2004s-25

The Impact of Sampling Frequency and Volatility Estimators on Change-Point Tests

Elena Andreou, Eric Ghysels

Série Scientifique
Scientific Series

Montréal
Mai 2004

CIRANO
Centre interuniversitaire de recherche
en analyse des organisations

The Impact of Sampling Frequency and Volatility Estimators on Change-Point Tests*

Elena Andreou[†], Eric Ghysels[‡]

Résumé / *Abstract*

Ce papier évalue la performance de plusieurs tests de changement structurel CUSUM et EDF pour la structure dynamique de la variance conditionelle et de la distribution conditionnelle. Nous étudions l'impact 1) de la fréquence des observations, 2) de l'utilisation des données de haute fréquence pour le calcul des variances conditionnelles et 3) de transformation des séries pour améliorer la puissance des tests.

Mots clés : tests de changement structurel, CUSUM, Kolmogov-Smirnov, GARCH, variation quadratique, 'power variation', données de haute fréquence.

The paper evaluates the performance of several recently proposed change-point tests applied to conditional variance dynamics and conditional distributions of asset returns. These are CUSUM-type tests for beta-mixing processes and EDF-based tests for the residuals of such nonlinear dependent processes. Hence the tests apply to the class of ARCH and SV type processes as well as data-driven volatility estimators using high-frequency data. It is shown that some of the high-frequency volatility estimators substantially improve the power of the structural breaks tests especially for detecting changes in the tail of the conditional distribution. Similarly, certain types of filtering and transformation of the returns process can improve the power of CUSUM statistics. We also explore the impact of sampling frequency on each of the test statistics.

Keywords: *Change-point tests, CUSUM, Kolmogorov-Smirnov, GARCH, quadratic variation, power variation, high-frequency data, location-scale distribution family.*

[†] Department of Economics, University of Cyprus.
[‡] Department of Economics, University of North Carolina and CIRANO, Gardner Hall CB 3305, Chapel Hill, NC 27599-3305, phone: (919) 966-5325, e-mail: eghysels@unc.edu.

1 Introduction

The importance of structural breaks in the conditional distribution of financial risk has been recognized by a number of authors. Bates (2000) shows that the distribution implicit in option prices since the crash of 1987 is substantially negatively skewed, in contrast to the essentially symmetric and slightly positively skewed lognormal distribution before the crash. Pástor and Stambaugh (2001) examine the equity premium and test for structural breaks in a long historical sample of US stock market returns. They document multiple breaks in the equity premium and find that economically sensible priors are important in estimating the equity premium as well as in identifying the most likely dates at which breaks occurred. Ghysels (1998) examines in a similar vein the time variation in conditional beta models and finds structural breaks in the beta dynamics. Timmerman (2001) also presents empirical evidence on the existence of structural breaks in the fundamentals process underlying US stock prices. In related work Andreou and Ghysels (2002b, 2003b) find multiple breaks in volatility dynamics and co-movements of FX and stock returns associated with the Asian and Russian financial crises. Ignoring change-points in financial time-series may also lead to spurious volatility persistence (see e.g. Diebold, 1986, Hendry, 1986) as well as a long memory (see e.g. Diebold and Inoue, 2001, Mikosch and Starica, 2003).

So far the arguments seem to point towards testing for breaks in the conditional mean and/or variance. In testing for change-points in the conditional variance it is important to acknowledge that ARCH-type processes are β-mixing (e.g. Carrasco and Chen, (2001), Davis and Mikosch (1998) among others) which precludes the application of many of the change-point tests developed for weakly dependent processes. This is because ϕ-mixing does not imply β-mixing. A number of recent studies, however, develop structural break tests for ARCH-type processes. Chu (1995) and Lundberg and Terasvirta (2002) propose LM type tests, de Lima (1998) considers a recursive BDS test for testing nonstationarities, Kokoszka and Leipus (1999, 2000) specify CUSUM tests for ARCH processes and Mikosch and Starica (2003) develop tests based on the data periodogram to detect breaks in the long memory and GARCH dynamics of asset returns.

In many circumstances the need to test for breaks in higher moments, or more precisely the tail behavior of the conditional distribution is very acute as well. Andreou and Ghysels (2003a) argue that testing for breaks in conditional distributions is important in the context of risk management since

extreme tail observations and structural breaks are two types of rare events. The former is one of the main focuses of financial risk management, the latter represents a fundamental shift in the distribution of risky outcomes. Andreou and Ghysels (2003a) therefore suggest to perform "quality control" for risk exposure in order to make sure that no breaks occur in the samples used to estimate risk management measures such as Value-at-Risk and Expected Shortfall.

In addition to CUSUM tests, tests for distributional change based on the residuals of nonlinear dynamic models have also been recently proposed. For instance, Empirical Distribution Function (EDF) tests for the residuals of ARCH-type models can be found in Horvath et al. (2001). Koul (1996, 2002) and Koul and Mukherjee (2002) generalize these results for more general nonlinear dynamic location-scale models and extend it to other types of estimators. These residual-based tests aim to examine breaks in the conditional distribution of the process. Giraitis et al. (1996) and Inoue (2001) also propose EDF statistics for strongly mixing processes. However, these are based on the actual process (as opposed to the residual process). Although residual-based EDF tests yield asymptotically distribution free Kolmogorov-Smirnov type tests when in general the same tests are applied to the actual process (not the residuals) their limiting distributions depend on the nuisance parameters and the use of bootstrap methods (Inoue, 2001). Using different methods Lavielle and Moulines (2000) and Lavielle (1999) propose penalized least squares minimum contrast estimators for strongly dependent and strongly mixing processes to detect multiple change points in the dynamics and the marginal distribution of a process, respectively. Last but by no means least, Quintos et al. (2001) recognize the importance of testing for breaks in the tails of financial series and use recursive and sequential Hill estimators.

Our objective is to examine further the CUSUM-type test in Kokoszka and Leipus (1999, 2000) and the EDF-based tests in Horvath et al. (2001) and Koul and Mukherjee (2002). Our analysis presents the following extensions:
First, for the EDF-type tests we consider different conditional variance estimators based on high-frequency data-driven volatility filters. Horvath et al. (2001) suggest least squares and QML GARCH estimates. On the other hand, Koul (2002) and Koul and Mukherjee (2002) address robust estimators such as M- and minimum distance (m.d.-) estimators. Limited simulation and empirical evidence exists regarding the performance of these tests for

2

either family of estimators except for Kokoszka and Teyssièrre (2002) who focus on the QMLE. Our results show that the choice of the volatility estimator is crucial for the power of the EDF tests. Instead of focusing on robust and QMLE estimators for specific GARCH parameterizations driven by innovation in the location-scale family, we choose to complement these tests with high-frequency estimators such as the Quadratic Variation (QV), the Power Variation (PV) and the Historical QV. These are high-frequency data driven volatility estimators discussed in recent work of Andersen et al. (2003), Andreou and Ghysels (2002a), Barndorff-Nielsen and Shephard (2002a, b, 2003a), among others. One of the main reasons is that these estimators are valid for general semimartingale models and can be considered as volatility filters for general types of GARCH and SV models under much more general conditions. The properties of these estimators and some of their advantages especially with respect to change-point tests are discussed in detail. As a primer to our results we show that the high-frequency PV normalized returns enjoys the highest power in detecting change points under all alternative volatility and distributional changes even for small samples, small changes and end-of-sample instability, compared with the other high-frequency and daily volatility estimators. One of the key factors of these EDF tests is the dependence structure of the ARCH scaled residuals. Horvath et al. (2001), Koul (2002) and Koul and Mukherjee (2002) show that despite the dependence of residuals in parametric models due to estimation error, one can still apply EDF-type tests. Andreou and Ghysels (2003b) use the asymptotic distribution theory of Barndorff-Nielsen and Shephard to argue that despite the presence of estimation error when high-frequency filters are used, one can use under certain conditions EDF tests. In this paper we present simulation evidence based on the Hong and Lee (2003) test showing that returns scaled by high-frequency volatility filters support the independence assumption under a GARCH generated process. In addition, it is also shown that the possibility of weak dependence in normalized returns due to temporal aggregation and estimation error does not affect seriously the performance of these tests.

Second, for the CUSUM procedure proposed in Kokoszka and Leipus (1999, 2000) we also find that high-frequency volatility filters perform better instead of the subordinate observed daily returns absolute and quadratic daily returns proposed by the original authors. These high-frequency volatility processes satisfy the mixing conditions required by the tests and represent more efficient approximations of the volatility since they are based on the

3

intra-day cumulative information. It is interesting that they also yield power in detecting not only small shifts in the volatility parameters but also (unconditional) distribution changes. As a final finding we note that the CUSUM statistic is standardized by a HAC type estimator and it is shown that the ARHAC estimator by Den Haan and Levin (1997) for the daily absolute and squared returns shares good asymptoric properties (Andreou and Ghysels, 2002b). However, for highly persistent high-frequency processes addressed here the choice parameters of the ARHAC estimator also depends on the level of aggregation, returns transformation and filtering of the process.

The paper is organized as follows. In section 2 we present the EDF-based tests that examine the hypothesis of distributional homogeneity based on the residuals of nonlinear time series models and extend this analysis to financial returns processes normalized by high-frequency volatility filters. In section 3 we consider the CUSUM-type test for assessing changes in the volatility structure or tail related processes of financial series. A final section concludes the paper.

2   EDF tests for distributional change

In discrete time we assume that financial data are generated by the following type of processes:

$$r_t = \mu_t + \sigma_t \varepsilon_t \tag{2.1}$$

where r_t is the return on an asset (or a portfolio of assets), σ_t is the conditional volatility, μ_t is the conditional mean and ε_t is an i. i. d. process. Parametric models within the class of ARCH and SV models can be nested into equation (2.1). The former class of models assumes that σ_t is a measurable function of past returns, whereas the latter assumes that volatility is latent (see Bollerslev et al. (1994) and Ghysels et al. (1996) for further discussions). It should be noted that the specification in (2.1) amounts to working with the location-scale family of distributions for the conditional distributions once the conditional mean and variance are determined. It is well known that in the class of the location-scale family of distributions the pivotal quantity is $\varepsilon = (r - \mu)/\sigma$ whose distribution $F(.)$ depends neither on the location μ nor the scale σ parameters. Some members of this family featured in finance are the Normal, Student's t, Generalized Error, Generalized Extreme Value (GEV) such as the Frechet distribution and the Stable distributions such as the Lévy distribution. It should also be noted that

(2.1) allows for the presence of a jump component, which we assume to be incorporated in ε_t. In the absence of leverage effects equation (2.1) can also be viewed as a discretization of continuous time jump-diffusion (see e. g. Ghysels et al. (1996) for further discussion).

For the general model (2.1) the conditional moments μ_t and σ_t need to be estimated. We will operate with both parametric and data-driven estimates $\hat{\sigma}_t$, or related processes and consider the normalized returns process $\hat{X}_t \equiv g[(r_t - \hat{\mu}_t)/\hat{\sigma}_t]$. To estimate $\hat{\sigma}_t$ one may follow two strategies. The first consists of assuming a parametric model, this typically being an ARCH-type model and hence formulate $\hat{\sigma}_t \equiv \hat{\sigma}_t(\mathbf{b})$ where $\mathbf{b}$ is a parameter vector governing the conditional mean and volatility. Horvath, Kokoszka and Teyssière (2001), Koul (2002) and Koul and Moukherjee (2002) provide recent contributions on the empirical process properties of $X_t(\hat{\mathbf{b}})^2$. The second strategy involves using high-frequency volatility filters. This is one of our objectives analyzed below.

In a first subsection we review the EDF tests for (semi)parametric models. In the second subsection we extend this test for returns normalized by HF volatility filters. In the last subsection we present simulation evidence for the historical performance of these tests for alternative high-frequency volatility estimators.

2.1   EDF tests for (semi)parametric models

Horvath, Kokoszka and Teyssière (2001) (henceforth HKT) and Koul (2002) show that unlike the residuals of ARMA processes (e.g. Bai, 1994), the residuals of the ARCH models yield sequential empirical processes that do not behave like asymptotically independent random variables. In particular they show that the asymptotic distribution involves among others a term depending on the unknown parameters of the model. However, in certain interesting cases, including the detection of changes in the distribution function of unobserved innovations, the sequential Empirical Distribution Function (EDF) tests yield asymptotically distribution free statistics (see e.g. Bai (2003)).

The discussion below holds for a Generalized ARCH model and for simplicity purposes we focus on the ARCH model (Engle, 1982) for the returns

process $\{r_t\}$ given by:[1]

$$r_t = \sigma_t \varepsilon_t \qquad (2.2)$$

$$\sigma_t^2 = b_0 + \sum_{j=1}^{p} b_j r_{t-j}^2 \qquad (2.3)$$

where the volatility $\{\sigma_t\}$ is assumed to be driven by an *i.i.d.* innovation process $\{\varepsilon_t \ -\infty < t < \infty\}$ with mean zero and unit variance and $b_j > 0$. The conditional distribution change-point analysis is based on the normalized returns process ε_t in (2.2) which is given by $X_t := (r_t/\sigma_t) := \varepsilon_t$. HKT assume the following two model conditions: (M1) The unknown density function f of ε_t exists and is continuous. (M2) The innovation process has a finite fourth moment, $E(\varepsilon_0^4) < \infty$, and the ARCH process satisfies the stationarity condition:[2]

$$E(\varepsilon_0^4)^{1/2} \sum_{1 \le j \le p} b_j < 1. \qquad (2.4)$$

Hence the ARCH equations have a unique *strictly* stationary solution such that $E(r_t^4) < \infty$ and the squares r_t^2 have a Volterra representation

$$r_t^2 = \sum_{l=0}^{\infty} \sum_{j_1,\dots,j_l}^{p} b_{j_1}...b_{j_l} \varepsilon_t^2 \varepsilon_{t-j_1}^2 ... \varepsilon_{t-j_l}^2.$$

Thus r_t^2 is a function of $\varepsilon_t, \varepsilon_{t-1}, ...$, and so it follows that $\{r_t^2\}$ is ergodic. It is also well-known that ARCH-type sequences are not only ergodic but mixing with geometric rate (Chen and Carrasco, 2001).

The properties of the normalized returns (or residual) empirical process intimately depend on the family of estimators of the parameter vector $\mathbf{b} = (b_0, b_1, ...b_p)$. They are assumed to satisfy the conditions of asymptotic linearity and $\sqrt{T}$-consistency given by:

$$\hat{b}_i - b_i = \frac{1}{T} \sum_{1 \le t \le T} l_i(\varepsilon_t^2) f_i(\varepsilon_{t-1}, \varepsilon_{t-2}, ...) + o(T^{-1/2}), 0 \le i \le p \qquad (2.5)$$

The functions l_i and f_i above are regular in the sense that:

$$E(l_i(\varepsilon_0^2)) = 0, \quad E\left[l_i(\varepsilon_0^2)\right]^2 < \infty, \quad E\left[f_i(\varepsilon_0, \varepsilon_1, ...)\right]^2 < \infty, 0 \le i \le p \qquad (2.6)$$

---

[1]In the discussion below we present more general nonlinear dynamic structures that incorporate various GARCH specifications and nonlinear AR models. We also suppress for simplicity the presence of a mean component in (2.2).

[2]It is worth noting that Inoue (2001), Giraitis et al. (1997), Mikosch and Starica (1999), Quintos et al. (2001) also require existence of the fourth moment of the *return process itself, rather than the residuals.* The above approach requires moment restrictions on ε_t which can be easier to evaluate if such an assumption is verifiable as opposed to the analogous condition on the strongly dependent process r_t.

Berkes et al. (2002) show that the above conditions hold for GARCH(p, q) processes.

The sequential (or two-time parameter) empirical process of the squared residual process $\hat{X}_t = r_t^2/\hat{\sigma}_t^2$ is defined by:

$$\hat{e}_T(x, s) = T^{1/2} s \left(\hat{F}_T(x, s) - F(x) \right)$$

where

$$\hat{F}_T(x, s) = \begin{cases} \frac{1}{Ts} \sum_{p < t \leq Ts} \mathbf{1}\{\hat{X}_t \leq x\} & if \quad p/T < s \leq 1 \\ \\ 0 & if \quad 0 \leq s \leq p/T. \end{cases}$$

Theorem 1.1 in Horvath et al. (2001, p.3) shows that under the above conditions this converges to a Kiefer process:

$$\hat{e}_T(x, s) \rightarrow \Gamma(x, s) \tag{2.7}$$

where $\Gamma(x, s)$ is the limiting Gaussian process with zero mean and covariance function Γ. This is equivalent to the convergenence $\hat{e}_T(F^{-1}(x), s) \rightarrow \Gamma(F^{-1}(x), s), 0 \leq x, s \leq 1$, in $D([0, 1] \times [0, 1])$. However, its covariance depends on several unknown parameters and functions involving the innovation process, its density function, the volatility process and their expected values. In particular it is shown in HKT (2001, Theorem 1.1) that

$$\sup_{0 \leq x < \infty} \sup_{0 \leq s \leq 1} \left| \hat{e}_T(x, s) - \left(e_T(x, s) + x f(x) s \sum_{1 < i \leq p} T^{1/2}(\hat{b}_i - b_i)\beta_i \right) \right| = o_p(1)$$

$$\tag{2.8}$$

and therefore the joint convergence of $e_T(x, s)$ and $\sqrt{T}(\hat{\mathbf{b}}_T - \mathbf{b})$ imply the result in (2.7).

The above results imply that Empirical Distribution Function (EDF) of the squared returns normalized by the estimated ARCH variance ($\hat{X}_t \equiv \hat{\varepsilon}_t^2$) can be used to study the distribution change-point problem just like the EDF tests for an *i.i.d.* process, the latter being a widely used and studied statistical problem (e.g. Csörgö and Horvath, 1997, section 2.6, Szyszkowicz, 1998). Therefore according to Horvath et al. (2001) the sequential two-parameter

EDF process, $w_T(x, s)$ for $\{\hat{\varepsilon}_t^2\}$ is:

$$w_T(x, s) = \begin{cases} 0 & 0 \le s \le p/T \\ \frac{[Ts](T-[Ts])}{T^{3/2}} \left(\hat{F}_T(x, s) - \hat{F}_T^*(x, s) \right) & p/T < s \le (T-1)/T \\ 0 & (T-1)/T < s \le 1, \end{cases} \tag{2.9}$$

where $\hat{F}_T^*(x, s) = 1/(T - Ts) \sum_{Ts < t \le T} \mathbf{1}\{\hat{\varepsilon}_t^2 \le x\}$, or equivalantly, $\hat{F}_T^*(x, s) = 1/(T - Ts) \sum_{Ts < t \le T} \mathbf{1}\{\hat{X}_t \le x\}$ so that it compares the EDF of $\hat{X}_{p+1}, ..., \hat{X}_{[Ts]}$ to that of $\hat{X}_{[Ts]+1}, ..., \hat{X}_T$. What is interesting about (2.9) is that the third term in (2.8), $xf(x)s \sum_{1 < i \le p} T^{1/2}(\hat{b}_i - b_i)\beta_i$, in the two-parameter process cancel out so that the process $\{w_T(x, s)\}$ converges to a tied-down Kiefer process. Hence the following well known statistics can be used to examine the distributional homogeneity hypothesis. The supremum statistic is:

$$\sup_{0 \le x < \infty, 0 \le s \le 1} |w_T(x, s)| \xrightarrow{d} \sup_{0 \le u, s \le 1} |K^*(u, s)| \tag{2.10}$$

where K^* is a "tied" down Kiefer process. This statistic has an asymptotic distribution equivalent to the Kolmogorov-Smirnov statistic and has been studied for time series models originally by Picard (1985). Similarly since F is continuous the quadratic statistic:

$$\int_0^1 \int_0^\infty w_T^2(x, s) d\hat{F}_T(x, 1) ds \xrightarrow{d} \int_0^1 \int_0^1 [K^*(u, s)]^2 du ds \tag{2.11}$$

has an asymptotic distribution equivalent to the Cramer-von Mises representation studied for instance in Blum et al. (1961).

Commonly used estimators for GARCH models admit the representation (2.5) e.g. conditional likelihood, pseudo maximum likelihood, conditional least squares estimators. Moreover, in general (2.6) can be considered as score functions which are assumed to be bounded, nondecreasing and real-valued, and can be related to robust estimators. Koul and Koul and Mukherjee (2002) extend this to a class of nonparametric estimators, namely M- and minimum distance (m.d.) estimators. In the context of this simple ARCH model Koul and Mukherjee (2002) show the asymptotic relative efficiency of an M-estimator for **b** relatively to the QMLE. In addition they prove the above limiting distribution of the weighted sequential statistic (2.9) for the residuals of a general location-scale dynamic model which is useful for change-point detection.

2.2 EDF tests and high-frequency volatility estimators

The above analysis considers nonlinear time series models where the structure of the conditional mean and variance needs to be a priori specified including a set of stationarity and moment conditions. Parametric or semiparametric nonlinear time series belong in this approach where for instance the conditional moments are specified whereas the distribution of the innovations is left unspecified and satisfies some mild conditions relating to continuity and decreasing tail behavior. Consider for instance the ARCH model restrictions (M1) and (M2) in the previous section. These conditions in the semiparametric model along with certain estimator properties allow us to derive the asymptotic distribution of the residual based EDF tests. In this section we will consider more general models that belong to the family of semimartingale models for financial asset price processes and again we shall focus on estimators of volatility. We shall consider estimators of the integrated volatility (discussed in more detail below).

We consider the general semimartingale process for r_t widely used in finance. A semimartingale is a process right continuous with left limits (a cadlag process) of the form:

$$r_t = r_0 + M_t + A_t$$

where r_0 is finite-valued and $\Im_0$-measurable, M is a local martingale and A some process of finite variation. Semimartingale models include the continous diffusions, jump-diffusions, hence stochastic volatility models as well as Lévy processes and most additive processes.[3] Take for instance the SV model for log prices $p^*(t)$:

$$p^*(t) = a^*(t) + \int_0^t \sigma^{1/2}(u)dw(u) \tag{2.12}$$

where the processes $\sigma^{1/2}$ and $a^*(t)$ are assumed to be stochastically independent of the Brownian motion w. The instantaneous or spot volatility is given by $\sigma^{1/2}$ and $a^*(t)$ represents the mean of the process. For example $a^*(t) = \mu t + \beta \sigma^{p*}(t)$ where $\sigma^{p*}(t) = \int_0^t \sigma^p(u)du$. The process $\sigma^{p*}(t)$ is called the integrated volatility of power $p > 0$. In view of the cadlag assumption $a^*(t)$ is a continuous local martingale (i.e. predictable) and the volatility can have deterministic effects, leverage, jumps, long memory, be non-stationary.

---

[3]The former are a special case of the latter in that they have not only independent but also stationary increments.

The idea that volatility can be precisely estimated using high-frequency data goes back at least to Merton (1980) and has been the subject of recent research given the availability of high-frequency financial data. High-frequency data volatility estimators computed as the sum of high-frequency intra-daily returns involve a discretization based on m intradaily returns and pertains to the increments in the quadratic variation of day t, $\int_{t-1}^{t} \sigma_u^2 du$. The Quadratic Variation is defined as $QV_{m,t} = \sum_{j=0}^{m-1} [r_{t-j/m}^{(m)}]^2$ and the Power Variation is defined as $PV_{m,t} = \sum_{j=0}^{m-1} |r_{t-j/m}^{(m)}|$. These processes have been studied extensively by Andersen et al. (2003), Barndorff-Nielsen and Shephard (2002a,b, 2003a), among others. In addition, Barndorff-Nielsen and Shephard (2003a) also consider the so called realized power variation where $p = 1$, i.e. the cumulative sum of absolute intra-daily returns is computed. We shall use the notation P^r to refer to a general specification, with special cases QV ($p = 2$) and PV ($p = 1$). $QV_{m,t}$ and $PV_{m,t}$ represent two estimators of the integrated volatility used in the remainder of the paper. The QV estimators have been used quite extensively since many years, see for example French, Schwert and Stambaugh (1987), Poterba and Summers (1987), Schwert (1989), Andersen and Bollerslev (1998), Andersen et al. (2003), Andreou and Ghysels (2002b), among others.

Barndorff-Nielsen and Shephard (2002a,b) show that the realized volatility $QV_{m,t}$ has an asymptotic distribution that follows a mixture of normal distributions (see equation (2.13) below), a result obtained using "fill-in" asymptotic arguments keeping the interval $[t-1, t]$ fixed and letting $m \to \infty$. Similarly, Barndorff-Nielsen and Shephard (2003a) show the realized power variation, with $r > 1/2$ has an asymptotic distribution:

$$\frac{\log[\eta_r^{-1} m^{1-r/2} P_{m,t}^r] - \log P_t^r}{\eta_r^{-1} m^{1-r/2} \sqrt{\frac{\eta_r \nu_r P_{m,t}^{2r}}{[\eta_{2r} m^{1-r/2} P_{m,t}^r]}}} \xrightarrow{L} N(0,1) \tag{2.13}$$

where $\eta_r = E[|u|^r]$ and $\nu_r = Var[|u|^r]$ with u standard Gaussian. Note that the case of $r = 2$ covers the quadratic variation and hence (2.13) generalizes earlier results on quadratic variation estimation, notably by Barndorff-Nielsen and Shephard (2002 a,b). These results show that $QV_{m,t}$ has $\sqrt{m}$ rate of convergence, that the limit is unaffected by the drift of the process a^*, that volatility dynamics do not play a role for the limit behavior, that the volatility process σ can be non-stationary, exhibit long-memory or include intra-day effects. Moreover, existence of the fourth moment is not required

for the asymptotic normality to hold since this implies that the stochastic denominator in (2.13) would not possess an unconditional mean. Last but not least the results show that $QV_{m,t} - \int_{(t-1)h}^{th} \sigma(u)^2 du$ has a mixed Gaussian limit implying that marginally it will have heavier tails than the normal. Simulation evidence presented by Barndorff-Nielsen and Shephard (2003a,b) shows that the log transformation of $QV_{m,t}$ and $PV_{m,t}$ has relatively better asymptotic properties for $m > 48$ and perform well under alternative volatility specifications (including long-memory, leverage).

Further advantages of high-frequency filters can be found in Andersen and Bollerslev (1998) although it is important to acknowledge that these estimators are not without caveats. These estimators are valid for actively traded markets. Also very little is known about the asymptotic properties of $QV_{m,t}$ when the return process is generated by a jump diffusion or when the diffusion features leverage. Barndorff-Nielsen and Shephard show via simulation that leverage does not affect the asymptotic properties. We use an alternative data-driven volatility that could deal with some of these issues. Recently, Alizadeh et al. (2002) suggest to use the daily range as a measure of volatility instead of the increments in quadratic variation. The range estimator was initially considered as an extreme value estimator for indepedent processes (e.g. Feller, 1951, Parkinson, 1980). The appeal of the daily range is that is measured without error, unlike the quadratic and power variation estimators discussed above. To proceed let us define the daily range, namely:

$$RA_{m,t} = (\sup_{t-1/m<\tau\leq t} p_\tau - \inf_{t-1/m<\tau\leq t} p_\tau)^2 \tag{2.14}$$

Let us also consider the stochastic process $X_t \equiv r_{m,t}^2/RA_{m,t}$.[4] Note that unlike the quadratic variation, we no longer hold the time interval fixed, instead we consider returns over a ever shrinking time interval and the associated range. Note also that we do not denote the process as $\hat{X}_t$ since in principle no estimation is involved. Obviously we can view the range as an estimator of volatility, yet we do not directly link it to the parameters of the underlying process, nor to its quadratic variation. Such links can only be established in some special cases as discussed by Alizadeh et al. (2002). We view the ratio $X_t \equiv r_{m,t}^2/RA_{m,t}$, as a process not involving parametric or non-parametric estimation. In practice we will have to assume that we sample over sufficiently small intervals, instead of sampling squared returns over

---

[4]The volatility estimator based on the range as it appears in Parkinson is $\hat{\sigma}^2 = RA_{m,t}/(4\ln 2)$.

some sufficiently small subintervals of $[t-1,t]$ to compute quadratic variation estimators. Such asymptotic analysis is reminiscent of the continuous record asymptotic theory of Foster and Nelson (1996). Using arguments similar to those of Foster and Nelson, Andreou and Ghysels (2002c) show that $X_t \equiv r_{m,t}^2/RA_{m,t}$ becomes an $i.i.d.$ random variable for m sufficiently large, or equivalently for sufficiently short sampling intervals. The basic intuition driving the results is that returns, $r_{m,t}/\sqrt{RA_{m,t}}$, over short intervals appear like approximately $i.i.d.$ with zero conditional mean and finite conditional variance and have regular tail behavior. Using the range, instead of a rolling sample estimator of instantaneous volatility, as advocated by Foster and Nelson, removes the requirement of a smooth volatility process, so that jumps can be accommodated. It should also be noted that we can accommodate leverage effects with the range.

The above high-frequency volatility filters have important implications for our change-point analysis for the following reasons:

(i) The semimartingale framework allows us to nest many of the volatility models proposed for modeling financial asset processes. Hence we do not have to a priori specify a given parametric model for the conditional variance before change-point analysis can be applied. This obviously avoids the possibility of spurious change-point detection due to misspecification in the functional form of the conditional moments or the dynamics of the process. In a parametric model misspecification of the conditional variance function would in general invalidate the consistent estimation. Hence under a misspecified parametric model the standardization would be incorrect, even asymptotically, which would influence the test properties. The above high-frequency $QV_{m,t}$ and $PV_{m,t}$ filters are shown to be consistent estimators for alternative semimartingale models as opposed to for instance QMLE or other estimators.

(ii) The derivation of the quadratic and power variation estimators of volatility and their limiting distributions as described above do not impose existence of the fourth moment as required by the QMLE GARCH estimates - a condition that is often challenged empirically. Related is the stylized fact of heavy tailed distributions for returns for which Hall and Yao (2003) show that the QMLE has a multivariate stable limiting distribution and exhibits poor rates of convergence. In contrast, under heavy tailed distribution such as Lévy processes the integrated volatilities of power p not only enjoy standard rates of convergence but also provide the flexibility for allowing the pth power to vary according to the jump activity (Woerner, 2002).

(iii) The information set of $QV_{m,t}$ and $PV_{m,t}$ exploits the intra-day in-

formation of usually a day. Andreou and Ghysels (2002a) consider high frequency estimators over a longer horizon of usually two days presented by the Historical Quadratic Variation (HQV) which is defined as the sum of m rolling QV estimates. The distinguishing feature of all high-frequency volatility estimators is that they do not involve aggregation or smoothing over long horizons like daily data driven estimators such as the 26-day rolling volatility in Foster and Nelson (1996) and the RiskMetrics (used by practitioners) or IGARCH which is essentially an exponentially weighted moving average filter. It is shown in the simulation section that these inter-day filters, as opposed to the intra-day ones, smooth out the change points in the volatility dynamics and tails and thereby yield poor power for the change point tests considered.

2.3   EDF tests and temporal dependence

As shown by Horvath et al. (2001), Koul (2002) and Koul and Mukherjee (2002), despite the dependence in residuals created by estimation error one can still apply EDF-based tests. It is important to note, however, that parametric specifications assume a *strong* GARCH process, using the terminology of Drost and Nijman (1993). More specifically, they impose that the (true) normalized returns (without estimation error) are *i.i.d.* for a particular sampling frequency. Strong GARCH models do not temporally aggregate, so that imposing an *i.i.d.* assumption at one frequency results in temporal dependence at another frequency. Consequently, parametric models come with a restrictive distributional assumption. Obviously, it is important to acknowledge that the process $\hat{X}_t \equiv r_{1,t}/\sqrt{QV_{m,t}}$ is not the same as ε_t from the estimation of (2.2)-(2.3) via say QMLE. Do high-frequency data volatility processes carry similar restrictions? Temporal dependence may indeed exist in $\hat{X}_t \equiv r_{1,t}/\sqrt{QV_{m,t}}$ or $\hat{X}_t \equiv r_{1,t}/\sqrt{PV_{m,t}}$ due to the following three sources: (i) measurement or estimation error, (ii) misspecification and (iii) temporal aggregation. Consequently the properties of residual-based tests may be affected depending on the degree of temporal dependence in $\hat{X}_t$. It turns out that there are some conditions required for returns normalized by high-frequency volatility estimators to be *i.i.d.* and this is the subject we discuss here.

Given the above properties of high-frequency volatility filters we consider the normalized returns $\hat{X}_t$ based on the $QV_{m,t}$. Andreou and Ghysels (2003b) use the result in (2.13) to show that $\hat{X}_t$ is *i.i.d.* if $P_{m,t}^4/[P_{m,t}^2]$ is *i.i.d.* This

process involves verifying that the normalized returns is independent if the intra-day sum of the fourth power of absolute returns scaled by $QV_{m,t}$ squared is also independent.

It is clear that some appealing features of high-frequency filters emerge here. For parametric specifications we have to *assume* as strong GARCH specification and carry out change-point tests. For high-frequency filters we have some *empirical processes* (namely $P_{m,t}^4/[P_{m,t}^2]$) that allows us to verify certain conditions that warrant the assumption *i.i.d.* and hence applicability of EDF tests.

As a first step towards establishing the dependence structure of normalized returns by high-frequency volatility filters we apply the Hong and Lee (2003) test which examines the following null hypotheses: (H1). Martingale difference, (H2). Serial correlation under heteroskedasticity, (H3). ARCH-in-mean under heteroskedasticity, (H4). Skewness-in-mean under heteroskedasticity, (H5). Kurtosis-in-mean under heteroskedasticity. This test is applied to normalized returns for alternative high-frequency volatility filters as well as to intraday normalized returns such as those at hourly and six-hourly frequency (discussed further below). Note that other temporal dependence can also be applied to $\hat{X}_t$. However, they all assume that the process is stationary. Therefore it is useful to investigate via a simulation study the Hong and Lee test results relating to the temporal dependence of $\hat{X}_t$. In the presence of no support against the null hypothesis the properties of the EDF test are not expected to be violated due to temporal dependence. However, if support is found against any of the null hypotheses (H1) to (H5) then the bootstrap approach can provide results regarding the effects of weak temporal dependence on the EDF change-point test. Simulation evidence is provided regarding the properties of these tests in the next section.

2.4   Simulation evidence for the EDF test

The objective is to examine the properties of the EDF-based test is examined for the squared normalized returns of a general family of ARCH processes and high-frequency estimators. This subsection is organized as follows: First, we outline the Monte Carlo design. Second, we present simulation results that, based on the Hong and Lee (2003), test that assess the temporal dependence structure of normalized returns. Third we examine via simulations the size and power of EDF tests for alternative normalized returns processes.

2.4.1 The Monte Carlo design

The simulated returns process $r_{(m),t}$ sampled at frequency $1/m$ is generated by a GARCH(1,1) model (e.g. Bollerslev et al. 1994):

$$\ln p_t - \ln p_{t-1/m} \equiv r_{(m),t} = \mu_{(m),t} + \sigma_{(m),t} \cdot z_{(m),t}$$
$$\sigma^2_{(m),t} = b_{0,(m)} + b_{1,(m)} r^2_{(m),t-1/m} + \gamma_{(m)} \sigma^2_{(m),t-1/m}, \quad t = 1, ..., T. \quad (2.15)$$

where $z_{(m),t}$ is $i.i.d.(0,1)$ and $\sigma^2_{(m),t}$ is the volatility process. The Data Generating Processes (DGP) at the 5-minute frequency are defined by a low persistent GARCH (DGP1) where $b_{0,(m)} = 0.1$, $b_{1,(m)} = 0.3$, $\gamma_{(m)} = 0.3$ and by a high persistent GARCH (DGP2) $b_{0,(m)} = 0.00044$, $b_{1,(m)} = 0.06341$, $\gamma_{(m)} = 0.93482$ which can be considered representative processes of financial asset returns. Note that at the 5-minute frequency $m = 288$ for the 24-hour traded markets. A one year sample is considered equivalent to $T_{days} = 250$ and $T_{5\min s} = 72000$ daily and 5-minute observations, respectively. The small sample choice is related not only to the evaluation of the finite sample properties of the aforementioned change-point tests but also to the financial regulators directive which recommends that the banking institutions use at least a year of historical data for evaluating their risk exposure (in an attempt to avoid extreme events and structural changes in long historical samples). Based on the GARCH process in (2.15) generated at the 5-minute interval we temporally aggregate returns at other frequencies such as daily and hourly without imposing the assumption that at those frequencies the data are driven by a GARCH process with $i.i.d.$ errors. The reason being that we do not wish to impose the strong GARCH process assumption (Drost and Nijman, 1993) at all frequencies given that these do not temporally aggregate. For instance, the daily returns and high-frequency volatility filters are based on a simple aggregation of the 5-minute returns generated from the DGP in (2.15) and are therefore not $i.i.d.$ by construction.

The process in (2.15) is driven by a homogeneous white noise with neither breaks in the conditional variance nor level shifts denotes the process under the null hypothesis.[5] The returns process is de-volatilized using methods discussed in section 2.2.[6] Under the alternative hypothesis the normalized

---

[5]It is assumed that $\mu_{(m),t} = 0$ for simplicity purposes and for focusing on scale and tail change point alternatives that are more challenging to detect as opposed to mean shifts.

[6]Some simulation evidence regarding the dependence structure of $\hat{X}_t = r_t/\hat{\sigma}_t$, under the null hypothesis, for alternative data driven estimators is reported in Andreou and

returns process is assumed to exhibit change-points and the following independent simulated processes are generated for the evaluation of the tests' power. We start by examining breaks in *the conditional variance dynamics* $(\sigma_{(m),t})$ which can be thought as permanent regime shifts in volatility at point πT (where $\pi = .8$ represents an end-of-sample instability). Such breaks may be due to an increase in the intercept, $b_{(m),t}$, or a shift in the volatility persistence, $b_{1,(m)} + \gamma_{(m)}$ or both. Second, we consider *a change in the tails* of the d.f. from $N(0,1)$ to $t(0,1;\nu)$ or from $t(0,1;\nu_1)$ to $t(0,1;\nu_2)$. For high-frequency processes we also consider a change point from a Lévy to a Cauchy innovation.[7] The distribution shift is an interesting alternative for financial stock returns due to the plethora of empirical evidence for different heavy-tailed distributions proposed to model stock returns. Also it is often common practice to fit such alternative distributions to asset returns assuming (and rarely testing) that the sample is homogeneous ignoring the possibility that certain distributional characteristics (such as heavy tails or asymmetries) may spuriously exist due to breaks. Similarly, spurious effects of volatility persistence and long memory in stock returns have been reported due to ignored structural change effects.

A number of alternative data-driven volatility filters are considered below which differ in terms of the estimation method, sampling frequency and information set. These high-frequency estimators, namely $QV_{m,t}$ and $PV_{m,t}$, $RA_{m,t}$ and $HQV_{m,t}$, were discussed in the previous section.[8] We complement the high-frequency filters with daily data-driven volatility estimators in order to compare the sampling frequency aspect and Kokoszka and Teyssièrre (2002) analysis. The daily frequency filters are: (i) The Exponentially Weighted Moving Average (EWMA) volatility or RiskMetrics (RM) is defined following the industry standard introduced by J.P. Morgan (see Riskmetrics Manual, 1995) as: $RM_t = \lambda RM_{t-1} + (1-\lambda) r_t^2$, $t = 1, ..., T_{days}$, where $\lambda = 0.94$ for daily data, r_t is the daily return and T_{days} is the number of trading days. This is an IGARCH process. (ii) The Rolling Volatility

---

Ghysels (2003b). The interested reader may refer to the aforementioned paper Tables 1 and 5 which present simulation evidence and Table 7 provides empirical results, regarding the distributional and dependence properties of data-driven de-volatilized returns using the Riskmetrics, Rolling Volatility and Quadratic Variation filters. Further analysis is worth pursuing for the Power Variation standardized returns.

[7]In all cases the simulated innovation process is a standardized process.

[8]We will henceforth often drop the m subscripts to facilitate the discussion and therefore refer to QV_t and PV_t, RA_t and HQV_t.

with a 26-day window is proposed in Foster and Nelson (1996). All volatility filters except the Rolling Volatility were also estimated using intraday sampling of 1- and 6-hours. Moreover, the intraday frequency is further exploited by examining transformations of block minima processes, denoted by $R\min$, involving the 1- and 6-hours block samples. This process is defined as $R\min = \min_{1\leq j\leq b}(r_{(t-1)b+j})$, where b is the block size.[9]

2.4.2 Simulation results for temporal dependence in normalized returns

The Hong and Lee (2003) test simulation results for daily as well as hourly normalized returns based on the aggregation of the 5-minute return process simulated as described above are reported in Table 1. The values refer to the simulated mean test statistic and its respective p-value for each of the temporal dependence null hypotheses (H1)-(H5). The results show that there is no support against the null hypothesis. The simulations reported use the Bartlett kernel. The second column in Table 1 presents the residuals of the GARCH model at the daily frequency which support (as expected) the null hypotheses in the Hong and Lee test. The same results are obtained for the daily and hourly Quadratic and Power variation normalized returns. Therefore at these frequencies there is simulation evidence that supports the null hypotheses of temporal indenpendence using the Hong and Lee test. These results are robust to (1) other kernel-based estimators such as the Parzen kernel, (2) choosing the bandwidth, (3) alternative intraday sampling frequencies such as half-daily, (4) lag lengths as well as to (5) transformations of the normalized returns process such as $\hat{X}_t^2$. The Hong and Lee test is also applied to the higher order process $\tilde{P}_{[t-1,t]}^{4,(m)}/[\tilde{P}_{[t-1,t]}^{2,(m)}]$ as it reflects the Barndorff-Nielsen and Shephard condition for temporal dependence in $QV_{(m),t}$ normalized returns. These results are reported in the last column of Table 1 and also do not present support against the null hypotheses. Therefore, it appears these normalized returns processes can be used as a representative control process for detecting change-points using the EDF based test.

---

[9]Block maxima processes can be treated in the same way. We focus on the block minima returns process since it has negative effects on portfolio returns which are of more fundamental concern to risk managers and investors.

2.4.3 Simulation results for the EDF test

We study the EDF-based change-point tests in Horvath et al. (2001) and Koul (2002) for $\hat{X}_t^2$ based on alternative volatility estimators especially using high-frequency data. In the spirit of the classical Kolmogorov-Smirnov statistic the weighted difference between the sequential EDFs:

$$
\sup_{1 \leq Ts \leq T} \sup_{x \in \mathbf{R}} \sqrt{T} \left(\frac{Ts}{T} \left(1 - \frac{Ts}{T} \right) \right) \left| \frac{1}{Ts} \sum_{t=1}^{Ts} \mathbf{1}\{\hat{X}_t^2 \leq x\} - \frac{1}{T - Ts} \sum_{t=Ts+1}^{T} \mathbf{1}\{\hat{X}_t^2 \leq x\} \right|
$$

(2.16)

is written in a less computationally intensive format as given by the second equality:

$$
KS = \sup_{1 \leq Ts \leq T} \sup_{x \in \mathbf{R}} \frac{1}{\sqrt{T}} \left| \sum_{t=1}^{Ts} \mathbf{1}\{\hat{X}_t^2 \leq x\} - \frac{Ts}{T} \sum_{t=1}^{T} \mathbf{1}\{\hat{X}_t^2 \leq x\} \right|
$$

(2.17)

which yields asymptotically a Kiefer process $\{K(x,s),\, 0 \leq x, s \leq 1\}$. Henceforth we refer to this test as KS.

The simulation results for the KS statistic are reported in Table 2. All tests are applied to daily normalized returns based on alternative volatility fitlers. The simulation results refer to the daily returns scaled by data-driven volatility fitlers using daily data (RM_t, RV_t) and high-frequency data $(QV_t, HQV_t, RA_t, PV_t)$. The two reported numbers for each case represent the mean number of rejections at the 5% and 10% nominal levels, respectively. The KS test appears to have mild size distortions due the small effective sample of $T_{days} = 200$ used (given that the first $0.2T_{days}$ observations were ignored due to the 26-day window involved in the Rolling Volatility and the erratic behavior due to the normalization at the beginning of the sample). The size distortions are relatively higher for PV_t normalized returns and for all high-frequency filters when generated by a Lévy process.[10]

It should be noted that the finite sample properties of the Kolmogorov-Smirnov test is found to exhibit relatively low power for the QMLE GARCH squared residuals (Kokoszka and Teyssière, 2002) which is verified here by alternative daily data-driven volatility filters based on the quadratic variation and using either daily or high-frequency estimators. In contrast, we find the interesting result that the Power Variation substantially improves the performance of the EDF test results for conditional variance and conditional

---

[10]The KS test for the Lévy process generated GARCH and RM_t and RV_t normalized returns yield more serious size distortions and are therefore not considered here.

distribution changes. The daily data-driven filters (RM_t, RV_t) have some power in detecting change points in the dynamics of the GARCH and changes in the degrees of freedom of the t distribution. We mention that the power for the de-volatilized returns by the RM_t is comparable with the QMLE GARCH estimators used in Kokoszka and Teyssière (2002).[11] In contrast, the daily returns scaled by the high-frequency filters QV_t, HQV_t, RA_t do not seem to have any power for the KS statistic in detecting any of change-point alternatives except the fat tailed alternative from Lévy to Cauchy. Note that increasing the sample size does not alter these results. These poor power performance of the KS is altered by the PV_t normalized returns which has very good power properties in detecting volatility parameter changes. In order to challenge the robustness of this result we simulate very small changes in the GARCH parameters that yield the reported results in Table 2.[12] In addition, the KS test also has good power properties for the PV_t scaled returns in detecting changes in the conditional distribution except for changes from Normal to t distribution.[13]

3 CUSUM tests for volatility shifts

So far we examined tests involving the process $\hat{X}_t \equiv g[(r_t)/\hat{\sigma}_t]$ using both parametric and data-driven estimators. We now change the object of interest, namely we focus on the volatility process σ_t. This implies that we examine a different type of null hypothesis, instead of studying change-points in the distribution of normalized returns r_t/σ_t we examine distributional changes in the volatility process. Naturally, we may expect that the tests proposed in this section also have power in detecting distributional shifts. Conversely, the EDF-based tests based for normalized returns are expected also to have power in detecting changes in the volatility dynamics. These issues are addressed by the Monte Carlo investigation. We provide a brief discussion of

---

[11]Note the differences involved are that RM_t is an IGARCH process with fixed parameters and that here we consider a smaller sample size and a break at the end of the sample.

[12]These parameter changes are from $\gamma_{1,0} = 0.99$ to $\gamma_{1,1} = 0.89$ for DGP1 and from $\gamma_{1,0} = 0.93$ to $\gamma_{1,1} = 0.83$ for DGP2 and also from $b_{0,0} = 0.000093$ to $b_{0,1} = 0.00093$ for DGP1 and from $b_{0,0} = 0.00044$ to $b_{0,1} = 0.0044$ for DGP2 at πT.

[13]Further simulation results for a larger sample size of $T_{days} = 500$ establish that the above results for PV_t normalized returns improve. For conciseness we do not report them here.

the Kokoszka and Leipus (2000) where the process monitored for homogeneity is $|r_{1,t}|^\delta$, $\delta = 1, 2$. Andreou and Ghysels (2002b) in the context of their empirical analysis apply this test to the $(H)QV_t$ type filters for the 5-minute YN/US\$ series. Since the process of interest $X_t = |r_t|^\delta$ for $\delta = 1, 2$ represents an observed measure of the variability of returns we may use high-frequency volatility filters e.g. $X_t = QV_t, HQV_t, PV_t$ and RA_t as discussed in the previous section which are locally smoothed filters of the quadratic and power variation. Recall that in the context of the SV and GARCH models $\{r_t\}$ represents a β-mixing process and that the measurable functions of mixing processes are mixing and of the same size (White, 1984, Theorem 3.49). Similarly the high-frequency returns process $\{r_{(m),t}\}$ generated by (2.12) is β-mixing and the high-frequency filters are $X_t = G(r_{(m),t}, .., r_{(m),t-\tau})$, for finite τ, are also β-mixing. Different considerations exist for the choice the transformation of the returns process. The absolute returns process is empirically found also to exhibit long memory properties (Ding et al., 1993). Similarly the absolute rather than the squared returns process is preferred for the study of the autocorrelation properties of heavy tailed ARCH processes when $E(r_t^4) < \infty$ (Davis and Mikosch, 1998, Mikosch and Starica, 2000). Woerner (2002) also shows that the quadratic variation will give more weigth to extreme observations like jumps and outliers as opposed to the absolute variation. Hence the choice of the rth power becomes an important point when detecting for changes in the volatility dynamics but especially the tail of the process since we would like a process that is robust to outliers in order to detect the actual change-points. We also consider the process of block minima $X_t = |R\min|^\delta$ $\delta = 1, 2$ where $R\min$ is used in extreme value theory for tail estimation. These are defined as $R\min = \min_{1 \le j \le b}(r_{(t-1)b+j})$, where b is the block size. The properties of this process for GARCH sequences are found in Mikosch and Starica (2000) and are expected to have power in detecting change points in the tails of the process since they focus on tail observations. The block maxima process can be thought in the same way.

In order to test for breaks in an ARCH(∞) Kokoszka and Leipus (1999, 2000) consider the following process:

$$U_T(k) = \left(1/\sqrt{T} \sum_{j=1}^k X_j - k/(T\sqrt{T}) \sum_{j=1}^T X_j \right) \tag{3.18}$$

where $0 < k < T$, $X_t = r_t^2$. The returns process $\{r_t\}$ follows an ARCH(∞) process, $r_t = u_t\sqrt{h_t}$, $h_t = b_0 + \sum_{j=1}^\infty b_j r_{t-j}^2$, $a \ge 0, b_j \ge 0, j = 1, 2$, with finite

fourth moment and errors that can be non-Gaussian. The CUSUM-type estimator $\hat{k}$ of a change point k^* is defined as:

$$\hat{k} = \min\{k : |U_T(k)| = \max_{1 \leq j \leq T} |U_T(j)|\} \qquad (3.19)$$

The estimate $\hat{k}$ is the point at which there is maximal sample evidence for a break in the squared returns process. In the presence of a break it is proved that $\hat{k}$ is a consistent estimator of the unknown change-point k^*. It is more convenient to state the results in terms of the estimator of τ^* of $\hat{\tau} = \hat{k}/T$ with $P\{|\tau^* - \hat{\tau}| > \varepsilon\} \leq C/(\delta\varepsilon^2\sqrt{T})$, where C is some positive constant and δ depends on the ARCH parameters and $|\tau^* - \hat{\tau}| = O_p(1/T)$ (Kokoszka and Leipus, 1998, 2000). Under the null hypothesis of no break:

$$U_T(k) \rightarrow_{D[0,1]} \sigma B(k) \qquad (3.20)$$

where $B(k)$ is a Brownian bridge and $\sigma^2 = \sum_{j=-\infty}^{\infty} Cov(X_j, X_0)$. Consequently, using an estimator $\hat{\sigma}$, one can establish that under the null:

$$\sup\{|U_T(k)|\}/\hat{\sigma} \rightarrow_{D[0,1]} \sup\{B(k) : k\epsilon[0,1]\} \qquad (3.21)$$

which establishes a Kolmogorov-Smirnov type asymptotic distribution.

The computation of the Kokoszka and Leipus (1999, 2000) test is relatively straightforward, with the exception of $\hat{\sigma}$ appearing in (3.21). The authors suggest to use a Heteroskedasticity and Autocorrelation Consistent (HAC) estimator applied to the X_j process. Andreou and Ghysels (2002b) experimented with a number of estimators in addition to the procedure of Den Haan and Levin (1997) who propose a HAC estimator without any kernel estimation, which is called the Autoregression Heteroskedasticity and Autocorrelation Consistent (ARHAC) estimator. This estimator has an advantage over any estimator which involves kernel estimation in that the circular problem associated with estimating the optimal bandwidth parameter can be avoided. Some further results relating this estimator are presented in the simulations section.

3.1 Simulation results for the CUSUM test

The simulation analysis of the CUSUM test is based on the simulation design described in section 2.4.1 and aims to extend the Kokoszka and Leipus (1999, 2000) CUSUM-type test in two directions. First we examine the performance

of the test for processes beyond the absolute and squared returns to alternative volatility processes that can also be used as subordinate, observed series for detecting change points in the conditional variance dynamics. The alternative volatility processes are found to be crucial for the power properties of the test. Second we evaluate whether the test has power in detecting not only scale but shape changes relating to tail changes in the innovation. Note that here we detect changes in the unconditional distribution. Other transformations of returns that relate to tail observations such as block minima processes (used in the estimation of extremes) also yield power. All these processes satisfy the mixing conditions required by the Kokoszka and Leipus test. When these processes are combined with the high-frequency data they are practically useful in that they multiply the size of the sample which is crucial for change-point tests and tail analysis. However, the stylized fact that the higher frequency yields a more persistent process may challenge the stationarity and moment conditions of the tests and their performance. In general it is expected that it may be more difficult to detect breaks in a highly persistent process. Hence we examine via simulations the tradeoff between the good size/power of the CUSUM test and the intraday frequency processes.

Tables 3-5 present the Kokoszka and Leipus CUSUM test simulation results defined in (3.18)-(3.21) and referred to as KL. These results are summarized as follows:

Daily volatility processes. The KL test was proposed for $|r_t|^\delta$ where obvious choices are $\delta = 1, 2$ to represent the absolute and quadratic variations as subordinate observed processes of volatility. The simulation results for these processes are reported in Table 3 where it is evident that these transformations and especially the absolute returns process have power in detecting changes in the volatility parameters whereas they have less power in detecting changes in the innovation distribution of the GARCH process. We enlarge the scope of the KL test for the daily high-frequency volatility filters $(QV_t, HQV_t, RA_t, PV_t)$. These results also appear in Table 3 for which the KL test appears to have good size properties and in some cases to be rather conservative especially for the high persistent 5-minute GARCH driven by Normal and Student's t innovations. The exceptions to this result are for the Lévy driven process where the test is undersized. Power appears to be good in detecting volatility dynamic changes as shown by the first three rows in the power panel for all processes and filters expect for the QV_t and HQV_t

when the process is persistent (even for a large drop in b_1).[14] The KL test is also evaluated for changes in the distribution of innovations (shown in the last three rows of Table 3) and it is found that the test has power in detecting changes in the tail of a Student's t distribution from 6 to 3 degrees of freedom using any of the high-frequency filters except the RA_t. The other relatively more subtle distributional alternatives (from Normal to $t(6)$ and from Lévy to Cauchy) are not detected by the KL test except when that is applied to the PV_t for the low persistence GARCH. There is an interesting comparison to be made in the simulations results for the PV_t that show how the persistence of the GARCH affects the power of the test in detecting distributional shifts. Overall the process PV_t filter seems to yield the best power in detecting conditional variance and distribution changes in a low persistent high-frequency GARCH process. These KL test results appear to be robust to the ARHAC parameters relating to the alternative criteria of lag length selection. Related results in the literature are by Woerner (2002) who shows that theoretically power variation is more robust relative to quadratic variation and Mikosch and Starica (2003) who find absolute returns to perform better than squared returns for estimating the autocorrelation function when the fourth moment is not finite.

Intradaily volatility processes. The properties of the KL are further investigated for high-frequency processes pertaining not only to intraday absolute and squared returns as well as hourly volatility filters but also for block minima processes. These processes also share the same mixing properties as the daily returns and daily volatility filters but they have the advantage that they increase the sample size to the intraday one which may turn out to important in detecting tail changes in the process. Yet, it is expected that intraday returns processes and their transformations may in some cases be more noisy and also exhibit more persistence. This point is related to the optimal choice of m. The results for the 5-minute GARCH with low and high persistence are reported in Tables 4 and 5, respectively. Table 4 shows that the KL test still appears rather conservative for hourly volatilities except for absolute transformations of the returns process and hourly block minima where the simulated size is very close to the nominal one. However, these processes

---

[14]It is worth noting that neither of the daily sampled filters RM_t and RV_t have any power in detecting volatility coefficient changes and hence these results are not reported in Table 3. One interpretation could be that these involve more smoothing and aggregation than in the absolute and squared returns (that form the basis of this CUSUM test) as well as the high-frequency filters presented here.

appear to be undersized when the GARCH model is driven by Lévy innovations. In Table 4 size distortions are observed for the 5-minute absolute and square transformations of the low persistence 5-minute GARCH process. For the high-persistence GARCH in Table 5 the KL seems to have very serious size distortions for the raw 5-minute transformations, the hourly range, the hourly absolute returns as well as hourly block minima transformations. For exposition purposes we show the results for the latter three processes. This result is not surprising for such approximate IGARCH process given the theoretical finite second moment condition required in the KL test. Since such a condition is likely to be violated by high-frequency processes we recommend that the raw transformations of absolute and squared returns suggested in KL be avoided for these processes. Instead it seems that some level of aggregation and smoothing as it is involved for instance in the 6-hour absolute returns and block minima as well as high-frequency filters will solve the size distortions of the KL test when applied to the observed 5-minute highly persistent returns process. Related is the evidence in Barndorff-Nielsen and Shephard (2003) which shows that for $m > 48$ the QV_t and PV_t asymptotic approximations become quite accurate.

With these aggregation and filtering considerations involved in the KL test for high-frequency processes we now turn to evaluate the test's power results. The overall picture sketched in Tables 4 and 5 suggests that the absolute 6-hour returns and block minima have good power for all alternatives except for the subtle change from Normal to $t(6)$ innovations in the low persistence GARCH. It is noteworthy that these processes capture the change from a Lévy to Cauchy innovation a result that is not surprising for the block minima process is essentially based on such extreme tail observations. Although the 1-hour and 6-hours Range and Quadratic Variation filters seem to be doing better in terms of capturing the Normal to the t tail changes, they have no power in differentiating between the two heavy alternatives (Lévy to Cauchy). The volatility filter that seems to yield the best power for the KL under all alternatives is the Power Variation for both high and low persistent GARCH processes and even for small samples of 1 and 6 hours. Comparing the KL test results for the PV_t (in Tables 3, 4 and Table 3) we conclude that although $m = 288$ in the daily PV_t as opposed to $m = 12$ and 72 in the 1- and 6-hours PV_t, respectively, the latter appear to yield better power for the KL (with no size distortions) under all alternatives considered. This points to an interesting tradeoff between intraday sample information for the estimation of PV_t and sample size considerations for change-point

24

tests. Finally we add that the KL statistic for the 6-hour filtered processes and the 1-hour PV_t and QV_t is robust to the ARHAC parameters as opposed to the 5-minute and 1-hour transformations (such as absolute returns, RM_t and RA_t) that highly sensitive to the alternative lag length choices (whether these are the AIC and SC or a fix lag length). This presents an additional reason in favor of using high-frequency volatility estimators instead of simply the observed squared and absolute returns in the KL test or filters that are too smooth (e.g. the hourly RM_t) or ignore most of the intraday information (e.g. the hourly RA_t).

Summarizing and combining the results from different change-point tests, simulated processes, alternative structural break hypotheses, volatility filters and returns transformations, we find that for the KL test the daily Power Variation yields better power (under all the alternatives hypotheses considered) followed by the absolute daily returns process. However, the KL test for intraday absolute returns for persistent 5-minute simulated GARCH processes as well as their transformations up to an hour suffer serious size distortions. Aggregating to 6-hour absolute returns intraday process as well as high-frequency volatility filters such as the Range, Quadratic and Power Variations enjoy good size and power properties even for high-persistent processes. The 6-hour absolute returns, block minima, absolute returns and Power Variation yield power under both volatility and distribution change points. These results are robust to the alternative parameter choices of the ARHAC estimator used for standardizing the sup of the CUSUM statistic. Comparatively the EDF test for normalized returns appears to yield less power in detecting change-points in GARCH processes using either daily or high-frequency data-driven volatility estimators except for the Power Variation scaled returns. Some explanation involved in this interesting result can be fact that the Power Variation attaches less weight to outliers as opposed to the Quadratic Variation and yields a more robust (or less noisy) process to jumps and therefore makes it possible to detect changes in the scale and shape of the process. Moreover, we find by observing the simulated Power Variation normalized returns process that undergoes volatility parameter changes as well as some tail changes that there is a dinstict mean shift in this process before and after the change-point. The ability of this transformation to interpret scale and shape changes into mean shifts may also explain partly why it enjoys better change-point power than other quadratic variation filters. It is well-known in the change-point literature that mean shifts are relatively easier to detect as opposed to variance or tail changes.

25

We conclude by acknowledging that little is known about the theoretical properties of the power variation normalized returns. Yet, the above positive simulation results regarding power variation and the normalized power variation returns for change point analysis encourages further research in the area.

4   Conclusions

The scope of the paper was to evaluate the performance of several recently proposed change-point tests applied to conditional variance dynamics and conditional distributions of asset returns. These are CUSUM-type tests for beta-mixing processes and EDF-based tests for the residuals of such nonlinear dependent processes. Hence the tests apply to the class of ARCH and SV type processes as well as data-driven volatility estimators using high-frequency data. It was shown that some of the high-frequency volatility estimators substantially improve the power of the structural breaks tests especially for detecting changes in the tail of the conditional distribution. First, for the EDF-type tests we consider different conditional variance estimators based on high-frequency data-driven volatility filters and find that using high-frequency data filters, particularly Power Variation, leads to improved power. We also examine the potential impact of estimation error due to the use of high-frequency data and find no evidence this should be an issue of concern. We also examine both daily and intra-daily sampling schemes, both yielding similar results.

Regarding CUSUM tests, we find that for the KL test the daily Power Variation yields better power (under all the alternatives hypotheses considered) followed by the absolute daily returns process. However, the KL test for intraday absolute returns for persistent 5-minute simulated GARCH processes as well as their transformations up to an hour suffer serious size distortions, whereas 6-hour absolute returns intraday process as well as high-frequency volatility filters such as the Range, Quadratic and Power Variations enjoy good size and power properties even for high-persistent processes. The 6-hour absolute returns, block minima, absolute returns and Power Variation yield power under both volatility and distribution change points. These results are robust to the alternative parameter choices of the ARHAC estimator used for standardizing the sup of the CUSUM statistic. Comparatively the EDF test for normalized returns appears to yield less power in detect-

ing change-points in GARCH processes using either daily or high-frequency data-driven volatility estimators except for the Power Variation scaled returns.

The paper enlarges the scope of the change-point CUSUM test as well as the residual based EDF tests in Horvath et al. (2001) for high-frequency volatility estimators and returns transformations for which the tests enjoy good power. Futher investigation into the properties of some of these volatility filters as well as the sample of extremes and their relation to change-point analysis of strongly dependent processes is one direction of future research. Others involve the application of the above procedures for financial risk management.

References

[1] Alizadeh, S., M. W. Brandt and F.X. Diebold (2002), "Range-based estimation of stochastic volatility models", *Journal of Finance*, 57, 1047-1091.

[2] Andersen, T. and T. Bollerslev (1998) "Answering the Skeptics: Yes, Standard Volatility Models Do Provide Accurate Forecasts", *International Economic Review*, 39, 885-905.

[3] Andersen, T.G., Bollerslev, T. and Diebold, F.X. (2003), " Parametric and Nonparametric Volatility Measurement," in L.P. Hansen and Y. Ait-Sahalia (eds.), Handbook of Financial Econometrics, Amsterdam: North-Holland, forthcoming.

[4] Andreou, E. and E. Ghysels (2002a), "Rolling sample volatility estimators: some new theoretical, simulation and empirical results" *Journal of Business and Economic Statistics*, 20, 3, 363-376.

[5] Andreou E. and E. Ghysels (2002b), "Detecting multiple breaks in financial market volatility dynamics", *Journal of Applied Econometrics*, 17, 5, 579-600.

[6] Andreou, E. and E. Ghysels, (2002c), "When do Jump Diffusions Approximate Returns? Nonparametric Test Procedures Based on Price Rigidity and Discreteness", Work in progress.

[7] Andreou, E. and E. Ghysels (2003a), "Quality Control for Financial Risk Management: Monitoring Disruptions in the Distribution of Risk Exposure", Discussion Paper, UNC.

[8] Andreou, E. and E. Ghysels (2003b), "Tests for breaks in the conditional co-movements of asset returns", *Statistica Sinica*, 13, 1045-1074.

[9] Andrews, D. W. K. (1991), "Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation," *Econometrica*, 59, 817-858.

[10] Bai, J. (1994), "Weak convergence of the sequential empirical processes of residuals of ARMA models", *The Annals of Statistics*, 22, 2051-2061.

[11] Bai, J. (2003), "Testing Parametric Conditional Distributions of Dynamic Models" *Review of Economics and Statistics* (forthcoming).

[12] Barndorff-Nielsen, O. (1963), "On the limit behaviour of extreme order statistics", *Annals of Mathematical Statistics*, 34, 992-1002.

[13] Barndorff-Nielsen, O. and N. Shephard (2001), "Non-Gaussian Ornstein–Uhlenbeck-based models and some of their uses in financial economics (with discussion), *Journal of the Royal Statistical Society*, Series B, 63, 167-241.

[14] Barndorff-Nielsen, O. and N. Shephard (2002a), "Econometric analysis of realized volatility and its use in estimating stochastic volatility models", *Journal of the Royal Statistical Society*, Series B, 64, 253-280.

[15] Barndorff-Nielsen, O. and N. Shephard (2002b) "Estimating quadratic variation using realised variance" *Journal of Applied Econometrics*, 17, 5, 457-477.

[16] Barndorff-Nielsen, O. and N. Shephard (2003a) "Realised power variation and stochastic volatility models", *Bernoulli* forthcoming.

[17] Barndorff-Nielsen, O. and N. Shephard (2003b) "How accurate is the asymptotic approximation to the distribution of realised volatility?" in D.W.K. Andrews, J. Powell, P. Ruud and J. Stock (ed.), *Identification and Inference for Econometric Models. A Festschrift for Tom Rothenberg*, Cambridge University Press, (forthcoming).

[18] Bates, D. (2000), "Post-'87 Crash Fears in the S&P 500 Futures Option Market", *Journal of Econometrics*, 94, 181-23

[19] Berkes I., L. Horvath and P. Kokoszka (2002), "GARCH processes: structure and estimation", *Bernoulli* forthcoming.

[20] Blum, J.R., Kiefer J. and M. Rosenblatt (1961) "Distribution free tests of independence based on the sample distribution function", *The Annals of Mathematical Statistics,* 32, 485-498.

[21] Bollerslev, T., R. F. Engle and D. B. Nelson (1994), "ARCH Models", in Robert F. Engle and Dan McFadden (eds.), *Handbook of Econometrics*, Volume IV, 2959-3038. Amsterdam: North-Holland.

[22] Carrasco, M. and X. Chen (2002), "Mixing and Moment Properties of Various GARCH and Stochastic Volatility Models", *Econometric Theory*, 17, 1, 17-39.

29

[23] Christoffersen P., J. Hahn and A. Inoue (2001), "Testing and Comparing Value-at-Risk Measures", *Journal of Empirical Finance*, 8, 325-342.

[24] Chu, C.-S. (1995), "Detecting parameter shift in GARCH models", *Econometric Reviews*, 14, 241-266.

[25] Csörgo, M. and L. Horváth (1997), *Limit Theorems in Change-Point Analysis*, Wiley, New York.

[26] Davis R.A. and T. Mikosch (1998), "The maximum of the periodogram of a non-Gaussian sequence", *Annals of Statistics*, 26, 2049-2080.

[27] Den Haan, W.J., and A. Levin (1997), "A Practioner's Guide to Robust Covariance Matrix Estimation" in C.R. Rao and G.S. Maddala (eds.) *Handbook of Statistics - Vol. 15*, 291-341.

[28] Diebold, F.X. (1986), "Modeling the persistence of conditional variances: A comment", *Econometric Reviews*, 5, 51-56.

[29] Diebold, F.X. and Inoue, A. (2001), "Long Memory and Structural Change", *Journal of Econometrics* 105, 131-159.

[30] de Lima P. (1998), "Nonlinearities and nonstationarities in stock returns", *Journal of Business and Economic Statistics*, 16, 2, 227-236.

[31] Ding, Z., Granger, C.W.J. and Engle, R.F. (1993) "A Long-memory Properties of Stock Market Returns and a New Model", *Journal of Empirical Finance*, 1, 83-106.

[32] Drost, F. C. and T. Nijman (1993), "Temporal Aggregation of GARCH Processes", *Econometrica*, 61, 909-727.

[33] Drost, F. C. and B. M.J. Werker (1996), "Closing the GARCH Gap: Continuous Time GARCH Modeling", *Journal of Econometrics*, 74, 31-57.

[34] Engle, R.F. (1982), "Autoregressive conditional heteroskedasticity with estimates of the variance of the United Kingdom inflation", *Econometrica*, 50, 987-1007.

[35] Feller W. (1951) "The asymptotic distribution of the range of sums of independent random variables", *Annals of Mathematical Statistics*, 22, 427-32.

[36] French, K. R., W. Schwert and R.F. Stambaugh (1987) "Expected stock returns and volatility", *Journal of Financial Economics*, 19, 3-29.

[37] Foster, D. and D. Nelson (1996) "Continuous Record Asymptotics for Rolling Sample Estimators", *Econometrica*, 64, 139-174.

[38] Ghysels, E. (1998), "On Stable Factor Structures in the Pricing of Risk: Do Time-Varying Betas Help or Hurt" *Journal of Finance* 53, 549-573.

[39] Ghysels, E. Harvey A. and E. Renault, (1996) "Stochastic Volatility", G. S. Maddala and C. R. Rao. eds., *Handbook of Statistics*, Vol. 14.

[40] Giraitis L., R. Leipus and D. Surgailis (1996), "The change-point problem for dependent observations", Journal of Statistical Planning and Inference, 53, 3, 297-310.

[41] Hall P. and Q. Yao (2003) "Inference in ARCH and GARCH models with heavy-tailed errors", *Econometrica*, forthcoming.

[42] Hendry, D. F. (1986), "An Excursion into Conditional Variance Land", *Econometric Reviews*, 5, 63-69.

[43] Hong, Y. and Y-J Lee (2003) "Generalized Spectral Tests for Conditional Mean Models in Time Series with Conditional Heteroskedasticity of Unknown Form", *Econometric Theory*, forthcoming.

[44] Horvath, L., P. Kokoszka and G. Teyssière (2001), "Empirical Process of the Squared Residuals of an ARCH sequence", *Annals of Statistics*, 29, 445-469.

[45] Hsieh (1991), Chaos and Nonlinear Dynamics: Application to Financial Markets, *Journal of Finance* 46, 1839-1877.

[46] Inoue, A. (2001), "Testing for Distributional Change in Time Series", *Econometric Theory*, 17, 156-187.

[47] Kokoszka, P. and R. Leipus (1999), "Testing for parameter changes in ARCH models", *Lithuanian Mathematical Journal* 39, 231-247.

[48] Kokoszka, P. and R. Leipus (2000), "Change-point estimation in ARCH models", *Bernoulli*, 6, 1-28.

[49] Kokoszka P. and G. Teyssière (2002) "Change-point detection in GARCH models: Asymptotic and boostrap tests", Core Discussion Paper, 65.

[50] Koul, H.L., (1996), "Asymptotics of some Estimators and Sequential Empiricals in Non-Linear Time Series", *Annals of Statistics*, 24, 380-404.

[51] Koul, H.L., (2002), Weighted Empirical Processes in Dynamic Nonlinear Models, 2nd ed., Lecture Notes in Statistics, Vol., 166, Springer.

[52] Koul H.L and K. Mukherjee (2002), "Some estimation procedures in ARCH models", Discussion Paper.

[53] Lamoureux, C.G. and W.D. Lastrapes (1990), "Persistence in variance, structural change and the GARCH model", *Journal of Business and Economic Statistics*, 8, 225-234.

[54] Lavielle, M. (1999) "Detection of multiple changes in a sequence of dependent variables", Stochastic Processes and their Applications, 83, 1, 79-102.

[55] Lavielle, M. and E. Moulines (2000), "Leat-squares Estimation of an Unknown Number of Shifts in Time Series", *Journal of Time Series Analysis*, 20, 33-60.

[56] Lundbergh, S. and T. Terasvirta (1998), "Evaluating GARCH models", *Journal of Econometrics*, 110, 417-435.

[57] Merton, R. C. (1980), "On Estimating the Expected Return on the Market: An Exploratory Investigation", *Journal of Financial Economics* 8, 323-361.

[58] Mikosch, T. and C. Starica, C. (2000), "Limit theory for the sample autocorrelation and extremes of a GARCH(1,1) process", *Annals of Statistics*, 28, 1427-1451,

[59] Mikosch, T. and C. Starica, C. (2003), "Change of structure in financial time series, long range dependence and the GARCH model", *Review of Economics and Statistics*, forthcoming.

[60] Parkinson M. (1980), "The extreme value method for estimating the variance of the rate of return", *Journal of Business*, 53, 61-5.

[61] Pástor, L. and R. Stambaugh (2001), "The equity premium puzzle and structural breaks", *Journal of Finance*, 56, 4, 1207-39.

[62] Picard D. (1985), "Testing and Estimating Change-Points in Time Series", *Advances in Applied Probability*, 17, 841-867.

[63] Poterba, J.M. and L. H. Summers, (1986), "The Persistence of Volatility and Stock Market Fluctuations", *American Economic Review*, 76, 1142-1151.

[64] Quintos C., Z. Fan and P.C.B. Phillips (2001), "Structural Change in Tail Behavior and the Asian Financial Crisis", *Review of Economic Studies*, 68, 3, 633-663.

[65] *Riskmetrics* Manual, (1995), J.P. Morgan Corporation, New York.

[66] Schwert, G.W. (1989), "Why does Stock Market Volatility Change Over Time?", *Journal of Finance*, 44, 1115-1154.

[67] Szyszkowich, B. (1998) "Weighted Sequential Empirical Type Processes with Application to Change-Point Problems", *Handbook of Statistics*, edited by N. Balakrishnan and C.R. Rao, Vol. 16, Ch. 20, 573-630.

[68] Timmerman, A. (2001), "Structural Breaks, Incomplete Information, and Stock Prices", *Journal of Business and Economic Statistics*, 299-315.

[69] White, H. (1984) *Asymptotic Theory for Econometricians*, Academic Press.

[70] Woerner J. (2002) "Variational sums and power variation: a unifying approach to model selection and estimation in semimartingale models", Discussion Paper, Oxford University.

Table 1: Hong and Lee (2003) test simulation results for normalized returns.

Normalized returns based on the following Volatilities

	$GARCH$	QV_{1d}	PV_{1d}	QV_{1h}	PV_{1h}	$\tilde{P}^{4,(m)}_{[t-1,t]}/[\tilde{P}^{2,(m)}_{[t-1,t]}]$
			Low-persistence GARCH			
		5-minute GARCH(1,1):	$b_0 = 0.1, b_1 = 0.3, \gamma = 0.3$			
Lags:	20	20	20	20	20	20
Hypotheses:						
H1	-0.009 (.514)	-0.185 (.569)	-0.177 (.562)	-0.074 (.548)	-0.045 (.533)	0.099 (.499)
H2	0.037 (.504)	-0.078 (.551)	-0.078 (.549)	-0.071 (.538)	0.048 (.524)	0.338 (.450)
H3	-0.053 (.532)	-0.052 (.532)	-0.022 (.538)	0.074 (.518)	0.066 (.502)	0.333 (.443)
H4	0.100 (.466)	-0.066 (.535)	0.041 (.505)	-0.022 (.530)	0.014 (.509)	0.301 (.439)
H5	-0.009 (.506)	-0.025 (.522)	0.009 (.509)	0.083 (.513)	-0.010 (.521)	0.250 (.443)

Notes: The normalized returns for various volatility filters are considered. The volatility filters $\tilde{Q}^{(m)}_{[t-1,t]}$ and $\tilde{P}^{(m)}_{[t-1,t]}$ use intra-daily five minute data and are defined respectively as $\sum_{j=0}^{m-1}[r^{(m)}_{t-j/m}]^2$ and $\sum_{j=0}^{m-1}|r^{(m)}_{t-j/m}|$ (see section 2.2 for details). The Hong and Lee (2003) test examines the various null hypotheses of temporal independence H1-H2 (discussed in the paper). The reported test statistics results use the Bartlett kernel (results with Parzen kernel are similar but not reported), with their respective p-values in the parenthesis. The tests are also performed for the squared de-volatilized returns and various lag lengths (from 5 and 20) and similar qualitative results are obtained.

Table 2: Kolmogorov Smirnov test for Daily Scaled Returns using daily and high-frequency volatility filters

Squared de-volatilized returns based on the following Volatilities

	RM_{1d}	RV_{26d}	QV_{1d}	HQV_{1d}	RA_{1d}	PV_{1d}
Low-persistence GARCH						
5-minute GARCH(1,1): $b_0 = 0.1, b_1 = 0.3, \gamma = 0.3$						
Size						
N(0,1)	0.10, 0.18	0.08, 0.12	0.08, 0.14	0.08, 0.12	0.10, 0.17	0.18, 0.22
t(0,1;6)	0.12, 0.20	0.10, 0.13	0.07, 0.10	0.08, 0.12	0.07, 0.14	0.19, 0.32
Lévy	-	-	0.02, 0.09	0.31, 0.38	0.08, 0.14	0.18, 0.20
Power						
$b_{0,1}$ to $b_{0,2}$	0.15, 0.24	0.13, 0.15	0.06, 0.12	0.11, 0.16	0.03,0.07	0.77, 0.90
$b_{1,1}$ to $b_{1,2}$	0.72, 0.81	0.54, 0.65	0.09, 0.17	0.11, 0.20	0.10, 0.19	1.00, 1.00
$b_{0,2}, b_{1,2}$	0.38, 0.47	0.26, 0.42	0.08, 0.16	0.11, 0.16	0.05, 0.13	0.27, 0.40
N(0,1) to t(6)	0.15, 0.23	0.09, 0.16	0.07, 0.12	0.05, 0.14	0.07, 0.16	0.18, 0.31
t(6) to t(3)	0.23, 0.32	0.23, 0.33	0.07, 0.14	0.10, 0.14	0.13, 0.23	0.37, 0.44
Lévy to Cauchy	-	-	1.00, 1.00	1.00, 1.00	1.00, 1.00	0.99, 0.99
High-persistence GARCH						
5-minute GARCH(1,1): $b_0 = 0.00044, b_1 = 0.93482, \gamma = 0.06341$						
Size						
N(0,1)	0.12, 0.19	0.09, 0.20	0.08, 0.11	0.07, 0.11	0.10, 0.16	0.12, 0.20
t(0,1;6)	0.10, 0.21	0.14, 0.21	0.08, 0.17	0.05, 0.10	0.07, 0.13	0.10, 0.17
Lévy	-	-	0.10, 0.22	0.20, 0.29	0.08, 0.14	0.12, 0.18
Power						
$b_{0,1}$ to $b_{0,2}$	0.12, 0.22	0.16, 0.28	0.09, 0.16	0.07,0.12	0.09, 0.15	0.81, 0.87
$b_{1,1}$ to $b_{1,2}$	0.28, 0.38	0.18, 0.24	0.09, 0.16	0.08,0.16	0.10, 0.18	1.00, 1.00
$b_{0,2}, b_{1,2}$	0.12, 0.26	0.10, 0.13	0.10, 0.15	0.05, 0.14	1.00, 1.00	
N(0,1) to t(6)	0.12, 0.25	0.17, 0.29	0.08, 0.15	0.11, 0.16	0.12, 0.17	0.11, 0.22
t(6) to t(3)	0.49, 0.55	0.30, 0.42	0.07, 0.15	0.04,0.09	0.09, 0.15	0.39, 0.56
Lévy to Cauchy	-	-	1.00, 1.00	1.00, 1.00	1.00, 1.00	1.00, 1.000

Notes: The critical values for the KS test are $c_5 = 0.775$ and $c_10 = 0.712$. The break point occurs at $k = 0.8T$ and sample=0.2T:T. The following coefficient details for the Data Generating Processes apply: (1) Persistent GARCH: 5minute GARCH parameters: $c_0 = 0.000093$ and $c_1 = 0.093$ and $b_0 = 0.99055$ and $b_1 = 0.5$. These processes are simulated for obtaining the high-frequency volatility filters (QV1, HQV1, Range). Non-persistent GARCH Daily GARCH parameters: $b_0 = 0.1$ and $c_1 = 0.2$ and $b_0 = 0.3$ and $b_1 = 0.6$. 5-minute GARCH parameters: $c_0 = 0.00044$ and $b_1 = 0.0044$ and $b_0 = 0.93482$ and $b_1 = 0.46$. For the PV_{1d} normalized returns we choose more challenging change points for the GARCH parameters as: $b_1 = 0.00093$ and $b_1 = 0.0044$, in the 5-minute persistent and non-persistent GARCH models. $b_1 = 0.89$ and $b_1 = 0.83$, in the 5-minute persistent and non-persistent GARCH models.

Table 3: Kokoszka and Leipus CUSUM-type test based for various intraday asset returns processes using daily volatility filters

	$(R_d)^2$	$|R_d|$	QV_{1d}	HQV_{1d}	RA_{1d}	PV_{1d}
			Low-persistence GARCH			
		5-minute GARCH(1,1):	$b_0 = 0.1, b_1 = 0.3, \gamma = 0.3$			
Size						
N(0,1)	0.03, 0.07	0.04, 0.07	0.04, 0.08	0.01, 0.05	0.06, 0.09	0.04, 0.11
t(0,1;6)	0.03, 0.07	0.05, 0.09	0.03, 0.08	0.01, 0.05	0.02, 0.05	0.05, 0.12
Lévy	0.00, 0.00	0.00, 0.02	0.00, 0.01	0.00, 0.00	0.00, 0.01	0.00, 0.01
Power						
$b_{0,1}$ to $b_{0,2}$	0.64, 0.75	0.58, 0.69	0.99, 0.99	0.60, 0.88	1.00, 1.00	1.00, 1.00
$b_{1,1}$ to $b_{1,2}$	0.99, 1.00	0.99, 0.99	0.99, 0.99	0.78, 0.89	1.00, 1.00	0.80, 0.96
$b_{0,2}, b_{1,2}$	0.99, 0.99	1.00, 1.00	0.95, 0.98	0.57, 0.72	1.00, 1.00	0.18, 0.41
N(0,1) to t(6)	0.04, 0.08	0.04, 0.09	0.07, 0.09	0.04, 0.08	0.08, 0.19	0.99, 1.00
t(6) to t(3)	0.04, 0.10	0.03, 0.10	0.29, 0.34	0.25, 0.32	0.10, 0.20	1.00, 1.00
Lévy to Cauchy	0.01, 0.02	0.19, 0.25	0.05, 0.07	0.03, 0.06	0.04, 0.07	0.66, 0.70
			High-persistence GARCH			
		5-minute GARCH(1,1):	$b_0 = 0.00044, b_1 = 0.93482, \gamma = 0.06341$			
Size						
N(0,1)	0.02, 0.06	0.04, 0.07	0.01, 0.02	0.00, 0.00	0.02, 0.04	0.00, 0.01
t(0,1;6)	0.02, 0.05	0.02, 0.06	0.00, 0.03	0.00, 0.01	0.01, 0.04	0.02, 0.03
Lévy	0.00, 0.00	0.00, 0.00	0.00, 0.00	0.02, 0.02	0.00, 0.02	0.01, 0.03
Power						
$b_{0,1}$ to $b_{0,2}$	0.78, 0.85	0.96, 0.98	0.78, 0.85	0.65, 0.77	0.79, 0.89	0.81, 0.89
$b_{1,1}$ to $b_{1,2}$	0.31, 0.48	0.95, 0.98	0.00, 0.01	0.01, 0.02	0.40, 0.54	0.78, 0.88
$b_{0,2}, b_{1,2}$	0.30, 0.44	0.88, 0.93	0.29, 0.42	0.18, 0.29	0.41, 0.53	0.66, 0.80
N(0,1) to t(6)	0.02, 0.04	0.05, 0.09	0.22, 0.36	0.13, 0.24	0.03, 0.05	0.02, 0.05
t(6) to t(3)	0.06, 0.12	0.29, 0.27	0.28, 0.44	0.20, 0.30	0.07, 0.12	0.45, 0.59
Lévy to Cauchy	0.03, 0.05	0.20, 0.27	0.04, 0.06	0.03, 0.05	0.06, 0.09	0.07, 0.08

Notes: All results reported pertain to $T = 250$ days or 72000 5-minutes. The break point occurs at $k = 0.8T$. The two values reported in each cell refer to the mean number of rejections for the 5% and 10% nominal levels, respectively. (2) Change-points for the GARCH parameters are defined as follows: For the low-persistence GARCH we consider $b_{0,2} = 2b_{0,1}$ and $b_{1,2} = 2b_{1,1}$. For the high-persistence GARCH we consider $b_{0,2} = 10b_{0,1}$ and $b_{0,2} = 0.5b_{0,1}$.

Table 4: Kokoszka and Leipus CUSUM-type test based for various intraday asset returns processes using intraday volatility filters

	$(R_{5m})^2$	$|R_{5m}|$	$(Rmin_{1h})^2$	$|R_{1h}|$	$|Rmin_{1h}|$	RA_{1h}
Size						
N(0,1)	0.09, 0.20	0.11, 0.22	0.03, 0.08	0.05, 0.10	0.03, 0.09	0.04, 0.09
t(0,1;6)	0.09, 0.19	0.14, 0.23	0.03, 0.07	0.05, 0.10	0.05, 0.09	0.02, 0.07
Lévy	0.00, 0.00	0.18, 0.23	0.02, 0.03	0.00, 0.01	0.03,0.05	0.00, 0.01
Power						
$b_{0,1}$ to $b_{0,2}$	1.00, 1.00	1.00, 1.00	1.00, 1.00	1.00, 1.00	1.00, 1.00	1.00, 1.00
$b_{1,1}$ to $b_{1,2}$	1.00, 1.00	1.00, 1.00	1.00, 1.00	1.00, 1.00	1.00, 1.00	1.00, 1.00
$b_{0,2}, b_{1,2}$	1.00, 1.00	1.00, 1.00	1.00, 1.00	1.00, 1.00	1.00, 1.00	1.00, 1.00
N(0,1) to t(6)	0.14, 0.25	1.00, 1.00	0.03, 0.08	0.08,0.15	0.09, 0.15	0.09, 0.15
t(6) to t(3)	0.32, 0.41	0.99, 0.99	0.13, 0.23	0.84, 0.91	0.69, 0.79	0.40, 0.51
Lévy to Cauchy	0.00, 0.00	0.65, 0.70	0.15, 0.20	0.56,0.60	0.55, 0.60	0.09, 0.13

	QV_{1h}	PV_{1h}	$|R_{6h}|$	$|Rmin_{6h}|$	RA_{6h}	QV_{6h}	PV_{6h}
Size							
N(0,1)	0.04, 0.09	0.04, 0.09	0.04, 0.08	0.05, 0.10	0.03, 0.08	0.03, 0.07	0.05,0.11
t(0,1;6)	0.03, 0.07	0.04, 0.11	0.03, 0.08	0.04, 0.09	0.03, 0.08	0.02, 0.05	0.03, 0.09
Lévy	0.00, 0.00	0.00, 0.02	0.00, 0.01	0.01, 0.02	0.00, 0.00	0.00, 0.00	0.01, 0.02
Power							
$b_{0,1}$ to $b_{0,2}$	1.00, 1.00	1.00, 1.00	1.00, 1.00	1.00, 1.00	1.00, 1.00	1.00, 1.00	1.00, 1.00
$b_{1,1}$ to $b_{1,2}$	1.00, 1.00	1.00, 1.00	1.00, 1.00	1.00, 1.00	1.00, 1.00	1.00, 1.00	1.00, 1.00
$b_{0,2}, b_{1,2}$	1.00, 1.00	1.00, 1.00	1.00, 1.00	1.00, 1.00	1.00, 1.00	1.00, 1.00	1.00, 1.00
N(0,1) to t(6)	0.05, 0.09	0.99, 1.00	0.16, 0.28	0.05,0.12	0.05, 0.08	0.37, 0.44	0.99, 1.00
t(6) to t(3)	0.28, 0.35	1.00, 1.00	0.39, 0.49	0.31, 0.45	0.21, 0.30	0.25, 0.32	1.00, 1.00
Lévy to Cauchy	0.03, 0.04	0.69, 0.71	0.53, 0.59	0.82, 0.85	0.04, 0.05	0.04, 0.05	0.63, 0.66

Notes: Results based on 5-minute GARCH(1,1): $b_0 = 0.1, b_1 = 0.3, \gamma = 0.3$. For the $(R_{5m})^2$, $|R_{5m}|$ and the Lévy process (size and power) we consider the ARHAC estimator with fix lag length=1. This option yields no size distortion for the above the 5-minute processes. For the rest of the processes the results are based on AIC lag length selection of the ARHAC estimator. The break point occurs at $k = 0.8T$.

Table 5: Kokoszka and Leipus CUSUM-type test based for various intraday asset returns processes using intraday volatility filters

	$(Rmin_{1h})^2$	$|Rmin_{1h}|$	$|R_{1h}|$	PV_{1h}
Size				
N(0,1)	0.24, 0.30	0.23, 0.32	0.03, 0.96	0.00, 0.00
t(0,1;6)	0.54, 0.68	0.24, 0.33	0.28, 0.38	0.01, 0.02
Lévy	0.00, 0.00	0.00, 0.01	0.01, 0.03	0.01, 0.02
Power				
$b_{0,1}$ to $b_{0,2}$	0.96, 0.99	1.00, 1.00	1.00, 1.00	0.99, 0.99
$b_{1,1}$ to $b_{1,2}$	1.00, 1.00	1.00, 1.00	1.00, 1.00	0.93, 0.99
$b_{0,2}, b_{1,2}$	0.88, 0.94	0.34, 0.48	0.43, 0.55	0.83, 0.96
N(0,1) to t(6)	0.65, 0.77	0.37, 0.49	0.93, 0.96	0.03, 0.25
t(6) to t(3)	0.90, 0.93	0.79, 0.85	0.31, 0.38	0.04, 0.14
Lévy to Cauchy	0.11, 0.15	0.60, 0.67	0.53, 0.60	0.68, 0.72

	$|R_{6h}|$	$|Rmin_{6h}|$	RA_{6h}	QV_{6h}	PV_{6h}
Size					
N(0,1)	0.00, 0.00	0.04, 0.06	0.09, 0.14	0.00, 0.02	0.00, 0.01
t(0,1;6)	0.16, 0.22	0.17, 0.24	0.07, 0.15	0.01, 0.02	0.00, 0.01
Lévy	0.00, 0.02	0.00, 0.02	0.00, 0.00	0.00, 0.00	0.01, 0.02
Power					
$b_{0,1}$ to $b_{0,2}$	0.99, 0.99	0.99, 0.99	0.90, 0.94	0.77, 0.84	0.91, 0.95
$b_{1,1}$ to $b_{1,2}$	1.00, 0.99	0.99, 0.00	0.70, 0.79	0.35, 0.50	
$b_{0,2}, b_{1,2}$	0.99, 0.99	0.98, 0.98	0.66, 0.76	0.27, 0.42	0.80, 0.86
N(0,1) to t(6)	0.92, 0.95	0.91, 0.94	0.36, 0.53	0.09, 0.17	0.57, 0.74
t(6) to t(3)	0.64, 0.73	0.69, 0.77	0.15, 0.24	0.06, 0.13	0.39, 0.52
Lévy to Cauchy	0.35, 0.44	0.82, 0.83	0.04, 0.05	0.02, 0.04	0.63, 0.68

Notes: Results based on 5-minute GARCH(1,1): $b_0 = 0.00044, b_1 = 0.93482, \gamma = 0.06341$. $T = 250$ days or 72000 5-minutes. The $|R_{5min}|$ and $(R_{5min})^2$ have distorted size for high-persistent IGARCH processes and hence are not reported. Note that this result is also valid for the hourly Range and Quadratic Variation. Note that this result is true for whatever parameter choice of the ARHAC estimator. (2) The ARHAC is estimated using the AIC and de-meaning the series for the following processes: RM_{1h}, $|R_{6h}|$, $Range_{6h}$, QV_{6h} and PV_{6h}. The tests are performed for the ARHAC with the fix lag length VAR(1) and no de-meaning for: $(Rmin_{1h})^2$, $|Rmin_{1h}|$, $|R_{1h}|$, PV_{1h}.