\theta^0\right)^4}$$

then the first term

$$\sum_{s,s'<t}E\left[\varepsilon_t^2b_t^2\left(r_1\right)\varepsilon_{s'}b_{s'}\left(r_2\right)\varepsilon_s b_s\left(r_2\right)\right]\leq T^2C\left[\frac{\sigma_{\varepsilon}^2\left(\sigma_{\varepsilon/z}^2\right)^2|r_2-r||r-r_1|}{\left(1-\theta^{0^2}\right)\left(1-\theta^0\right)}+2\frac{\sigma_{\varepsilon}^2\left(\sigma_{\varepsilon/z}^2\right)^2|r_2-r||r-r_1|^2}{\left(1-\theta^0\right)^3}\right]+$$

$$TT^{\alpha2}\sigma_{\varepsilon}^2\sigma_{\varepsilon/z}^4\frac{|r_1-r||r-r_2|}{\left(1-\theta^0\right)^4}$$

Then we can define $F\left(r_2\right)=C_1r_2$, which is non decreasing and continuous, and with $\alpha<1/2$,

$$\sum_{s,s'<t}E\left[\varepsilon_t^2b_t^2\left(r_1\right)\varepsilon_{s'}b_{s'}\left(r_2\right)\varepsilon_s b_s\left(r_2\right)\right]\leq T^2\left[F\left(r_1\right)-F\left(r_2\right)\right]^2$$

The second summand is equal. The third one,

$$\sum_{t',s'<t=s}E\left[\varepsilon_t^2b_t\left(r_1\right)b_t\left(r_2\right)\varepsilon_{t'}b_{t'}\left(r_1\right)\varepsilon_{s'}b_{s'}\left(r_2\right)\right]=TE\left[\varepsilon_t b_t\left(r_1\right)\sum_{s'=1}^{t-1}\varepsilon_{s'}b_{s'}\left(r_2\right)\varepsilon_t b_t\left(r_2\right)\sum_{t'=1}^{t-1}\varepsilon_{t'}b_{t'}\left(r_1\right)\right]$$

using the Cauchy-Schwartz' inequality

$$T\left\{E\left[\varepsilon_t^2 b_t^2\left(r_1\right)\left(\sum_{s'=1}^{t-1}\varepsilon_{s'}b_{s'}\left(r_2\right)\right)^2\right]\right\}^{1/2}\left\{E\left[\varepsilon_t^2 b_t^2\left(r_2\right)\left(\sum_{t'=1}^{t-1}\varepsilon_{t'}b_{t'}\left(r_1\right)\right)^2\right]\right\}^{1/2}$$

here we can use the same bound as in the first and second summand. The fourth summand

$$\sum_{s'<s=t=t'} E\left[\varepsilon_t^3 b_t^2\left(r_1\right)b_t\left(r_2\right)\varepsilon_{s'}b_{s'}\left(r_2\right)\right] \leq TE\left[\varepsilon_t^2 b_t\left(r_1\right)b_t\left(r_2\right)\varepsilon_t b_t\left(r_1\right)\sum_{s'=1}^{t-1}\varepsilon_{s'}b_{s'}\left(r_2\right)\right] \leq$$

$$T\left\{E\left[\varepsilon_t^2 b_t^2\left(r_1\right)\varepsilon_t^2 b_t^2\left(r_2\right)\right]\right\}^{1/2}\left\{E\left[\varepsilon_t^2 b_t^2\left(r_1\right)\left(\sum_{s'=1}^{t-1}\varepsilon_{s'}b_{s'}\left(r_2\right)\right)^2\right]\right\}^{1/2}$$

as in equation (7)

$$E\left[\varepsilon_t^2 b_t^2\left(r_1\right)\varepsilon_t^2 b_t^2\left(r_2\right)\right] \leq \sigma_\varepsilon^4\sigma_{\varepsilon/z}^4\frac{\left|r_1-r\right|\left|r-r_2\right|}{\left(1-\theta^0\right)^4}$$

and with the bound of the first summand we have the fourth and fith summand. For the last summand we use this last inequality.

Then we have a Chentsov's Criterion for tightness. Now we only have to prove that

$$\lim_{T\to\infty} P\left[\sup_r T^{-1/2}\sum_{t=1}^{T}V_t\left(r\right) > \varepsilon\right] = 0$$

$$\sup_r T^{-1/2}\sum_{t=1}^{T}V_t\left(r\right) \leq \theta^{0^{T^\alpha}}T^{-1/2}\sum_{t=1}^{T}\left|\varepsilon_t\right|\sum_{j=1}^{t}\theta^{0^{j-1}}\left|\varepsilon_{t-T^\alpha-j}\right|$$

Now using that

$$\theta^{0^{T^\alpha}}T^{-1/2}E\left(\sum_{t=1}^{T}\left|\varepsilon_t\right|\sum_{j=1}^{t}\theta^{0^{j-1}}\left|\varepsilon_{t-T^\alpha-j}\right|\right) \leq \theta^{0^{T^\alpha}}T^{1/2}\frac{E\left|\varepsilon_t\right|}{1-\theta^0} \to 0$$

we are done. ■

9 References

Andrews, D.W.K (1994): "Empirical process methods in econometrics." In R.F. Engle and D.L. McFadden (eds.), *Handbook of Econometrics*, Vol. IV. Elsevier Science, 2248-2296.

Andrews, D.W.K, and Ploberger, W. (1994): "Optimal tests when a nuisance parameter is present only under the alternative ." Econometrica, 62, 1383-1414.

Beveridge, S and Nelson, C.R. (1981): "A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the 'Business Cycle'." Journal of Monetary Economics 7, 151-174.

Blanchard, O. and Quah, D. (1989): "The dynamic effects of Aggregate Demand and Supply disturbances." The American Economic Review, 79, 655-673.

Chan, K.S (1993): "Consistency and Limiting Distribution of the Least Squares Estimator of a Threshold Autoregressive Model". The Annals of Statistics 21, 520-533.

Davidson, J. (1994): *Stochastic Limit Theory*. Oxford University Press.

Davies, R. B. (1977): "Hypothesis testing when a nuisance parameter is present only under the alternative." Biometrika, 64, 247-254.

Davies, R. B. (1987): "Hypothesis testing when a nuisance parameter is present only under the alternative." Biometrika, 74, 33-43.

Elwood, S. (1998):"Is the Persistence of Shocks to Output Asymmetric?". Journal of Monetary Economics, 41, 411-426.

Engle, Robert F. and Smith, Aaron D. (1999): "Stochastic Permanent Breaks". Review of Economics and Statistics 84, 553.

Faust, J. (1998): "On the Robustness of the Identified VAR conclusions about money." Carnegie Rochester Conference Series on Public Policy, vol. 49, 207-244.

Gooijer, Jan G. De (1998): "On Threshold Moving-average Models". Journal of Time Series Analysis 19, n^{o}1, 1-18.

Granger, C.W.J and Andersen, Allan.(1978): "On the invertibility of time series models". Stochastic Processes and their Applications 8, 87-92.

Guay, A. and Scaillet, O. (2003):"Indirect Inference, Nuisance Parameter, and Threshold Moving Average Models". Journal of Business & Economics Statistics, Vol. 21, 122-132.

Hallin, M. (1980): "Invertibility and Generalized Invertibility of Time Series Models ". J. R. Statist. Soc. B, 42, 210-212.

Hansen, Bruce E. (1996):"Inference when a Nuisance Parameter is not Identified under the Null Hypothesis". Econometrica 64, n^{o}2, 413-430.

Hansen, Bruce E. (2000):"Sample splitting and threshold estimation". Econometrica, Vol. 68, n^{o}3, pag 575-603.

Harvey, Andrew C. (1989): *Forecasting, structural time series models and the Kalman filter.* Cambridge University Prest.

Koop, G., Pesaran, H., Potter, S. (1996): "Impulse response analysis in nonlinear multivariate models". Journal of Econometrics 74, 119-148.

Maravall, A. (1995): *Handbook of applied econometrics* Vol. I. M. Hashem Pesaran and Michael R. Wickens. Blackwell.

Pöcher, M. and Prucha, I.R.(1997): *Dynamic Nonlinear Econometric Models. Asymptotic Theory.* Springer.

Potter, S. (2000):"Nonlinear impulse response functions". Journal of Economics Dynamics & Control, 24, 1425-1446.

Tong, H. (1990): *Non-linear Time series. A Dynamical System approach.* Oxford: Oxford University Press.

Uhlig, H. (1997): "What are the Effects of Monetary Policy? Results from and Agnostic Identification Procedure," draft, Tilburg University, Center for Economic Research.

Watson, Mark W. (1986): "Univariate detrending methods with stochastic trends". Journal of Monetary Economics 18, 49-75.

Wecker, William E.(1981): "Asymmetric Time Series". Journal of the American Statistical Association 76, n^{o}373, 16-21.