New CUSUM Based Ratio Tests for Parameter Constancy: With application to variance stability*

Uwe Hassler[a], Mehdi Hosseinkouchack[a],

Paulo M.M. Rodrigues[b] and Antonio Rubia[c]

[a] Goethe University Frankfurt

[b] Banco de Portugal, NOVA School of Business and Economics, Universidade Nova de Lisboa

[c] University of Alicante

February 14, 2017

Abstract

New CUSUM based tests for parameter constancy are introduced following a class of ratio tests recently proposed by Hassler and Hosseinkouchack (2016), which are based on a Karhunen-Loeve expansion. Theoretical results are presented and an in-depth Monte Carlo analysis is performed to evaluate the finite sample performance of the procedure and a comparison with available approaches, such as the tests proposed by Berkes et al. (2006) and Shao and Zhang (2010), is provided. The new tests present superior finite sample size properties and interesting power behaviour. An empirical application to the Bitcoin Coindesk index is also included.

*We thank the participants of the CFE conference in Seville, December 2016, for useful comments and suggestions. Paulo M. M. Rodrigues is funded by Portuguese National Funds through FCT - Fundação para a Ciência e Tecnologia under the project Ref. UID/ECO/00124/2013.

1 Introduction

Model stability is important for prediction, econometric inference and adequate interpretation of regression results (Hansen, 1992). Since the seminal work of Page (1955) there has been great interest in the theory and applications of change point analysis in various fields; see e.g. Csörgő and Horváth (1997) and Perron (2006) for interesting overviews. One popular method for its ease of usage has been the CUSUM test and several of its variants. The CUSUM test was proposed by Brown, Durbin and Evans (1975) and is used to detect model instability of general form, both in the mean and in the variance.

In recent literature, variance homogeneity in particular has received considerable attention; see, for example, Inclán and Tiao (1994), Loretan and Phillips (1994), Aggarwal, Inclán and Leal (1999), Berkes et al. (2006), Lee and Park (2001), Andreou and Ghysels (2002), Sansó, Aragó and Carrion (2004), Deng and Perron (2008b), Rapach and Strauss (2008), Cavaliere and Taylor (2008), Rodrigues and Rubia (2011), among others. Several of the approaches proposed in this literature belong to the CUSUM of squares test family. One popular approach belonging to the latter class was introduced by Inclán and Tiao (1994) who proposed a statistic directed to test for changes in the unconditional variance of a stochastic process. However, it is built on the restrictive assumptions that the disturbances are independent and Gaussian distributed, which is an important limitation for applications on economic and financial time series. In fact, Sansó et al. (2004) and Rodrigues and Rubia (2011) showed that, for stochastic processes which are not mesokurtic and do not have a constant conditional variance, the distribution of the Inclán and Tiao (1994) test depends on nuisance parameters (i.e. is non pivotal). Andreou and Ghysels (2002) illustrate experimentally the pervasive effects of persistent volatility on CUSUM tests and Rodrigues and Rubia (2011) formally discuss the implications of sample contamination on the asymptotic properties of CUSUM-type tests. To overcome these problems, new procedures which belong to the CUSUM-type family were proposed which have an asymptotic distribution free of nuisance parameters. In particular, the non-parametric generalization proposed by Berkes et al. (2006) has proven very useful. The non-parametric correction introduced ensures invariance and renders consistent estimates of the break-date.

In this paper we propose a new test statistic which builds on the variance ratio approach recently introduced by Hassler and Hosseinkouchack (2016). We analyse the asymptotic behaviour of our new proposal and evaluate its finite sample performance through an in-depth Monte Carlo analysis. A comparison to the improved version of the Inclán and Tiao (1994) approach proposed by Berkes et al. (2006), which is suitable for time series which display dependence and conditional heteroscedasticity; as well as to the recently introduced test by Shao and Zhang (2010) which also displays robustness against features which are commonly found in economic and financial time series is also provided. Overall, the results indicate that our new procedure displays superior finite sample size performance and interesting power behavior comparatively to currently available tests, thus making it an interesting and useful contribution to this literature.

The rest of the paper is organized as follows. Section 2 outlines the CUSUM type test statistics which are analysed in this article and introduces the new test procedure. Section 3 presents a Monte Carlo analysis of the finite sample properties of the tests and compares their size and power performance. Section 4 addresses the variance homogeneity hypothesis of the Bitcoin series and Section 5 summarizes and concludes.

2 CUSUM type tests

In this section we introduce the test for parameter constancy which is an application of a principle proposed by Hassler and Hosseinkouchack [HH] (2016). The test is based on the ratio of quadratic forms of weighted CUSUM quantities and can be used to test for changes in the mean as well as in the variance. We also describe briefly the improved version of the Inclán and Tiao (1994) approach proposed by Berkes et al. (2006) (see also Sansó et al., 2004), which is suitable for series which display dependence and conditional heteroscedasticity; as well as the recently proposed procedure by Shao and Zhang (2010) which is also robust against these features.

2.1 The mean change case

To illustrate the CUSUM type approaches that will be analysed in this paper let us look first at the case of testing for a mean change. Consider under the null hypothesis a stationary, univariate time series of length T, $\{y_t\}_{t=1}^{T}$, with constant mean; see

Assumption 1 below. Moreover, define the array of normalized partial sums, which correspond to the CUSUM quantities or form the CUSUM process (see Brown, Durbin and Evans, 1975) as,

$$S_k := \frac{\sum_{t=1}^{k}(y_t - \bar{y})}{\sqrt{T}} = \frac{1}{\sqrt{T}}\left(\sum_{t=1}^{k} y_t - \frac{k}{T}\sum_{t=1}^{T} y_t\right). \tag{1}$$

All tests building on the array[1] $\{S_k\}_{k=1}^{T}$ defined in (1) are said to fall into the class of CUSUM tests. Under familiar assumptions we have a functional central limit theorem holding under the null hypothesis, i.e., for $T \to \infty$ it follows that,

$$S_{\lfloor sT \rfloor} \Rightarrow \omega BB(s), \quad 0 \leq s \leq 1, \tag{2}$$

where $BB(s) := W(s) - sW(1)$ is a standard Brownian bridge, $W(s)$ a standard Wiener process, and the scaling parameter ω is a finite positive constant. When testing for a constant mean we consider the following assumptions.

Assumption 1 *Let $\{y_t\}_{t=1}^{T}$ be a weakly stationary process with mean $\mu := E(y_t)$, autocovariances $\gamma_y(h) := Cov(y_t, y_{t+h})$, and finite and positive long-run variance:*

$$0 < \omega_y^2 := \sum_{h=-\infty}^{\infty} \gamma_y(h) < \infty. \tag{3}$$

Further, assume that (2) holds with $\omega = \omega_y$.

Remark 1 *More technical assumptions such that (2) holds for $\{y_t\}$ are given e.g. in Phillips and Perron (1988) or Phillips and Solo (1992). More generally, in the computation of (1) we may consider regression residuals, $\widehat{u}_t := y_t - \widehat{\beta}' x_t$ instead of $y_t - \bar{y}$. The null hypothesis is that β is constant over the sample, while this does not hold true under the alternative. Ploberger and Krämer (1992) stress that a CUSUM test does "not specify a particular pattern of possible coefficient variation. We rather view the CUSUM procedures as pure significance tests, with power against various alternatives". With $x_t = 1$, we have $\widehat{u}_t := y_t - \overline{y}$, which is the case on which, without loss of generality, we will focus here for simplicity, such that no further assumptions with respect to $\{x_t\}$ have to be made.*

[1] For brevity we write S_k instead of $S_{k,T}$.

If the long-run variance, ω^2, was known, one could use (1) to compute the infeasible max statistic in the tradition of Kolmogorov-Smirnov with the following limiting null distribution in the case of a constant mean (or constant parameter vector β):

$$\mathcal{M} := \max_{1\leq k\leq T} \frac{|S_k|}{\omega} \;\Rightarrow\; \sup_{0\leq s\leq 1} |BB(s)|\,, \tag{4}$$

and for which asymptotic critical values are available. In practice, since the max test in (4) is infeasible, an estimate of ω is required. Thus, considering a weakly consistent estimator, $\widehat{\omega}^2 \xrightarrow{p} \omega^2$ (see e.g. Liu and Wu, 2010), we can then define the statistic $\widehat{\mathcal{M}}$ where the limiting distribution remains unchanged under the null hypothesis, i.e., as $T \to \infty$,

$$\widehat{\mathcal{M}} := \max_{1\leq k\leq T} \frac{|S_k|}{\widehat{\omega}} \;\Rightarrow\; \sup_{0\leq s\leq 1} |BB(s)|; \tag{5}$$

see e.g. Aue and Horváth (2013) for details.

To circumvent the difficulties associated with the estimation of ω, we adopt a proposal by HH and consider the ratio of quadratic forms of weighted CUSUM quantities S_k. For illustration of the procedure let BB_θ denote a Brownian bridge with random drift, such as,

$$BB_\theta(s) := BB(s) + \theta\, s(1-s)\, Z\,, \tag{6}$$

where Z is a standard normal variable independent of BB. Under the null hypothesis, H_0: $\theta_0 = 0$, and under the alternative H_1: $\theta = \theta_1$, where θ_1 is chosen such that the asymptotic power equals 50% at θ_1. HH (Prop. 2) determine the eigenfunctions and eigenvalues, $f_{j,\theta_h}(s)$ and λ_{j,θ_h}, respectively, of the autocovariance kernel of BB_{θ_h} under the null and the alternative hypotheses, $h \in \{0,1\}$. This eigenstructure is used to compute q weighted sums of the CUSUM process S_k, i.e.,

$$X_{j,\theta_h} := \sum_{k=1}^{T} \left[\int_{(k-1)/T}^{k/T} f_{j,\theta_h}(s)\, ds \right] S_k\,, \quad j = 1, \ldots, q\,. \tag{7}$$

Hence, the variance ratio test advocated by HH under the null hypothesis ($\theta_0 = 0$) is

$$\mathcal{VR}_q = \frac{\sum_{j=1}^{q} \lambda_{j,\theta_0} X_{j,\theta_0}^2}{\sum_{j=1}^{q} \lambda_{j,\theta_1} X_{j,\theta_1}^2}\,, \tag{8}$$

with

$$\sum_{j=1}^{q} \lambda_{j,\theta_h} X_{j,\theta_h}^2 \ \Rightarrow\ \omega^2 \mathcal{L}_{\theta_h}(\theta)\,, \tag{9}$$

where the limit $\mathcal{L}_{\theta_h}(\theta)$ in (9) is a quadratic form characterized in HH (eq. (22)) which depends on the chosen weights given θ_h, and on the true unknown parameter value θ. Consequently, the nuisance parameter ω^2 present in (9) cancels from $\mathcal{VR}_q$ asymptotically, without having to be estimated, i.e.,

$$\mathcal{VR}_q \ \Rightarrow\ \mathcal{L}_q(\theta) = \frac{\mathcal{L}_{\theta_0}(\theta)}{\mathcal{L}_{\theta_1}(\theta)} \tag{10}$$

The limiting cumulative distribution function of $\mathcal{L}_q(\theta)$ in (10) can be computed from HH (Thm.1), which provides critical values under the null hypothesis as well as asymptotic power curves under the alternative. This results in a family of tests depending on q such that all have correct asymptotic size. The parameter q is assumed to be finite. HH stress that a recommendable choice of q balancing finite sample size distortions and power depends on the specific testing problem; for our new CUSUM variance ratio test this will be discussed below.

In terms of consistency consider the following theorem.

Theorem 1 *Considering a break at $\lfloor \tau T \rfloor$, such that before the break $E(y_t) = \mu_1$ and after the break $E(y_t) = \mu_2$, with $\mu_1 \neq \mu_2$; and*

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{\lfloor rT \rfloor} (y_t - \mu_1) \Rightarrow \omega_y W(r) \ \ for\ r \leq \tau, \tag{11}$$

$$\frac{1}{\sqrt{T}} \sum_{t=\lfloor \tau T \rfloor + 1}^{\lfloor rT \rfloor} (y_t - \mu_2) \Rightarrow \omega_y \left(W(r) - W(\tau) \right) \ \ for\ r > \tau, \tag{12}$$

it follows that,

$$\mathcal{VR}_q \xrightarrow{p} R(q,\tau),$$

where

$$R(q,\tau) = \frac{\sum_{j=1}^{q} \lambda_{j,\theta_0} \left(\int_0^1 f_{j,\theta_0}(r) \left(\min(r,\tau) - r\tau \right) dr \right)^2}{\sum_{j=1}^{q} \lambda_{j,\theta_1} \left(\int_0^1 f_{j,\theta_1}(r) \left(\min(r,\tau) - r\tau \right) dr \right)^2}.$$

Proof: See Appendix.

Remark 2 *Let $cv_{q,\alpha}$ be the α level critical value of the limiting null distribution of $\mathcal{VR}_q$, that is $\Pr\{\mathcal{L}_q(\theta_0) > cv_{q,\alpha}\} = \alpha$. Then, $\mathcal{VR}_q$ rejects the null of no break at a nominal size of α if $R(q,\tau) > cv_{q,\alpha}$. $R(q,\tau)$ is deterministic and depends on τ and q. Therefore, the consistency of $\mathcal{VR}_q$ depends on τ, q and α, that is for some $\tau \in [\tau_{q,\alpha}, 1-\tau_{q,\alpha}]$ we have $\Pr\{\mathcal{VR}_q > cv_{q,\alpha}\} \to 1$ as $T \to \infty$. The specific values for θ_1 and $\tau_{q,\alpha}$ for different values of q and $\alpha = 0.01, 0.05$ and 0.10 are summarized in Table 1. The values for θ_1 for each q are determined using the approach of King (1980) in the framework laid out by Hassler and Hosseinkouchack (2016).*

Table 1: Asymptotic critical values and corresponding values for θ_1 for $\mathcal{VR}_q$ and the corresponding values for $\tau_{q,\alpha}$.

q	5	6	7	8	9	10	50
θ_1	6.2594	5.8792	5.6475	5.4935	5.3832	5.3003	4.8205
$cv_{q,0.01}$	4.5742	3.3746	2.7492	2.3739	2.1265	1.9525	1.1277
$cv_{q,0.05}$	2.5741	2.0998	1.8400	1.6776	1.5670	1.4870	1.0723
$cv_{q,0.10}$	1.9764	1.7017	1.5457	1.4457	1.3764	1.3256	1.0505
$\tau_{q,0.01}$	–	0.40	0.32	0.28	0.25	0.22	0.06
$\tau_{q,0.05}$	0.28	0.22	0.19	0.16	0.14	0.13	0.04
$\tau_{q,0.10}$	0.20	0.16	0.14	0.12	0.10	0.09	0.04

For an illustration of the interpretation of the results in Table 1, note that, for example, considering a 1% significance level with $q = 9$, the test will detect breaks if they occur between the first and last quarter of the sample. On the other hand, if a more conservative significance level is considered, say 10%, with $q = 9$ the test will detect breaks if they occur between the first and the last 10 % observations of the sample.

To complete the discussion of this section we include one further test which was recently proposed by Shao and Zhang (2010) and which displays two interesting features: First and similarly to the Karhunen-Loeve based ratio in (8), it is also a ratio such that the long-run variance cancels without having to be estimated; second, it is directed against the alternative of a break in parameters. This test also makes use of the CUSUM process S_k in (1), and requires the following variance-type term V_k:

$$V_k := \frac{1}{T^2}\left[\sum_{t=1}^{k}\left\{\sum_{j=1}^{t} y_j - \frac{t}{k}\sum_{j=1}^{k} y_j\right\}^2 + \sum_{t=k+1}^{T}\left\{\sum_{j=t}^{T} y_j - \frac{T-t+1}{T-k}\sum_{j=k+1}^{T} y_j\right\}^2\right]. \quad (13)$$

Thus, the test statistics suggested by Shao and Zhang (2010) is,

$$\mathcal{SZ} := \max_{1\leq k<T} \frac{S_k^2}{V_k} \,. \tag{14}$$

Under the null hypothesis of no break, $S^2_{\lfloor sT \rfloor}$ and $V_{\lfloor sT \rfloor}$ converge to limiting processes which are both scaled by the long-run variance $\omega^2 = \omega_y^2$, and which hence cancels from $\mathcal{SZ}$ asymptotically; see Shao and Zhang (2010, Thm. 3.1). The denominator V_k is tailored such as to detect "the one change point alternative"; see Shao and Zhang (2010, p.1230). Breakpoint determination will be addressed in subsection 2.3.

2.2 The variance change case

In our application we wish to test the return series r_t for constant unconditional variance. This amounts to testing either the squared returns, r_t^2, or the absolute returns, $|r_t|$, for a constant mean. Note that this can be generalized as suggested in Remark 1. Hence, in what follows for simplicity of presentation we only consider $y_t := r_t^2$ (but the results also hold if alternatively $y_t := |r_t|$ is used), where the returns satisfy the following assumption.

Assumption 2 *Let $\{r_t\}_{t=1}^T$, be a weakly stationary process with mean $E(r_t) = 0$, $E(r_t^2) = \sigma^2$, finite fourth moments $E(r_t^4) = \kappa\sigma^4$ and $Var(r_t^2) = (\kappa - 1)\sigma^4$. We assume that the long-run variance of the squares is finite and positive, i.e.,*

$$0 < \omega_2^2 := Var(r_t^2) + 2\sum_{h=1}^{\infty} E\left[(r_t^2 - \sigma^2)(r_{t+h}^2 - \sigma^2)\right] < \infty \,.$$

Further, we also maintain that,

$$\frac{1}{\sqrt{T}}\left(\sum_{t=1}^{\lfloor sT \rfloor} r_t^2 - \frac{\lfloor sT \rfloor}{T}\sum_{t=1}^{T} r_t^2\right) \Rightarrow \omega_2 BB(s)\,, \quad 0 \leq s \leq 1\,. \tag{15}$$

Assumption 2 implies that (2) holds with $\omega = \omega_2$. Assumption 2 is met e.g. by all stationary GARCH processes with finite fourth moments (see Berkes et al., 2006). More general technical, sufficient conditions for (15) are established by Sansó, Aragó, Carrion (2004, Prop. 3). Consequently, we can apply the procedures of Section 2.1 to test the null hypothesis that σ^2 is constant over the sample. We only have to replace y_t by r_t^2 in (1) when computing S_k to get three tests for constant variance. First, $\widehat{\omega}$ in (5) is replaced by a consistent estimator of ω_2; then the new $\widehat{\mathcal{M}}$ amounts to the test statistic called κ_2 in Sansó, Aragó and Carrion (2004). Second, $\mathcal{VR}_q$ is computed from the new S_k in terms of r_t^2. Third, y_t is replaced by r_t^2 when computing V_k from (13) entering the Shao-Zhang statistic in (14). Then, the limiting distribution characterized in the previous section continues to hold under the null hypothesis of constant unconditional variance.

Remark 3 *Note that the result of Theorem 1 also holds in this case. In other words, when the interest lies on testing for a break in the variance of a zero mean process, say r_t, then the results of Theorem 1 hold under $\frac{1}{\sqrt{T}}\sum_{t=1}^{\lfloor rT\rfloor}(r_t^2-\sigma_1^2) \Rightarrow \omega_{r^2,1}W(r)$ for $r \leq \tau$ and $\frac{1}{\sqrt{T}}\sum_{t=\lfloor \tau T\rfloor+1}^{\lfloor rT\rfloor}(r_t^2-\sigma_2^2) \Rightarrow \omega_{r^2,2}\left(W(r)-W(\tau)\right)$ for $r > \tau$, where before and after the break we have that $\mathrm{Var}(r_t)=\sigma_1^2$ and $\mathrm{Var}(r_t)=\sigma_2^2$, respectively, with corresponding long-run variances $\omega_{r^2,1}^2$ and $\omega_{r^2,2}^2$.*

2.3 Iterative breakpoint detection

We now consider the specific alternative with a break in parameter (μ when testing for a break in the mean or σ^2 when testing for a break in the variance). The break is assumed to occur after a certain fraction of the data, at time $\lfloor \tau T\rfloor$. Following Bai and Perron (1998), one can use a sequential testing procedure to estimate the number of breaks and the breakpoints at the same time; see also Bai (1994, 1997). Such an iterated CUSUM procedure has also been spelled out for the case of variance testing by Inclán and Tiao (1994). The first step consists of a test over the full sample. If $\mathcal{VR}_q$ is significant, then we determine the first break point $\lfloor \tau T\rfloor$ by estimating τ. To that end one determines the strongest contribution of the CUSUM terms as,

$$\widehat{\tau} = \frac{1}{T}\underset{1\leq k\leq T}{\operatorname{argmax}}\, S_k^2\,. \tag{16}$$

Then one splits the sample in data from 1 to $\lfloor \widehat{\tau}T\rfloor$ and from $\lfloor \widehat{\tau}T\rfloor+1$ to T. In the second step, $\mathcal{VR}_q$ is applied to both subsamples, where one may either use the asymptotic critical values as before, or simulated finite sample critical values as suggested by Sansó, Aragó and Carrion (2004) to improve the performance. If the test is not significant in both subsamples, one concludes that there is only one break. If $\mathcal{VR}_q$ is significant, one proceeds with the corresponding subsample as in step 1. This iterative procedure is repeated until no further rejection occurs.

An analogous iterated CUSUM procedure can be applied with $\mathcal{SZ}$, where this tests is in fact settled in the framework of parameter breaks. Since the denominator V_k also depends on k, the change point estimation differs from $\widehat{\tau}$ in that

$$\widetilde{\tau} = \frac{1}{T} \operatorname*{argmax}_{1 \leq k < T} \frac{S_k^2}{V_k}\,. \tag{17}$$

Except for this modification, the iterated CUSUM procedure can be implemented with $\mathcal{SZ}$ as summarized above.

If the true alternative is a break in parameters, then the question naturally arises whether $\widetilde{\tau}$ is superior to $\widehat{\tau}$ in determining the true unknown breakpoint(s). Furthermore, it is also important to investigate whether $\mathcal{SZ}$ outperforms $\widehat{\mathcal{M}}$ or $\mathcal{VR}_q$ in terms of power and size in finite samples. This brings us to the next section.

3 Monte Carlo evidence

Although the test introduced in this paper can be applied to test for mean breaks as well as for variance breaks, in the Monte Carlo analysis developed in this section will only focus on the performance of the tests when testing for constant variance. The data generation process (DGP) considered is,

$$r_t = \begin{cases} \sigma_{1,t}\eta_t, \text{ for } t = 1, 2, ..., \lfloor \tau T \rfloor \\ \sigma_{2,t}\eta_t, \text{ for } t = \lfloor \tau T \rfloor + 1, ..., T, \end{cases} \tag{18}$$

where $\lfloor \tau T \rfloor$ is the break date, $\eta_t \sim iid\mathcal{N}(0, 1)$ and

$$\begin{cases} \sigma_{1,t}^2 = \gamma + \alpha r_{t-1}^2 + \beta \sigma_{1,t-1}^2 \\ \sigma_{2,t}^2 = (\gamma + \delta_1) + \alpha r_{t-1}^2 + (\beta + \delta_2)\sigma_{2,t-1}^2, \end{cases} \tag{19}$$

with $(\gamma, \beta) \in \{(0.4, 0.5), (0.2, 0.7), (0.1, 0.8), (0.05, 0.85), (0.01, 0.89)\}$, $\alpha = 0.1$, $\delta_1, \delta_2 \in \{-0.1, 0, 0.1\}$, $\tau \in \{0.3, 0.5\}$ and $T \in \{500, 1000, 3000\}$. When $\delta_1 = \delta_2 = 0$, there will be no break in the parameters, while when $\delta_1 \neq 0$ or $\delta_2 \neq 0$ a break will be present. To provide a general picture on the empirical size ($\delta_1 = \delta_2 = 0$) for different tests we also set $(\alpha + \beta) \in [.80, 0.99]$ with $\alpha \in [0.01, 0.20]$ and $\gamma = 1 - \alpha - \beta$ for $T \in \{500, 1000\}$. Moreover, we replace (19) by (20), below, to shed light on the empirical local power of the tests we consider in this paper:

$$\begin{cases} \sigma_{1,t}^2 = \gamma + \alpha r_{t-1}^2 + \beta \sigma_{1,t-1}^2 \\ \sigma_{2,t}^2 = (\gamma + c/\sqrt{T}) + \alpha r_{t-1}^2 + \beta \sigma_{2,t-1}^2, \end{cases} \tag{20}$$

for $(\alpha, \beta, \gamma) = (0.01, 0.94, 0.05)$ where $c \in [-1, 1]$ and $(\alpha, \beta, \gamma) \in \{\ (0.01, 0.94, 0.05), (0.10, 0.89, 0.01)\ \}$ where $c \in [0, 1]$. For the empirical local power analysis we set $T \in \{500, 1000\}$ and $\tau \in \{0.3, 0.4, 0.5, 0.6, 0.7\}$. All results are based on 25,000 replications and the nominal size is 5%.

In Table 2 we present the empirical rejection frequencies of the three tests of section 2, the $\widehat{\mathcal{M}}$, $\mathcal{VR}_q$, for $q = 5, 6, 7$ and $\mathcal{SZ}$ tests. In particular, we consider the test statistic given in equation (4) in which the long-run variance is estimated either using a deterministic bandwidth given by $\lfloor 4(T/100)^{(2/9)} \rfloor$ or the bandwidth choice is data-driven as suggested by Andrews (1991); we denote these as $\widehat{\mathcal{M}}_1$ and $\widehat{\mathcal{M}}_2$, respectively. For both $\widehat{\mathcal{M}}_1$ and $\widehat{\mathcal{M}}_2$ we employed a Bartlett kernel as suggested by Berkes et al. (2006).

[Please insert Table 2 about here]

From the first panel of Table 2, where $\delta_1 = \delta_2 = 0$, we observe that $\widehat{\mathcal{M}}_1$ and $\widehat{\mathcal{M}}_2$ only present adequate empirical size when the underlying GARCH process is not very persistent ($\alpha = 0.1$ and $\beta = 0.5$). As the GARCH persistence increases, $\beta > 0.7$, the empirical size of these tests deteriorates considerably, a feature that does not seem to ameliorate as T increases (note that for $T = 3000$ we still observe empirical size distortions of 94% and 47% for $\widehat{\mathcal{M}}_1$ and $\widehat{\mathcal{M}}_2$, respectively, in the most extreme case considered ($\omega = 0.1$, $\alpha = 0.1$ and $\beta = 0.89$)). The $\mathcal{SZ}$ test displays considerably improved size performance comparatively to $\widehat{\mathcal{M}}_1$ and $\widehat{\mathcal{M}}_2$ and its performance seems to improve as the sample size increases. However, an overall best empirical size performance is displayed by the $\mathcal{VR}_q$ test. Although some size distortions are still observed in the smaller sample ($T = 500$) as the sample size increases the empirical size is essentially the same as the 5% nominal size considered. Further, we observe that the size performance of $\mathcal{VR}_q$ slightly worsens as q increases when the sample size is small, $T = 500$. These slight size distortions disappear almost completely when $T = 1000$. It should, however, be noted that increasing q is at the same time associated with stronger rejection rates when a break is present.

[Please insert Figures 3 - 14 about here]

In Figures 3 - 8 we present the empirical size results for $T = 500$ while we set $T = 1000$ in Figures 9 - 14, for data generated based on (18) and (19) under $\delta_1 = \delta_2 = 0$, $(\alpha+\beta) \in [0.8; 0.99]$, $\alpha \in [0.01, 0.2]$ and $\gamma = 1-\alpha-\beta$. The conclusions are qualitatively the same as those drawn from the first panel of Table 2.

[Please insert Figures 15 - 20 about here]

In Figures 15 - 17 we present the empirical local power for $T = 500$ while we set $T = 1000$ in Figures 18 - 20, when the data are generated based on (18) and (20) for different values of α, β and $\gamma = 1 - \alpha - \beta$ and when c is either in $[-1, 1]$ or $[0, 1]$. From these Figures we learn for $\mathcal{VR}_q$ that, in general and as mentioned earlier, increasing q leads to more power at the cost of some size distortions when the sample size is small ($T = 500$). Further, we observe that when the tests are relatively correctly sized (as in Figures 15 and 18), then the most powerful test is the conventional $\widehat{\mathcal{M}}$ followed by $\mathcal{SZ}$ and then by $\mathcal{VR}_q$, for $q = 7, 6$ and 5. Furthermore, we observe that increasing the sample size leads, as expected, to a general increase in power and improvements for size.

4 Empirical Application

In this section we provide an application of the $\widehat{\mathcal{M}}$, $\mathcal{VR}_q$, $q = 5, 6$, and 7 and $\mathcal{SZ}$ test procedures introduced in Section 2 to returns computed from the closing price of the Bitcoin (BTC) index [2]. BTC is a cryptocurrency (or virtual money) derived from mathematical cryptography that was introduced by Satoshi Nakamoto (Nakamoto, 2008), and which had its first transaction in January 2009. BTC has since become

[2] For an historical overview of events that marked the development of BTC see https://bitcoinhelp.net/know/more/price-chart-history.

popular because of its unique features: i) there is no central authority guaranteeing or having control over it, i.e., BTC is fully decentralized; and ii) it depends on a sophisticated protocol which limits its supply.

BTC is becoming a potential alternative to standard fiat currencies (e.g., US dollar, the Euro or Japanese Yen) because of its various advantages, such as, low or no fees, a controlled and known algorithm for currency creation, and an informational transparency for all transactions. BTC attracted increased attention when its exchange rate to the US dollar breached the $1000 level (with a maximum of $1242 per BTC at the Mt. Gox market) in late November and early December 2013. However, the development of BTC has also undergone serious drawbacks, of which the most famous are the closing of the Silk Road website on October 2013 and closing of the Mt. Gox market. The former was a website where illegal operations were uncovered thanks to anonymity and the latter was the disappearance, on February 2014, of the Mt. Gox platform which negotiated around 80% of the transactions in BTC, causing a loss of 750,000 BTCs ($440 million).

Since its introduction in 2009, the BTC value has been rapidly increasing, registering a steady growth in popularity despite its volatile development in terms of its US dollar exchange rate. Regardless of whether BTC is regarded as a form of payment, an alternative currency, or a financial asset and speculative investment (Yermack, 2013; Bouoiyour and Selmi, 2015a and 2015b), its time series properties need to be well understood. The few studies that look into the time series properties of this series are the recent papers by Bouri et al. (2016), which analyses the long memory properties of BTC volatility, and the paper by Hencic and Gouriéroux (2015), who investigate the suitability of a noncausal autoregressive model for the BTC/USD exchange rate.

The recent acceptance of BTC as a commodity and financial product by the US Commodity Futures Trading Commission makes the BTC volatility dynamics of crucial importance in any potential derivatives pricing and trading context. However, modeling price volatility and its persistence must account for nonlinearity and structural breaks particularly given the evidence of booms and busts observed in the BTC market (Cheah and Fry, 2015; Harris and Sollis, 2005, Bouri, et al., 2016).

4.1 Data

Our sample consists of daily closing prices for the BTC Coindesk Index from July 19th, 2010 to July 18th, 2016. From Figure 1 we observe that BTC prices, appear relatively stable until early 2013, but increase dramatically in late 2013 (the closing value in 04/12/2013 was $1,147.2457). However, as also evidenced in Figure 2 through the sub-period plots, even in the earliest years of BTC existence prices rose considerably.

Figure 1 also shows that BTC experienced several jumps and a large number of swings between 2010 to 2016. Although BTC's price has displayed considerable variation since its creation in 2009, during the first half of 2015, it gradually became less volatile comparatively to previous years. Note that after a period of great volatility during 2013 and 2014 (where prices grew from a value of less than $20 in January 2013, to over $1,100 in December 2013, and fell to under $400 in December 2014), BTC prices seem, to some extent, to have stabilized in the first half of 2015. However, from the second half of 2015 until the end of the sample a growing trend seems to have emerged again.

Moreover, from Figure 1 it is clear that BTC experienced periods of exuberance. In particular, the April 2013 spike in the BTC series is striking and corresponds to the maximum value achieved by BTC since its inception, and the equally drastic correction of its price in the subsequent period is also noteworthy, a phenomena possibly driven by demand. The large price increase in 2013 has been attributed to the speculative interests in the currency by citizens of troubled EU economies as an alternative to their traditionally bank deposits. This increased interest was, for instance, motivated by the EU bailout of Cyprus which imposed the seizing of percentages of ordinary citizens' bank savings (Bouoiyour and Selmi, 2015b).

The great variability in the BTC price has led to the development of the speculative nature of this virtual currency. Volatility is an important property of any financial asset. It is the basic measure of risk to which investors are exposed to when buying an asset. For this reason it is crucial for BTC investor to assess the potential threats arising from BTC volatility. Hence, testing for breaks in the variance is an important issue as it may help trace the risk profile of this series.

Table 3 presents descriptive statistics for the complete sample considered (19/07/2010

- 18/07/2016) in our analysis, as well as for two sub-periods. The splitting point of the sample coincides with the closing of the Silk Road website.

Table 3: Descriptive statistics of BTC returns

	mean	median	st.dev.	skewness	kurtosis
19/07/2010 - 18/07/2016	0.41%	0.16%	0.0613	-0.3439	14.8429
19/07/2010 - 03/10/2013	0.62%	0.22%	0.0726	-0.4110	12.3629
04/10/2013 - 18/07/2016	0.17%	0.13%	0.0447	-0.1404	14.0562

The first observation we can make from these results are the heavy tailed features of the sample (kurtosis=14.8429) and negative asymmetry (skewness=-0.3439), which is also evidenced by the contrast between the mean and median daily returns.

The division of the sample into the two sub-periods considered serves only to illustrate that in fact the behaviour of the series may not have been constant throughout its development. We observe, for instance, that mean and median daily returns, standard deviation (st. dev.) as well as skewness were larger in the first sub-period than in the second. Only kurtosis displays a larger value in the second sub-period. Hence, it will be useful to apply the variance breaks tests previously discussed to infer whether there were breaks in the varinace and if so, use the iterative procedure described above to determine the break dates.

4.2 Testing for Variance Breaks

To test for unconditional variance breaks we first estimate an autoregressive (AR(p)) model of order p for the returns and use the resulting residuals, $\hat{r}_t^*$, to set up the necessary partial sums to compute the tests. The lag order p is determined based on the Schwarz information criteria considering a maximum lag order estimated by Schwert's rule.

The partial sums were formed from squared (sq) demeaned $\hat{r}_t^*$, $f(\bar{\hat{r}}_t^*) := (\bar{\hat{r}}_t^*)^2$, where $\bar{\hat{r}}_t^* := (\hat{r}_t^* - \bar{\hat{r}}^*)$, and absolute (abs) demeaned $\hat{r}_t^*$, $f(\bar{\hat{r}}_t^*) := |\bar{\hat{r}}_t^*|$,

$$S_k = T^{-1/2} \sum_{t=1}^{k} \left(f(\bar{\hat{r}}_t^*) - \frac{1}{T} \sum_{t=1}^{T} f(\bar{\hat{r}}_t^*) \right)$$

and the $\widehat{\mathcal{M}}$ test discussed in (4), the $\mathcal{VR}_q$ test introduced in (7) and the Shao-Zhang test presented in (9) were computed.

Table 4: $\mathcal{VR}_q$ test results

	$q = 5$	$q = 6$	$q = 7$
$\mathcal{VR}_{q_{abs}}$	3.2104	2.3783	2.2849
$\mathcal{VR}_{q_{sq}}$	3.6819	2.6330	2.4942

Note that application of the $\mathcal{SZ}$ test provides the following results $\mathcal{SZ}_{sq} = 35.4519$ and $\mathcal{SZ}_{abs} = 32.7089$.

From Table 4 and the results for the $\mathcal{SZ}$ test, we observe that the $\mathcal{VR}_q$ procedure discussed in this paper as well as $\mathcal{SZ}$ reject the null hypothesis, thus suggesting that there is evidence of variance breaks in the data. Thus, given the rejection of the null hypothesis by these tests, we proceed with the application of the $\mathcal{M}$, $\mathcal{SZ}$ and $\mathcal{VR}_q$ variance break tests iteratively.

Table 5: Variance Breaks in BTC returns

squared demeaned returns

$\widehat{\mathcal{M}}$		$\mathcal{SZ}$		$\mathcal{VR}_5$		$\mathcal{VR}_6$		$\mathcal{VR}_7$	
N^o of breaks = 3		N^o of breaks = 1		N^o of breaks = 1		N^o of breaks = 1		N^o of breaks = 1	
Break positions		Break positions		Break positions		Break positions		Break positions	
548	15/01/2012	549	18/01/2012	549	18/01/2012	549	18/01/2012	549	18/01/2012
973	17/03/2013	-	-	-	-	-	-	-	-
1369	17/04/2014	-	-	-	-	-	-	-	-

absolute demeaned returns

$\widehat{\mathcal{M}}$		$\mathcal{SZ}$		$\mathcal{VR}_5$		$\mathcal{VR}_6$		$\mathcal{VR}_7$	
N^o of breaks = 3		N^o of breaks = 1		N^o of breaks = 1		N^o of breaks = 1		N^o of breaks = 1	
Break positions		Break positions		Break positions		Break positions		Break positions	
585	23/02/2012	584	22/02/2012	584	22/02/2012	584	22/02/2012	584	22/02/2012
959	03/03/2013	-	-	-	-	-	-	-	-
1371	19/04/2014	-	-	-	-	-	-	-	-

Regarding the first break date (15/01/2012) detected by the $\widehat{\mathcal{M}}$ test, and the (18/01/20012) detected by the $\mathcal{VR}_q, q = 5, 6, 7$ and $\mathcal{SZ}$ tests for the squared demeaned returns, it is likely picking up the aftermath of the great BTC bubble of 2011 which had its peak in June 2011 and registered a price decline of 94% until November 2011. Prices only peaked again in 2013. Interestingly, the second and third break dates of the shifts detected by the $\widehat{\mathcal{M}}$ statistic are close to April 2013 (when BTC reached its maximum) and February 2014 (the disappearance of Mt Gox).

$\mathcal{VR}_q$ and $\mathcal{SZ}$ suggest the presence of only one variance break, both when the squares and the absolute values of the returns are considered, although the break dates change depending on which transformation of the returns is used.

5 Concluding remarks

In this paper a new testing framework for parameter constancy is introduced. The procedures introduced follow a class of ratio tests recently proposed by Hassler and Hosseinkouchack (2016), which are based on a Karhunen-Loeve expansion. Consistency of the procedure is shown theoretically and an in-depth Monte Carlo analysis illustrates the interesting finite sample size properties of the test, which displays superior size behavior than the tests proposed by Berkes et al. (2006) and Shao and Zhang (2010). An empirical application to the Bitcoin Coindesk index is also considered and the results suggest that BTC observed a single variance break in early 2012.

References

Aggarwal, R., Inclán, C. and Leal, R. (1999). Volatility in emerging stock markets, Journal of Financial and Quantitative Analysis 34, 33-55.

Andreou, E. and Ghysels, E. (2002). Detecting multiple breaks dynamics, Journal of Applied Econometrics 17, 579-600.

Andreou, E. and Ghysels, E. (2004). The impact of sampling frequency and volatility estimators and change point test, Journal of Financial Econometrics 2, 290-318.

Aue, A. and L. Horváth (2013) Structural breaks in time series, Journal of Time Series Analysis 34, 1-16.

Bai, J. (1994) Least squares estimation of a shift in linear processes. Journal of Time Series Analysis 15, 453-472.

Bai, J. (1997). Estimating multiple breaks one at a time. Econometric Theory 13, 315-52.

Bai, J. and P. Perron (1998). Estimating and testing linear models with multiple structural changes. Econometrica 66, 47-78.

Bouoiyour, J. and Selmi, R. (2015a) Bitcoin Price: Is it really that New Round of Volatility can be on way? MPRA Paper No. 65580.

Bouoiyour, J., Selmi, R., (2015b) Greece withdraws from Euro and runs on Bitcoin; April Fools Prank or Serious Possibility? Working paper CATT, University of Pau.

Bouri, E., L.A. Gil-Alana, R. Gupta, D. Roubaud (2016) Modelling Long Memory Volatility in the Bitcoin Market: Evidence of Persistence and Structural Breaks, University of Pretoria, Department of Economics Working Paper Series.

Brown, R. L., J. Durbin, and J. M. Evans (1975). Techniques for testing the constancy of regression relationships over time (with discussion). Journal of the Royal Statistical Society. Series B (Methodological) 37, 149-192.

Cavaliere, G. and A.M.R. Taylor (2008). Time-change unit root tests for time series with nonstationary volatility. Journal of Time Series Analysis 29, 300-330.

Cheah, E.T. and J. Fry (2015) Speculative bubbles in Bitcoin markets? An empirical investigation into the fundamental value of Bitcoin, Economics Letters 130, 32-36.

Chen, G., Choi, Y. and Zhou, Z. (2005). 'Nonparametric estimation of structural change points in volatility models for time series', Journal of Econometrics 16, 79-114.

Csörgő, M. and Horváth, (1997) Limit theorems in change-point analysis. Wiley Series in Probability and Mathematical Statistics. Chichester: John Wiley & Sons.

Deng, A. and Perron, P. (2008a). 'The limit distribution of the CUSUM of squares test under general mixing conditions', Econometric Theory 24, 809-822.

Deng, A. and Perron, P. (2008b). 'Anon-local perspective on the power properties of the CUSUM and CUSUM of squares tests for structural change', Journal of Econometrics 142, 212-240.

Hansen, B. (1992) Testing for Parameter Instability in Linear Models, Journal of Policy Modeling 14, 517-533.

Hassler, U. and M. Hosseinkouchack (2016) Ratio Tests under Limiting Normality.

Hencic, A. and C. Gouriéroux (2014) Noncausal Autoregressive Model in Application to Bitcoin/USD Exchange Rates, in Econometrics of Risk, Volume 583 of the series Studies in Computational Intelligence, 17-40.

Inclán, C. and Tiao, C. J. (1994). 'Use of cumulative sums of squares for retrospective detection of changes of variance', Journal of the American Statiscal Association 89, 913-923.

Kokoszka, P. and Leipus, R. (2000). 'Change-point estimation in ARCH models', Bernoulli 6, 513-539.

Krämer, W., Ploberger, W. and Ah, W. (1988) Testing for structural change in dynamic models, Econometrica 56, 1355-1369.

Kristoufek, L. (2015) What Are the Main Drivers of the Bitcoin Price? Evidence from Wavelet Coherence Analysis, PLoS One. 2015; 10(4): e0123923.

Lee, S. and S. Park (2001) The CUSUM of squares test for scale changes in infinite order moving average processes. Scandinavian Journal of Statistics 28, 625-644.

Liu, W. and Wu, W. B. (2010) Asymptotics of spectral density estimates. Econometric Theory 26, 1218-1245.

Loretan, M. and P.C.B. Phillips, 1994. Testing the covariance stationarity of heavy-tailed time series. Journal of Empirical Finance 1, 211-248.

McConnell, M. and Perez-Quirós, G. (2000). 'Output fluctuations in the United States: what has changed since the early 1980s', American Economic Review 90, 1464-1476.

Nakamoto, S. (2008). Bitcoin: A Peer-to-Peer Electronic Cash System.

Pagan, A. R. and Schwert, G.W. (1990). 'Testing for covariance stationarity in stock market data', Economics Letters 33, 165-170.

Page, E. (1955). A Test for a Change in a Parameter Occurring at an Unknown Point. Biometrika 42, 523-527.

Perron, P. (2006) Dealing with Structural Breaks, in Palgrave Handbook of Econometrics, Vol. 1: Econometric Theory, K. Patterson and T.C. Mills (eds.), Palgrave Macmillan, 278-352.

Phillips, P.C.B. and P. Perron (1988) Testing for a Unit Root in Time Series Regression. Biometrica 75, 335-346.

Phillips, P.C.B. and V. Solo (1992). Asymptotics for linear processes. Annals of Statistics 20, 971-1001.

Ploberger, W. and W. Krämer (1992). The CUSUM test with OLS residuals. Econometrica 60, 271-285.

Rapach, D. and J.K. Strauss, 2008. Structural breaks and GARCH models of exchange rate volatility. Journal of Applied Econometrics 23, 65-90.

Rodrigues, P.M.M. and A. Rubia (2011) The effects of additive outliers and measurement errors when testing for structural breaks in variance, Oxford Bulletin of Economics and Statistics 73, 449-468.

Sanso, A., Arago, V. and Carrion, J.L. (2004) Testing for Change in the Unconditional Variance of Financial Time Series, Revista de Economiá Financiera 4, 32-53.

Sensier, M. and van Dijk, D. (2004). Testing for volatility changes in US macroeconomic time series, Review of Economics and Statistics 86, 833-839.

Shao, X. and X. Zhang (2010). Testing for change points in time series. Journal of the American Statistical Association 105, 1228-1240.

Yermack, D. (2013). Is Bitcoin a Real Currency? An economic appraisal. NBER Working Paper No. 19747.

Velde, F.R. (2013) Bitcoin: A primer. Essays on Issues, The Federal Reserve Bank of Chicago No 317, December 2013.

A APPENDIX

Proof of Theorem 1

First note that

$$\frac{1}{T}\sum_{t=1}^{\lfloor rT \rfloor} y_t \xrightarrow{p} \begin{cases} r\mu_1 & \text{for } r \leq \tau \\ \tau\mu_1 + (r-\tau)\,\mu_2 & \text{for } r > \tau, \end{cases}$$

from which we obtain

$$\overline{y} \xrightarrow{p} \tau\mu_1 + (1-\tau)\,\mu_2,$$

and therefore

$$\frac{1}{T}\sum_{t=1}^{\lfloor rT \rfloor} (y_t - \overline{y}) \xrightarrow{p} (\mu_1 - \mu_2)\,(\min(r,\tau) - r\tau)\,.$$

This in turn implies that

$$\begin{aligned} \frac{1}{\sqrt{T}} X_T(r) &= \frac{1}{T}\sum_{t=1}^{\lfloor rT \rfloor} (y_t - \overline{y}) \\ &\xrightarrow{p} (\mu_1 - \mu_2)\,(\min(r,\tau) - r\tau) \end{aligned}$$

Now since

$$X_{j,\theta_h} = \int f_{j,\theta_h}(r)\, X_T(r)\, dr,$$

the result follows.

B Tables and figures

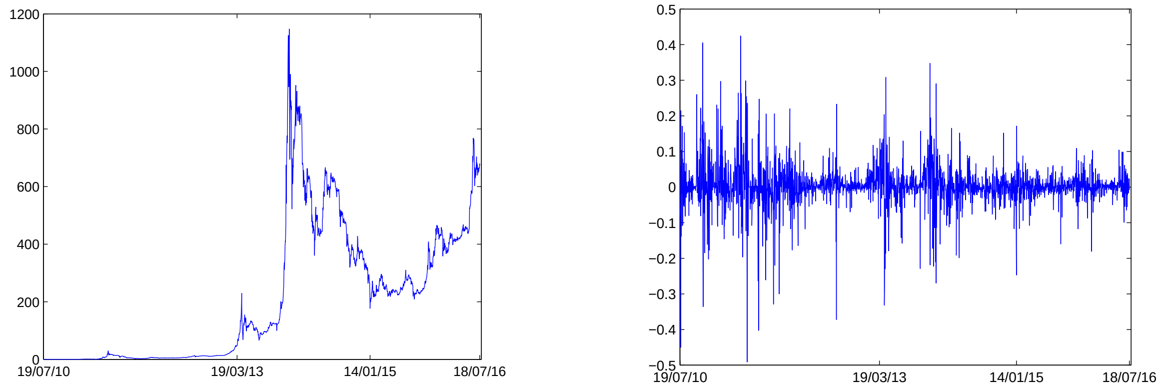

Figure 1: Bitcoin level and return

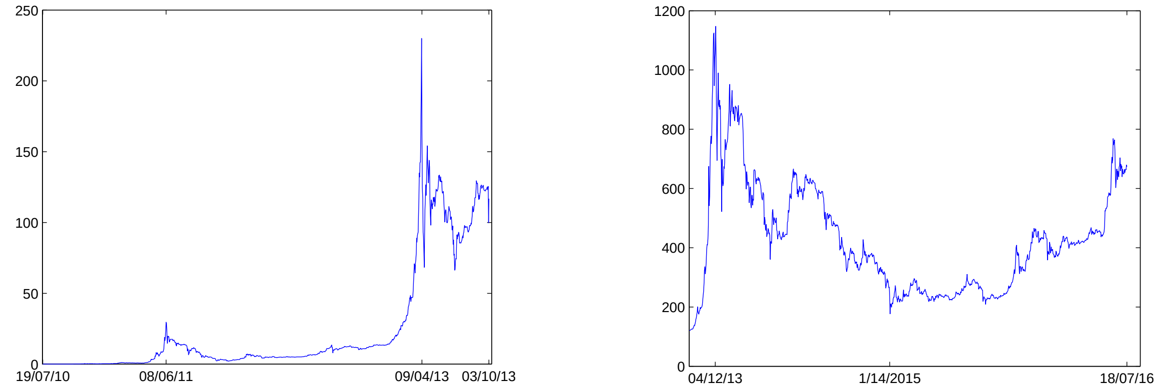

Figure 2: Sub-periods I and II

Table 2: Empirical rejection frequencies under the null and alternative hypotheses of the $\widehat{\mathcal{M}}$, $\mathcal{VR}_q$ and $\mathcal{SZ}$ tests.

			$T=500$						$T=1000$						$T=3000$					
ω	α	β	$\widehat{\mathcal{M}}_1$	$\widehat{\mathcal{M}}_2$	$\mathcal{VR}_5$	$\mathcal{VR}_6$	$\mathcal{VR}_7$	$\mathcal{SZ}$	$\widehat{\mathcal{M}}_1$	$\widehat{\mathcal{M}}_2$	$\mathcal{VR}_5$	$\mathcal{VR}_6$	$\mathcal{VR}_7$	$\mathcal{SZ}$	$\widehat{\mathcal{M}}_1$	$\widehat{\mathcal{M}}_2$	$\mathcal{VR}_5$	$\mathcal{VR}_6$	$\mathcal{VR}_7$	$\mathcal{SZ}$
0.40	0.10	0.50	0.05	0.05	0.05	0.05	0.05	0.06	0.06	0.05	0.05	0.05	0.05	0.06	0.06	0.05	0.05	0.05	0.05	0.05
0.20	0.10	0.70	0.11	0.07	0.05	0.05	0.05	0.07	0.11	0.06	0.05	0.05	0.05	0.06	0.11	0.06	0.05	0.05	0.06	0.05
0.10	0.10	0.80	0.24	0.13	0.05	0.05	0.05	0.08	0.25	0.11	0.05	0.05	0.05	0.07	0.25	0.08	0.05	0.05	0.05	0.06
0.05	0.10	0.85	0.44	0.25	0.05	0.06	0.06	0.12	0.49	0.21	0.05	0.05	0.05	0.09	0.52	0.14	0.05	0.05	0.05	0.07
0.01	0.10	0.89	0.79	0.58	0.08	0.10	0.13	0.28	0.88	0.57	0.05	0.06	0.07	0.21	0.94	0.47	0.04	0.05	0.05	0.13
									$\tau=0.5$											
0.40	0.10	$0.50 \mapsto 0.60$			0.18	0.20	0.22	0.28			0.29	0.34	0.36	0.47			0.54	0.64	0.69	0.87
0.20	0.10	$0.70 \mapsto 0.80$			0.37	0.43	0.46	0.62			0.51	0.60	0.65	0.83			0.74	0.87	0.91	0.99
0.10	0.10	$0.80 \mapsto 0.90$			0.45	0.55	0.61	0.72			0.37	0.45	0.51	0.64			0.31	0.37	0.41	0.57
0.05	0.10	$0.85 \mapsto 0.75$			0.43	0.51	0.55	0.72			0.54	0.64	0.69	0.85			0.74	0.86	0.91	0.98
0.01	0.10	$0.89 \mapsto 0.79$			0.51	0.61	0.67	0.75			0.49	0.57	0.63	0.73			0.52	0.60	0.65	0.77
$0.40 \mapsto 0.50$	0.10	0.50			0.14	0.16	0.17	0.21			0.22	0.25	0.27	0.34			0.43	0.51	0.56	0.73
$0.20 \mapsto 0.30$	0.10	0.70			0.24	0.27	0.29	0.39			0.37	0.43	0.46	0.62			0.63	0.75	0.80	0.95
$0.10 \mapsto 0.20$	0.10	0.80			0.34	0.40	0.43	0.58			0.48	0.57	0.62	0.79			0.71	0.84	0.89	0.98
$0.05 \mapsto 0.15$	0.10	0.85			0.39	0.46	0.50	0.66			0.50	0.59	0.64	0.80			0.70	0.82	0.87	0.97
$0.01 \mapsto 0.11$	0.10	0.89			0.40	0.49	0.55	0.69			0.40	0.48	0.53	0.68			0.48	0.56	0.60	0.74
									$\tau=0.3$											
0.40	0.10	$0.50 \mapsto 0.60$			0.13	0.14	0.15	0.16			0.18	0.21	0.23	0.28			0.31	0.38	0.44	0.63
0.20	0.10	$0.70 \mapsto 0.80$			0.21	0.24	0.26	0.31			0.28	0.33	0.38	0.48			0.39	0.51	0.61	0.85
0.10	0.10	$0.80 \mapsto 0.90$			0.31	0.38	0.43	0.42			0.22	0.27	0.31	0.34			0.16	0.18	0.2	0.25
0.05	0.10	$0.85 \mapsto 0.75$			0.26	0.41	0.52	0.86			0.31	0.54	0.68	0.96			0.43	0.77	0.9	1
0.01	0.10	$0.89 \mapsto 0.79$			0.28	0.59	0.74	0.97			0.28	0.58	0.72	0.97			0.32	0.63	0.75	0.97
$0.40 \mapsto 0.50$	0.10	0.50			0.11	0.11	0.12	0.13			0.15	0.16	0.18	0.2			0.27	0.32	0.37	0.49
$0.20 \mapsto 0.30$	0.10	0.70			0.16	0.18	0.19	0.22			0.22	0.26	0.29	0.36			0.34	0.44	0.53	0.76
$0.10 \mapsto 0.20$	0.10	0.80			0.2	0.23	0.26	0.29			0.27	0.32	0.37	0.47			0.38	0.5	0.6	0.84
$0.05 \mapsto 0.15$	0.10	0.85			0.22	0.25	0.29	0.33			0.27	0.32	0.37	0.45			0.37	0.48	0.57	0.77
$0.01 \mapsto 0.11$	0.10	0.89			0.25	0.3	0.35	0.36			0.22	0.26	0.3	0.32			0.25	0.29	0.33	0.35

Note: $\widehat{\mathcal{M}}_1$ and $\widehat{\mathcal{M}}_2$ represent the test statistic given in equation (5), where for $\widehat{\mathcal{M}}_1$ we estimate the long-run variance using a fixed band-width $[4(T/100)^{(2/9)}]$ while for $\widehat{\mathcal{M}}_2$ we use a random mechanism to choose the bandwidth. For both $\widehat{\mathcal{M}}_1$ and $\widehat{\mathcal{M}}_2$ we employed a Bartlett kernel. $\mathcal{VR}_q$ represent the variance ration test by HH, which is given under equation (8). Finally, $\mathcal{SZ}$ represents the test statistic by Shao and Zhang (2010) which is given under equation (14). The number of replications for MC is 25,000.

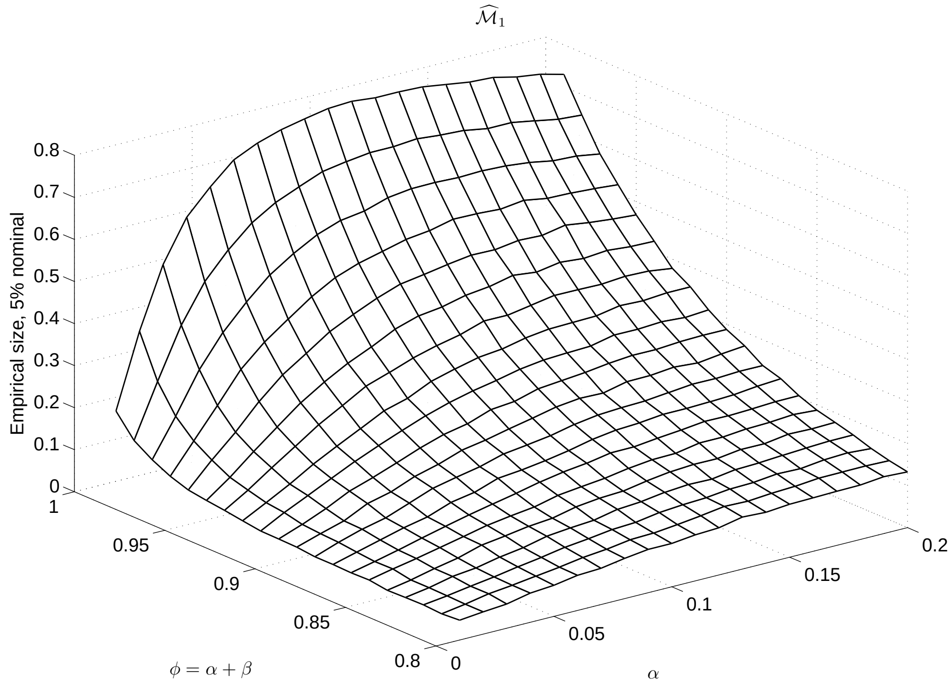

Figure 3: Size properties of $\widehat{\mathcal{M}}_1$ when $T = 500$ and the length of MC is $25,000$. The DGP is outlined under (18) and (19) when $\delta_1 = \delta_2 = 0$ and $\gamma = 1 - \alpha - \beta$. See the notes to Table 2 for further details.

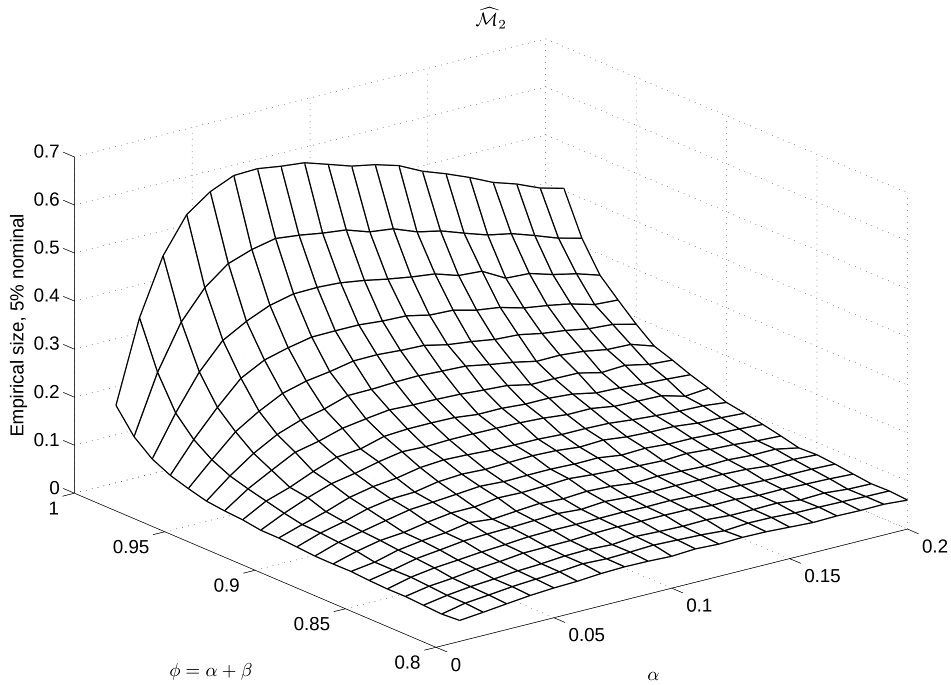

Figure 4: Size properties of $\widehat{\mathcal{M}}_2$ when $T = 500$ and the length of MC is $25,000$. The DGP is outlined under (18) and (19) when $\delta_1 = \delta_2 = 0$ and $\gamma = 1 - \alpha - \beta$. See the notes to Table 2 for further details.

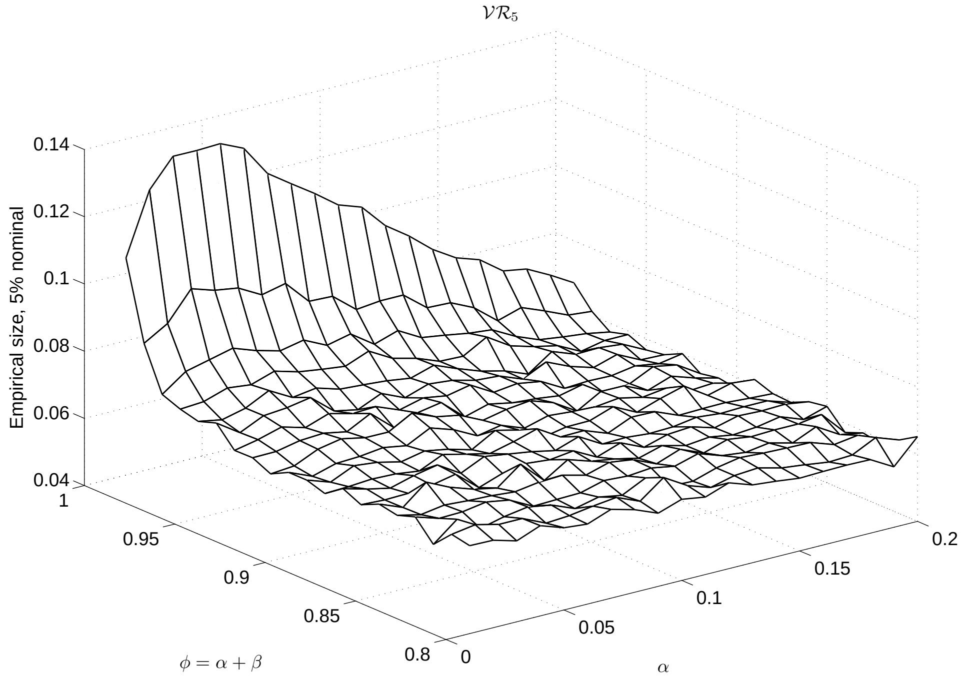

Figure 5: Size properties of $\mathcal{VR}_5$ when $T = 500$ and the length of MC is $25,000$. The DGP is outlined under (18) and (19) when $\delta_1 = \delta_2 = 0$ and $\gamma = 1 - \alpha - \beta$. See the notes to Table 2 for further details.

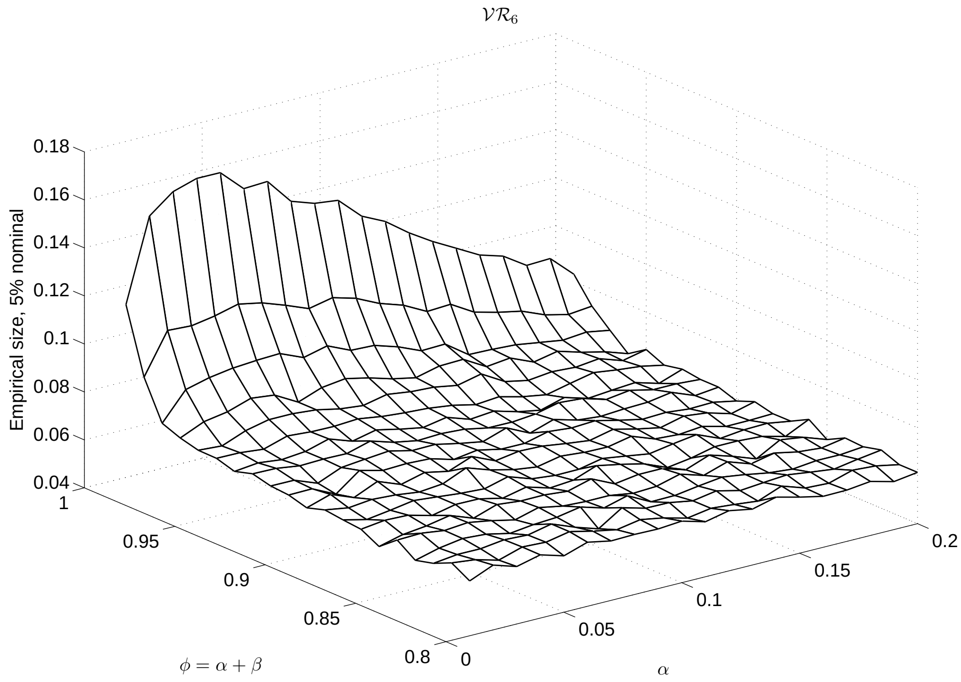

Figure 6: Size properties of $\mathcal{VR}_6$ when $T = 500$ and the length of MC is $25,000$. The DGP is outlined under (18) and (19) when $\delta_1 = \delta_2 = 0$ and $\gamma = 1 - \alpha - \beta$. See the notes to Table 2 for further details.

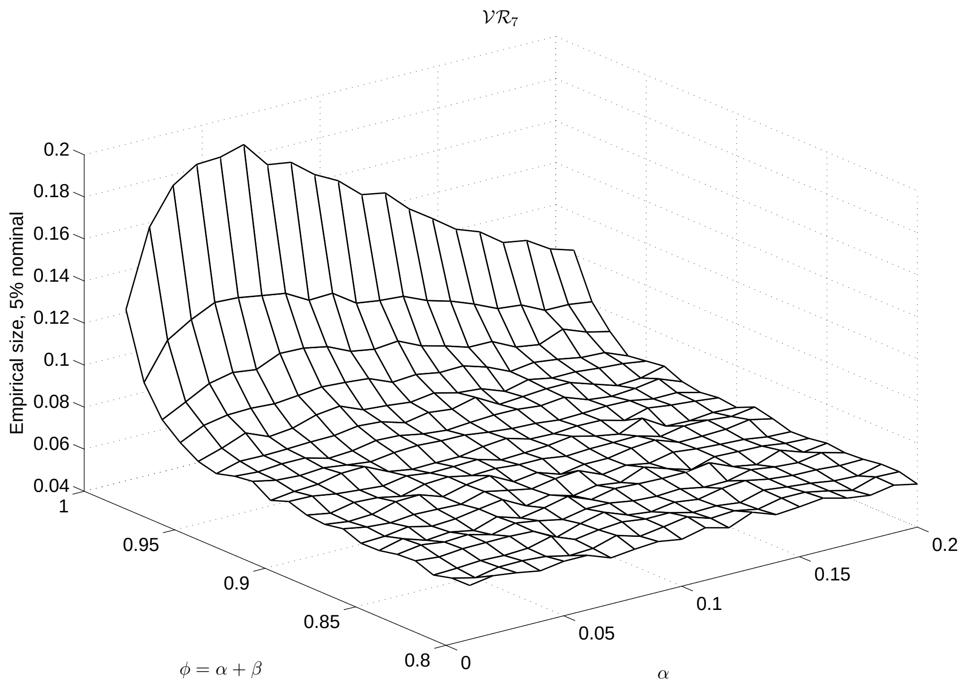

Figure 7: Size properties of $\mathcal{VR}_7$ when $T = 500$ and the length of MC is $25,000$. The DGP is outlined under (18) and (19) when $\delta_1 = \delta_2 = 0$ and $\gamma = 1 - \alpha - \beta$. See the notes to Table 2 for further details.

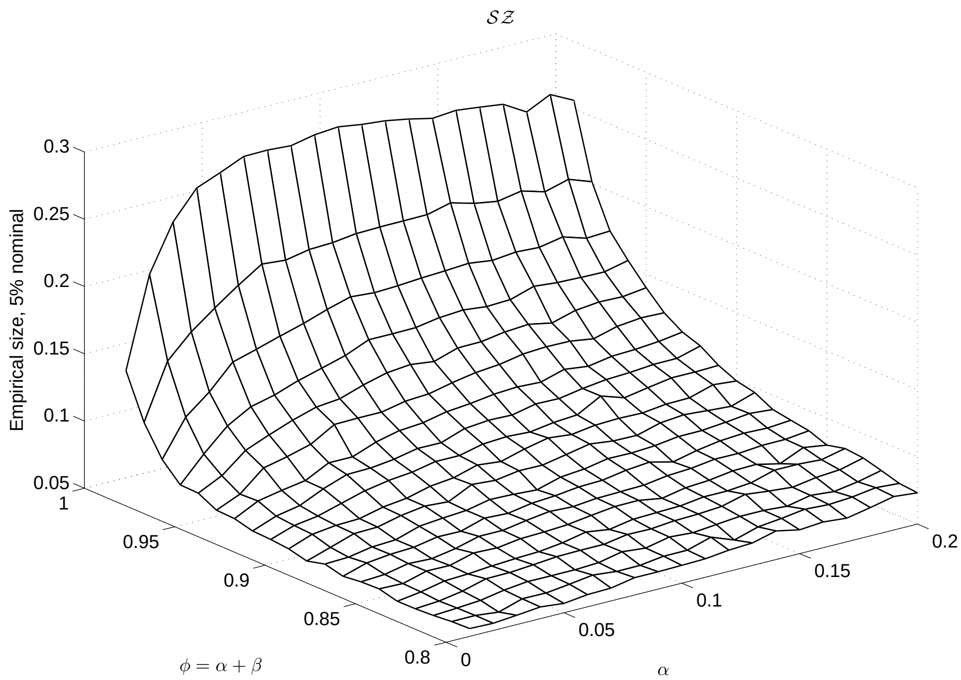

Figure 8: Size properties of $\mathcal{SZ}$ when $T = 500$ and the length of MC is $25,000$. The DGP is outlined under (18) and (19) when $\delta_1 = \delta_2 = 0$ and $\gamma = 1 - \alpha - \beta$. See the notes to Table 2 for further details.

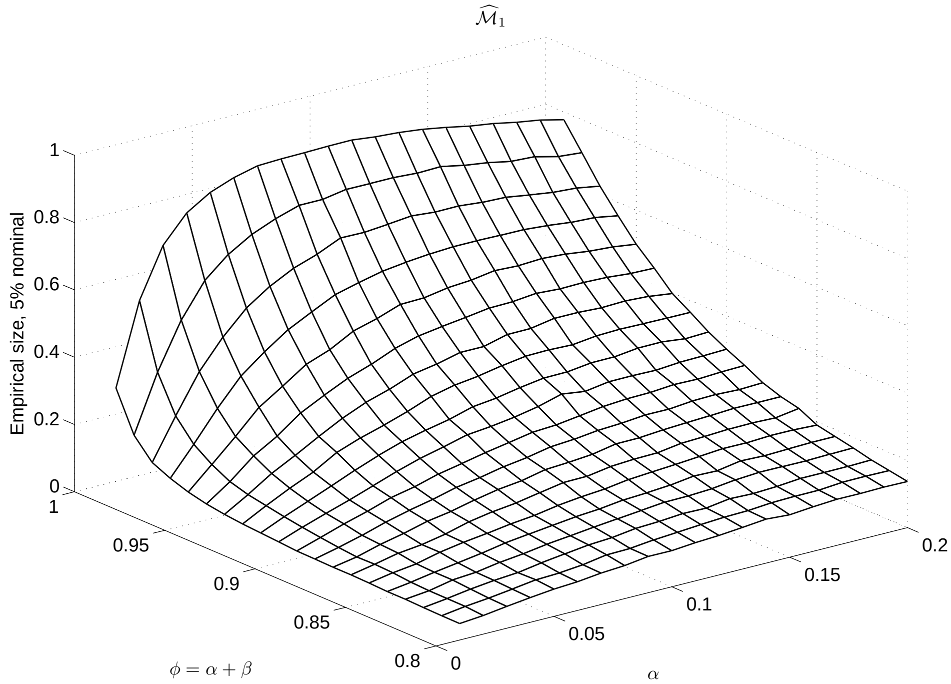

Figure 9: Size properties of $\widehat{\mathcal{M}}_1$ when $T = 1,000$ and the length of MC is $25,000$. The DGP is outlined under (18) and (19) when $\delta_1 = \delta_2 = 0$ and $\gamma = 1 - \alpha - \beta$. See the notes to Table 2 for further details.

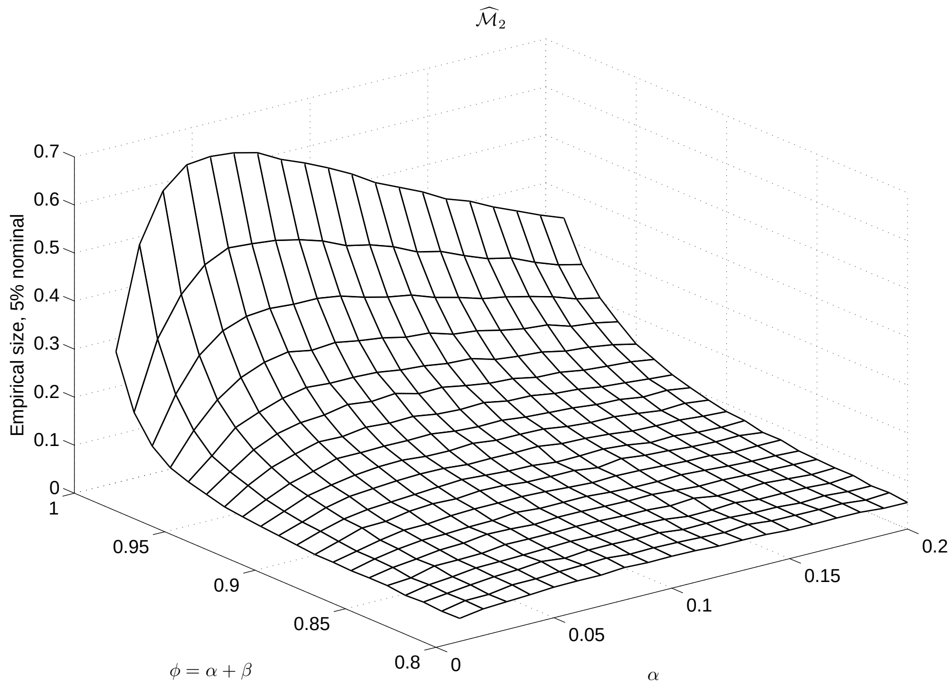

Figure 10: Size properties of $\widehat{\mathcal{M}}_2$ when $T = 1,000$ and the length of MC is $25,000$. The DGP is outlined under (18) and (19) when $\delta_1 = \delta_2 = 0$ and $\gamma = 1 - \alpha - \beta$. See the notes to Table 2 for further details.

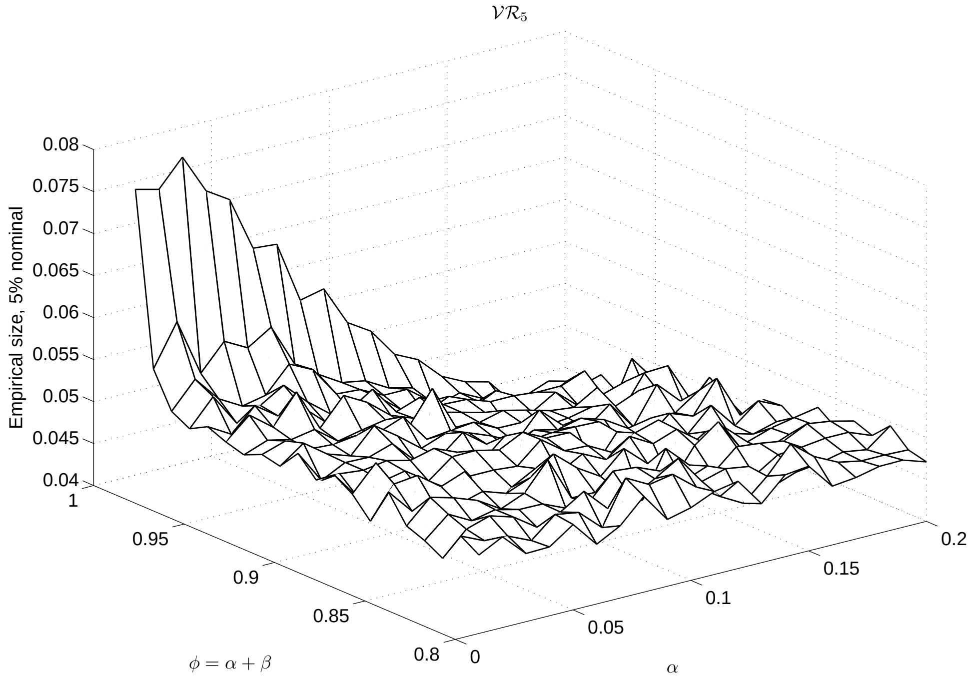

Figure 11: Size properties of $\mathcal{VR}_5$ when $T = 1,000$ and the length of MC is $25,000$. The DGP is outlined under (18) and (19) when $\delta_1 = \delta_2 = 0$ and $\gamma = 1 - \alpha - \beta$. See the notes to Table 2 for further details.

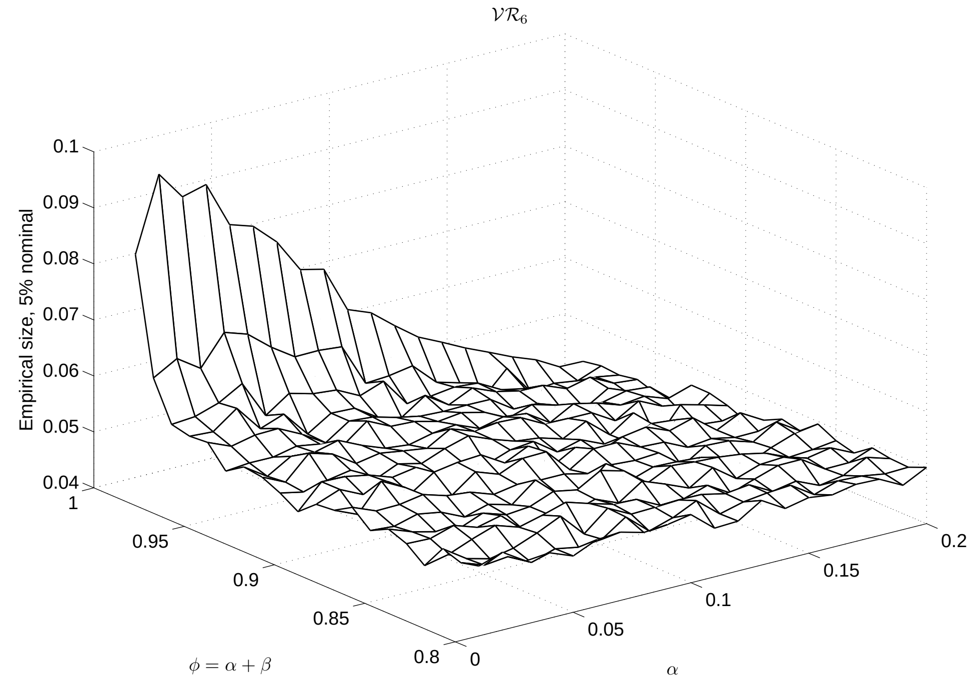

Figure 12: Size properties of $\mathcal{VR}_6$ when $T = 1,000$ and the length of MC is $25,000$. The DGP is outlined under (18) and (19) when $\delta_1 = \delta_2 = 0$ and $\gamma = 1 - \alpha - \beta$. See the notes to Table 2 for further details.

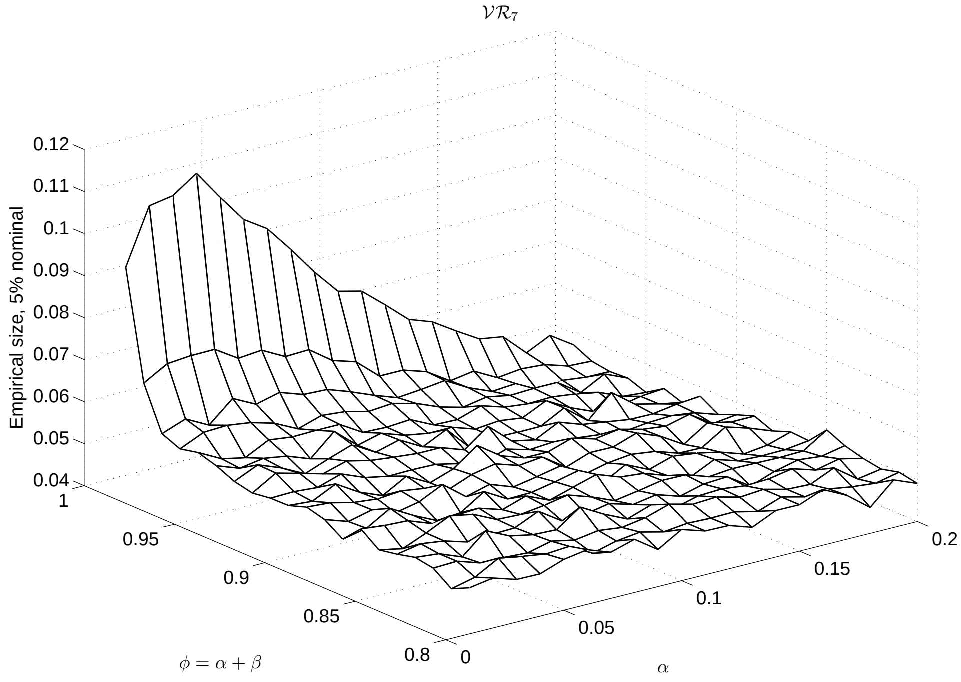

Figure 13: Size properties of $\mathcal{VR}_7$ when $T = 1,000$ and the length of MC is $25,000$. The DGP is outlined under (18) and (19) when $\delta_1 = \delta_2 = 0$ and $\gamma = 1 - \alpha - \beta$. See the notes to Table 2 for further details.

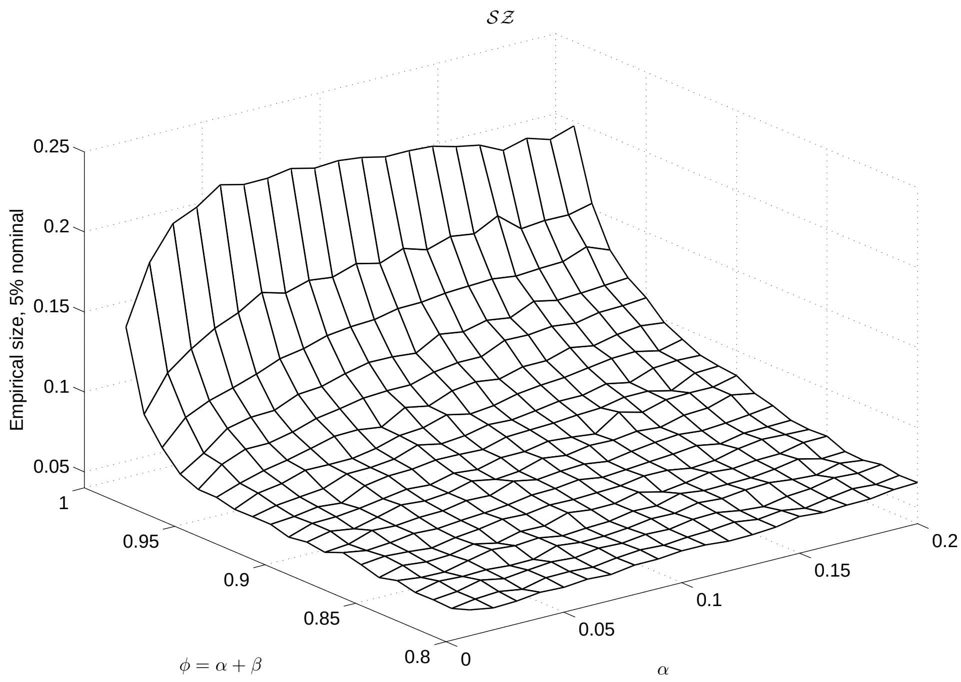

Figure 14: Size properties of $\mathcal{SZ}$ when $T = 1,000$ and the length of MC is $25,000$. The DGP is outlined under (18) and (19) when $\delta_1 = \delta_2 = 0$ and $\gamma = 1 - \alpha - \beta$. See the notes to Table 2 for further details.

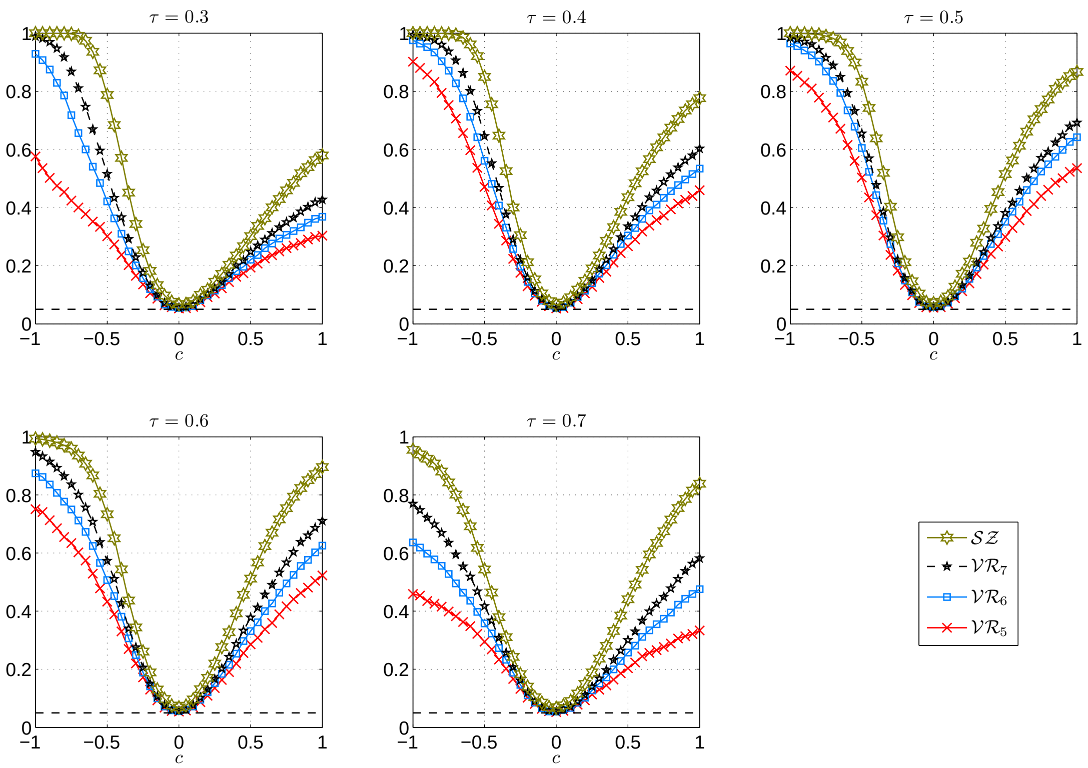

Figure 15: Empirical power for different test for a nominal 5% size, when $T = 500$. The DGP is as outlined under (18) and (20) when $\alpha = 0.01$, $\beta = 0.94$ ($\phi = 0.95$) and $\gamma = 0.05$ with $25,000$ replications for the MC. See the notes to Table 2 for further details.

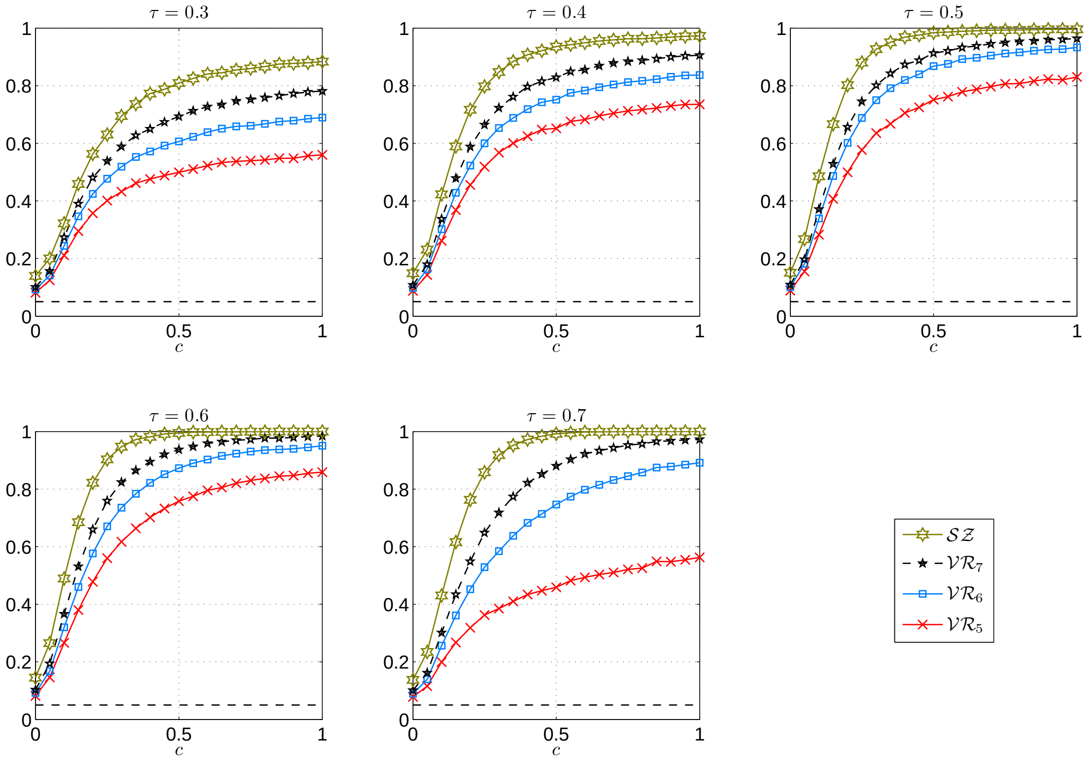

Figure 16: Empirical power for different test for a nominal 5% size, when $T = 500$. The DGP is as outlined under (18) and (20) when $\alpha = 0.01$, $\beta = 0.98$ ($\phi = 0.99$) and $\gamma = 0.01$ with 25,000 replications for the MC. See the notes to Table 2 for further details.

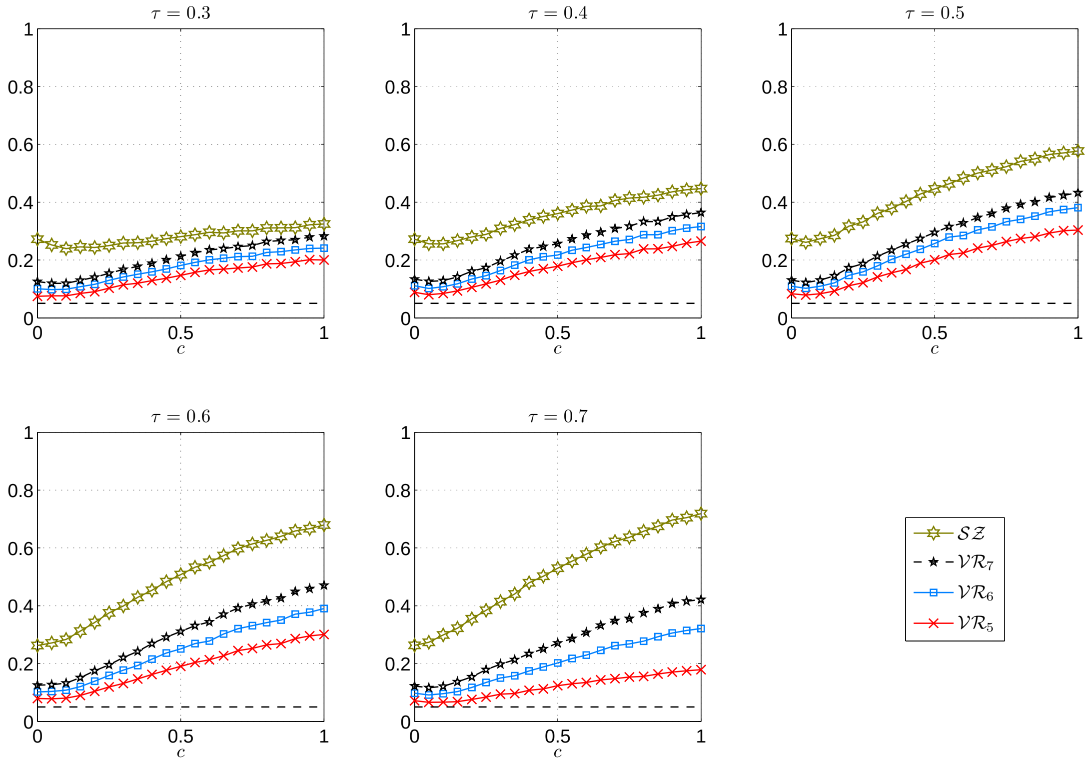

Figure 17: Empirical power for different test for a nominal 5% size, when $T = 500$. The DGP is as outlined under (18) and (20) when $\alpha = 0.1$, $\beta = 0.89$ ($\phi = 0.99$) and $\gamma = 0.01$ with 25,000 replications for the MC. See the notes to Table 2 for further details.

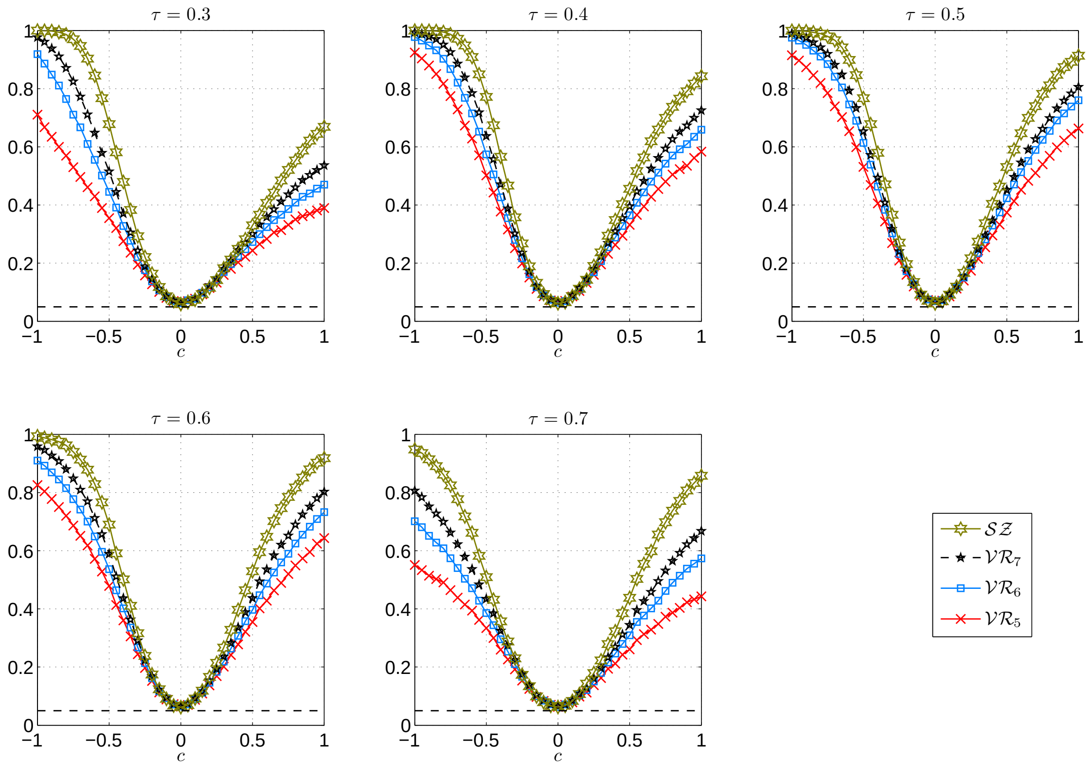

Figure 18: Empirical power for different test for a nominal 5% size, when $T = 1000$. The DGP is as outlined under (18) and (20) when $\alpha = 0.01$, $\beta = 0.94$ ($\phi = 0.95$) and $\gamma = 0.05$ with 25,000 replications for the MC. See the notes to Table 2 for further details.

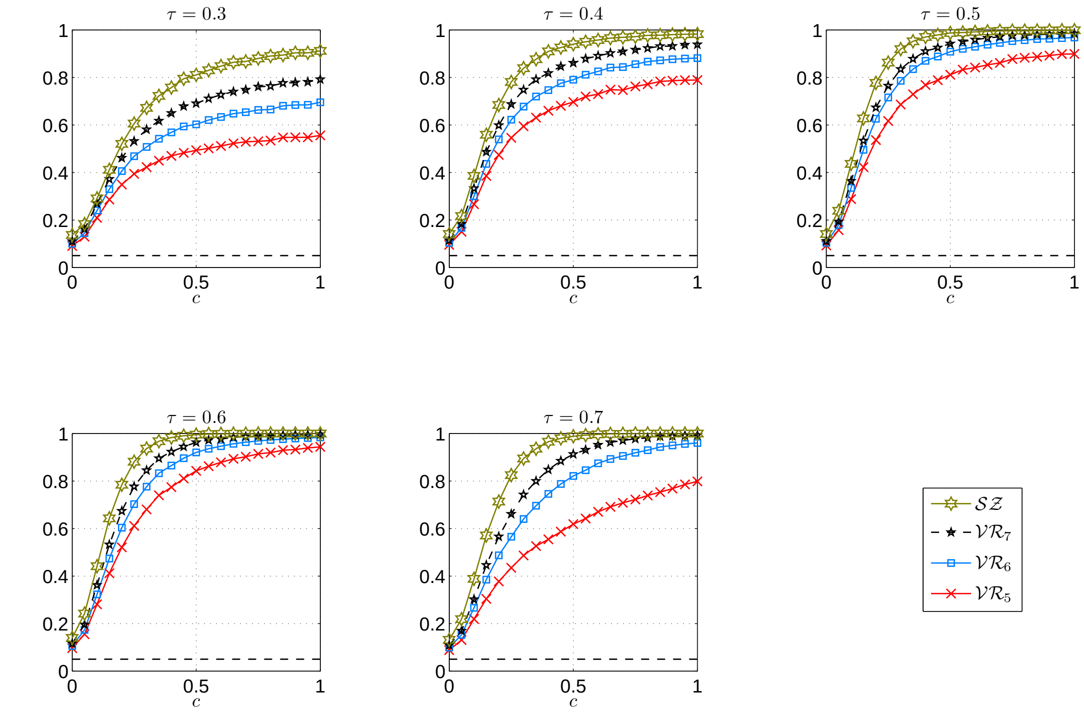

Figure 19: Empirical power for different test for a nominal 5% size, when $T = 1000$. The DGP is as outlined under (18) and (20) when $\alpha = 0.01$, $\beta = 0.98$ ($\phi = 0.99$) and $\gamma = 0.01$ with 25,000 replications for the MC. See the notes to Table 2 for further details.

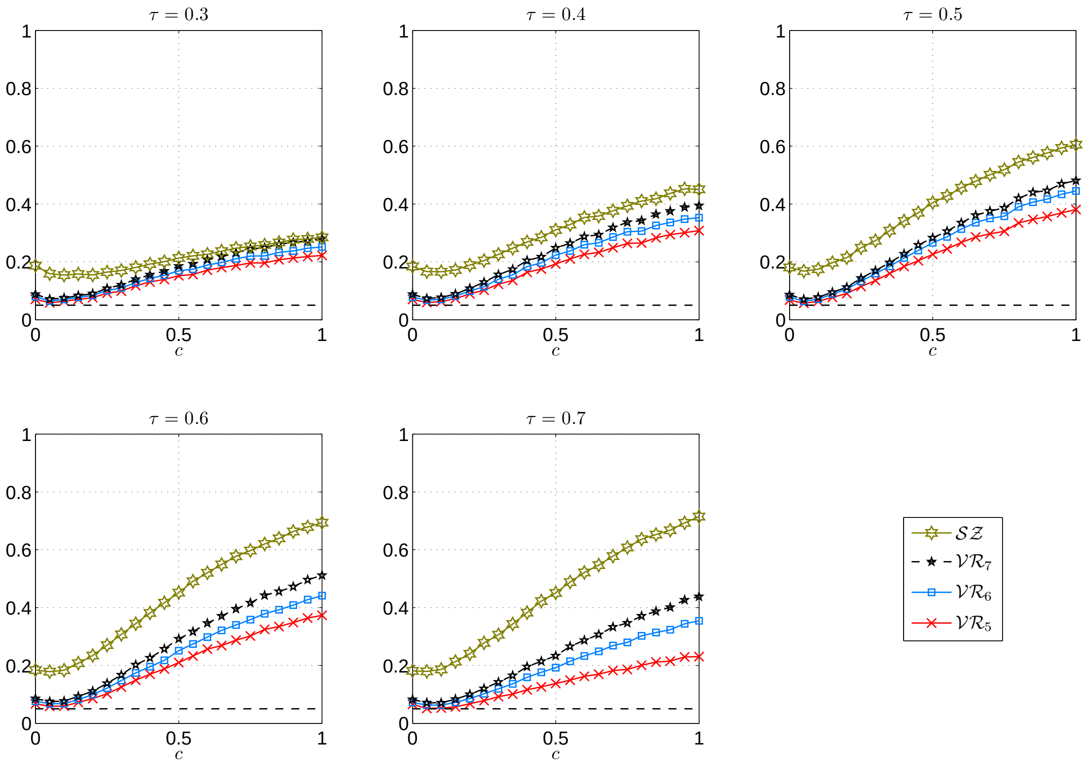

Figure 20: Empirical power for different test for a nominal 5% size, when $T = 1000$. The DGP is as outlined under (18) and (20) when $\alpha = 0.1$, $\beta = 0.89$ ($\phi = 0.99$) and $\gamma = 0.01$ with 25,000 replications for the MC. See the notes to Table 2 for further details.