Large shocks vs. small shocks.
(Or does size matter? May be so.)

Jesús Gonzalo*, Oscar Martínez

Department of Economics, U. Carlos III de Madrid, 28903 Getafe, Madrid, Spain

Abstract

What are the shocks that drive economic fluctuations? The answer to this question requires as a first step solving the shock identification issue. This paper proposes a new identification scheme based on two aspects: the long run effect of the shock (permanent or transitory), and the size of the shock (Large or small). This is done by using a threshold integrated moving average model (TIMA) previously introduced in the literature by the authors. Based on this model we develop a testing strategy to determine whether Large and small shocks have different long run effects, as well as whether one of them is purely transitory. The paper analyzes the impulse response function of both types of shocks, and provides the asymptotic results sufficient to implement the above testing strategy. Based on these results we develop a new nonlinear permanent transitory decomposition, that is applied to US stock prices to analyze the quality of the stock market, and to US GNP to investigate the asymmetric behavior of its shocks.
© 2005 Elsevier B.V. All rights reserved.

*Corresponding author.
E-mail address:* jesus.gonzalo@uc3m.es (J. Gonzalo).
URL: http://www.eco.uc3m.es/english/staff/contact_info/jgonzalo.html.

1. Introduction

What are the shocks that drive the economy? Where does the persistence present in most of the economic time series come from? Is the persistence of output shocks symmetric or asymmetric? What fraction of output variation is due to supply or demand shocks? All the attempts to answer these questions pass through the critical decision of shock identification.

Traditionally this has been done by decomposing the analyzed variable into unobserved permanent and transitory components (see Watson, 1986). Some examples are: (1) Beveridge and Nelson (1981) decomposition (B–N hereafter), where the permanent component is a random walk and there is perfect correlation between permanent and transitory shocks due to the fact that there is only one shock, the permanent one; (2) Unobserved component models with uncorrelated components (UC–0) (see Harvey, 1985; Clark, 1987), where the permanent component is also assumed to be a random walk but its innovation is uncorrelated with the one in the transitory term. Note that not all the ARIMA models admit an UC–0 decomposition. This problem can be solved eliminating the random walk constraint on the permanent component. In this case we have an extra identification problem that is overcome by imposing an ad hoc smooth condition on the permanent term (see the ''canonical'' decomposition in Pierce, 1979). All these decompositions present two basic problems that are not solved in the literature: (i) at every time period t there is always a permanent shock, and (ii) none of the assumptions on permanent and transitory components are testable. The approach proposed in this paper tries to solve the identification issue without encountering these problems.

Our proposal is based on two pillars. First, as in the above mentioned identification attempts we accept that there are permanent and transitory shocks, and secondly, shocks behave differently in the long-run according to their own characteristics (sign, size, etc.) or some characteristics of the economy. The first premise, as is well known (see Quah, 1992) is not enough to identify the permanent and transitory shocks of a univariate time series. This paper shows that with our second assumption we cannot only identify the permanent and transitory shocks of a single economic variable, but we can test whether in fact these shocks are transitory or permanent. In order to implement our proposal we use a class of threshold models, the threshold integrated moving average models (TIMA) previously introduced in the literature by Gonzalo and Martínez (2004). These are models with a unit root in the autoregressive part to capture persistence, and a threshold design in the moving average side to allow for asymmetries. The threshold variable can represent any characteristic of the shocks (large or small, positive or negative, etc.) or of the economy (expansion or recession, inflation or deflation, etc.) we are interested on. This threshold design is able to capture the possibility that any of these characteristics may have different long run effects. By allowing the existence of a unit root in some of the moving average regime, we permit that the characteristics triggering those regimes have only a transitory effect.

In principle, TIMA models can deal with different shock characteristics, but in this paper we will focus on shock size as in the StopBreak model of Engle and Smith

(1999). Our maintained hypothesis is that large shocks will have permanent effects while small shocks will produce only transitory effects. This hypothesis is based on the assumption that certain times series are less likely to mean revert after a large shock than after a small one. For example in the stock market, investors may perceive large shocks as containing informational content and small shocks as mere noise. In macroeconomics, an unsolved issue is whether economic fluctuations are due to an accumulation of small shocks or instead mostly to infrequent large shocks. A positive aspect of our framework is that we are able to test our maintained hypothesis.

The identification of permanent and transitory shocks with our procedure, allows us to construct an orthogonal nonlinear permanent–transitory decomposition of the original variable. We also obtain a nonlinear B–N decomposition that while sharing part of the spirit of both, the standard linear B–N and the UC–0 decompositions, its behavior lies in between.

Threshold moving average (TMA) models have already been considered in Wecker (1981) and in De Gooijer (1998). Both works are centered on presenting the new TMA model and on analyzing some of the moment properties in detail. They both assume normality, the threshold parameter to be known and equal to zero, and they do not present asymptotic results. These can be found in Guay and Scaillet (2003) using indirect inference. One part of our paper can be considered complementary to this one in the sense of also developing asymptotic results; but in our case using Hansen (1996), González and Gonzalo (1998), and Caner and Hansen (2001) approach to threshold models.

The paper finishes with two applications of our TIMA Shock-Size model: first, to stock prices, where following Hasbrouck (1993) the size of the transitory component can be considered a measure of the quality or efficiency of the stock market; and second to GNP, where the majority of the research that investigates the size of its permanent component has implicitly imposed symmetry, and the minority that allows for asymmetries driven by the sign of the shocks has not reached any final conclusion yet (see Elwood, 1998).

The rest of the paper is structured as follows. Section 2 introduces the TIMA Shock-Size model and analyzes in detail its impulse response function (IRF). Section 3 presents the asymptotic theory results needed for testing the hypothesis of interests. In Section 4 we define two new nonlinear permanent–transitory decompositions and compare them with the existing linear decompositions. Sections 5 and 6 present two empirical applications of our TIMA Shock-Size model, one to measure the quality of the stock market and the other to analyze whether or not the persistence of shocks to GNP is asymmetric. Finally, Section 7 draws some concluding remarks. The appendix contains technical derivations and proofs of the results in the main text.

2. TIMA Shock-Size models

In this section we lay out the basis of the TIMA Shock-Size model. In order to better understand this model, we start with a more general structure, the

autoregressive threshold integrated moving average ARTIMA $(p, m, 1, 1)$ model introduced in Gonzalo and Martínez (2004),

$$\Phi_p(L)(1 - L)y_t = \mu + \sum_{j=1}^{m+1} (1 - \theta_j L)\varepsilon_t 1(r_{j-1} < z_t \leqslant r_j), \tag{1}$$

where L is the lag operator, $\Phi_p(L) = (1 - \phi_1 L - \cdots - \phi_p L^p)$ with all its roots outside the unit circle, $1(.)$ denotes the indicator function, and z_t the threshold variable that triggers the regime switches. The random error term ε_t is a real *i.i.d* sequence with $E|\varepsilon_t|^{2\gamma} < \infty$ for some $\gamma > 2$. The threshold parameters denoted $(r_1, \ldots, r_m)$ with $r_0 = -\infty$, $r_{m+1} = \infty$ are such that $r_i \in R_m \ \forall i = 1, \ldots, m$ with $R_m = \{(r_1, \ldots, r_m) : -\infty <$

$\underline{r} < r_1 < \cdots < r_m < \bar{r} < \infty\}$. Thus we require all threshold parameters to lie in the bounded subset $[\underline{r}, \bar{r}]$ of the threshold variable sample space.

Some examples of threshold variables z_t are: (i) ε_t, for instance when the sign is the shock characteristic that triggers the regime switches (as in Wecker, 1981; Elwood, 1998; Guay and Scaillet, 2003); (ii) $(1 - L)y_t$ or any other economic variable that is strictly stationary and ergodic (as in De Gooijer, 1998); and (iii) $|\varepsilon_t|$, the shock's size.

In this paper we are only concerned with the asymmetry produced by the size of a shock ($z_t = |\varepsilon_t|$). To present the main results of ARTIMA models producing such asymmetry (ARTIMA Shock-Size), and without loss of generality, we will use a simpler version of model (1). A version with no autoregressive part ($\Phi_p(L) = 1$) and only two regimes (the case of more regimes can be handled as in Gonzalo and Pitarakis, 2002) that is denoted TIMA Shock-Size model:

$$(1 - L)y_t = \mu + \begin{cases} \varepsilon_t - \theta_1 \varepsilon_{t-1} & \text{if} \quad |\varepsilon_{t-1}| > r, \\ \varepsilon_t - \theta_2 \varepsilon_{t-1} & \text{if} \quad |\varepsilon_{t-1}| \leqslant r, \end{cases} \tag{2}$$

or in a more compact way

$$(1 - L)y_t = \mu + \varepsilon_t - \theta_{t-1}\varepsilon_{t-1}, \tag{3}$$

with $\theta_{t-1} = \theta_1$ if $|\varepsilon_{t-1}| > r$, and equal to θ_2 otherwise. This TIMA model is a time varying moving average model, and in that sense can be seen as the threshold version or approximation of the StopBreak model of Engle and Smith (1999). The main difference is that in the latter all the shocks are permanent (a shock is transitory if it is equal to zero), while in the former the existence of both permanent and transitory shocks is allowed.

Assumptions.

A.0. ε_t iid $(0, \sigma_\varepsilon)$, with a uniformly continuous density function $0 < f_\varepsilon < \infty$, and $E(\varepsilon_t^{2\gamma}) < \infty$ with $\gamma > 2$.

A.1. $|\theta_1^0 - \theta_2^0| = h^0 > 0$.

A.2. $\bar{E} = [E(|\theta(\varepsilon_{t-1})|) + rhf^m] < 1$ and $|\theta_1| < |\theta_2|$, where $h = |\theta_1 - \theta_2|$, and $f^m = \max_e(f_\varepsilon(-r + e) + f_\varepsilon(r + e))$ with e defined in the support of ε_t.

A.3. The vector of the true parameter values $\vartheta^0 = (\mu^0, \theta_1^0, \theta_2^0, r^0) \in \Theta = [-\overline{\mu}, \overline{\mu}] \times [-1 + \delta, 1 - \delta] \times [-1 + \delta, 1] \times [\underline{r}, \overline{r}]$ s.t.
$\overline{E}^* = [E^*(\theta_1, \theta_2, r) + rhf^{m*}] < 1$ and $|\theta_1| < |\theta_2| \quad \forall \vartheta \in \Theta$, where
$E^*(\theta_1, \theta_2, r) = |\theta_1|(1 - \overline{p}(r)) + |\theta_2|\overline{p}(r)$, $\overline{p}(r) \geqslant \sup_k P(|\varepsilon_t + k| < r)$,
$f^{m*} = 4\max_e f_\varepsilon(e)$, $\delta > 0$, and $0 < \underline{r} < \overline{r} < \infty$.

A.4. $E[\varepsilon_t / \mathfrak{I}_{t-1}] = 0$ and $E[\varepsilon_{t+j} 1(|\varepsilon_{t+j}| > r) / \mathfrak{I}_{t-1}] = 0$ for $j \geqslant 0$, with *sigma-field* $\mathfrak{I}_t \stackrel{\text{def}}{=} \{(\varepsilon_j), j \leqslant t\}$.

A.5. $\theta_i = 1$ for some $i = 1, 2$.

Assumptions A.0 A.1 are standard identification assumptions in threshold models and empirical processes (see Bai, 1994). Assumption A.2 is a sufficient condition for invertibility. We can distinguish two different parts in it. In the first one, $E(|\theta(\varepsilon_{t-1})|)$ must be less than 1. This takes care of the case of overdifferencing, especially when we allow a unit root in the MA part. Non-invertibility is not only a problem of overdifferencing, but a problem of nonlinearity too. The second part, rhf^m, takes care of the degree of nonlinearity, measured as a product of the *gap* (rh) and its upper limit probability f^m. A.3 describes the parametric space of TIMA Shock-Size models (partially related to the invertibility condition) and constitutes a sufficient requirement for the asymptotic results obtained in the paper. Finally, assumptions A.4 and A.5 are helpful assumptions to interpret the outcomes of our model, i.e. to prove the existence of an orthogonal permanent transitory decomposition in terms of the IRF. A.4 is a symmetry condition of the conditional distribution of the shocks that can be relaxed. It allows the identification of permanent and transitory shocks based on the IRF. Otherwise, the IRF is unknown and has to be estimated (see for example, Clarida and Taylor, 2003). A.5 (together with A.4) guarantees the existence of pure transitory shocks.

2.1. Impulse response function

To analyze in detail the asymmetric persistence of the shocks present in a TIMA model we have to study the behavior of its IRF. This function measures the effect of a perturbation at time t in the sample path $\{y_{t+k}\}_{k\ 0}^{\infty}$. If this effect on y_{t+k} does not vanish when $k \to \infty$ we say that the shock is persistent. With linear models there is a general consensus about the relation of the IRF and the persistence of the shock. However with nonlinear models three main aspects of the time series come up to determine the definition of the IRF and its relationship with persistence. These aspects are the history of the series at time $t - 1$, the shock ε_t and future shocks. To capture all these three new aspects we use the Generalized impulse response function (*GIRF*) introduced and defined in Koop et al. (1996) and in Potter (2000) as

$$GIRF(k, \varepsilon_t, w_{t-1}) = E[y_{t+k} / \varepsilon_t, w_{t-1}] - E[y_{t+k} / w_{t-1}] \quad \text{for } k = 0, 1, 2, \dots, \quad (4)$$

where w_t is the sample history of the process until time t. According to this definition, a shock is persistent if the effect of knowing it on the conditional expectation of y_{t+k} (given the past) does not vanish when $k \to \infty$.

In order to show that TIMA Shock-Size models can admit both type of shocks, permanent as well as transitory, we use the following result.

Result 1. *Let y_t follow the TIMA process* (2). *Under assumption* A.4, *the GIRF of y_t is given by*

$$GIRF(k, \varepsilon_t, w_{t-1}) = (1 - \theta_t)\varepsilon_t.$$

From this result is clear that the shock ε_t is transitory if $\theta_t = 1$ (assumption A.5). It is worth noting that it is the size of ε_t, that determines the persistent or transitory effect of ε_t. It is also important to notice that when A.4 does not hold the *GIRF* may be different from $(1 - \theta_t)\varepsilon_t$. This is the case, for instance, when the threshold variable is the sign of the shock. In this situation, TIMA model will not have any transitory shock.

To better understand the behavior of the *GIRF* and the main differences between our TIMA model and other linear and nonlinear models, we consider four examples for which we calculate their theoretical *GIRF*. In all these examples, the shocks ε_t are assumed to be *iid*. A simulation exercise where the different *GIRF*'s are compared is shown in Fig. 1.

Example 1 (*Linear model*). In this example we consider the standard general linear model,

$$y_t = \sum_{j=0}^{\infty} \theta_j \varepsilon_{t-j},$$
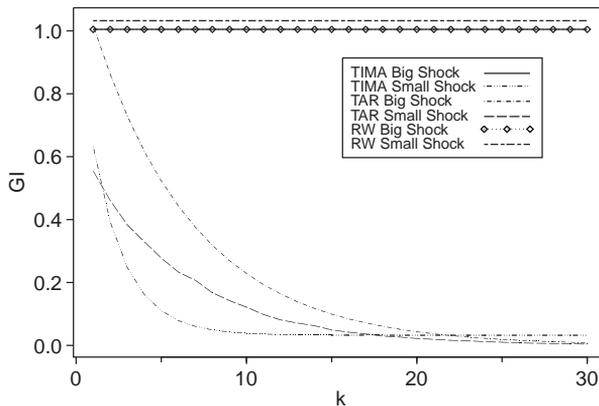
Fig. 1. Mean of the estimated *GIRF* (divided by ε_t). The mean of the *GIRF* is obtained conditional on the shock size ($|\varepsilon_t| \leqslant 0.5$ and $|\varepsilon_t| > 0.5$). The mean is calculated over 1000 Monte Carlo replications for 200 values of $\varepsilon_t \equiv N(0,1)$. We select a random history w_{t-1} with $t = 201$.

with $\theta_0 = 1$. Then

$$y_{t+k} = \sum_{j\,0}^{k-1} \theta_j \varepsilon_{t+k-j} + \theta_k \varepsilon_t + \sum_{j\,1}^{\infty} \theta_{k+j} \varepsilon_{t-j}.$$

Therefore

$$GIRF(k, \varepsilon_t, w_{t-1}) = \theta_k \varepsilon_t.$$

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 $\{\varepsilon_t\}_{-\infty}^{\infty}$, are transitory. In the second one, all of them are permanent. The long run properties of the shock, ε_t, only depend on the parameter θ_k, they do not depend either on ε_t or on the past. There is no asymmetric persistence. For the simulation and comparison exercise of Fig. 1, we use as an example of linear model a random walk (RW)

$$y_t = y_{t-1} + \varepsilon_t.$$

This model produces the following *GIRF*:

$$GIRF(k, \varepsilon_t, w_{t-1}) = \varepsilon_t,$$

where clearly all the shocks are permanent.

Example 2 (*TAR model*). In this example we consider the following TAR model

$$y_t = \phi_1 y_{t-1} 1(|\varepsilon_{t-1}| > r) + \phi_2 y_{t-1} 1(|\varepsilon_{t-1}| \leqslant r) + \varepsilon_t.$$

Rewriting this model at time $t + k$

$$y_{t+k} = \sum_{j\,0}^{\infty} \theta_{t+k,j} \varepsilon_{t+k-j}$$

$$= \sum_{j\,0}^{k-1} \theta_{t+k,j} \varepsilon_{t+k-j} + \theta_{t+k,k} \varepsilon_t + \sum_{j\,k+1}^{\infty} \theta_{t+k,j} \varepsilon_{t+k-j},$$

with $\theta_{t,j} = \prod_{i\,1}^{j} \phi_{t-i}$, $\phi_{t-i} = \phi_1 1(|\varepsilon_{t-i}| > r) + \phi_2 1(|\varepsilon_{t-i}| \leqslant r)$ for $j \geqslant 1$, and $\theta_{t,0} = 1$. Therefore

$$GIRF(k, \varepsilon_t, w_{t-1}) = \sum_{j\,0}^{k-1} [E(\theta_{t+k,j} \varepsilon_{t+k-j}/w_{t-1}, \varepsilon_t) - E(\theta_{t+k,j} \varepsilon_{t+k-j}/w_{t-1})]$$

$$+ E(\theta_{t+k,k}/\varepsilon_t, w_{t-1}) \varepsilon_t - E(\theta_{t+k,k} \varepsilon_t/w_{t-1}). \tag{5}$$

For TAR models the *GIRF* is nonlinear and it depends on the past of ε_t. In general, the problem is to obtain the expectations involved in the right-hand side of (5). This problem can be solved by simulation (see Koop et al., 1996; Potter, 2000). For the particular TAR model of this example, it can be proved that under assumption A.4,

$$GIRF(k, \varepsilon_t, w_{t-1}) = \phi_t \prod_{j\,t+1}^{t+k-1} E(\phi_j/\varepsilon_t, w_{t-1}) \varepsilon_t.$$

Assuming the parameters (ϕ_1, ϕ_2) satisfy some stationarity and ergodicity conditions (see Petrucelli and Woodford, 1984), the above *GIRF* tends to 0 when $k \to \infty$. Therefore although we can obtain asymmetric behavior in the effect of the shocks, all of them are transitory.

For the simulation and comparison study, we use the following TARUR model (introduced by González and Gonzalo, 1998) with a unit root in the regime corresponding to large shocks,

$$y_t = \begin{cases} y_{t-1} + \varepsilon_t & \text{if} \quad |\varepsilon_{t-1}| > 0.5, \\ 0.6 y_{t-1} + \varepsilon_t & \text{if} \quad |\varepsilon_{t-1}| \leqslant 0.5. \end{cases}$$

Its *GIRF* is

$$GIRF(k, \varepsilon_t, w_{t-1}) = \begin{cases} [\mathrm{E}(\phi_t)]^{k-1} \varepsilon_t & \text{if} \quad |\varepsilon_t| > 0.5, \\ 0.6 [\mathrm{E}(\phi_t)]^{k-1} \varepsilon_t & \text{if} \quad |\varepsilon_t| \leqslant 0.5, \end{cases}$$

where $\mathrm{E}(\phi_t) < 1$. Note that in this case all shocks are transitory, although the effect depends on the shock's size.

Example 3 (*TIMA model*). In this example we analyze our TIMA Shock-Size model (2) with *GIRF* described in Result 1. For the simulation and comparison study of Fig. 1, and in order to compare closely to the TAR model of Example 2, we use the following ARTIMA (1,1,1,1) model:

$$(1 - 0.6L)(1 - L)y_t = \begin{cases} \varepsilon_t - 0.6\varepsilon_{t-1} & \text{if} \quad |\varepsilon_{t-1}| > 0.5, \\ \varepsilon_t - \varepsilon_{t-1} & \text{if} \quad |\varepsilon_{t-1}| \leqslant 0.5. \end{cases}$$

From Result 1 it is obvious that the *GIRF* of this model is

$$GIRF(k, \varepsilon_t, w_{t-1}) = \begin{cases} \varepsilon_t & \text{if} \quad |\varepsilon_t| > 0.5, \\ (0.6)^{k-1} \varepsilon_t & \text{if} \quad |\varepsilon_t| \leqslant 0.5. \end{cases}$$

If $|\varepsilon_t| \leqslant 0.5$ (small shock), ε_t is transitory. When $|\varepsilon_t| > 0.5$ (large shock), ε_t is permanent. Recall that although the threshold in t depends on ε_{t-1}, it is the size of ε_t which determines the type of persistence effect.

Examples 2 and 3 are very similar. In both examples, y_t behaves like a random walk in the big shock regime, and like a stationary process in the small shock regime. Nevertheless the *GIRF* is totally different. Comparing the three theoretical *GIRFs* it is observed that: (i) in the RW case, there is no asymmetric behavior, *GIRF* is the same for all shocks; (ii) in the TAR case, the size of the shock affects its persistence (asymmetric behavior), but all of them are transitory; and (iii) for the ARTIMA Shock-Size model, *GIRF* is equal to the one produced by a random walk when shocks are large, and closer to the *GIRF* produced by the TAR model when shocks are small.

Fig. 1 shows the results of a simulation exercise, where the *GIRFs* (divided by ε_t) of Examples 1–3 have been estimated. To estimate the *GIRF* we generate 1000 future paths of $\{\varepsilon_{t+j}\}_{j-1}^{30}$ for each 200 realizations of ε_t, and only one history w_{t-1}. The *GIRF*

is obtained as the sample mean of these 1000 future paths conditional on the shock size. Fig. 1 summarizes graphically the results obtained in Examples 1–3.

Finally we compare the TIMA Shock-Size model (2) with the StopBreak model of Engle and Smith (1999).

Example 4 (*StopBreak model*). In its simplest form, the StopBreak process is

$$y_t = m_t + \varepsilon_t,$$

where m_t is a time-varying conditional mean which is updated via

$$m_t = m_{t-1} + q_{t-1}\varepsilon_{t-1},$$

with $q_t = q(\varepsilon_t)$. In particular, Engle and Smith (1999) consider the following $q(.)$ function

$$q(\gamma)_t = \frac{\varepsilon_t^2}{\gamma + \varepsilon_t^2}, \quad \gamma > 0.$$

The *GIRF* function for the StopBreak model (assuming $E(q_t\varepsilon_t) = 0$) is

$$GIRF(k, \varepsilon_t, w_{t-1}) = q_t\varepsilon_t.$$

Fig. 2 compares graphically the *GIRF* (divided by ε_t) of a StopBreak process ($\gamma = 5$), with the one of a TIMA Shock-Size model ($\theta_1 = 0$, $\theta_2 = 1$ and $r = 1$). From this figure, it is clear that in the StopBreak process all the shocks are permanent, while in the TIMA Shock-Size model both type of shocks (permanent and transitory) are able to co-exist.
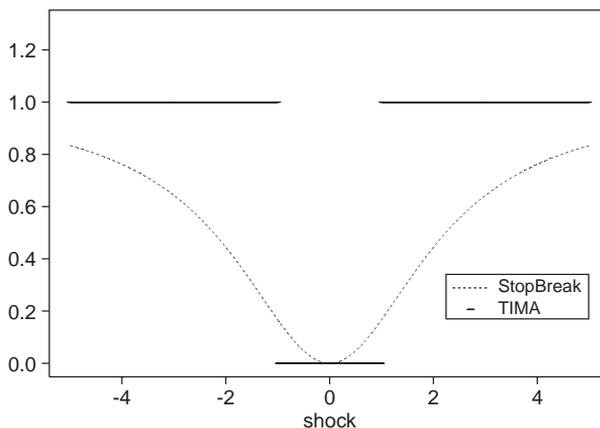
Fig. 2. *GIRF* of StopBreak ($\gamma = 5$) and TIMA ($\theta_1 = 0, \theta_2 = 1, r = 1$) models (divided by ε_t). StopBreak model: $y_t = m_t + \varepsilon_t$, with $m_t = m_{t-1} + q_{t-1}(\gamma)\varepsilon_{t-1}$ and $q_t = \varepsilon_t^2/(\gamma + \varepsilon_t^2)$. TIMA Shock-Size: $(1 - L)y_t = \varepsilon_t - \theta_1\varepsilon_{t-1}1(|\varepsilon_{t-1}| > r) - \theta_2\varepsilon_{t-1}1(|\varepsilon_{t-1}| \leqslant r)$.

3. Estimation and inference of TIMA Shock-Size models

In TIMA, as well as in TAR models, it is necessary to distinguish between the cases of observable and unobservable threshold variable. When the threshold variable is observable (for instance, z_t $(1 - L)y_t$), it can be proved (see Gonzalo and Martínez (2004)) that the least squares (LS) estimators of the TIMA model parameters behave asymptotically like the corresponding LS estimators of the discontinuous TAR model parameters: $T^{1/2}$ consistent for $\widehat{\mu}, \widehat{\theta}_1, \widehat{\theta}_2$ and T for $\widehat{r}$. The latter speed of convergence is due to the kind of discontinuity in r, present in both models via the indicator function.

In the TIMA Shock-Size model, the situation becomes different because the threshold variable ε_t is unobservable, and therefore has to be estimated. The model is discontinuous in all the parameters, again via the indicator function. This is the case, even if the threshold parameter r is known. The discontinuity implies that the rate of convergence for all the parameters becomes now T^{-1}. In more detail, the (joint) estimation of the parameter vector ϑ $(\mu, \theta_1, \theta_2, r)$ is carried out by conditional least squares (CLS) (see Chan, 1993). For simplicity, without lack of generality we assume μ 0. First we minimize

$$Q_T(\vartheta) \quad \sum_{t\ 1}^{T} e_t^2(\vartheta),$$

with

$$e_t(\vartheta) \quad \theta_{t-1} e_{t-1}(\vartheta) + x_t, \quad e_0 \quad 0,$$

$$\theta_{t-1} \quad \begin{cases} \theta_1 & \text{if} \quad |e_{t-1}| > r, \\ \theta_2 & \text{if} \quad |e_{t-1}| \leqslant r, \end{cases}$$

and x_t $(1 - L)y_t$. For a given r, the LS solution is $\widehat{\theta}_1(r)$ and $\widehat{\theta}_2(r)$. Second, we minimize $Q_T(\widehat{\theta}_1(r), \widehat{\theta}_2(r), r)$, obtaining $\widehat{r}$. The CLS estimator of ϑ (θ_1, θ_2, r) is

$$\widehat{\vartheta} \quad (\widehat{\theta}_1, \widehat{\theta}_2, \widehat{r}) \equiv (\widehat{\theta}_1(\widehat{r}), \widehat{\theta}_2(\widehat{r}), \widehat{r}).$$

The rate of convergence of $\widehat{\vartheta}$ is provided in the next result.

Result 2. Let y_t follow the TIMA process (2). Under A.0, A.1 and A.3, $\widehat{\theta}_{i,T}$ $\theta_i^0 + O_p(T^{-1})$ for i 1, 2, and $\widehat{r}_T$ $r_0 + O_p(T^{-1})$.

In TIMA Shock-Size models, all the parameters enter into the indicator function through $e_t(\vartheta)$, and therefore there is a discontinuity in r, as well as in the slope parameters θ_i (i 1, 2).

When the threshold variable is observable the asymptotic distribution of $\widehat{\theta}_1, \widehat{\theta}_2$ is obtained in Gonzalo and Martínez (2004). For the TIMA Shock-Size case we have not been able yet to obtain the asymptotic distribution directly. This is a very difficult task and we recognize that it is beyond the scope of this paper. Therefore we propose an alternative way of obtaining the asymptotic distribution of $\widehat{\theta}$ (θ_1, θ_2). This alternative procedure consists on two steps. In a first step, the TIMA

Shock-Size model is estimated by CLS as previously described. In a second step, we consider ε_t and r^0 to be known and equal to the CLS estimates of the first step. Basically, this can be done because in the first step everything is $T-$consistent. In more detail, in the second step we obtain $\widehat{\theta}(\widehat{\vartheta})$ by minimizing

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

with

$$e_t(\theta, \widehat{\vartheta}) = \theta_1 1(|\widehat{e}_{t-1}| > \widehat{r})\widehat{e}_{t-1} + \theta_2 1(|\widehat{e}_{t-1}| \leqslant \widehat{r})\widehat{e}_{t-1} + x_t.$$

The asymptotic normality of $\widehat{\theta}(\widehat{\vartheta})$, is derived from the fact that $(\widehat{\theta}(\widehat{\vartheta}) - \theta^0) = O_p(T^{-1/2})$, together with Result 2. This is established in the next result.

Result 3. *Let y_t follow the TIMA process* (2). *Under A.0, A.1 and A.3,*

$$T^{1/2}(\widehat{\theta}(\widehat{\vartheta}) - \theta^0) \xrightarrow{\ d\ } N(0, \Omega),$$

where

$$\Omega = E[\varepsilon_t^2] \begin{bmatrix} E(\varepsilon_{t-1}^2 1(|\varepsilon_{t-1}| > r^0)) & 0 \\ 0 & E(\varepsilon_{t-1}^2 1(|\varepsilon_{t-1}| \leqslant r^0)) \end{bmatrix}^{-1}.$$

From this result it is easy to test whether a shock is transitory or not. This is done by testing if one of the θ parameters is equal to one, for instant, via a standard $t - test$.

3.1. A testing strategy

In order to be able to detect whether the persistence of shocks is asymmetric or not, and to discover the existence of pure transitory shocks, we propose a simple testing strategy inside the framework of TIMA models. This testing strategy follows the lines of González and Gonzalo (1998), Caner and Hansen (2001), and Gonzalo and Montesinos (2002) for TAR models with unit roots. It consists on testing first for the existence of a threshold effect

$$H_0 : \theta_1 = \theta_2$$
$$H_a : \theta_1 \neq \theta_2, \tag{6}$$

and second, on testing for a unit root in one of the moving average regimes (e.g. the second regime)

$$H_0 : \theta_2 = 1$$
$$H_a : \theta_2 \neq 1. \tag{7}$$

There are many ways of carrying out a test for the first hypothesis. In this paper, we test the null of linearity versus the alternative of a threshold effect, in a similar fashion as Engle and Smith (1999) test linearity versus a StopBreak process. Under

the null hypothesis of linearity we fit an ARIMA$(p, 1, q)$ model to y_t

$$\Phi_p(L)(1 - L)y_t = \mu + \Theta_q(L)\varepsilon_t, \tag{8}$$

and use the residuals, $\widehat{\varepsilon}_t$, to construct the following auxiliary regression

$$\widehat{\varepsilon}_t = \alpha\widehat{\varepsilon}_{t-1}1(|\widehat{\varepsilon}_{t-1}| < r) + u_t. \tag{9}$$

Under the null, $\alpha = 0$, and under the alternative, $\alpha \neq 0$. This can be tested by a $t - test$. The only problem is that r is not identified under the null, so we proceed as in Davis (1977, 1987) taking the *sup* of the absolute value of the *t*-statistic over all values of $r \in [\underline{r}, \overline{r}]$. The asymptotic distribution is provided in the next result, where $\Rightarrow$ denote weak convergence with respect to the Skorohod metric. Following Wooldridge (1994), let $\widehat{\lambda} = (\widehat{\mu}, \widehat{\phi}_1, \ldots, \widehat{\phi}_p, \widehat{\theta}_1, \ldots, \widehat{\theta}_q)$ be the LS estimator of $\lambda^0 = (\mu^0, \phi_1^0, \ldots, \phi_p^0, \theta_1^0, \ldots, \theta_q^0)$ from model (8), defined as

$$\widehat{\lambda} = \arg\min_{\lambda} \sum_{t\ 1}^{T} e_t^2(\lambda) = \arg\min_{\lambda} \sum_{t\ 1}^{T} q_t(w_t, \lambda),$$

with $e_t(\lambda)$ the error term from model (8) for a given value of the parameter λ. Then, from Theorem 4.4 in Wooldridge (1994), it is easy to prove that under A.0,

$$T^{1/2}(\widehat{\lambda}_i - \lambda_i^0) = T^{-1/2}H_{i.}^{-1}(\lambda^0)\sum_{t\ 1}^{T} s_t(\lambda^0) + o_p(1), \quad \text{for} \quad i = 0, \ldots, p + q,$$

where $s_t(\lambda)$ is the score vector, and $H_i(\lambda^0)$ the *i*-row of $H(\lambda^0) = \lim_{T\to\infty} T^{-1}\sum_{t\ 1}^{T} E[h_t(\lambda^0)]$ with $h_t(\lambda)$ the Hessian matrix.

Result 4. *Let $t_{\alpha(r)}$ be the t-ratio of the parameter α in the auxiliary regression* (9) *for a given r. Under the null of linearity plus assumption A.0,*

$$t_{\alpha(r)} \Rightarrow D(r) \qquad in \quad [\underline{r}, \overline{r}],$$

with

$$D(r) \equiv \left[D_1(r) + \sum_{i\ 0}^{p+q} V_{\lambda_i}\Sigma_\varepsilon(r)\right][\Sigma_\varepsilon(r)\sigma_\varepsilon^2]^{-1/2},$$

where $D_1(r)$ is a continuous gaussian process with covariance function

$$M_1(r_1, r_2) = E(\varepsilon_t^2)E(\varepsilon_{t-1}^2 1(|\varepsilon_{t-1}| \leqslant \min(r_1, r_2))),$$

V_{λ_i} *a random variable with distribution given by the asymptotic normal distribution of $T^{1/2}(\widehat{\lambda}_i - \lambda_i^0)$, and $\Sigma_\varepsilon(r) = E(\varepsilon_t^2 1(|\varepsilon_t| \leqslant r))$. Finally, to complete the definition of $D(r)$, $Cov(D_1(r), V_{\lambda_i}) = H_{i.}^{-1}(\lambda^0)\lim_{T\to\infty} T^{-1}\sum_{t\ 1}^{T} E[\varepsilon_t\varepsilon_{t-1}1(|\varepsilon_{t-1}| \leqslant r)s_t(\lambda^0)]$.*
Then, by the continuous mapping theorem,

$$\sup_{r\in[\underline{r},\overline{r}]} |t_{\alpha(r)}| \xrightarrow{d} \sup_{r\in[\underline{r},\overline{r}]} |D(r)|.$$

The asymptotic distribution of $\sup_{r\in[\underline{r},\overline{r}]}|D(r)|$ depends on nuisance parameters. Therefore to obtain the p-values of the test we propose a simple bootstrap method. This is a model-based bootstrap (see Davison and Hinkley, 1997; Section 8.2.2), that is next described in a computer algorithm format.

Algorithm 1 generates bootstrap samples $\{y_t^\star\}_{t\,1}^T$ from model (8) and calculates the bootstrap approximation of the distribution of $\sup_{r\in[\underline{r},\overline{r}]}|t_{\alpha(r)}|$.

Algorithm 1 (*Model-based Bootstrap procedure*).

1. $l = 1$.
2.   *Generate* $\{\varepsilon_t^\star\}_{t\,1}^T$ *resampling from* $\{\widehat{\varepsilon}_t - \overline{\widehat{\varepsilon}}\}_{t\,1}^T$, *with* $\overline{\widehat{\varepsilon}} = T^{-1}\sum_{t\,1}^T \widehat{\varepsilon}_t$.
3.   *Generate* $\{y_t^\star\}_{t\,1}^T$ *from* $\widehat{\Phi}_p(L)(1 - L)y_t^\star = \widehat{\mu} + \widehat{\Theta}_q(L)\varepsilon_t^\star$.
4.   *From model* (8), *with bootstrap sample* $\{y_t^\star\}_{t\,1}^T$, *obtain* $\{\widehat{\varepsilon}_t^\star\}_{t\,1}^T$ *by LS.*
5.   *From model* (9), *with* $\{\widehat{\varepsilon}_t^\star\}_{t\,1}^T$, *estimate* $\alpha^\star(r)$ *by LS and compute* $\overline{t}_\alpha^l = \sup_{r\in[\underline{r},\overline{r}]}|t_{\alpha^\star(r)}|$.
6. $l = l + 1$. *Go to step 2 while* $l \leqslant B$.
7.   *Estimate the p-value, p_v, from the bootstrap approximation,*

$$p_v = \frac{1}{B}\sum_{l\,1}^B 1\left(\overline{t}_\alpha^l > \sup_{r\in[\underline{r},\overline{r}]}|t_{\alpha(r)}|\right).$$

If the null of linearity is rejected, then we proceed to test the existence of transitory shocks by constructing the t-statistic of $\theta_2 = 1$ in model (2). According to Result 3, this statistic follows a standard normal distribution.

4. Permanent–transitory decomposition

For both, theoretical and empirical reasons, it is frequently desirable to decompose an economic time series into the sum of unobservable permanent and temporary components that generate the series. Most of the literature has focused on decomposing linear models, see Watson (1986) and more recently Mortely et al. (2001). As the latter authors note, this implicitly imposes the restriction that the transitory component (related to business cycle) is symmetric, whereas recent research suggests the opposite (see for example, Neftci, 1984; Hamilton, 1989; Beaudry and Koop, 1993). In this section, we propose two new nonlinear permanent and transitory (*P–T* hereafter) decompositions based on the ARTIMA Shock-Size model.

First we present a modified version of Quah (1992)'s definition of a *P–T* decomposition.

Definition 1. Let Y be an integrated sequence. A permanent–transitory ($P–T$) decomposition for Y is a pair (Y^P, Y^T) such that:

(i) Y^P is integrated and Y^T is covariance stationary;
(ii) Y^P does not have any transitory shock;

(iii) $Var(\triangle Y^P)$ and $Var(\triangle Y^T)$ are strictly positive, with $\triangle = (1 - L)$;
(iv) $Y_t = Y_t^P + Y_t^T$.

Further, if

(v) $\triangle Y^P$ is uncorrelated with Y^T at all leads and lags,

then the P–T decomposition is said to be orthogonal.

The new condition (ii) guarantees that the permanent component is formed only by permanent shocks. It is worth mentioning that the Beveridge and Nelson decomposition does not satisfy condition (v), because its permanent and transitory shocks are perfectly correlated.

Given the following ARTIMA Shock-Size model,

$$\Phi(L)(1 - L)y_t = \mu + (1 - \theta_1 L)\varepsilon_t^L + (1 - \theta_2 L)\varepsilon_t^s, \tag{10}$$

with $\theta_2 = 1$, the first new P–T decomposition proposed is

$$y_t = y_t^P + y_t^T,$$
$$y_t^P = \Phi(1)^{-1}\mu + y_{t-1}^P + \Phi(L)^{-1}(1 - \theta_1 L)\varepsilon_t^L, \tag{11}$$

$$y_t^T = \Phi(L)^{-1}\varepsilon_t^s, \tag{12}$$

where the permanent component is formed only by permanent shocks ($\varepsilon_t^L = \varepsilon_t 1(|\varepsilon_t| > r)$), and the transitory component, only by transitory shocks ($\varepsilon_t^s = \varepsilon_t 1(|\varepsilon_t| \leqslant r)$). It is obvious that this nonlinear decomposition is an orthogonal decomposition according to Definition 1. An important characteristic of this new P–T decomposition is that its existence can be tested, by testing $\theta_2 = 1$. Besides, it allows the possibility of a richer dynamic structure in the permanent component than the one produced by a pure random walk.

There are situations where economic theory suggests the permanent component be a random walk (see the application to stock prices). For these cases we propose to apply B–N directly to our ARTIMA Shock-Size model. The method developed by Beveridge and Nelson (1981) consists on defining the permanent component as the long-horizon level forecast of the original series, or the part that remains after all transitory dynamics have worked themselves out. More precisely,

$$y_t = y_t^P + y_t^T,$$
$$y_t^P = \lim_{k \to \infty} [E(y_{t+k}|y_t, y_{t-1}, \ldots) - kE(y_t - y_{t-1})],$$
$$y_t^T = y_t - y_t^P.$$

To apply this to the ARTIMA Shock-Size model (10), we first rewrite it as

$$y_t = \Phi(1)^{-1}\mu + y_{t-1} + C_L(L)\varepsilon_t^L + C_s(L)\varepsilon_t^s,$$

where $C_L(L) = \Phi^{-1}(L)(1 - \theta_1 L)$ and $C_s(L) = \Phi^{-1}(L)(1 - \theta_2 L)$. Second, we use the well-known algebraic result (see Phillips and Solo, 1992) by which a lag polynomial

$C(L)$ can be written as $C(L) = C(1) + (1 - L)\widetilde{C}(L)$. Then the corresponding B–N decomposition can easily be obtained as follows:

$$y_t^P = \Phi(1)^{-1}\mu + y_{t-1}^P + C_L(1)\varepsilon_t^L + C_s(1)\varepsilon_t^s, \tag{13}$$

$$y_t^T = \widetilde{C}_L(L)\varepsilon_t^L + \widetilde{C}_s(L)\varepsilon_t^s. \tag{14}$$

This decomposition constitutes our second proposal of a nonlinear P–T decomposition. For the simple TIMA model (2),

$$y_t^P = \mu + y_{t-1}^P + (1 - \theta_1)\varepsilon_t^L + (1 - \theta_2)\varepsilon_t^s,$$
$$y_t^T = -\theta_1\varepsilon_t^L - \theta_2\varepsilon_t^s.$$

When the small shocks are transitory ($\theta_2 = 1$) the nonlinear B–N becomes

$$y_t^P = \mu + y_{t-1}^P + (1 - \theta_1)\varepsilon_t^L, \tag{15}$$

$$y_t^T = -\theta_1\varepsilon_t^L - \varepsilon_t^s. \tag{16}$$

The big advantage of this nonlinear B–N decomposition, as the following result shows, does not come from the permanent component but from the transitory one. The above examples show that neither the shocks of the permanent and transitory components are perfectly correlated nor that all the shocks at time t are permanent. In this way we overcome the two main drawbacks of the linear B–N decomposition (more examples of nonlinear B–N decompositions can be found in Clarida and Taylor, 2003).

The next result establishes the relationship between the B–N and UC–0 decompositions applied to the Wold (1938) representation of a TIMA Shock-Size process, and the nonlinear B–N decomposition (15)–(16). Note that, in order for UC–0 to exist, the parameter θ_1 must satisfy $0 \leqslant \theta_1 \leqslant 1$.

Result 5. *Let y_t be the* TIMA *model* (2), *then*

$$V^{BN}(\Delta y_t^P) = V^{\text{TIMA}}(\Delta y_t^P) = V^{UC\ 0}(\Delta y_t^P),$$
$$V^{BN}(y_t^T) \leqslant V^{\text{TIMA}}(y_t^T) \leqslant V^{UC\ 0}(y_t^T),$$

where V^{BN} and $V^{UC\ 0}$ are the variances of the components of the B–N and UC–0 decompositions applied to the Wold's representation of $(1 - L)y_t$, and V^{TIMA} is the variance of the components forming the nonlinear B–N decomposition of the TIMA *model. Strict inequality is obtained if $0 < \theta_1 < 1$.*

As it is well known, (see Quah, 1992), the variance of the innovations of the random walk component is always the same (equal to the height of the spectral density of $(1 - L)y_t$ at frequency zero), independently of the used decomposition. However, for $0 < \theta_1 < 1$, the variance of the transitory component of each decomposition is different. In particular, the variance of the transitory component of the nonlinear B–N decomposition for the TIMA model (2) lies always between the corresponding variances of the UC–0 and the linear B–N decompositions. This fact has important consequences for the applications analyzed in the next two sections.
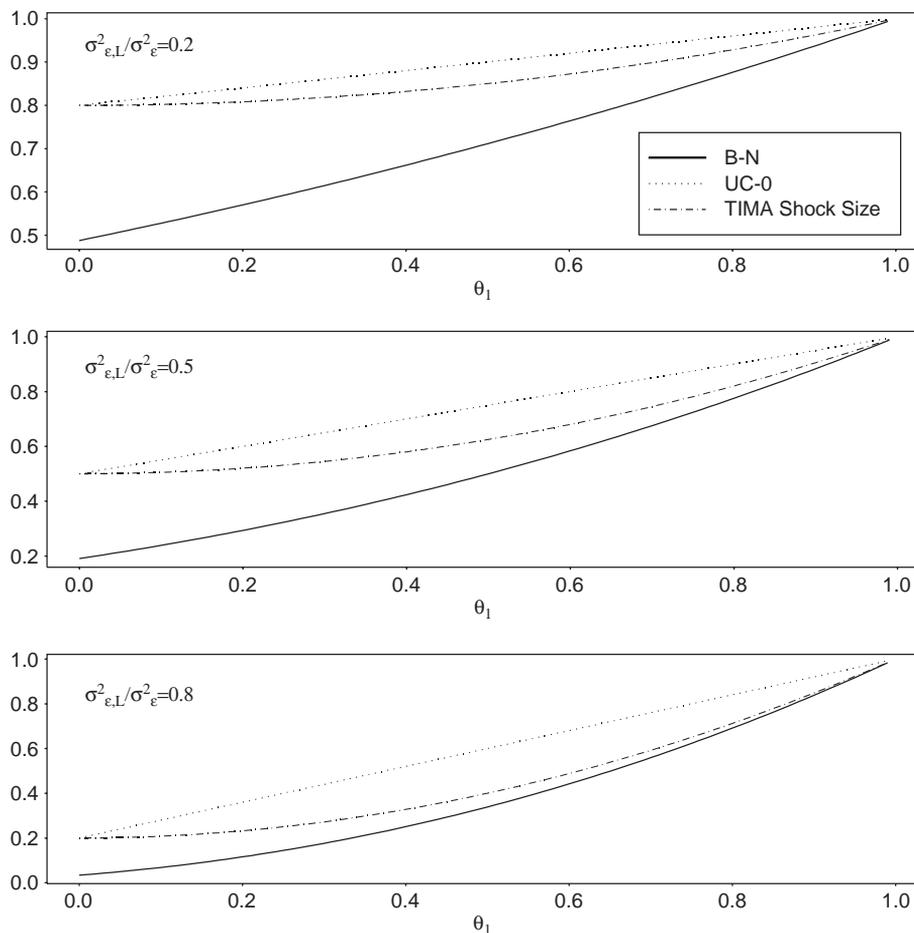
Fig. 3. Variance of the transitory component of different P T decompositions of a TIMA Shock-Size Process. $\sigma^2_{\bar{\varepsilon},L} = \mathrm{E}(\varepsilon^2_t 1(|\varepsilon_t| > r))$ and $\sigma^2_{\bar{\varepsilon}_t} = \mathrm{E}(\varepsilon^2_t)$. $\theta_2 = 1$ and $\varepsilon_t \equiv \mathrm{N}(0, 1)$.

In Fig. 3 we illustrate graphically the inequalities of Result 5 for $\theta_2 = 1$, and different values of $\theta_1 \in [0, 1]$ and of the ratio $\mathrm{E}(\varepsilon^2_t 1(|\varepsilon_t| > r))/\mathrm{E}(\varepsilon^2_t)$.

5. Application to stock prices: a measure of the stock market quality

In this section we apply ARTIMA Shock-Size models to measure the deviations between actual transaction and implicit efficient prices, following the methodology developed by Hasbrouck (1993). Measurement of this difference arises in financial market analysis in two important contexts. First, with the purpose of evaluating the

broker's performance, and second, for a comparative analysis of the market regulatory structures. Hasbrouck's approach is based on decomposing the transaction price p_t, into two components, the efficient price m_t, and the pricing error s_t, and on measuring the relative size of s_t. In more detail Hasbrouck's model takes the logarithm of the observed transaction price at time t, as the sum of two components:

$$p_t = m_t + s_t.$$

The element m_t (efficient price) is assumed to follow a random walk,

$$m_t = m_{t-1} + \varepsilon_t,$$

and s_t (pricing error) is a zero-mean covariance-stationary process that is modelled as

$$s_t = \beta\varepsilon_t + \eta_t,$$

where η_t is uncorrelated with ε_t. Given that the permanent component is identified with a random walk, the correct P–T decomposition depends only on the assumption about the correlation structure of both components. If the shocks of both components are perfectly correlated ($\eta_t = 0$), we are in the B–N scenario. If both shocks are uncorrelated ($\beta = 0$), UC–0 is the decomposition to be used. Unfortunately, the two identification restrictions are obviously not testable, and as it is shown in the previous section they imply different dispersion of the pricing error. Besides, economic theory does not argue conclusively in favor of any of the identification restrictions.

Roll (1984) bid-ask spread model corresponds to $\beta = 0$ and $\eta_t = \pm(spread)/2$. Roll proves that the bid-ask spread generates a first order autocorrelation different from 0 in the increments of the transaction prices, and this correlation explains the existence of the pricing error. The perfect correlation restriction is a particular case of Glosten (1987) model. In this model, the spread is partially due to the asymmetric information revealed in the trade. With no nontrade public information, it is obtained $\beta > 0$ and $\eta_t = 0$.

Starting from these two models, we develop a more complete model which takes into account several issues not considered in any of the previous ones. We assume the market description of Glosten (1987), that is, two kinds of agents, uninformed and informed, with a market maker. The first aspect to be considered is that when there is a transaction cost, not all the new information will be translated into the transaction prices. Only the new information which implies a profit greater than the transaction cost will be translated into the transaction price. In other words, the shocks that drive the efficient price component must be ''big'' shocks to the transaction price. The second aspect is that the transactions of the uninformed agents cannot generate big inefficient changes in the transaction prices, because the informed traders will arbitrate the situation. This implies that the shocks driving the pricing error component must be "small" shocks to the transaction price. Based on these two aspects and taking Glosten model as starting point we propose the following

pricing specification

$$p_t = m_t + s_t,$$
$$m_t = m_{t-1} + 1(|\varepsilon_t| > r)\varepsilon_t,$$
$$s_t = 1(|\varepsilon_t| \leqslant r)\varepsilon_t.$$

The parameter r is identified as the spread or transaction cost, and only the shocks greater than this cost will affect the efficient price m_t. From this specification the following representation for the transaction prices is obtained

$$(1 - L)p_t = \begin{cases} \varepsilon_t & \text{if} \quad |\varepsilon_t| > r, \\ \varepsilon_t - \varepsilon_{t-1} & \text{if} \quad |\varepsilon_t| \leqslant r. \end{cases} \tag{17}$$

Model (17) is a TIMA Shock-Size model. At this point, it is worth mentioning again that although we only can observe one shock, ε_t, in the structural model there are two type of shocks. One related to new information and therefore driving the efficient price process, and another coming from the uninformed traders and therefore of transitory nature. Using the reasoning described above, these two types of structural shocks translate into the transaction price process in terms of "big" and "small" shocks. Within this TIMA model we are able to test the existence of these two type of structural shocks.

Fig. 4 represents the logarithm of the S&P500 daily series. Table 1 shows the Least Squares estimates of the linear ARIMA and nonlinear ARTIMA model fitted to S&P500 daily series. Parameter estimates of the TIMA Shock-Size model were obtained by numerically minimizing the squared estimated error using the FORTRAM optimization procedure DBCPOL. The moving average parameter estimates of the TIMA model are clearly different. Following the testing strategy described in Section 3.1, we test the validity of this model in two steps. In the first one, the null hypothesis of linearity is tested against the alternative of threshold
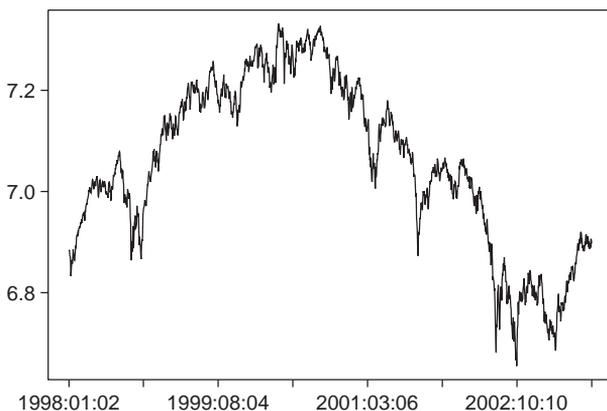
Fig. 4. Logarithm of S&P500, 2 January 1998 29 July 2003.

Table 1
Least squares estimates of ARIMA(0,1,1) and ARTIMA(0,1,1,1) models for stock prices

Parameters	θ	θ_1	θ_2	r	σ_ε	AIC
ARIMA$(0,1,1)$	0.0176 (0.0267)				0.0136	5.7521
ARTIMA(0,1,1,1)		0.0131 (0.0268)	0.9717 (0.3632)	0.0035	0.0135	5.7564

Note: In brackets are the corresponding standard errors. The ARTIMA parameter estimates correspond to the second step procedure described in Section 3.

Table 2
Testing linearity versus threshold Shock-Size effect

Auxiliary regression
$\widehat{\varepsilon}_t = \alpha_0 + \alpha_1\widehat{\varepsilon}_{t-1} + \alpha_2\widehat{\varepsilon}_{t-1}1(|\widehat{\varepsilon}_{t-1}|\leqslant r)$

$\sup_r |t_{\alpha_2}|$	p-value	$\widehat{\alpha}_0(\bar{r})$	$\widehat{\alpha}_1(\bar{r})$	$\widehat{\alpha}_2(\bar{r})$	$\bar{r}$
2.213	0.067	0.000 (0.0003)	0.004 (0.0268)	0.806 (0.3645)	0.0035

Note: $\widehat{\varepsilon}_t$ are the residuals of the null linear model. Bootstrap p-value is obtained from 1000 replications. In brackets are the corresponding standard errors.

Table 3
Variance decomposition of different $P\ T$ decompositions for stock prices

	$B\ N$ ARIMA(0,1,1)	$B\ N$ ARTIMA(0,1,1,1)	$UC\ 0$ ARIMA(0,1,1)
$\sigma_{\Delta m_t} \times 10^2$	1.338	1.370	1.339
$\sigma_{s_t} \times 10^2$	0.024	0.074	0.176

Note: $p_t = m_t + s_t$, with m_t a random walk and s_t the transitory component. $\sigma_{\Delta m_t}$ and σ_{s_t} are the standard deviation of Δm_t and s_t, respectively.

effect. If we reject this null, in a second step we test the null hypothesis of a unit root in the MA regime driven by small shocks.

To test linearity, we fit an ARIMA(0,1,1) to the logarithm of the series, construct the following auxiliary regression with the estimated residuals,

$$\widehat{\varepsilon}_t = \alpha_0 + \alpha_1\widehat{\varepsilon}_{t-1} + \alpha_2\widehat{\varepsilon}_{t-1}1(|\widehat{\varepsilon}_{t-1}|\leqslant r) + u_t,$$

and test the significance of $\alpha_2(r)$. Given that under the null hypothesis the parameter r is unidentified, the test proposed is $\sup_r|t_{\alpha_2(r)}|$, where $t_{\alpha_2(r)}$ is the *t-ratio* for the null of $\alpha_2(r) = 0$. The results of this test are shown in Table 2.

Linearity is rejected at a 10% significant level (*p-value* = 0.067). From Table 1 the hypothesis of a unit root in the MA side corresponding to the regime of small shocks cannot be rejected at the usual significance levels. With these two results we conclude

that the analyzed $S\&P500$ data do not reject the proposed TIMA Shock-Size model. Therefore in order to measure the size of the pricing error component s_t, we obtain the nonlinear B–N decomposition of the estimated TIMA model. To do that we impose $\theta_2 = 1$. The results are in Table 3. Notice that, as the theory suggests, the contribution of the random walk component is practically the same in all the P–T decompositions, and the sizes of the transitory components satisfy the inequality established in Result 5.

Besides this measure of the quality of the stock market, the TIMA Shock-Size model provides some extra information. For example, the estimated value of the threshold parameter $\hat{r} = 0.00283$ (imposing $\theta_2 = 1$) implies that the estimated transaction cost is 0.283 percent of the stock price. This figure is very similar to that obtained by Hasbrouck (1993) for intra-day data. Also, the percentage of pure transitory shocks can be estimated (23.65%). The periods where these shocks occurred can be identified in this context with periods of non-outstanding information.

6. Application to GNP: asymmetric persistence of shocks to output

The goal of this section is to answer empirically an important question about output fluctuations: Is the persistence of shocks to output asymmetric? This question has been analyzed by Beaudry and Koop (1993), Hess and Iwata (1997) and Elwood (1998) among others. All these authors deal with asymmetries produced by the sign of the shocks, and there is not a clear conclusion. The first authors give an affirmative answer about the existence of asymmetries, while the others contradict this result. We will use our TIMA Shock-Size model to conclude that shocks with different sizes have different long run effects on output. Small shocks will be transitory, while large shocks will be permanent.

In this section, indirectly we will also try to answer the question of which type of shocks are responsible for economic fluctuations. This question has been analyzed in detail by Blanchard and Quah (1989), Cochrane (1994), and many other authors. Cochrane (1994) remarks two key issues: (i) shock identification and (ii) endogenous shocks. About the latter, he demands new theoretical models, probably nonlinear, that are able to explain the fluctuation of the business cycle in terms of endogenous shocks. TIMA Shock-Size models are perfect candidates to handle both issues simultaneously.

Fig. 5 represents the logarithm of quarterly real GNP for the period 1947:01-2003:03. Table 4 shows the least square estimates of several linear ARIMA models and a nonlinear ARTIMA model fitted to GNP. Following Campbell and Mankiw (1987), we have considered all ARIMA models with p and q less than 3. For space considerations, Table 4 only presents the results for the ARIMA(1,1,0), ARIMA(0,1,1) and ARIMA(1,1,1). The ARTIMA Shock-Size fitted model is

$$(1 - \phi L)(1 - L)y_t = \mu + (1 - \theta_1 L)\varepsilon_t^L + (1 - \theta_2 L)\varepsilon_t^s.$$

Table 4 shows that all linear models fit the data very similarly. According to the AIC criteria the best linear model is the ARIMA(1,1,0), however to eliminate a few correlations in the residuals, we select the ARIMA(1,1,1) model for our testing
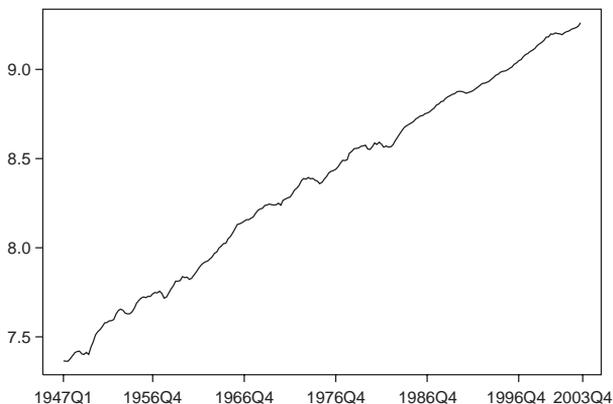
Fig. 5. Log of real GNP, 1947:01-2003:03.

Table 4
Least square estimates of ARIMA and ARTIMA models for real GNP

Parameter	ARIMA(1,1,0)	ARIMA(0,1,1)	ARIMA(1,1,1)	ARTIMA(1,1,1,1)
ϕ	0.335 (0.062)		0.453 (0.171)	0.647 (0.182)
θ		0.260 (0.064)	0.131 (0.191)	
θ_1				0.320 (0.196)
θ_2				1.210 (0.396)
$r \times 10^2$				0.477
$\sigma_\varepsilon \times 10^2$	0.956	0.969	0.950	0.941
AIC	6.452	6.424	6.447	6.475

Note: In brackets are the corresponding standard errors. The ARTIMA parameter estimates correspond to the second step procedure described in Section 3.

strategy. This testing strategy is the one suggested in Section 3.1 and used in the previous application. To test the null of linearity versus the alternative of a threshold effect, we construct the following auxiliary regression with the residuals of the linear ARIMA (1,1,1)

$$\widehat{\varepsilon}_t = \alpha_0 + \alpha_1 \widehat{\varepsilon}_{t-1} + \alpha_2 \widehat{\varepsilon}_{t-1} 1(|\widehat{\varepsilon}_{t-1}| \leqslant r) + u_t.$$

In this regression we test the significance of $\alpha_2(r)$ with the $\sup_r |t_{\alpha_2(r)}|$, where $t_{\alpha_2(r)}$ is the *t-ratio* for the null of $\alpha_2(r) = 0$. The results of this test are shown in Table 5. Clearly the null of linearity is rejected with a *p*-value smaller than 0.05. This implies that the size of the shocks is able to generate asymmetries in the persistence of shocks to *GNP*.

In order to test whether or not small shocks are transitory, we test $\theta_2 = 1$. From Result 3, this is done directly from last column of Table 4, concluding that the null of a unit root in the MA side cannot be rejected. Therefore the *GNP* data do not reject that large shocks are persistent while small shocks are only transitory.

Imposing $\theta_2 = 1$, we calculate the nonlinear orthogonal P–T decomposition ((11)–(12)) proposed in Section 4. This becomes

$$(1 - L)y_t^P = \frac{\mu}{1 - \phi} + \frac{(1 - \theta_1 L)}{(1 - \phi L)} \, \varepsilon_t^L,$$

$$y_t^T = \frac{1}{(1 - \phi L)} \, \varepsilon_t^s,$$

Table 5
Testing linearity versus threshold Shock-Size effect

Auxiliary regression
$\widehat{\varepsilon}_t = \alpha_0 + \alpha_1 \widehat{\varepsilon}_{t-1} + \alpha_2 \widehat{\varepsilon}_{t-1} 1(|\widehat{\varepsilon}_{t-1}| \leqslant r)$

$\sup_r |t_{\alpha_2}|$	p-value	$\widehat{\alpha}_0(\bar{r})$	$\widehat{\alpha}_1(\bar{r})$	$\widehat{\alpha}_2(\bar{r})$	$\bar{r}$
2.362	0.045	0.000 (0.0006)	0.021 (0.0676)	0.789 (0.3363)	0.004

Note: $\widehat{\varepsilon}_t$ are the residuals of the null linear model. Bootstrap p-value is obtained from 1000 replications. In brackets are the corresponding standard errors.
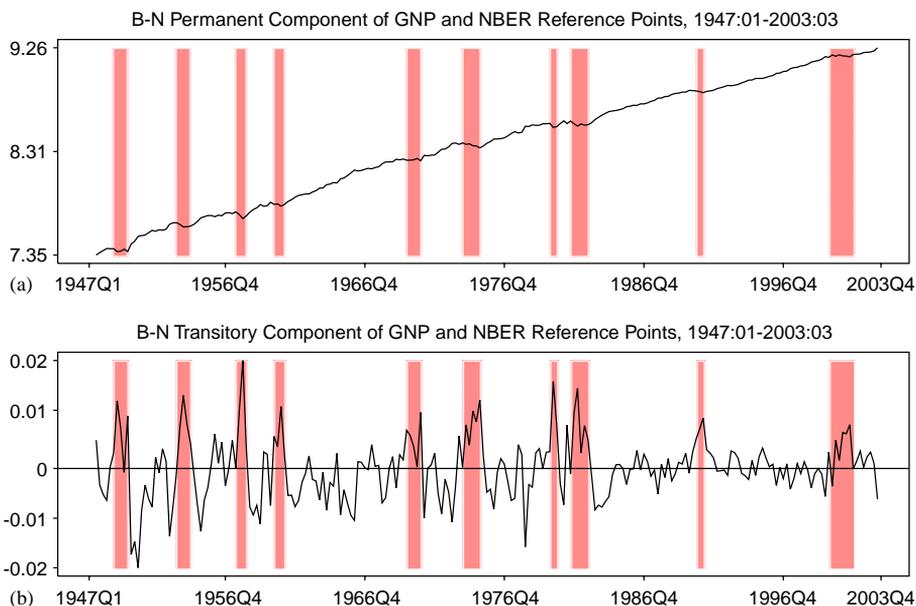
Fig. 6. *B*–*N* decomposition of real *GNP* based on an ARIMA(1,1,1): (a) *B*–*N* Permanent component of GNP and NBER reference points, 1947:01-2003:03; (b) *B*–*N* Transitory Component of GNP and NBER Reference Points, 1947:01-2003:03.
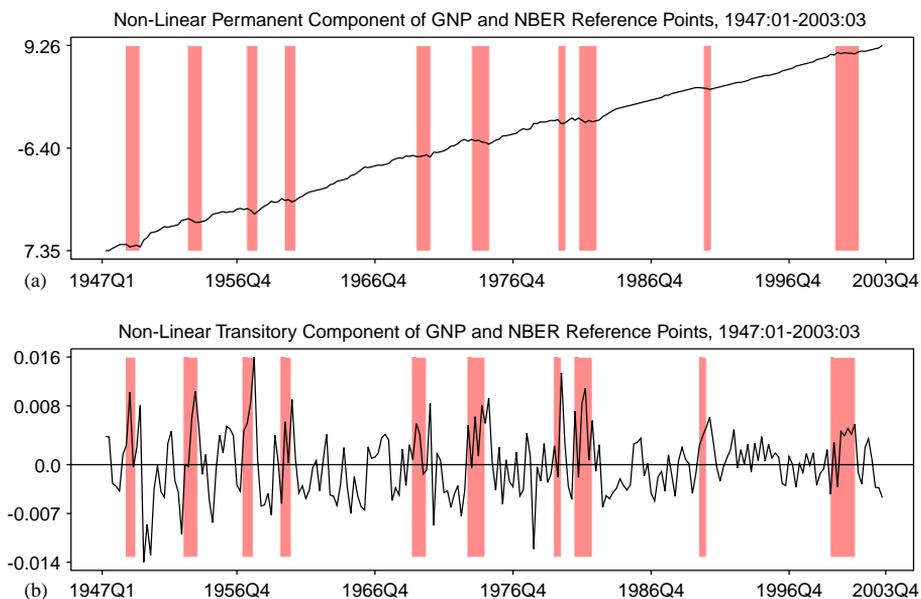
Fig. 7. *B N* decomposition of real GNP based on an ARTIMA(1,1,1,1): (a) Nonlinear permanent component of GNP and NBER reference points, 1947:01-2003:03; (b) Nonlinear transitory component of GNP and NBER reference points, 1947:01-2003:03.

where ε_t^s corresponds to shocks smaller than $\hat{r} = 0.0046$. This implies that shocks smaller than approximately $\sigma_\varepsilon/2$ are pure transitory. They amount to 51% of the *GNP* shocks (the remain 49% are permanent).

Figs. 6–8 represent graphically the *P–T* decompositions discussed in Section 4, together with the NBER reference points. Numerically this comparison is summarized in Table 6. Both linear and nonlinear *B–N* decompositions produce very similar results (the differences are due to estimation issues). However, the orthogonal nonlinear *P–T* decomposition is clearly different. First, the transitory component is less volatile, and second, the permanent component is smoother.

To finish, it is worthwhile to mention the close link between the large shocks from the ARTIMA (1,1,1,1) model and the NBER turning points. From Fig. 9 it can be seen that peaks and troughs are related to large shocks of opposite signs. Situations where output leaves a trough are generated by a positive large shock, and the opposite occurs when the output abandons a peak.

7. Conclusion

In this paper we introduce a new class of simple nonlinear models (TIMA) with the aim of being able to identify the shocks of a dynamic system. This identification

Non-Linear Permanent Component of GNP and NBER Reference Points, 1947:01-2003:03

Non-Linear Transitory Component of GNP and NBER Reference Points, 1947:01-2003:03
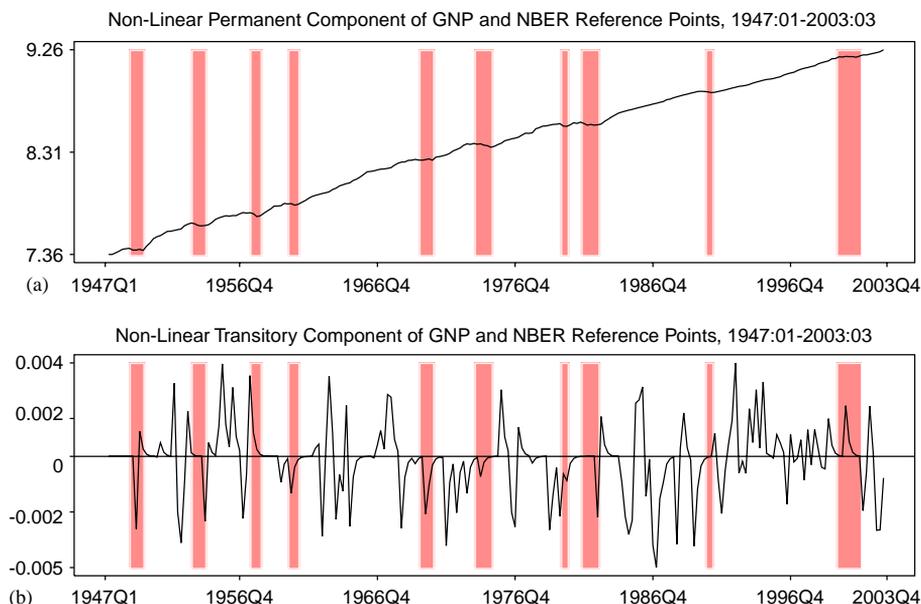
Fig. 8. Orthogonal decomposition of real GNP based on an ARTIMA(1,1,1,1): (a) Nonlinear permanent component of GNP and NBER reference points, 1947:01-2003:03; (b) Nonlinear transitory component of *GNP* and NBER reference points, 1947:01-2003:03.

Table 6
Variance decomposition of different *P T* decomposition for real GNP

	$\sigma^2_{\Delta y_t^P} \times 10^3$	$\sigma^2_{y_t^T} \times 10^3$
B N ARIMA(1,1,1)	0.2288	0.0398
B N ARTIMA(1,1,1,1)	0.1825	0.0229
Orthogonal ARTIMA(1,1,1,1)	0.0982	0.0032

Note: $y_t = y_t^P + y_t^T$. In the *B N* decomposition y_t^P is a random walk. In the orthogonal ARTIMA decomposition y_t^P is formed only by permanent shocks, s.t $E(\Delta y_t^P, y_t^T) = 0$ (see (11) and (12)). $\sigma^2_{\Delta y_t^P}$ and $\sigma^2_{y_t^T}$ are the variance of Δy_t^P and y_t^T, respectively.

is based on the long run effects of the shocks (permanent or transitory) and on their sizes (Large or small). The latter is a special case rather than a necessity. It could be any other shock or economic characteristic. TIMA models provide a nice framework to test if a shock with a given characteristic (in this paper its size) has a permanent or a transitory effect. Once shocks are identified we construct two new nonlinear permanent–transitory decomposition, that are compared with the standard linear *B*–*N* and *UC*–0 permanent–transitory decompositions available in the literature.
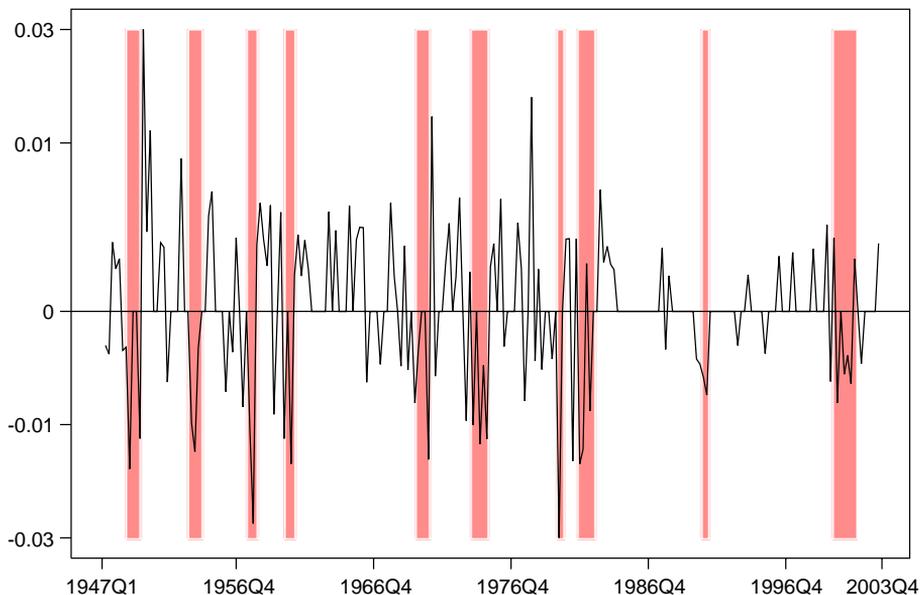
Fig. 9. Large shocks and the NBER reference points, 1947:01-2003:03.

Two applications of these new nonlinear P–T decompositions are shown in the paper. The first one studies the quality of the stock market by measuring the importance of the transitory component (pricing error) in the total transaction price. The second one studies the asymmetries in the persistence of the GNP shocks. While the literature has been inconclusive to this respect when the asymmetry comes from the shock sign, we find a clearly affirmative response when the asymmetry is generated by the shock size.

Extensions to a multivariate framework is under current research by the authors.

Acknowledgements

Appendix

Proof of Result 1. This result is easily proved by rewriting model (2) at time $t + k$ as

$$y_{t+k} = \frac{1}{(1-L)}\mu + \frac{1-\theta_1 L}{(1-L)}\varepsilon_{t+k-1}1(|\varepsilon_{t+k-1}| > r)$$
$$+ \frac{1-\theta_2 L}{(1-L)}\varepsilon_{t+k-1}1(|\varepsilon_{t+k-1}| \leqslant r). \tag{18}$$

Conditioning on w_{t-1}, the past history of the process (included ε_0), and using (18), the *GIRF* has the following expression:

$$GIRF(k, \varepsilon_t, w_{t-1}) = (1-\theta_1)\{\varepsilon_t 1(|\varepsilon_t| > r) - \mathrm{E}[\varepsilon_t 1(|\varepsilon_t| > r)|w_{t-1}]\}$$
$$+ (1-\theta_2)\{\varepsilon_t 1(|\varepsilon_t| \leqslant r) - \mathrm{E}[\varepsilon_t 1(|\varepsilon_t| \leqslant r)|w_{t-1}]\}$$
$$+ \sum_{i=0}^{k-2}(1-\theta_1)\{\mathrm{E}[\varepsilon_{t+k-1-i} 1(|\varepsilon_{t+k-1-i}| > r)|\varepsilon_t, w_{t-1}]$$
$$- \mathrm{E}[\varepsilon_{t+k-1-i} 1(|\varepsilon_{t+k-1-i}| > r)|w_{t-1}]\}$$
$$+ \sum_{i=0}^{k-2}(1-\theta_2)\{\mathrm{E}[\varepsilon_{t+k-1-i} 1(|\varepsilon_{t+k-1-i}| \leqslant r)|\varepsilon_t, w_{t-1}]$$
$$- \mathrm{E}[\varepsilon_{t+k-1-i} 1(|\varepsilon_{t+k-1-i}| \leqslant r)|w_{t-1}]\}.$$

Under Assumption A.4, we obtain that

$$\mathrm{E}[\varepsilon_{t+j} 1(|\varepsilon_{t+j}| > r)|w_{t-1}] = \mathrm{E}[\varepsilon_{t+j+1} 1(|\varepsilon_{t+j+1}| > r)|\varepsilon_t, w_{t-1}] = 0,$$

for $j \geqslant 0$. Hence,

$$GIRF(k, \varepsilon_t, w_{t-1}) = (1-\theta_t)\varepsilon_t = \begin{cases} (1-\theta_1)\varepsilon_t & \text{if} \quad \theta_t = \theta_1 \\ (1-\theta_2)\varepsilon_t & \text{if} \quad \theta_t = \theta_2. \end{cases} \qquad \square$$

Proofs of Results 2 and 3. They can be found in Gonzalo and Martínez (2004) and are available upon request. $\square$

Proof of Result 4. For simplicity and without loss of generality, this result is proved for the case of $p = 1$, $q = 1$ and $\mu = 0$. Part of the proof makes use of some of the results in Bai (1994).

Let us write the ARMA(1,1) representation of $x_t = (1-L)y_t$ as $\varepsilon_t = x_t - \phi_0 x_{t-1} - \theta_0 \varepsilon_{t-1}$ and the corresponding residuals as $\widehat{\varepsilon}_t = x_t - \widehat{\phi} x_{t-1} - \widehat{\theta}\widehat{\varepsilon}_{t-1}$, where $(\widehat{\phi}, \widehat{\theta})$ are the LS estimators of the parameters (ϕ_0, θ_0). These estimators satisfy that $T^{1/2}(\widehat{\phi} - \phi_0)$ and $T^{1/2}(\widehat{\theta} - \theta_0)$ are $O_p(1)$.

Subtracting the errors from the residuals, making use of $\widehat{\varepsilon}_0 = 0$, and by repeated substitution we obtain

$$\widehat{\varepsilon}_t - \varepsilon_t = (-1)^{t-1}\widehat{\theta}^t \varepsilon_0 - (\widehat{\phi} - \phi_0) \sum_{j\,0}^{t-1} (-1)^j \widehat{\theta}^j x_{t-1-j} - (\widehat{\theta} - \theta_0) \sum_{j\,0}^{t-1} (-1)^j \widehat{\theta}^j \varepsilon_{t-1-j}.$$

(19)

Denote $\hat{\rho} = (-\widehat{\theta}, \sqrt{T}(\widehat{\phi} - \phi_0), \sqrt{T}(\widehat{\theta} - \theta_0))$ and $\rho = (u, v, w) \in R^3$. Define

$$\Lambda_{\rho,t} = u^t \varepsilon_0 + T^{-1/2}\left(v \sum_{j\,0}^{t-1} u^j x_{t-1-j} + w \sum_{j\,0}^{t-1} u^j \varepsilon_{t-1-j}\right) = u^t \varepsilon_0 + T^{-1/2} \xi_{\rho t}.$$

(20)

From (9) and the definition of $t_{\alpha(r)}$

$$t_{\alpha(r)} = \frac{\widehat{\alpha}(r)}{\widehat{\sigma}_\alpha} = T^{-1/2} \sum_{t\,1}^{T} \widehat{\varepsilon}_t \widehat{\varepsilon}_{t-1}^s(r) \left[T^{-1} \sum_{t\,1}^{T} (\widehat{\varepsilon}_{t-1}^s(r))^2 T^{-1} \sum_{t\,1}^{T} \widehat{u}_t^2(r)\right]^{-1/2},$$

(21)

where $\widehat{\varepsilon}_t^s(r) = \widehat{\varepsilon}_t 1(|\widehat{\varepsilon}_t| \leqslant r)$ and $\widehat{u}_t(r) = \widehat{\varepsilon}_t - \widehat{\alpha}(r)\widehat{\varepsilon}_{t-1}^s(r)$.

The proof basically consists on decomposing the RHS term of (21) in several pieces, and analyzing the asymptotic behavior of these pieces. For the first term, $T^{-1/2}\sum_{t\,1}^{T} \widehat{\varepsilon}_t \widehat{\varepsilon}_{t-1}^s(r)$, we prove that,

$$T^{-1/2} \sum_{t\,1}^{T} \widehat{\varepsilon}_t \widehat{\varepsilon}_{t-1}^s(r) = X_{1,T}(r) + X_{2,T} X_{3,T}(r) + o_p(1),$$

where the $o_p(1)$ is uniform for all $r \in [\underline{r}, \overline{r}]$ and

$$X_{1,T}(r) \Rightarrow D_1(r),$$

$$X_{2,T} \xrightarrow{d} (V_\phi + V_\theta),$$

$$X_{3,T}(r) \xrightarrow{p} \Sigma_\varepsilon(r) \text{ uniformly in } r,$$

with V_ϕ and V_θ normal random variables with distribution, the asymptotic distribution of $T^{1/2}(\widehat{\phi} - \phi_0)$ and $T^{1/2}(\widehat{\theta} - \theta_0)$, respectively. For the other term in (21), we prove that

$$T^{-1} \sum_{t\,1}^{T} (\widehat{\varepsilon}_{t-1}^s(r))^2 T^{-1} \sum_{t\,1}^{T} \widehat{u}_t^2(r) = X_{4,T}(r) + o_p(1),$$

where

$$X_{4,T}(r) \xrightarrow{p} \Sigma_\varepsilon(r)\sigma_\varepsilon^2,$$

uniformly in r.

We start with $T^{-1/2}\sum_{t\,2}^{T}\widehat{\varepsilon}_t\widehat{\varepsilon}_{t-1}^{s}(r)$. Define

$$\Upsilon_{x,t}(\widehat{\theta})\quad\sum_{j\,0}^{t-1}(-1)^j\widehat{\theta}^j x_{t-1-j},$$

$$\Upsilon_{\varepsilon,t}(\widehat{\theta})\quad\sum_{j\,0}^{t-1}(-1)^j\widehat{\theta}^j \varepsilon_{t-1-j}.$$

From (19), $T^{-1/2}\sum_{t\,2}^{T}\widehat{\varepsilon}_t\widehat{\varepsilon}_{t-1}^{s}(r)$ can be written as

$$T^{-1/2}\sum_{t\,1}^{T}\widehat{\varepsilon}_t\widehat{\varepsilon}_{t-1}^{s}(r)\quad T^{-1/2}\sum_{t\,1}^{T}\varepsilon_t\varepsilon_{t-1}1(|\widehat{\varepsilon}_{t-1}|\leqslant r)$$

$$-(\widehat{\phi}-\phi_0)T^{-1/2}\sum_{t\,1}^{T}\varepsilon_t\Upsilon_{x,t-1}(\widehat{\theta})1(|\widehat{\varepsilon}_{t-1}|\leqslant r)$$

$$(\widehat{\theta}\quad\theta_0)T^{-1/2}\sum_{t\,1}^{T}\varepsilon_t\Upsilon_{\varepsilon,t-1}(\widehat{\theta})1(|\widehat{\varepsilon}_{t-1}|\leqslant r)\quad(\widehat{\phi}\quad\phi_0)T^{-1/2}$$

$$\times\sum_{t\,1}^{T}\varepsilon_{t-1}\Upsilon_{x,t}(\widehat{\theta})1(|\widehat{\varepsilon}_{t-1}|\leqslant r)-(\widehat{\theta}-\theta_0)T^{-1/2}$$

$$\times\sum_{t\,1}^{T}\varepsilon_{t-1}\Upsilon_{\varepsilon,t}(\widehat{\theta})1(|\widehat{\varepsilon}_{t-1}|\leqslant r)+R_{1,t}(\widehat{\phi},\widehat{\theta},r). \tag{22}$$

The first term of the RHS of (22) can be written as

$$T^{-1/2}\sum_{t\,1}^{T}\varepsilon_t\varepsilon_{t-1}1(|\widehat{\varepsilon}_{t-1}|\leqslant r)\quad T^{-1/2}\sum_{t\,1}^{T}\varepsilon_t\varepsilon_{t-1}1(|\varepsilon_{t-1}|\leqslant r)+T^{-1/2}$$

$$\times\sum_{t\,1}^{T}\varepsilon_t\varepsilon_{t-1}(1(|\widehat{\varepsilon}_{t-1}|\leqslant r)-1(|\varepsilon_{t-1}|\leqslant r)). \tag{23}$$

The Lemma 3.1 of Koul and Stute (1999) establishes that

$$T^{-1/2}\sum_{t\,1}^{T}\varepsilon_t\varepsilon_{t-1}1(|\varepsilon_{t-1}|\leqslant r)\Rightarrow D_1(r),$$

where $D_1(r)$ is a gaussian process with a covariance function given by

$$M_1(r_1,r_2)\quad E(\varepsilon_t^2)E[\varepsilon_{t-1}^2 1(|\varepsilon_{t-1}|\leqslant\min(r_1,r_2))].$$

For the last term of the RHS of (23) it suffices to prove that

$$\sup_{r\in[-\bar{r},\bar{r}]}\left|T^{-1/2}\sum_{t\,1}^{T}\varepsilon_t\varepsilon_{t-1}(1(\widehat{\varepsilon}_{t-1}<r)-1(\varepsilon_{t-1}<r))\right|\quad o_p(1).$$

Given that $\varepsilon_t\varepsilon_{t-1}\quad\varepsilon_t\varepsilon_{t-1}^{+}+\varepsilon_t\varepsilon_{t-1}^{-}$, where $\varepsilon_t^{+}\quad\varepsilon_t 1(\varepsilon_t>0)$ and $\varepsilon_t^{-}\quad\varepsilon_t 1(\varepsilon_t<0)$, it is enough to prove that

$$\sup_{r\in[-\bar{r},\bar{r}]}\left|T^{-1/2}\sum_{t\,1}^{T}\varepsilon_t\varepsilon_{t-1}^{+}(1(\widehat{\varepsilon}_{t-1}<r)-1(\varepsilon_{t-1}<r))\right|\quad o_p(1), \tag{24}$$

since the case of $\varepsilon_t \varepsilon_{t-1}^-$ is similar. For that, we use the proof of Theorem 1 in Bai (1994). Define

$$\mathsf{F}(r) = \mathrm{E}[\varepsilon_{t-1}^+ 1(\varepsilon_{t-1} < r)],$$

with $\mathsf{F}(-\infty) = 0$ and $\mathsf{F}(\infty) = K_{\mathsf{F}} < \infty$. Then, using (19) and (20)

$$T^{-1/2} \sum_{t\,1}^{T} \varepsilon_t \varepsilon_{t-1}^+ (1(\widehat{\varepsilon}_{t-1} < r) - 1(\varepsilon_{t-1} < r))$$

$$= T^{-1/2} \sum_{t\,1}^{T} \varepsilon_t \varepsilon_{t-1}^+ (1(\varepsilon_{t-1} < r + \varLambda_{\rho,t-1}) - 1(\varepsilon_{t-1} < r))$$

$$= \widehat{K}_T(r) - K_T(r).$$

To study this process, $\widehat{K}_T(r) - K_T(r)$, it suffices to study the auxiliary process,

$$G_T(r, \rho) = T^{-1/2} \sum_{t\,1}^{T} \varepsilon_t \varepsilon_{t-1}^+ (1(\varepsilon_{t-1} < r + \varLambda_{\rho,t-1}) - 1(\varepsilon_{t-1} < r)).$$

Expression (24) is implied by the following result:

$$\sup_{\rho \in D_b} \sup_{r \in [-\bar{r}, \bar{r}]} |G_T(r, \rho)| = \mathrm{o_p}(1) \quad \text{for every } b > 0, \tag{25}$$

where $D_b = [-\bar{\theta}, \bar{\theta}] \times [b, b]^2$ with $0 < \bar{\theta} < 1$ and $b > 0$. Define

$$Z_T(r, \rho) = T^{-1/2} \sum_{t\,1}^{T} \{\varepsilon_t[\varepsilon_{t-1}^+ 1(\varepsilon_{t-1} < r + \varLambda_{\rho,t-1}) - \mathsf{F}(r + \varLambda_{\rho,t-1})]$$

$$- \varepsilon_t[\varepsilon_{t-1}^+ 1(\varepsilon_{t-1} < r) - \mathsf{F}(r)]\},$$

$$\varPi_T(r, \rho) = T^{-1/2} \sum_{t\,1}^{T} \varepsilon_t[\mathsf{F}(r + \varLambda_{\rho,t-1}) - \mathsf{F}(r)].$$

Then, $|G_T(r, \rho)| \leqslant |Z_T(r, \rho)| + |\varPi_T(r, \rho)|$. To show (25), it suffices to prove the following two propositions.

Proposition 5. *Under Assumption A.0, we have*

$$\sup_{\rho \in D_b} \sup_{r \in [-\bar{r}, \bar{r}]} |Z_T(r, \rho)| = \mathrm{o_p}(1) \quad \text{for every } b > 0.$$

Proposition 6. *Under Assumption A.0, we have*

$$\sup_{\rho \in D_b} \sup_{r \in [-\bar{r}, \bar{r}]} |\varPi_T(r, \rho)| = \mathrm{o_p}(1) \quad \text{for every } b > 0.$$

The proof of both propositions follows Bai (1994) and therefore we do not show it in full detail. We will focus only on those details that are different. Given

that $\varepsilon_t = \varepsilon_t^+ + \varepsilon_t^-$, it is enough to prove Proposition 5 for

$$Z_T(r,\rho) = T^{-1/2} \sum_{t\,1}^{T} \{\varepsilon_t^+ [\varepsilon_{t-1}^+ 1(\varepsilon_{t-1} < r + \Lambda_{\rho,t-1}) - \mathsf{F}(r + \Lambda_{\rho,t-1})]$$
$$- \varepsilon_t^+ [\varepsilon_{t-1}^+ 1(\varepsilon_{t-1} < r) - \mathsf{F}(r)]\}.$$

The case ε_t^- is similar. Note that $\varepsilon_t^+ \varepsilon_{t-1}^+ 1(\varepsilon_{t-1} < r + \Lambda_{\rho,t-1})$ and $\varepsilon_t^+ \mathsf{F}(r)$ maintain the non-decreasing monotonicity of the indicator function used in Bai (1994).

Proof of Proposition 5. Define $\eta_t = C \sum_{j\,0}^{t-1} \tau^j (|x_{t-j-1}| + |\varepsilon_{t-j-1}|)$ for some $C > 0$ and $\tau \in (0,1)$, and for every $\lambda \in R$,

$$\widetilde{Z}_T(r,\rho,\lambda) = T^{-1/2} \sum_{t\,1}^{T} \varepsilon_t^+ \{[\varepsilon_{t-1}^+ 1(\varepsilon_{t-1} < r + \Gamma_{t-1}(\rho,\lambda))$$
$$- \mathsf{F}(r + \Gamma_{t-1}(\rho,\lambda))] - [\varepsilon_{t-1}^+ 1(\varepsilon_{t-1} < r) + \mathsf{F}(r)]\},$$

where $\Gamma_t(\rho,\lambda) = u^t \varepsilon_0 + \lambda t \overline{\theta}^{t-1} |\varepsilon_0| + T^{-1/2} \xi_{\rho t} + \lambda T^{-1/2} \eta_t$. Note that $\widetilde{Z}_T(r,\rho,0) = Z_T(r,\rho)$.

Using the compactness of D_b, this set can be partitioned into a finite number of subsets such that the diameter of each one is less than ζ. Denote these subsets by $\triangle_1, \triangle_2, \ldots \triangle_{m(\zeta)}$. Fix k and consider $\triangle_k$ and $\rho_k = (u_k, v_k, w_k) \in \triangle_k$. Proceeding as Bai (1994), choose C large enough and $\tau \in (0,1)$ such that

$$|\Lambda_{\rho,t} - \Lambda_{\rho_k,t}| \leqslant \zeta t \overline{\theta}^{t-1} |\varepsilon_0| + \zeta T^{-1/2} \eta_t \quad \text{for all } \rho \in \triangle_k.$$

With this inequality and the monotonicity of the functions involved in $Z_T(r,\rho)$, it can be proved that

$$Z_T(r,\rho) \leqslant \widetilde{Z}_T(r,\rho_k,\zeta) + T^{-1/2} \sum_{t\,1}^{T} \varepsilon_t^+ [\mathsf{F}(r + \Gamma_{t-1}(\rho_k,\zeta)) - \mathsf{F}(r + \Lambda_{\rho,t-1})],$$

for all $\rho \in \triangle_k$, and the reverse inequality with ζ replaced by $-\zeta$. Now the following lemma is needed.

Lemma 2. *Under Assumption A.0, for every given $\rho = (u, v, w) \in D_b$ and every $\lambda \in R$, we have*

(a) $\quad T^{-1/2} \sum_{t\,1}^{T} \varepsilon_t^+ (|u^t \varepsilon_0| + t \overline{\theta}^{t-1} |\lambda \varepsilon_0|) = o_p(1),$

(b) $\quad T^{-1/2} \max_{1 \leqslant t \leqslant T} \varepsilon_t^+ (|\xi_{\rho t}| + |\lambda \eta_t|) = o_p(1),$

(c) $\quad T^{-1} \sum_{t\,1}^{T} \varepsilon_t^+ (|\xi_{\rho t}| + |\lambda \eta_t|) = O_p(1).$

This lemma can be proved in a similar way as Lemma 1 in Bai (1994).

Applying Lemma 2 to the next inequality, that has been obtained from a mean value theorem,

$$T^{-1/2} \left| \sum_{t-1}^{T} \varepsilon_t^+ [\mathsf{F}(r + \Gamma_{t-1}(\rho_k \pm \zeta)) - \mathsf{F}(r + \Lambda_{\rho,t-1})] \right|$$

$$\leqslant 2\zeta f_e^* T^{-1/2} \sum_{t-1}^{T} \varepsilon_t^+ ((t-1)\overline{\theta}^{t-2}|\varepsilon_0| + T^{-1/2}\eta_{t-1}), \quad \text{for all } r \text{ and } \rho \in D_b,$$

with $f_e^* = \max_e e f_\varepsilon(e) < K_{fe} < \infty$, it is straightforward to show that

$$\sup_{\rho \in D_b} \sup_{r \in [-\overline{r},\overline{r}]} T^{-1/2} \left| \sum_{t-1}^{T} \varepsilon_t^+ [\mathsf{F}(r + \Gamma_{t-1}(\rho_k \pm \zeta)) - \mathsf{F}(r + \Lambda_{\rho,t-1})] \right| = \zeta O_p(1),$$

where the $O_p(1)$ is uniform in r and $\rho \in D_b$. Therefore,

$$\sup_{\rho \in D_b} \sup_{r \in [-\overline{r},\overline{r}]} |Z_T(r,\rho)| \leqslant \max_{k \leqslant m(\zeta)} \sup_{r \in [-\overline{r},\overline{r}]} |\widetilde{Z}_T(r,\rho_k,\zeta)|$$

$$+ \max_{k \leqslant m(\zeta)} \sup_{r \in [-\overline{r},\overline{r}]} |\widetilde{Z}_T(r,\rho_k,-\zeta)| + \zeta O_p(1). \tag{26}$$

Then, $\zeta O_p(1)$ can be arbitrary small in probability, so Proposition 5 follows if the first two terms of the RHS of (26) are $o_p(1)$. This will be true if

$$\sup_{r \in [-\overline{r},\overline{r}]} |\widetilde{Z}_T(r,\rho,\zeta)| = o_p(1) \quad \text{for every} \rho, \zeta. \tag{27}$$

For that Lemma 3 is needed.

Lemma 3. *For every $\ell \in (0, 1/2)$, every $\rho = (u,v,w) \in D_b$ and every $\lambda \in R$,*

$$\sup_{(r_1,r_2) \in B_{T,\ell}} T^{-1/2} \sum_{t-1}^{T} \varepsilon_t^+ |\mathsf{F}(r_2 + \Gamma_{t-1}(\rho,\lambda)) - \mathsf{F}(r_1 + \Gamma_{t-1}(\rho,\lambda))| = o_p(1),$$

where $B_{T,\ell} = \{(r_1, r_2) \in R \times R; |\mathsf{F}(r_1) - \mathsf{F}(r_2)| \leqslant K_{\mathsf{F},1} T^{-1/2-\ell}\}$ for some $K_{\mathsf{F},1} < \infty$.

Given the uniform continuity and boundedness of our $\mathsf{F}(r)$, the proof of this lemma is analogous to that of Lemma 2 of Bai (1994).

Now divide the real line into $N(T) = [T^{-1/2+\ell}] + 1$ parts, with $\ell \in (0, 1/2)$, by points $-\overline{r} = r_0 < r_1 < \cdots < r_{N(T)} = \overline{r}$ with $\mathsf{F}(r_i) = i N(T)^{-1} K_{\mathsf{F},1}$. Since $\varepsilon_t^+ \varepsilon_{t-1}^+ 1(\varepsilon_{t-1} < r)$ and $\varepsilon_t^+ \mathsf{F}(r)$ are non-decreasing, for $r_k < r < r_{k+1}$, we have

$$|\widetilde{Z}_T(r,\rho,\zeta)| \leqslant |\widetilde{Z}_T(r_{k+1},\rho,\zeta)| + T^{-1/2} \sum_{t-1}^{T} \varepsilon_t^+ [\mathsf{F}(r_{k+1} + \Gamma_{t-1}) - \mathsf{F}(r + \Gamma_{t-1})]$$

$$+ T^{-1/2} \sum_{t-1}^{T} \{\varepsilon_t^+ [\varepsilon_{t-1}^+ 1(\varepsilon_{t-1} < r_{k+1}) - \mathsf{F}(r_{k+1})]$$

$$- \varepsilon_t^+ [\varepsilon_{t-1}^+ 1(\varepsilon_{t-1} < r) - \mathsf{F}(r)]\},$$

and the reverse inequality with r_{k+1} replaced by r_k. Therefore,

$$\sup_{r\in[-\bar{r},\bar{r}]} |\widetilde{Z}_T(r,\rho,\zeta)| \leqslant \max_k |\widetilde{Z}_T(r_k,\rho,\zeta)|$$

$$+ \max_k T^{-1/2} \left| \sum_t^T \varepsilon_t^+ [\mathsf{F}(r_k + \varGamma_{t-1}) - \mathsf{F}(r + \varGamma_{t-1})] \right|$$

$$+ \sup_{|g-g'| \leqslant K_{\mathsf{F}}N(T)^{-1}} T^{-1/2} \sum_t^T \{\varepsilon_t^+ [\varepsilon_{t-1}^+ 1(\varepsilon_{t-1} < \mathsf{F}^{-1}(g)) - g]$$

$$- \varepsilon_t^+ [\varepsilon_{t-1}^+ 1(\varepsilon_{t-1} < \mathsf{F}^{-1}(g')) - g']\}. \tag{28}$$

For the third summand of the RHS of expression (28), it is needed to prove that $T^{-1/2} \sum_t^T {}_1 \varepsilon_t^+ [\varepsilon_{t-1}^+ 1(\varepsilon_{t-1} < \mathsf{F}^{-1}(g)) - g]$ is tight. This process can be written as

$$T^{-1/2} \sum_t^T \varepsilon_t^+ [\varepsilon_{t-1}^+ 1(\varepsilon_{t-1} < F^{-1}(g)) - g]$$

$$= T^{-1/2} \sum_t^T (\varepsilon_t^+ - \mathrm{E}(\varepsilon_t^+))[\varepsilon_{t-1}^+ 1(\varepsilon_{t-1} < \mathsf{F}^{-1}(g)) - g]$$

$$+ T^{-1/2} \sum_t^T \mathrm{E}(\varepsilon_t^+)[\varepsilon_{t-1}^+ 1(\varepsilon_{t-1} < \mathsf{F}^{-1}(g)) - g].$$

The tightness of the first and second summand can be proved by using Lemma 3.1 of Koul and Stute (1999). Then, the $o_p(1)$ of both summands follows from the tightness of the process and the fact that $K_{\mathsf{F}^{-1}}N(T)^{-1} = o_p(1)$. The convergence in probability to 0 of the second summand of the RHS of (28) follows from Lemma 3.

Finally, $\widetilde{Z}_T(r_k,\rho,\zeta)$ can be expressed as

$$\widetilde{Z}_T(r_k,\rho,\zeta) \quad T^{-1/2} \sum_t^T (\varepsilon_t^+ - \mathrm{E}(\varepsilon_t^+))d_{t-1}(r_k,\rho,\zeta) + T^{-1/2}\mathrm{E}(\varepsilon_t^+) \sum_t^T d_{t-1}(r_k,\rho,\zeta),$$

with $d_{t-1}(r_k,\rho,\zeta) = [\varepsilon_{t-1}^+ 1(\varepsilon_{t-1} < r_k + \varGamma_{t-1}(\rho,\zeta)) - \mathsf{F}(r_k + \varGamma_{t-1}(\rho,\zeta)) - \varepsilon_{t-1}^+ 1(\varepsilon_{t-1} < r_k) + \mathsf{F}(r_k)]$. Then, the processes involved in both sums are martingale arrays, and we can apply directly Bai (1994) to obtain that

$$\max_k |\widetilde{Z}_T(r_k,\rho,\zeta)| = o_p(1). \qquad \square$$

Proof of Proposition 6. This proof is analogous to Proposition 2 of Bai (1994), and it is thus omitted. $\square$

With this, result (24) is proved. Then, getting back into expression (23), we have proved

$$T^{-1/2} \sum_t^T \varepsilon_t \varepsilon_{t-1} 1(|\widehat{\varepsilon}_{t-1}| \leqslant r) \Rightarrow D_1(r) + o_p(1),$$

where the $o_p(1)$ term is uniform in r.

For the rest of the terms in the RHS of (22), using that

$$\mathrm{E}\left[\sup_u \Upsilon^2_{x,t}(u)\right] \leqslant K_{\Upsilon x} < \infty,$$

$$\mathrm{E}\left[\sup_u \Upsilon^2_{\varepsilon,t}(u)\right] \leqslant K_{\Upsilon \dot\varepsilon} < \infty,$$

and that ε_t is *iid*, it can be established, in a similar way as (23) and (24), that

$$T^{-1} \sum_{t\ 1}^{T} \varepsilon_t \Upsilon_{x,t-1}(\widehat{\theta}) \mathbf{1}(|\widehat{\varepsilon}_{t-1}| \leqslant r) = \mathrm{o_p}(1),$$

$$T^{-1} \sum_{t\ 1}^{T} \varepsilon_t \Upsilon_{\varepsilon,t-1}(\widehat{\theta}) \mathbf{1}(|\widehat{\varepsilon}_{t-1}| \leqslant r) = \mathrm{o_p}(1),$$

$$T^{-1} \sum_{t\ 1}^{T} \varepsilon_{t-1} \Upsilon_{x,t}(\widehat{\theta}) \mathbf{1}(|\widehat{\varepsilon}_{t-1}| \leqslant r) = T^{-1} \sum_{t\ 1}^{T} \varepsilon_{t-1}^2 \mathbf{1}(|\varepsilon_{t-1}| \leqslant r) + \mathrm{o_p}(1),$$

$$T^{-1} \sum_{t\ 1}^{T} \varepsilon_{t-1} \Upsilon_{\varepsilon,t}(\widehat{\theta}) \mathbf{1}(|\widehat{\varepsilon}_{t-1}| \leqslant r) = T^{-1} \sum_{t\ 1}^{T} \varepsilon_{t-1}^2 \mathbf{1}(|\varepsilon_{t-1}| \leqslant r) + \mathrm{o_p}(1),$$

$$T^{-1} \sum_{t\ 1}^{T} \varepsilon_{t-1}^2 \mathbf{1}(|\varepsilon_{t-1}| \leqslant r) \overset{p}{\longrightarrow} \mathrm{E}(\varepsilon_{t-1}^2 \mathbf{1}(|\varepsilon_{t-1}| \leqslant r)) = \Sigma_\varepsilon(r) \quad \text{uniformly in } r \in [\underline{r}, \overline{r}],$$

$$\sup_{r\in[\underline{r},\overline{r}]} R_{1,t}(\widehat{\phi}, \widehat{\theta}, r) = \mathrm{o_p}(1),$$

where the $\mathrm{o_p}(1)$ terms are uniformly in r. Now, given that $T^{1/2}[(\widehat{\phi} - \phi_0),(\widehat{\theta} - \theta_0)]$ converges in distribution to (V_ϕ, V_θ), which is a multivariate normal random variable with mean 0 and variance covariance matrix $M_{\phi,\theta}$, it is straightforward to prove that

$$T^{-1/2} \sum_{t\ 1}^{T} \widehat{\varepsilon}_t \widehat{\varepsilon}^s_{t-1}(r) \Rightarrow D_1(r) + (V_\phi + V_\theta)\Sigma_\varepsilon(r), \tag{29}$$

with $\quad Cov(D_1(r), V_\phi) = H_{1.}^{-1}(\lambda^0)\lim_{T\to\infty} T^{-1}\sum_{t\ 1}^{T} \mathrm{E}[\varepsilon_t\varepsilon_{t-1}\mathbf{1}(|\varepsilon_{t-1}| \leqslant r)s_t(\lambda^0)] \quad$ and $Cov(D_1(r), V_\theta) = H_{2.}^{-1}(\lambda^0)\lim_{T\to\infty} T^{-1}\sum_{t\ 1}^{T} \mathrm{E}[\varepsilon_t\varepsilon_{t-1}\mathbf{1}(|\varepsilon_{t-1}| \leqslant r)s_t(\lambda^0)]$.

Finally, it is necessary to prove that

$$T^{-1} \sum_{t\ 1}^{T} (\widehat{\varepsilon}^s_{t-1}(r))^2 \overset{p}{\longrightarrow} \Sigma_\varepsilon(r), \tag{30}$$

$$T^{-1} \sum_{t\ 1}^{T} \widehat{u}^2_t(r) \overset{p}{\longrightarrow} \sigma^2_\varepsilon, \tag{31}$$

uniformly in $r \in [\underline{r}, \overline{r}]$ with $0 < \Sigma_\varepsilon(r) < \infty$ and $0 < \sigma^2_\varepsilon < \infty$.

The proofs of (30) and (31) are tedious and similar to the above proofs. So, we only present an outline of them. For (30) note that $(\widehat{\varepsilon}_t^s(r))^2$ can be written as

$$
\begin{aligned}
(\widehat{\varepsilon}_t^s(r))^2 &= [\varepsilon_t + (\varepsilon_t - \widehat{\varepsilon}_t)]^2 1(|\widehat{\varepsilon}_t| \leqslant r) \\
&= \varepsilon_t^2 1(|\varepsilon_t| \leqslant r) + \varepsilon_t^2 [1(|\widehat{\varepsilon}_t| \leqslant r) - 1(|\varepsilon_t| \leqslant r)] \\
&\quad + (\varepsilon_t - \widehat{\varepsilon}_t)^2 1(|\widehat{\varepsilon}_t| \leqslant r) + 2\varepsilon_t(\varepsilon_t - \widehat{\varepsilon}_t) 1(|\widehat{\varepsilon}_t| \leqslant r).
\end{aligned}
$$

About the first term, $T^{-1} \sum_{t\ 1}^T \varepsilon_t^2 1(|\varepsilon_t| \leqslant r)$, given Assumption A.0, it is straightforward to show that

$$
T^{-1} \sum_{t\ 1}^T \varepsilon_t^2 1(|\varepsilon_t| \leqslant r) \xrightarrow{p} \Sigma_\varepsilon(r),
$$

uniformly in $r \in [\underline{r}, \overline{r}]$ with $0 < \Sigma_\varepsilon(r) < \infty$. The rest of terms can be handled as $T^{-1} \sum_{t\ 1}^T \varepsilon_t^+ \varepsilon_{t-1}^+ (1(\widehat{\varepsilon}_{t-1} < r) - 1(\varepsilon_{t-1} < r))$, $R_{1,t}(\widehat{\phi}, \widehat{\theta}, r)$ and $T^{-1} \sum_{t\ 1}^T \varepsilon_t \Upsilon_{x,t-1}(\widehat{\theta}) 1 (|\varepsilon_{t-1}| < r)$, respectively. Once (30) is proved, it is straightforward to show that

$$
\widehat{\alpha}(r) \xrightarrow{p} 0, \tag{32}
$$

uniformly in $r \in [\underline{r}, \overline{r}]$. From (30) and (32), it is easy to obtain (31).

Result 4,

$$
\sup_{r \in [\underline{r}, \overline{r}]} |t_{\alpha(r)}| \xrightarrow{d} \sup_{r \in [\underline{r}, \overline{r}]} |D(r)|,
$$

with

$$
D(r) \equiv [D_1(r) + (V_\phi + V_\theta)\Sigma_\varepsilon(r)][\Sigma_\varepsilon(r)\sigma_\varepsilon^2]^{-1/2},
$$

follows directly from the continuous mapping theorem and (29)–(31). $\square$

Proof of Result 5. For simplicity and without lack of generality we present the proof for the case of $\theta_2 = 1$.

Let $\sigma_x^2 = V(x_t)$, and $\rho_{x1} = Cov(x_t, x_{t-1})$ with $x_t = (1 - L)y_t$. If y_t follows the TIMA model (2), then,

$$
\begin{aligned}
\sigma_x^2 &= 2\sigma_\varepsilon^2 - (1 - \theta_1^2)\sigma_{\varepsilon,L}^2, \\
\rho_{x1} &= -\sigma_\varepsilon^2 + (1 - \theta_1)\sigma_{\varepsilon,L}^2,
\end{aligned} \tag{33}
$$

where $\sigma_{\varepsilon,L}^2 = E[(\varepsilon_t^L)^2]$. From Eqs. (15) and (16),

$$
\begin{aligned}
V^{\mathrm{TIMA}}(\Delta y_t^P) &= \sigma_x^2 + 2\rho_{x1}, \\
V^{\mathrm{TIMA}}(y_t^T) &= -\rho_{x1} + (\theta_1^2 - \theta_1)\sigma_{\varepsilon,L}^2.
\end{aligned}
$$

To obtain the corresponding results for linear B–N decomposition, we use the Wold representation of the TIMA model,

$$
(1 - L)y_t = \mu + u_t - \theta u_{t-1}, \tag{34}
$$

with u_t white noise, $(E(u_t) = 0, \sigma_u^2 = E(u_t^2))$. Then,

$$V^{BN}(\Delta y_t^P) = \sigma_x^2 + 2\rho_{x1},$$
$$V^{BN}(y_t^T) = -\theta\rho_{x1}.$$

Finally, proceeding in a similar way, the UC–0 decomposition of (34) is

$$(1 - L)y_t^P = \mu + \eta_t^P,$$
$$y_t^T = \eta_t^T,$$

with $E[\eta_t^P] = E[\eta_t^T] = 0$, $E[(\eta_t^P)^2] = \sigma_{\eta^P}^2$, $E[(\eta_t^T)^2] = \sigma_{\eta^T}^2$, $E(\eta_t^P, \eta_t^T) = 0$, and such that $\sigma_{\eta^P}^2 = \sigma_x^2 + 2\rho_{x1}$ and $\sigma_{\eta^T}^2 = -\rho_{x1}$. Therefore

$$V^{UC\ 0}(\Delta y_t^P) = \sigma_x^2 + 2\rho_{x1},$$
$$V^{UC\ 0}(y_t^T) = -\rho_{x1}.$$

Then, the $Var(\Delta y_t^P)$ is the same for the three decompositions. With respect to the $Var(y_t^T)$, using the fact that the parameters θ and θ_1 are less than 1, it is straightforward to show the largest one is obtained in the case of the UC–0 decomposition. To prove that $V^{BN}(y_t^T)$ is less or equal than $V^{TIMA}(y_t^T)$, it is enough to prove that $\sigma_u^2 \geqslant \sigma_\varepsilon^2$, since

$$V^{BN}(y_t^T) = \sigma_x^2 - \sigma_u^2,$$
$$V^{TIMA}(y_t^T) = \sigma_x^2 - \sigma_\varepsilon^2.$$

From the Wold representation (34),

$$\sigma_u^2 = \frac{\sigma_x^2 \pm \sqrt{(\sigma_x^2)^2 - 4\rho_{x1}^2}}{2}.$$

Given that $\rho_{x1} = -\theta\sigma_u^2$ and $|\theta| < 1$, the valid solution is

$$\sigma_u^2 = \frac{\sigma_x^2 + \sqrt{(\sigma_x^2)^2 - 4\rho_{x1}^2}}{2}.$$

Substituting σ_x^2 and ρ_{x1} in terms of θ_1, σ_ε^2 and $\sigma_{\varepsilon,L}^2$ from (33), we have

$$\sigma_x^4 - 4(\rho_{x1})^2 = \sigma_{\varepsilon,L}^4(1 - \theta_1^2)^2 + (\sigma_\varepsilon^2\sigma_{\varepsilon,L}^2 - \sigma_{\varepsilon,L}^4)(2 - 2\theta_1)^2.$$

Then, by using that $\sigma_x^2 = 2\sigma_\varepsilon^2 - (1 - \theta_1^2)\sigma_{\varepsilon,L}^2$ and $\sigma_\varepsilon^2 > \sigma_{\varepsilon,L}^2$ it is obtained that $\sigma_u^2 \geqslant \sigma_\varepsilon^2$ and the result follows. □

References

Bai, J., 1994. Weak convergence of the sequential empirical processes of residuals in ARMA models. The Annals of Statistics 22, 2051 2061.

Beaudry, P., Koop, G., 1993. Do recessions permanently change output? Journal of Monetary Economics 31, 149 164.

Beveridge, S., 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., Quah, D., 1989. The dynamic effects of aggregate demand and supply disturbances. The American Economic Review 79, 655 673.

Campbell, J., Mankiw, N.G., 1987. Are output fluctuations transitory? Quarterly Journal of Economics 102, 857 880.

Caner, M., Hansen, B.E., 2001. Threshold autoregressions with a unit root. Econometrica 69, 1555 1597.

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.

Clarida, R.H., Taylor, M.P., 2003. Nonlinear permanent temporary decompositions in macroeconomics and finance. The Economic Journal 113, C125 C139.

Clark, P.K., 1987. The cyclical component of U.S. economic activity. Quarterly Journal of Economics 102, 797 814.

Cochrane, J.H., 1994. Shocks. Carnegie Rochester Conference Series on Public Policy 41, 295 364.

Davis, 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.

Davison, A.C., Hinkley, D.V., 1997. Bootstrap Methods and their Application. Cambridge University Press, Cambridge.

De Gooijer, J.G., 1998. On threshold moving-average models. Journal of Time Series Analysis 19, 1 18.

Elwood, S., 1998. Is the persistence of shocks to output asymmetric? Journal of Monetary Economics 41, 411 426.

Engle, R.F., Smith, A.D., 1999. Stochastic Permanent Breaks. Review of Economics and Statistics 84, 553 574.

Glosten, L.R., 1987. Components of the Bid-Ask spread and the statistical properties of transaction prices. Journal of Finance 42, 1293 1307.

González, M., Gonzalo, J., 1998. Threshold unit root processes. Manuscript, Universidad Carlos III.

Gonzalo, J., Martínez, O., 2004. Threshold Integrated Moving Average. Manuscript, Universidad Carlos III.

Gonzalo, J., Montesinos, R., 2002. Threshold Stochastic Unit Root models. Manuscript, Universidad Carlos III.

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

Guay, A., Scaillet, O., 2003. Indirect inference, nuisance parameter and threshold moving average models. Journal of Business and Economic Statistics 21, 122 132.

Hamilton, J.D., 1989. A new approach to the economic analysis of nonstationary time series and the business cycle. Econometrica 57, 357 384.

Hansen, B.E., 1996. Inference when a nuisance parameter is not identified under the null hypothesis. Econometrica 64, 413 430.

Harvey, A.C., 1985. Trends and cycles in macroeconomics time series. Journal of Business and Economic Statistics 3, 216 227.

Hasbrouck, J., 1993. Assessing the quality of a security market: a new approach to transaction-cost measurement. The Review of Financial Studies 6, 191 212.

Hess, G.D., Iwata, S., 1997. Asymmetric persistence in GDP? A deeper look at depth. Journal of Monetary Economics 40, 535 554.

Koop, G., Pesaran, M.H., Potter, S.M., 1996. Impulse response analysis in nonlinear multivariate models. Journal of Econometrics 74, 119 147.

Koul, H.L., Stute, W., 1999. Nonparametric model checks for time series. The Annals of Statistics 27, 204 236.

Mortely, J.C., Nelson, C.R., Zivot, E., 2001. Why are Beveridge Nelson and unobserved-component decomposition of GDP so different? Review of Economics and Statistics 85, 235 244.

Neftci, S.N., 1984. Are economic time series asymmetric over the business cycle? Journal of Political Economy 92, 307–328.

Petrucelli, J., Woodford, S., 1984. A Threshold AR(1) model. Journal of Applied Probability 21, 270–286.

Phillips, P.C.B., Solo, V., 1992. Asymptotics for linear processes. The Annals of Statistics 20, 971–1001.

Pierce, D.A., 1979. Signal extraction error in nonstationary time series. The Annals of Statistics 7, 1303–1320.

Potter, S.M., 2000. Nonlinear impulse response functions. Journal of Economic Dynamics and Control 24, 1425–1446.

Quah, D., 1992. The relative importance of permanent and transitory components: identification and some theoretical bounds. Econometrica 60, 107–118.

Roll, R., 1984. A Simple model of the implicit bid–ask spread in an efficient market. Journal of Finance 39, 1127–1446.

Watson, M.W., 1986. Univariate detrending methods with stochastic trends. Journal of Monetary Economics 18, 49–75.

Wecker, W.E., 1981. Asymmetric time series. Journal of the American Statistical Association 76, 16–21.

Wold, H., 1938. A Study in the Analysis of Stationary Time Series. Almqvist & Wiksells, Uppsala.

Wooldridge, J.M., 1994. Estimation and inference for dependent processes. In: Engle, R.F., McFadden, D. (Eds.), Handbook of Econometrics, vol. 4. North-Holland, Amsterdam, pp. 2641–2738.