Powerful Tests for Structural Changes in Volatility*

KE-LI XU†
Texas A&M University

January 31, 2011

Abstract

The CUSUM test (when applied to squared series) is routinely used to detect structural change in volatility in a nonparametric way. We show that the test does not suffer from the problem of nonmonotonic power, unlike the CUSUM test for the changing mean. We then propose a modified test that improves the power to a fixed inflation factor for iid case and to an order of magnitude for mds case (allowing for weak dependence in the second moment).

1 Introduction

Ignorance of structural change in the second moment leads to the spurious long-range dependence in volatility and the integrated GARCH effect [Mikosch and Stårică (2004) and Hillebrand (2005)].

2 Testing for changing mean: a revisit

Consider

\[ y_t = \theta_t + e_t, \]
\[ e_t = s_t v_t \]

where \( v_t \) is stationary and weakly dependent with \( \text{E} v_t = 0, \text{E} v_t^2 = 1 \) and autocovariances \( \gamma_l (l = 0, 1, 2, \ldots) \), and \( \theta_t \) and \( s_t \) are deterministic. We observe \( y_t \) and we are interested in whether there is structural change in the mean of \( y_t \).

Let \( \tilde{e}_t = y_t - \overline{y}_t = e_t - \overline{e} + \theta_t - \overline{\theta} \) be the demeaned residuals. The test (focusing on CUSUM test) is defined as \( Q = \max_{1 \leq K \leq n} |n^{-1/2} \sum_{t=1}^{K} \tilde{e}_t|/\overline{\omega} \), where \( \overline{\omega}^2 = \overline{\gamma}_0 + \)...

---

*A very preliminary draft. Please do not circulate.
†Department of Economics, Texas A&M University, 3063 Allen, 4228 TAMU, College Station, TX 77843-4228, USA. Email: keli.xu@econmail.tamu.edu.
Indeed, for $2 \sum_{t=1}^{m} k(l/m) \tilde{\gamma}_t$ with $\tilde{\gamma}_t = n^{-1} \sum_{i=t+1}^{n} \hat{e}_i \hat{e}_{t-i}$, $\tilde{\gamma}_t$ consistently estimates $\gamma_t$ (times a constant involving $s$) under the null (even when $s_t$ is time varying), but not under the alternative (incorrectly centered). Data dependent truncation (Bartlett) $m = 1.1447 \left( \frac{\delta^2}{(1-\rho^2)^2} n \right)^{1/3}$, where $\hat{\rho} = \tilde{\gamma}_1/\tilde{\gamma}_0$.

The size of $Q$ is generally robust to NV, but the power depends on the behavior of $s_t$.

As is well known, $Q$ may have nonmonotonic power. The main cause is the behavior of $\hat{\rho}$ (which is caused by incorrect centering in $\tilde{\gamma}_t$) when $\theta_t$ is not constant. We start by revisiting Juhl and Xiao’s results when $s_t$ is constant and then examine the sensitivity to time-varying $s_t$.

Consider the behavior of $Q$ under the alternative $H_A : \theta_t = \theta(t/n) = \theta_0 + g(t/n) \eta$, where $\eta = O(n^b)$, $b \geq 0$. This is the asymptotics that should be used in studying non-monotonic power.

When $s_t$ is time-invariant (Juhl and Xiao), $\hat{\rho} \rightarrow 1$. Indeed, for $l = 0$ and 1,

\[
\tilde{\gamma}_l = n^{-1} \sum_{t=l+1}^{n} \hat{e}_i \hat{e}_{t-l} = n^{-1} \sum_{t=l+1}^{n} (e_t - \bar{e} + \theta_t - \bar{\theta})(e_{t-l} - \bar{e} + \theta_{t-l} - \bar{\theta}) \rightarrow \gamma_l + n^{-1/2} \eta X + \eta^2 \int (g(r) - \bar{g})^2
\]

where the third term is the dominant term and is the same for $\tilde{\gamma}_0$ and $\tilde{\gamma}_1$. Thus $\hat{\rho} \rightarrow 1$. Also $\hat{\rho} - 1 = O_p(n^{-1/2} \eta^{-1})$. If $\eta = n^{1/2}$, then $\hat{\rho} - 1 = O_p(n^{-1})$ and $m = O(n)$; a faster rate than under the null. We can show that

\[
n^{-1/2} \sum_{t=1}^{K} \hat{e}_i = n^{-1/2} \sum_{t=1}^{K} (e_t - \bar{e} + \theta_t - \bar{\theta}) = O(n^{-1/2} K \eta) = O(n^{1/2} \eta) \quad (3)
\]

\[
\hat{\omega}^2 = O(m \eta^2). \quad (4)
\]

So $Q = O_p(1)$ leading to nonmonotonic power.

Now consider the case when $s_t = \theta_t$ (which will be relevant below); we are interested in the behavior of the test for changing mean when there is also structural change in variance. In this case,

\[
\hat{\rho} \rightarrow \frac{\gamma_1}{\gamma_0} + \frac{[1 - (\int g)^2/\int g^2]}{[1 - (\int g)^2/\int g^2]} \neq 1
\]

Indeed, for $l = 0$ and 1,

\[
\tilde{\gamma}_l \rightarrow \eta^2 \int g^2 \gamma_l + \eta^2 \left[ \int g^2 - (\int g)^2 \right]
\]

where the dominant term depends on $l$. Thus $m = O(n^{1/3})$; the same rate as under the null. We can show that (3) and (4) are still true. So $Q = O_p(n^{1/3})$ leading to monotonic power.
Consider a shift in mean: \( \theta_t = \theta([nr]) \), where \( \theta(r) = \theta^0 + (\theta^1 - \theta^0)1_{r \geq r}, r \in (0, 1] \). We fix \( \theta^0 \) and let \( \theta^1 \) increase unboundedly, so \( \eta = (\theta^1 - \theta^0) \) and \( g(r) = 1_{r \geq r} \). In this case

\[
\hat{\rho} \to \frac{\gamma_1 + \tau}{\gamma_0 + \tau}.
\]

In general, consider the case when \( s_t \) is time-varying, \( s_t = s(t/n) = s^0 + g(t/n)\eta \), where \( \eta = O(n^b) \), \( b \geq 0 \). Using the similar arguments above, we have \( \hat{\rho} \to 1 \) if \( b < b \) but the convergence rate is slower if \( b > 0 \). Thus monotonic power.

On the other hand, \( \hat{\rho} \to \gamma_1/\gamma_0 \neq 1 \) and \( \hat{\gamma}_1 = O(\eta^2) \) if \( b \geq b \). (3) and (4) become

\[
n^{-1/2} \sum_{t=1}^{K} \hat{e}_t = O(\eta + n^{1/2}\eta),
\]

\[
\tilde{\omega}^2 = O(m\eta^2).
\]

So

\[
Q = \begin{cases} 
O_p(1/\sqrt{m}) = O_p(n^{-1/6}) , & \text{if } n^{-1/2}\eta^{-1}\eta \to \infty \text{ (having nonmonotonic power)} \\
O_p((\sqrt{m})/(\sqrt{m}\eta)) = O_p(n^{1/3}\eta/\eta) , & \text{if } n^{-1/2}\eta^{-1}\eta \to 0
\end{cases}
\]

So whether \( Q \) diverges depends the relationship between \( b \) and \( b ^{'} \):

<table>
<thead>
<tr>
<th>( b )</th>
<th>( b &lt; 0 )</th>
<th>( b = 0 )</th>
<th>( 0 &lt; b &lt; b )</th>
<th>( b = b )</th>
<th>( b &lt; b &lt; b + 1/3 )</th>
<th>( b \geq b + 1/3 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>monotonic power?</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>plim( \hat{\rho} )</td>
<td>( 1 )</td>
<td>( 1 )</td>
<td>( 1 )</td>
<td>( \frac{\gamma_1 + \tau}{\gamma_0 + \tau} )</td>
<td>( \gamma_1/\gamma_0 )</td>
<td>( \gamma_1/\gamma_0 )</td>
</tr>
</tbody>
</table>

When there is large structural change in variance, nonmonotonic power happens even if the data dependent truncation is not used.

In summary, a medium structural change in variance (as well) eliminates the nonmonotonic power problem, but a large one would generate the problem again.

2.1 Simulations

Consider model (1-2) where

\[
s_t = \theta_t^n,
\]

\[
v_t = \rho v_t + w_t, w_t \sim N(0, V_w),
\]

with \( \alpha = 0, 0.5, 1, 1.2, 2 \), corresponding no, small, medium, medium, and large shits in variance. We select \( V_w = 1 - \rho^2 \) such that \( \text{Var}(v_t) = 1 \) (identification of \( s_t \)). Set \( \theta_t \) as the step function above.

For values of \( \alpha, \hat{\rho} \) will go to 1 (approaching fast), 1 (approaching slowly), \( \frac{\gamma_1 + \tau}{\gamma_0 + \tau} \), \( \rho \) and \( \rho \), respectively, according the theories we develop earlier. Nonmonotonic power happens when \( \alpha = 0 \) and 2.

So

\[
\hat{\rho} \to \frac{\gamma_1 + \tau}{\gamma_0 + \tau} = \frac{\rho + \tau}{1 + \tau}.
\]
3 Testing for structural change in volatility

The main motivation of the paper comes from consideration of testing for structural change in volatility. It is well known that the stylized facts of long-range dependent or IGARCH effects in stock returns can be explained by non-constant unconditional variance. So it is important to test for it.

Consider the following model for a martingale difference sequence $u_t$ (e.g. log returns):

$$ u_t = \sigma_t \varepsilon_t, \quad (5) $$

where $\sigma_t$ is a deterministic function of $t$ and $\varepsilon_t$ is mds such that $E\varepsilon_t = 0$ and $E\varepsilon_t^2 = 1$. In (5), $\sigma_t$ accounts for nonstationary unconditional variances and $\varepsilon_t$ accounts for conditional heteroskedasticity. The variance $\sigma_t^2$ satisfies $\sigma_t^2 = \sigma^2(t/n)$ where $\sigma^2(r)$ is nonstochastic cadlag function on $(0,1]$ with a finite number of points of discontinuity. We also assume $\sigma^2(r) > 0$ and is twice differentiable except at the points of discontinuity with the second derivative function satisfying a (uniform) first-order Lipschitz condition.

The model (5) is quite general including regime-switching GARCH model.

The hypothesis of interest:

$$ H_0 : \sigma_t^2 \equiv \sigma^2 $$

$$ H_A : \sigma_t^2 \text{ is not constant over } t. $$

Rewrite the model (5) as

$$ u_t^2 = \sigma_t^2 + e_t, \quad (6) $$

where $e_t = \sigma_t^2(\varepsilon_t^2 - 1)$. (6) can be nested in the model (1) with $s_t = \theta_t$, when applied to $y_t = u_t^2$. That is, under the alternative, the same structural change occurs in both the mean and the variance. The CUSUM test for structural change in $\sigma_t^2$, denoted by $\tilde{Q}_1$, was used by Loretan and Phillips (1994), Andreou and Ghysels (2002), Santo et al. (2004), Deng and Perron (2008), Rapach and Strauss (2008), among others.

The results in Section 2 explain the monotonic power of $\tilde{Q}_1$ when an automatic truncation number is used.

3.1 An improved CUSUM test

Let $C_K = \sum_{i=1}^{K} u_i^2$ and $D_K = C_K - (K/n)C_n$ for $K = 1, \ldots, n$. The CUSUM test is defined as

$$ \tilde{Q}_1 = \max_{1 \leq K \leq n} n^{-1/2} |D_K|/\tilde{\omega}, $$

where $\tilde{\omega}^2$ is the estimator of $\omega^2 = \lim_{n \to \infty} Var(n^{-1/2} \sum_{i=1}^{n} (u_i^2 - \sigma_i^2)) = LRV(\varepsilon_t^2 - 1) \int_0^1 \sigma^4(r)$. $\omega^2$ reduces to the long run variance of $u_t^2 - \sigma_t^2$ if $\sigma^2(\cdot)$ is a constant.

Since $\sigma_t^2$ is not directly observable, we estimate it by $\tilde{\sigma}_t^2$. We consider

$$ \tilde{\sigma}_t^2 = \sum_{i=1}^{n} w_{ti} u_i^2, \quad (7) $$

4
and \( w_{ti} = \left( \sum_{i=1}^{n} K \left( \frac{t-i}{nh} \right) \right)^{-1} K \left( \frac{t-i}{nh} \right) \).

\( \tilde{\omega}^2 \) is thus defined by

\[
\tilde{\omega}^2 = n^{-1} \sum_{t=1}^{n} (u_t^2 - \tilde{\sigma}_t^2)^2 + 2n^{-1} \sum_{l=1}^{m} k(l/m) \sum_{t=l+1}^{n} (u_t^2 - \tilde{\sigma}_t^2)(u_{t-l}^2 - \tilde{\sigma}_{t-l}^2).
\]

Definition of \( \tilde{\omega}^2 \) under \( H_A \) is not obvious since it takes the form of a LRV estimator and it seems to rely on weak stationarity of \( u_t^2 - \sigma_t^2 \). We shall show below that \( \tilde{\omega}^2 \) is consistent for \( \omega^2 \) under both the null and alternative hypotheses.

**Assumption 1.** (i) \( \{\varepsilon_t^2 - 1\} \) is a real zero-mean, strictly stationary mixing process with \( E|\varepsilon_t^2 - 1|^p < \infty \) for some \( p > 2 \) and with mixing coefficients \( \{\alpha_k\} \) satisfying \( \sum_{k=0}^{\infty} \alpha_k^2(1/v-1/p) \) for some \( v \in (2, 4] \), \( v \leq p \). Furthermore, the long run variance \( \omega^2 = \sum_{l=-\infty}^{\infty} E(\varepsilon_t^2 - 1)(\varepsilon_{t-l}^2 - 1) \) is strictly positive and finite.

(ii) \( |k(x)| \leq 1 \) for all \( x \) on the real line, \( k(x) = k(-x) \), \( k(0) = 1 \); \( k(x) \) is continuous at zero and for almost all \( x \); \( |k(x)| \leq \overline{k}(x) \) where \( \overline{k}(x) \) is a nonincreasing function such that \( \int_{-\infty}^{\infty} |x\overline{k}(x)|dx < \infty \); \( m \to \infty \) as \( n \to \infty \) and \( m = o(n^\vartheta) \), \( \vartheta \leq 1/2 - 1/v \) where \( v \) is defined in Assumption 1(i).

(iii) \( K(x) \) is a symmetric, bounded kernel function and has compact support; The function \( |x|^jK(x) \) satisfies a (uniform) first-order Lipschitz condition, where \( 0 \leq j \leq 3 \); \( h + 1/(nh) \to 0 \) and \( nh^5 = O(1) \) as \( n \to \infty \).

**Theorem 1.** Assume \( m[h^2 + \sqrt{\ln n/(nh)}] \to 0 \).

(i) Under \( H_0 \)

\[
Q_1 \Rightarrow \sup_{0 \leq r \leq 1} |\overline{W}(r)|,
\]

where \( \overline{W}(r) = W(r) - rW(1) \) is a Brownian bridge.

(ii) Under \( H_A \)

\[
Q_1 = O_p \left( \left( \frac{n}{m + \sqrt{\ln n/(nh)}} \right)^{1/2} \right).
\]

If we assume \( m[h + \sqrt{\ln n/(nh)}] \to 0 \), we have \( n^{-1/2}Q_1 \Rightarrow \sup_{s \in [0,1]} |\int_0^s \sigma^2 - s \int \sigma^2|/\omega \).

Theorem 1 remains true under the diverging alternative.

The CUSUM test \( Q_1 \) is numerically equivalent to Juhl and Xiao’s (2009) test of structural change in mean applied to \( u_t^2 \).

Our results extend theirs by allowing for structural change in the error \( e_t \) which is ruled out in their study.

\( Q_1 \) is more powerful than the widely used test \( \tilde{Q}_1 \). Under \( H_0 \), \( \tilde{\omega}^2 \Rightarrow \omega^2 \) and \( \tilde{Q}_1 \to BB \).

Under \( H_A \), however, \( \tilde{\omega}^2 \) diverges (not only inconsistently estimates \( \omega^2 \)):

\[
m^{-1}\tilde{\omega}^2 = O_p(1).
\]

\(^1\)It is not appropriate to apply Juhl and Xiao’s (2009) model (i.e. the additive model with a time varying mean and an error term capturing serial correlation) to \( u_t^2 \) since restrictive assumption has to be imposed to guarantee positivity. The multiplicative model (5) we adopt here is more natural for time varying volatility (Granger and Starica).
Thus
\[ \tilde{Q}_1 = O_p((n/m)^{1/2}). \]  

\( Q_1 \) improves \( \tilde{Q}_1 \) in that it utilizes the information in the alternative hypothesis so that the long run fourth moment estimator does not diverge. It is powerful for a general class of nonstationary volatility models because of the nonparametric nature of the maintained hypothesis.

### 3.2 The case of independence

If we assume \( \varepsilon_t \) is iid, the CUSUMSQ test can be simplified. Here we follow the same notation system as above. Keep in mind that the same notation may have different meaning under different models.

The CUSUMSQ test \( Q_2 \) is defined the same as \( Q_1 \) except that \( \tilde{\omega}^2 \) is simplified as
\[ \tilde{\omega}^2 = n^{-1} \sum_{t=1}^{n} (u_t^2 - \tilde{\sigma}_t^2)^2. \]

Theorem 2. Assume \( \varepsilon_t \) is iid such that \( E|\varepsilon_t|^{4+\delta} < \infty \) and \( E\varepsilon_t^4 = \kappa_4 \), and that Assumption 1 (iii) holds.

(i) Under \( H_0 \) and \( H_A \),
\[ \tilde{\omega}^2 \to \omega^2, \]  
where \( \omega^2 = \lim_{n \to \infty} n^{-1} \sum_{t=1}^{n} (u_t^2 - \sigma_t^2)^2 = (\kappa_4 - 1) \int_0^1 \sigma^4(r)dr. \)

(ii) Under \( H_0 \)
\[ Q_2 \Rightarrow \sup_{0 \leq r \leq 1} |\bar{W}(r)|, \]

(iii) Under \( H_A \)
\[ n^{-1/2}Q_2 \Rightarrow (\kappa_4 - 1)^{-1/2} \Psi, \]
where \( \Psi = \left( \int_0^1 \sigma^4 \right)^{-1/2} \sup_{s \in [0,1]} \left| \int_0^s \sigma^2 - s \int \sigma^2 \right|. \]

The such simplified version of \( \tilde{Q}_1 \), denoted by \( \tilde{Q}_2 \), has been proposed in the literature; see the papers cited above. Precisely, \( Q_2 \) is defined the same as \( Q_2 \) except that \( \tilde{\omega}^2 = n^{-1} \sum_{t=1}^{n} (u_t^2 - \tilde{\sigma}_t^2)^2 \) is used instead of \( \tilde{\omega}^2 \). Under the maintained hypothesis,
\[ \tilde{\omega}^2 \to \left( \kappa_4 - \left( \int_0^1 \sigma^2(r)dr \right)^2 / \int_0^1 \sigma^4(r)dr \right) \int_0^1 \sigma^4(r)dr. \]  

Note that \( \tilde{\omega}^2 \) consistently estimates \( \omega^2 \) under \( H_0 \) but does not under \( H_A \). This induces some power loss. Under \( H_A \),
\[ n^{-1/2}Q_2 \Rightarrow \left( \kappa_4 - \left( \int \sigma^2 \right)^2 / \int \sigma^4 \right)^{-1/2} \Psi. \]  

By Cauchy–Schwarz inequality, \( \left( \int_0^1 \sigma^2(r)dr \right)^2 \leq \int_0^1 \sigma^4(r)dr \) and \( Q_2 \) has the higher power against alternatives than \( Q_2 \).
Power gain: 

\[
\sqrt{\kappa_4 - \left( \int_0^1 \sigma^2(r) dr \right)^2 / \int_0^1 \sigma^4(r) dr} / \kappa_4 - 1 \geq 1
\]

3.2.1 Implication of Gaussianity

Assuming that \( \varepsilon_t \) is iid \( N(0, 1) \), Inclan and Tiao’s (1994) test is defined as

\[
IT = \max_{1 \leq K \leq n} \sqrt{n/2}|D_K|/C_n.
\]

For this test, the nuisance parameters are estimated under the maintained hypothesis.

4 Simulations

Model 1 : \( u_t = \sigma_t \varepsilon_t, \varepsilon_t = (\alpha_0 + \alpha_1 \varepsilon_{t-1}^2)^{1/2} \eta_t, \eta_t \sim N(0, 1) \)

Model 2 : \( u_t = \sigma_t \varepsilon_t, \varepsilon_t = (\alpha_0 + \alpha_1 \varepsilon_{t-1}^2)^{1/2} \eta_t, \eta_t \sim t(5) \cdot \sqrt{0.6} \)

In both cases, \( \alpha_0 = 1 - \alpha_1 \) such that \( u_t \) has unit unconditional variance. Set \( \alpha_1 \in \{0.1, 0.5\} \).

\( \sigma_t = \sigma(t/n), \sigma(r)^2 = \sigma_0^2 + (\sigma_t^2 - \sigma_0^2)1_{r \geq \tau}, r \in [0, 1]. \)

Let \( \delta^2 = \sigma_t^2/\sigma_0^2 \) and set \( \sigma_0^2 = 1. \)

We let \( \tau = 0.5 \) and \( \delta \) ranges from 1/5 to 5. When \( \delta > 1 \), there is a variance shift from \( \sigma_0^2 \) to \( \sigma_1^2 \) at time \([n\tau]\).

\[ n = 100, m = n^{1/3}, h_i = [1 + (i - 1)/2]n^{-1/5}, \] where \( i = 1, 2, 3, 4. \)

Data-dependent truncation gives the similar results (report later)

5 Application

Appendix: Proofs

Let \( C \) be a generic constant, which can be different at different places.

Proof of Theorem 1. We shall first show that

\[
\hat{\omega}^2 = \omega^2 + O_p(m[h^2 + \sqrt{\ln n/(nh)}]), \text{ under } H_0
\]

\[
\hat{\omega}^2 = \omega^2 + O_p(m[h + \sqrt{\ln n/(nh)}]), \text{ under } H_A.
\]

Consider (14) first. For \( 0 \leq l \leq m \), write \( \hat{\gamma}_l = n^{-1} \sum_{t=l+1}^n (u_t^2 - \hat{\sigma}_t^2)(u_{t-l}^2 - \hat{\sigma}_{t-l}^2) \) then \( \hat{\omega}^2 = \hat{\gamma}_0 + 2 \sum_{l=1}^m k(l, m)\hat{\gamma}_l. \) Note that \( \hat{\gamma}_l = n^{-1} \sum_{t=l+1}^n [u_t^2 - \sigma_t^2 + (\sigma_t^2 - \hat{\sigma}_t^2)](u_{t-l}^2 - \sigma_{t-l}^2 + (\sigma_{t-l}^2 - \hat{\sigma}_{t-l}^2)) = S_1 + S_2 + S_3 + S_4 \) where the terms \( S_j \) (\( j = 1, \ldots, 4 \)) are defined below.

Let \( \gamma_l = \psi_l \int \sigma^4, \) where \( \psi_l = E(\varepsilon_{l}^2 - 1)(\varepsilon_{l-1}^2 - 1) \) is the \( l \)-th autocovariance of \( \varepsilon_{l}^2 - 1. \)
It can be shown that

\begin{align}
S_1 &= n^{-1} \sum_{t=1}^{n} (u_t^2 - \sigma_t^2)(u_{t-1}^2 - \sigma_{t-1}^2) \xrightarrow{p} \gamma_t \\
S_2 &= n^{-1} \sum_{t=1}^{n} (u_t^2 - \sigma_t^2)(\sigma_{t-1}^2 - \tilde{\sigma}_{t-1}^2) = O_p(h + \sqrt{\ln n/(nh)}) \\
S_3 &= n^{-1} \sum_{t=1}^{n} (\sigma_t^2 - \tilde{\sigma}_t^2)(u_{t-1}^2 - \sigma_{t-1}^2) = O_p(h + \sqrt{\ln n/(nh)}) \\
S_4 &= n^{-1} \sum_{t=1}^{n} (\sigma_t^2 - \tilde{\sigma}_t^2)(\sigma_{t-1}^2 - \tilde{\sigma}_{t-1}^2) = O_p(h + \ln n/(nh))
\end{align}

(15) follows from Cavaliere (2004, Lemma 4, p.284).

Now consider (16). Let \( N = \{l+1, 2, \cdots, n\} \) and \( N^* = \{t \in N : (t-l)/n \in (r^*-h, r^*+h) \} \) where \( r^* \) is a discontinuity point of \( \sigma(\cdot) \). Since \( \sigma(\cdot) \) has only finite number of discontinuities, the cardinality of \( N^* \) is \( O(nh) \). So

\begin{align}
S_2 &= n^{-1} \sum_{t \in N^*} (u_t^2 - \sigma_t^2)(\sigma_{t-1}^2 - \tilde{\sigma}_{t-1}^2) + n^{-1} \sum_{t \in N \setminus N^*} (u_t^2 - \sigma_t^2)(\sigma_{t-1}^2 - \tilde{\sigma}_{t-1}^2) \\
&= O_p(h) + O_p(\sqrt{\ln n/(nh)} + h^2) \\
&= O_p(h + \sqrt{\ln n/(nh)})
\end{align}

by the facts that \( n^{-1} |\sum_{t \in N \setminus N^*} (u_t^2 - \sigma_t^2)(\sigma_{t-1}^2 - \tilde{\sigma}_{t-1}^2)| \leq \sup_{t \in N \setminus N^*} |\sigma_{t-1}^2 - \tilde{\sigma}_{t-1}^2| \cdot n^{-1} \sum_{t \in N \setminus N^*} |u_t^2 - \sigma_t^2| \) and \( \sup_{t \in N \setminus N^*} |\sigma_{t-1}^2 - \tilde{\sigma}_{t-1}^2| = \sup_{r \in (0,1) \setminus D^*} |\sigma_{[nr]-t}^2 - \tilde{\sigma}_{[nr]-t}^2| = O_p(\sqrt{\ln n/(nh)} + h^2) \)

where \( D^* = \{r \in (0,1) : r \in (r^*-h, r^*+h) \} \) and \( r^* \) is a discontinuity point of \( \sigma(\cdot) \). Here we have used the fact that \( \tilde{\sigma}_{[nr]}^2 \) is consistent for \( \sigma^2(r) \) on \( r \in (0,1) \setminus D^* \) and \( \tilde{\sigma}_{[nr]}^2 \) converges to a weighted average of \( \sigma^2(r+) \) and \( \sigma^2(r-) \) on \( r \in D^* \) and also the uniform convergence rate of \( \tilde{\sigma}_{[nr]}^2 \) on \( r \in (0,1) \setminus D^* \). Similarly we can show (17) and (18).

Given the results above on \( S_j \), we have \( \tilde{\gamma}_t = \gamma_t + O_p(h + \sqrt{\ln n/(nh)}) \). Hence (14) holds.

Note that under \( H_0 \), the first term in (19) disappears so (13) holds.

(i) Under \( H_0 \), we shall show that (see the proof of (ii) below)

\[ n^{-1/2} D_K \Rightarrow \omega \cdot W(r). \]

Then (i) follows from (13) under \( H_0 \).

(ii) We shall show that under \( H_A \)

\[ n^{-1} D_K \Rightarrow \int_s^1 \sigma^2(r)dr - s \int_0^s \sigma^2(r)dr. \]

Let \( \zeta_t = u_t^2 - \sigma_t^2 \). It follows from Cavaliere (2004, Lemma 1, p.265) that [noting that \( \varepsilon_t^2 - 1 \) satisfies his Assumption E (p.262)]

\[ n^{-1/2} \sum_{t=1}^{K} \zeta_t = n^{-1/2} \sum_{t=1}^{K} \sigma_t^2(\varepsilon_t^2 - 1) \Rightarrow \omega W_\sigma(s), \]
where \( W_\sigma(s) = [\int_0^1 \sigma^4(r) dr]^{-1/2} \int_0^s \sigma^2(r) dW(r) \) is a time-deformed Brownian motion (e.g. Davidson, 1994). \( W_\sigma(s) \) reduces to a standard Brownian motion if \( \sigma(\cdot) \) is a constant. So

\[
n^{-1/2} \sum_{t=1}^K \zeta_t - Kn^{-3/2} \sum_{t=1}^n \zeta_t \Rightarrow \omega W(s) - \omega W(1) := \omega \overline{W}_\sigma(s),
\]

(22)

where \( \overline{W}_\sigma(s) = W_\sigma(s) - sW_\sigma(1) \). Under \( H_A \), rewrite the left-hand side of (22) as

\[
n^{-1/2} \sum_{t=1}^K \zeta_t - Kn^{-3/2} \sum_{t=1}^n \zeta_t = n^{-1/2} D_K - n^{-1/2} \sum_{t=1}^K \sigma_t^2 + Kn^{-3/2} \sum_{t=1}^n \sigma_t^2,
\]

[note that the last two terms on the right-hand side cancel under \( H_0 \) and (20) can be obtained] which gives

\[
n^{-1} D_K = n^{-1/2} \left( n^{-1/2} \sum_{t=1}^K \zeta_t - Kn^{-3/2} \sum_{t=1}^n \zeta_t \right) + n^{-1} \sum_{t=1}^K \sigma_t^2 - Kn^{-2} \sum_{t=1}^n \sigma_t^2
\]

\[
= O_p(n^{-1/2}) + n^{-1} \sum_{t=1}^K \sigma_t^2 - Kn^{-2} \sum_{t=1}^n \sigma_t^2
\]

\[
\Rightarrow \int_0^s \sigma^2(r) dr - s \int_0^1 \sigma^2(r) dr.
\]

So (21) holds. Then (ii) follows from (21) and (14).

**Proof of (8) and (9).** Define \( \tilde{\gamma}_l = n^{-1} \sum_{t=l+1}^n (u_{l+t}^2 - \tilde{\sigma}^2) \), then \( \tilde{\omega}^2 = \tilde{\omega}_0^2 + 2 \sum_{t=1}^m k(l/m) \tilde{\gamma}_l \). First we shall show that for any fixed \( l \), where \( 0 \leq l \leq m \) and \( l = [ns] \), we have

\[
\tilde{\gamma}_l \overset{p}{\to} \psi_l \int_s^1 \sigma^2(r) \sigma^2(r-s) dr + R_s
\]

(23)

where \( \psi_l = E(\varepsilon_{l+1}^2 - 1)(\varepsilon_{l-t}^2 - 1) \) and \( R_s = \int_s^1 \sigma^2(r) \sigma^2(r-s) dr - (1-s)(\int \sigma^2) \int_s^1 \sigma^2 - (1-s)(\int \sigma^2) \int_0^{1-s} \sigma^2 - (1-s)(\int \sigma^2)^2 \). Note that \( \gamma_l \to 0 \) by the \( \alpha \)-mixing assumption and \( R_s \to \int \sigma^4 - (\int \sigma^2)^2 \) as \( s \to 0 \), so \( \tilde{\gamma}_l \) is dominated by the term \( R_s \) as \( n \to \infty \) (thereby \( m \to \infty \) and \( s \to 0 \)). Having (23),

\[
m^{-1} \tilde{\omega}_l^2 = m^{-1} \tilde{\omega}_0^2 + 2m^{-1} \sum_{l=1}^m k(l/m) \tilde{\gamma}_l = O_p(1)
\]

as \( n \to \infty \) and (8) follows. Now it remains to show (23).

Rewrite \( \tilde{\gamma}_l \) as \( \tilde{\gamma}_l = n^{-1} \sum_{t=l+1}^n (u_{l+t}^2 - \tilde{\sigma}^2) = n^{-1} \sum_{t=l+1}^n [u_{l+t}^2 - \sigma_t^2 + (\sigma_t^2 - \tilde{\sigma}^2)] = T_1 + T_2 + T_3 + T_4 \) where the terms \( T_i \) \( (i = 1, \cdots, 4) \) are defined below. It can be shown that for any fixed \( l \)
\[ T_1 = n^{-1} \sum_{t=l+1}^{n} (u_t^2 - \sigma_t^2)(u_{t-l}^2 - \sigma_{t-l}^2) \xrightarrow{p} \psi_l \int_s^1 \sigma^2(r)\sigma^2(r-s)dr \]  
(24)

\[ T_2 = n^{-1} \sum_{t=l+1}^{n} (u_t^2 - \sigma_t^2)(\sigma_{t-l}^2 - \bar{\sigma}^2) \xrightarrow{p} 0 \]  
(25)

\[ T_3 = n^{-1} \sum_{t=l+1}^{n} (\sigma_t^2 - \bar{\sigma}^2)(u_{t-l}^2 - \sigma_{t-l}^2) \xrightarrow{p} 0 \]  
(26)

\[ T_4 = n^{-1} \sum_{t=l+1}^{n} (\sigma_t^2 - \bar{\sigma}^2)(\sigma_{t-l}^2 - \bar{\sigma}^2) \xrightarrow{p} \int_s^1 \sigma^2(r)\sigma^2(r-s)dr - 
\]
\[ -(1-s)(\int_s^1 \sigma^2) \int_s^1 \sigma^2 - (1-s)(\int_s^0 \sigma^2) \int_0^{1-s} \sigma^2 - (1-s)(\int_0^s \sigma^2)^2. \]  
(27)

Consider (24) first. By LLN for mixing sequences,

\[ T_1 = n^{-1} \sum_{t=l+1}^{n} (u_t^2 - \sigma_t^2)(u_{t-l}^2 - \sigma_{t-l}^2) \]
\[ = n^{-1} \sum_{t=l+1}^{n} E(u_t^2 - \sigma_t^2)(u_{t-l}^2 - \sigma_{t-l}^2) + o_p(1) \]
\[ = n^{-1} \sum_{t=l+1}^{n} \sigma_t^2 \sigma_{t-l}^2 E(\varepsilon_t^2 - 1)(\varepsilon_{t-l}^2 - 1) + o_p(1) \]

by stationarity
\[ \psi_l n^{-1} \sum_{t=l+1}^{n} \sigma_t^2 \sigma_{t-l} + o_p(1) \xrightarrow{p} \psi_l \int_s^1 \sigma^2(r)\sigma^2(r-s)dr \]

where we have used the fact that
\[ n^{-1} \sum_{t=l+1}^{n} \sigma_t^2 \sigma_{t-l} = \sum_{t=l+1}^{n} \int_{t/n}^{(t+1)/n} \sigma^2([nr]/n)\sigma^2(([nr] - [ns])/n)dr \]
\[ = \int_{(l+1)/n}^{(n+1)/n} \sigma^2([nr]/n)\sigma^2(([nr] - [ns])/n)dr \]
\[ \rightarrow \int_s^1 \sigma^2(r)\sigma^2(r-s)dr, \]

as \( n \to \infty \) for fixed \( l \). Thus (24) holds. Similarly we can show (25), (26) and (27). Therefore (23) holds. Lastly (9) follows from (21) and (8).

Proof of Theorem 2. Note that \( \hat{\rho} - 1 = (\bar{\gamma}_1 - \bar{\gamma}_0)/\bar{\gamma}_0 \). We can show that \( \bar{\gamma}_0 = \eta^2 \int g^4 - (\int g^2)^2 + o_p(\eta^2) \) and \( \bar{\gamma}_1 - \bar{\gamma}_0 = \gamma_1 - \gamma_0 + O_p(\eta^2/n) = \eta^2(\psi_1 - \psi_0) \int g^4 + o_p(\eta^2) \).
References


Table 1: Power of the CUSUMSQ tests $Q_1$ and $\widetilde{Q}_1$ under Model 1 when $\tau = 0.5$ (nominal size: 0.05)

<table>
<thead>
<tr>
<th>$\tau$</th>
<th>$\delta$</th>
<th>$\alpha_1 = 0.1$</th>
<th>$\alpha_1 = 0.5$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$\widetilde{Q}_1$</td>
<td>$Q_1$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$h_1$</td>
<td>$h_2$</td>
</tr>
<tr>
<td>0.5</td>
<td>1</td>
<td>0.044</td>
<td>0.087</td>
</tr>
<tr>
<td>1.2</td>
<td>0.168</td>
<td>0.242</td>
<td>0.238</td>
</tr>
<tr>
<td>1.5</td>
<td>0.563</td>
<td>0.691</td>
<td>0.654</td>
</tr>
<tr>
<td>2</td>
<td>0.895</td>
<td>0.957</td>
<td>0.953</td>
</tr>
<tr>
<td>2.5</td>
<td>0.979</td>
<td>0.996</td>
<td>0.989</td>
</tr>
<tr>
<td>1.2$^{-1}$</td>
<td>0.127</td>
<td>0.236</td>
<td>0.190</td>
</tr>
<tr>
<td>1.5$^{-1}$</td>
<td>0.464</td>
<td>0.609</td>
<td>0.605</td>
</tr>
<tr>
<td>2$^{-1}$</td>
<td>0.876</td>
<td>0.951</td>
<td>0.926</td>
</tr>
<tr>
<td>2.5$^{-1}$</td>
<td>0.978</td>
<td>0.996</td>
<td>0.989</td>
</tr>
</tbody>
</table>
Table 2: Power of the CUSUMSQ tests $\tilde{Q}_1$ and $\tilde{Q}_1$ under Model 1 when $\tau = 0.8$ (nominal size: 0.05)

<table>
<thead>
<tr>
<th>$\tau$</th>
<th>$\delta$</th>
<th>$h_1$</th>
<th>$h_2$</th>
<th>$h_3$</th>
<th>$h_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.8</td>
<td>1</td>
<td>0.044</td>
<td>0.087</td>
<td>0.073</td>
<td>0.062</td>
</tr>
<tr>
<td></td>
<td>1.2</td>
<td>0.129</td>
<td>0.204</td>
<td>0.162</td>
<td>0.146</td>
</tr>
<tr>
<td></td>
<td>1.5</td>
<td>0.371</td>
<td>0.514</td>
<td>0.474</td>
<td>0.434</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.793</td>
<td>0.890</td>
<td>0.845</td>
<td>0.830</td>
</tr>
<tr>
<td></td>
<td>2.5</td>
<td>0.936</td>
<td>0.975</td>
<td>0.969</td>
<td>0.960</td>
</tr>
<tr>
<td></td>
<td>1.2$^{-1}$</td>
<td>0.057</td>
<td>0.121</td>
<td>0.073</td>
<td>0.066</td>
</tr>
<tr>
<td></td>
<td>1.5$^{-1}$</td>
<td>0.108</td>
<td>0.203</td>
<td>0.167</td>
<td>0.139</td>
</tr>
<tr>
<td></td>
<td>2$^{-1}$</td>
<td>0.193</td>
<td>0.345</td>
<td>0.313</td>
<td>0.259</td>
</tr>
<tr>
<td></td>
<td>2.5$^{-1}$</td>
<td>0.287</td>
<td>0.488</td>
<td>0.415</td>
<td>0.380</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\alpha_1 = 0.5$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.054</td>
<td>0.095</td>
<td>0.081</td>
<td>0.072</td>
<td>0.065</td>
</tr>
<tr>
<td>1.5</td>
<td>0.218</td>
<td>0.336</td>
<td>0.309</td>
<td>0.276</td>
<td>0.273</td>
</tr>
<tr>
<td>2</td>
<td>0.496</td>
<td>0.639</td>
<td>0.583</td>
<td>0.555</td>
<td>0.536</td>
</tr>
<tr>
<td>3</td>
<td>0.796</td>
<td>0.898</td>
<td>0.856</td>
<td>0.841</td>
<td>0.826</td>
</tr>
<tr>
<td>5</td>
<td>0.903</td>
<td>0.965</td>
<td>0.948</td>
<td>0.940</td>
<td>0.934</td>
</tr>
<tr>
<td>1.5$^{-1}$</td>
<td>0.059</td>
<td>0.116</td>
<td>0.096</td>
<td>0.090</td>
<td>0.086</td>
</tr>
<tr>
<td>2$^{-1}$</td>
<td>0.086</td>
<td>0.171</td>
<td>0.157</td>
<td>0.122</td>
<td>0.125</td>
</tr>
<tr>
<td>3$^{-1}$</td>
<td>0.129</td>
<td>0.255</td>
<td>0.219</td>
<td>0.190</td>
<td>0.158</td>
</tr>
<tr>
<td>5$^{-1}$</td>
<td>0.179</td>
<td>0.345</td>
<td>0.278</td>
<td>0.231</td>
<td>0.204</td>
</tr>
</tbody>
</table>
Table 3: Power of the CUSUMSQ tests $Q_1$ and $\tilde{Q}_1$ under Model 2 when $\tau = 0.5$ (nominal size: 0.05)

<table>
<thead>
<tr>
<th>$\tau$</th>
<th>$\delta$</th>
<th>$h_1$</th>
<th>$h_2$</th>
<th>$h_3$</th>
<th>$h_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\alpha_1 = 0.1$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.5</td>
<td>1</td>
<td>0.034</td>
<td>0.069</td>
<td>0.054</td>
<td>0.049</td>
</tr>
<tr>
<td>1.2</td>
<td>0.088</td>
<td>0.149</td>
<td>0.134</td>
<td>0.117</td>
<td>0.113</td>
</tr>
<tr>
<td>1.5</td>
<td>0.263</td>
<td>0.390</td>
<td>0.360</td>
<td>0.343</td>
<td>0.319</td>
</tr>
<tr>
<td>2</td>
<td>0.601</td>
<td>0.734</td>
<td>0.677</td>
<td>0.663</td>
<td>0.643</td>
</tr>
<tr>
<td>4</td>
<td>0.859</td>
<td>0.922</td>
<td>0.897</td>
<td>0.880</td>
<td>0.880</td>
</tr>
<tr>
<td></td>
<td>$\alpha_1 = 0.5$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.044</td>
<td>0.089</td>
<td>0.073</td>
<td>0.051</td>
<td>0.053</td>
</tr>
<tr>
<td>1.5</td>
<td>0.176</td>
<td>0.293</td>
<td>0.261</td>
<td>0.242</td>
<td>0.196</td>
</tr>
<tr>
<td>2</td>
<td>0.364</td>
<td>0.497</td>
<td>0.477</td>
<td>0.436</td>
<td>0.429</td>
</tr>
<tr>
<td>3</td>
<td>0.590</td>
<td>0.716</td>
<td>0.681</td>
<td>0.662</td>
<td>0.633</td>
</tr>
<tr>
<td>5</td>
<td>0.719</td>
<td>0.820</td>
<td>0.773</td>
<td>0.772</td>
<td>0.744</td>
</tr>
</tbody>
</table>