Faculty of Law, Economics and Finance

# On residual-based CUSUM test in X-ARCH models

Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of Master in Quantitative Economics and Finance

*Author:*
Yevhen Perehuda

*Supervisor:*
Assistant Professor
Benjamin Holcblat

August 2023

# Abstract

# Contents

# 1 Introduction

Time series data, due to its temporal nature, often exhibits changing patterns and structural shifts. Detecting these changes is of paramount importance in various fields ranging from finance to environmental studies, as it can significantly influence forecasting, policy making, and strategic decisions. One of the longstanding challenges in time series analysis is accurately detecting structural breaks in the presence of various error structures, especially when the errors exhibit autoregressive conditional heteroscedasticity (ARCH) patterns.

The model in Equation 1 provides a general setting for a regression model with ARCH errors:

$$\begin{aligned} y_t &= \mu + \lambda' X_{T,t-1} + u_t, \\ u_t &= \sigma_t \epsilon_t, \\ \sigma_t^2 &= \omega + \alpha u_{t-1}^2. \end{aligned} \tag{1}$$

where and $\epsilon_t$ is a sequence of length $n$ of independent and identically distributed (IID) random variables with zero mean, $\sigma_t$ - time-varying conditional volatility and $X_{T,t-1}$ - covariates, which are time-dependent and can, also rely on sample $T$. subsequently, $y_T, t$ becomes T-dependent too.

Our research aims to provide a comprehensive study on the Cumulative Sum (CUSUM) test applied to regression models that have ARCH errors X-ARCH models. Firstly, we delve into the theoretical foundations of the residual-based CUSUM test for detecting mean shift proposed by (Grønneberg & Holcblat 2019*a*). Our exploration is rooted in the assumption that the ARCH process is stationary and ergodic, allowing us to derive robust results and implications. Secondly, our empirical analysis includes a simulation study that examines the power and size properties of the tests, providing insights into its efficacy under various scenarios. Furthermore, a real data analysis of the yen/dollar exchange rate returns adds an empirical flavour to our study, bridging the gap between theory and practice.

Pioneering works like that of (Page 1955) and (Brown et al. 1975) have set the stage for detecting change-points in time series. However, the intersection of CUSUM tests with ARCH models remains sparsely explored. While studies such as those by (Kim et al. 2000) and (Lee et al. 2003) have broached the topic, an explicit test for mean change in regression models with ARCH errors remains largely untouched. Our study bridges this gap, providing a comprehensive guide and test for detecting structural breaks in such models

The previous general results on CUSUM tet often require IID (or martingale difference) errors (i.e. no weak white noise or autocorrelation). However, work (Grønneberg & Holcblat 2019*a*) provides a low-tech and versatile toolbox consisting of theoretical foundations that allow the construction of new residual-based CUSUM tests. The underlying assumptions of such tests were designed to be very general in order to make them universal. It provides the flexibility to capture shifts in parameters influencing the first moment of the distribution. Such adaptability means our test can effectively pick up on changes that have pronounced effects on the central tendencies of many time series. Particularly, the test is not restricted to errors that are strictly IID. It comfortably accommodates a broader spectrum of error structures, such as those shaped by ARCH and generalized autoregressive conditional heteroskedasticity (GARCH) effects.

Yet, it's also essential to understand the limitations of our proposed approach. While the investigated CUSUM test has many applications, it might not be as efficient in detecting alterations in parameters affecting the second moment in ARCH models. However, its prowess in identifying changes impacting the first moment remains very strong.

The remainder of this paper is structured as follows: Section 2 offers a brief review of the existing literature on the topic. Section 3 presents the residual-based CUSUM test's theoretical underpinnings and slightly modifies and adapts general theory to the needs of our model. Section 4 showcases our simulation study, elucidating the test's empirical properties. In Section 5, we conduct a real data analysis, emphasizing our test's practical applications. Finally, Section 6 concludes the paper, summarizing our findings and pointing towards potential areas for future research.

# 2 Literature review

The seminal work of (Page 1955) introduced the problem of testing the change-point in time-series. Page's work has since become foundational, inspiring a plethora of research for the tests of structural breaks. Such a test can be formulated as checking the null hypothesis that parameters of the time-series models stay constant and alternative one in case of changing at some unknown point. Change-point estimation finds diverse applications. This can include monitoring of intensive-care patients (Fried & Imhoff 2004), climatic time-series (Gallagher et al. 2013), detecting financial bubbles (Phillips et al. 2011, 2015).

Due to (De Prado 2018) structural break tests can be broadly categorized into cumulative sum (CUSUM) tests (Brown et al. 1975), which assess if cumulative forecasting errors deviate from white noise, and explosiveness tests, which examine the process for signs of exponential growth or collapse, considering different functional forms like right-tail unit-root tests (Phillips et al. 2011) and sub/super-martingale tests (Phillips et al. 2015). Among them, the CUSUM test is a prevalent choice as a diagnostic tool due to its simplicity in understanding and straightforward implementation.

CUSUM tests can be divided into estimation-based, score-based, and residuals-based groups. The estimation-based tests lean on parameter estimators (Lee et al. 2003), while the score vector-based ones use likelihood function scores (Berkes et al. 2004, Oh & Lee 2018*b*). The residuals-based tests, which are used most widely, prioritize model residuals, gaining prominence for their adaptability and aptitude in identifying shifts in series' mean or variance. These can take into the count either recursive (Brown et al. 1975, Krämer et al. 1988, Otto & Breitung 2023) or standard (Ploberger & Krämer 1992, 1996, Grønneberg & Holcblat 2019*a*, Lee et al. 2004) residuals, when the latter ones, are, practically, standard residuals are preferred over recursive residuals because they are easier to compute, eliminating the need for repeated calculations. Regarding testing and estimating multiple structural breakpoints, see (Andreou & Ghysels 2002), (Bai & Perron 1998) and (Inclan & Tiao 1994).

In the analysis of financial time series, autoregressive conditional heteroscedastic (ARCH) models (Engle 1982) are frequently used to fit return series. (Bollerslev 1986) extended the model to the GARCH framework, allowing for a more generalized approach to capturing the dynamics of volatility. Over the years, the basic GARCH model has undergone numerous modifications to better capture the complexities in financial data. These include the introduction of the EGARCH (Exponential GARCH) model by (Nelson 1991) which allows for asymmetric responses of volatility to shocks, and the TGARCH (Threshold GARCH) which also considers the leverage effect. Additionally, the introduction of the IGARCH (Integrated GARCH) model recognizes the persistence in volatility by constraining the sum of ARCH and GARCH parameters to be one. Such continual advancements reflect the industry's effort to model and predict financial volatilities with increasing precision.

In the domain of intersection of the CUSUM tests with the realm of ARCH and GARCH models, (Inclan & Tiao 1994) were pioneers in proposing the CUSUM test, primarily devised to detect and pinpoint variance shifts within IID samples. Subsequently, (Kim et al. 2000) adapted the CUSUM test to the GARCH(1,1) models. Their rationale was based on the notion that since variance is contingent upon GARCH parameters, any change in these parameters should manifest as variance alterations. While their theoretical grounding was sound, in practice, the CUSUM test exhibited significant size distortions and reduced power. Addressing these limitations, (Lee et al. 2003) introduced an enhanced version of the CUSUM test grounded on residuals. These residuals were defined as the square of the observations normalized by their estimated conditional variances. It's noteworthy that while their study claimed applicability to regression with ARCH models, their simulation studies only revolved around GARCH(1, 1) models with constant means. Although, the simulation study for "pure" ARCH has been done by (Deng & Perron 2008), where they considered the CUSUM of squares test in a model with general mixing assumptions.

Subsequent studies by (Lee & Lee 2014) explored AR-ARCH models, and research by (Oh & Lee 2018*a*,*b*) delved into the intricacies of ARMA-GARCH. Yet, to the best of our understanding, no explicit tests have been formulated for a regression model with covariates and ARCH errors, termed the X-ARCH model. Importantly, it should be highlighted that the CUSUM tests discussed in this and the previous paragraph targeted structural breaks affecting the second moment - the variance - and not the first moment, which is the mean. The reason for that was that these tests operated over squares of residuals.

However, the scenario took a promising turn with (Grønneberg & Holcblat 2019*a*) work. They unveiled a CUSUM statistic adept at detecting mean shifts in regressions, even when confronted with significant error variations, for, autocorrelated ones. This innovation holds profound implications, especially for a vast array of models under the ARMAX-GARCH umbrella. It is relevant exceptionally in the context of the X-ARCH model, filling a gap in the literature.

# 3 Residual-based CUSUM test

Before embarking on our simulation journey, it's imperative to lay a solid theoretical groundwork. The essence of any simulation lies in the accurate derivation of statistics and the identification of the specific partial-sum process of residuals. This necessitates a deep dive into the theory that governs these processes.

One of the standout features of the approach is the adoption of relatively mild assumptions, which not only ensures robustness but also enhances the applicability. As it was already mentioned, our theoretical framework is heavily based on (Grønneberg & Holcblat 2019*a*) but with modifications tailored to the X-ARCH context for relevant and precise simulations. For those interested in the details of our derivations, the supplementary work of (Grønneberg & Holcblat 2019*b*) includes proofs and extended discussions. It provides a comprehensive guide with deeper insights into the assumptions, theorems, and lemmas that form the foundation of our study.

The X-ARCH residuals are defined as $(\hat{u}_t)_{t=-q+1}^{T}$. For $t \in \{1 \dots T\}$,

$$\hat{u}_t := (y_t - \hat{\mu}) - \hat{\lambda}' X_{T,t-1}, \tag{2}$$

and for $t \leqslant 0, \hat{u}_t = 0$.

Based on them we can define the average-corrected error and the average-corrected residuals:

$$v_{t,T} := u_t - \frac{1}{T}\sum_{j=1}^{T} u_j, \quad \text{and} \quad \hat{v}_{t,T} := \hat{u}_t - \frac{1}{T}\sum_{j=1}^{T} \hat{u}_j.$$

Our main focus is on the following partial-sum processes:

$$\hat{\mathfrak{U}}_T(s) := \frac{1}{T}\sum_{t=1}^{\lfloor Ts \rfloor} \hat{u}_t, \quad \mathfrak{U}_T(s) := \frac{1}{T}\sum_{t=1}^{\lfloor Ts \rfloor} u_t,$$

$$\hat{\mathfrak{V}}_T(s) := \frac{1}{T}\sum_{t=1}^{\lfloor Ts \rfloor} \hat{v}_{t,T}, \quad \text{and} \quad \mathfrak{V}_T(s) := \frac{1}{T}\sum_{t=1}^{\lfloor Ts \rfloor} v_{t,T},$$

where for all $a \in \mathbf{R}$, $\lfloor a \rfloor := \max\{n \in \mathbf{Z} : n \leqslant a\}$.

**Assumption 1** ($O_P\left(T^{-\frac{1}{2}}\right)$-consistency of X-ARCH parameters). *Let $\hat{\mu}$ and $\hat{\lambda}$ be the respective estimators of $\mu$ and $\lambda$ s.t. (a) $\sqrt{T}(\hat{\mu} - \mu) = O_P(1)$ and (b) $\sqrt{T}(\hat{\lambda} - \lambda) = O_P(1)$*

We don't make specific assumptions about the estimators of the X-ARCH parameters other than they are $O_P(T^{-\frac{1}{2}})$ away from their population values. Moreover, it is essential only for parameters that affect X part. This means that the results are generally valid regardless of the estimation method chosen, if the parameters are correctly identified.

**Assumption 2** (Error term $u_t$). *(a) For a constant $\epsilon_u > 0, \sup_{t\in\mathbf{Z}} \mathbb{E}|u_t|^{1+\epsilon_u} < \infty$. (b)* $\sup_{s\in[0,1]} \left|\frac{1}{T}\sum_{t=1}^{\lfloor Ts\rfloor} u_t\right| = o_P(1)$.

Assumption 2(a) is pivotal in confirming the presence of specific power series in B applied to the error term $u_t$ in terms of almost sure convergence. This is evident from Lemma 7(i) as mentioned in (Grønneberg & Holcblat 2019*b*). Furthermore, Assumption 2(a) paves the way for the application of the Phillips–Solo device methodology (Phillips & Solo 1992). It's worth noting that Assumption 2(a) is relatively mild, primarily because it necessitates the existence of the first moment. This stands in contrast to the work by (Lee et al. 2003, Deng & Perron 2008, Oh & Lee 2018*a*), which demands the existence of the fourth moment, making our assumption more accommodating and less restrictive.

One of the standout features of Assumption 2(b) is its allowance for both heteroscedasticity (conditional and unconditional) and various forms of time-dependence, including autocorrelation. Given the inherent heteroscedastic nature of numerous financial and economic time series and the challenges in negating autocorrelation in errors, this assumption's generality is one of the main breakthrough parts. Many models often produce errors that deviate from being just IID or martingale differences and this underscores the broad applicability and relevance of Assumption 2(b) in diverse scenarios. By Assumption 3(b), together with the Phillips–Solo device, ensures the asymptotic vanishing of the partial-sum average process of power series of the error.

**Lemma 1.** *Under Assumptions 1 and 2, w.p.a.1 as $T \to \infty$:*

(a) $\sup_{s\in[0,1]}\left|\sqrt{T}\left[\hat{\mathfrak{U}}_T(s) - \mathfrak{U}_T(s)\right] + s\sqrt{T}(\hat{\mu}-\mu) + \frac{1}{\sqrt{T}}\sum_{t=1}^{\lfloor Ts\rfloor}(\hat{\lambda}-\lambda)' X_{T,t-1}\right| = o_P(1)$

(b) $\sup_{s\in[0,1]}\left|\sqrt{T}\left[\hat{\mathfrak{B}}_T(s) - \mathfrak{B}_T(s)\right] - \frac{1}{\sqrt{T}}\sum_{t=1}^{\lfloor Ts\rfloor}\left[(\hat{\lambda}-\lambda)' X_{T,t-1} - \frac{1}{T}(\hat{\lambda}-\lambda)'\sum_{t=1}^{T} X_{T,t-1}\right]\right| = o_P(1)$,

Lemma 1 shows how partial-sum processes for standard and average-corrected residuals can be characterized without assumptions on the covariates. Based on this lemma we find the gap between error and residuals appearing asymptotically and after get rid of nuisance terms with help of additional assumptions.

**Assumption 3** (Covariates $X_t$). *(a) For a constant $\epsilon_X > 0$, for all $l \in \{1, d_\lambda\}$,* $\sup_{(T,t)\in\mathbf{N}\times\mathbf{Z}} \mathbb{E}\left|X_{t,l}\right|^{1+\epsilon_X} < \infty$. *(b)* $\sup_{s\in[0,1]} \mid \frac{1}{T}\times \sum_{t=1}^{\lfloor Ts\rfloor}\left(X_{T,t-1} - \mathbb{E}X_{T,t-1}\right) \mid = o_P(1)$. *(c) For all $i \in [\![1,p]\!]$,* $\sup_{s\in[0,1]} \mid \left[\frac{1}{T}\times \sum_{t=1}^{\lfloor Ts\rfloor} \mathbb{E}X_{T,t-1-i}\right] - s\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}X_{T,t-1-i} \mid = o(1)$.

Assumptions 3(a) and (b)-(c) collectively serve as the covariate counterparts to Assumption 2(a) and (b) respectively. Thus, Assumption 3(a) also facilitates the use of the Phillips–Solo device, especially when analyzing partial-sums of $X_{T,t-1}$. Assumption 3(d) imposes constraints on the term that is deducted in Assumption 3(c). As highlighted in (Grønneberg & Holcblat 2019*b*), the majority of processes discussed in time-series literature adhere to Assumptions 3(c) and (d).

**Theorem 1.** *Under Assumptions 1, 2 and 3(a)-(b):*

(a) $\sup_{s\in[0,1]} \mid \frac{1}{\sqrt{T}}\sum_{t=1}^{\lfloor Ts\rfloor}(\hat{\lambda}-\lambda)'X_{T,t-1} - \sqrt{T}(\hat{\lambda}-\lambda)'\frac{1}{T}\sum_{t=1}^{\lfloor Ts\rfloor}\mathbb{E}X_{T,t-1} \mid = o_P(1)$ and hence $\sup_{s\in[0,1]} \mid \sqrt{T}\left[\hat{\mathfrak{U}}_T(s) - \mathfrak{U}_T(s)\right] - \sqrt{T}(\hat{\lambda}-\lambda)'\frac{1}{T}\sum_{t=1}^{\lfloor Ts\rfloor}\mathbb{E}X_{T,t-1} + s\sqrt{T}(\hat{\mu}-\mu) \mid = o_P(1)$.

*Under the additional Assumption 3(c):*

(b) $\sup_{s\in[0,1]}\left|\frac{1}{T}\sum_{t=1}^{\lfloor Ts\rfloor}\mathbb{E}X_{T,t-1} - s\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}X_{T,t-1}\right| = o(1)$ and, for all $i \in [1,p]$, $\sup_{s\in[0,1]}\left|\frac{1}{T}\sum_{t=1}^{\lfloor Ts\rfloor}\mathbb{E}X_{T,t-1-i} - s\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}X_{T,t-1-i}\right| = o(1)$; and

(c) $\sqrt{T}\sup_{s\in[0,1]}\left|\hat{\mathfrak{V}}_T(s) - \mathfrak{V}_T(s)\right| = o_P(1)$.

The first part of Theorem 1 provides insights into the asymptotic difference between the scaled residual partial-sum process from its error counterpart. The subsequent part, Theorem 1(ii), together with satisfying Assumption 4(d) allow to mask the covariate nuisance term, while analysing the partial-sum process of *average-corrected residuals*. Again the authors used Phillips–Solo device for proof (Grønneberg & Holcblat 2019*b*). Meanwhile, Theorem 1(ii.b) presents the most straightforward approach to derive a pivotal statistic for residual-based CUSUM tests, especially when the covariates don't have a zero mean. However, in order to derive a pivotal statistic in conventional scenarios, it's imperative that $\sigma_u$ should be consistently estimated.

**Proposition 1.** *Suppose a system of X-ARCH model, fulfilling Assumptions 1 and 2. Denote the ith element of the covariates in the model with $X_{t-1,i}$, and the error term with $u_t$. If $\sup_{t\in\mathbf{Z}}\mathbb{E}|u_t|^2 < \infty$ and $\sup_{t\in\mathbf{Z}}\mathbb{E}X_{t-1,i}^2 < \infty$, then*

$$\hat{\sigma}_{u,T} = \sigma_{u,T} + o_P(1), \tag{3}$$

*where $\sigma_{u,T}$ is the empirical standard deviation of the error $u_t$ and $\sigma_{u,T}$ the empirical standard deviation of the residuals $\hat{u}_t$.*

Proposition 1 identifies conditions that imply the consistency of empirical residual-based standard deviation, and its convergence to error-based one. The only difference with the same Proposition in (Grønneberg & Holcblat 2019*a*) is that we consider the univariate case and use, consequently, variance and standard deviation instead of a covariance matrix. Proposition 1 combined with Theorem 1(ii) provides pivot CUSUM statistics when used in conjunction with a statistical functional, such as the supremum.

**Corollary 1.** *Assume that the following conditions hold:*

*(a)* $\hat{\sigma}_{u,T} = \sigma_{u,T} + o_P(1)$,.

*(b) We have process convergence* $\sqrt{T}\mathfrak{U}_T(s) \xrightarrow[T\to\infty]{\mathscr{L}} \sigma_u^{1/2}B(s)$*, where B is Brownian motion on the unit interval* $[0,1]$.

*Then, under the assumptions of Proposition 1 and Assumptions 3(a)-(c) we have process convergence:*

$$\hat{\sigma}_{u,T}^{-1/2}\sqrt{T}\hat{\mathfrak{B}}_T(s) \xrightarrow[T\to\infty]{\mathscr{L}} B^{\circ}(s), \tag{4}$$

*where* $\hat{\sigma_u}$ *is the empirical standard deviation and* $B^{\circ}$ *denotes a Brownian bridge process on the unit root interval* $[0,1]$.

After applying supremum on both sides, we get:

$$\sup_{s\in[0,1]} \left|\hat{\sigma}_u^{-1}\sqrt{T}\hat{\mathfrak{B}}_T(s)\right| \xrightarrow[T\to\infty]{\mathscr{L}} \sup_{s\in[0,1]} |B^{\circ}(s)| \tag{5}$$

Condition (b) aligns with a functional central limit theorem (Theorem 27.14 in (Davidson 1994)). Meanwhile, Condition (a) simply requires that standard empirical covariances approach the actual covariance. So, Corollary 1 directly infers the asymptotic distribution of new residual-based CUSUM-statistics. Its limiting distribution is well known and was formalized by (Billingsley 1968):

$$F(x) = 1 - 2\sum_{i=1}^{\infty}(-1)^{i+1}exp(2i^2x^2) \tag{6}$$

The critical values for $\alpha = 0.01, 0.05, 0.10$ are 1.63, 1.36 and 1.22 respectively. These, as well as other values, are provided in the Appendix, referring to (Smirnov 1948). Sometimes in the literature, it can be found as such called "Kolmogorov" distribution.

**Remark 1.** In our analysis, we observed a limitation of the discussed residual-based CUSUM test. Despite the mild assumptions, making the test a versatile toolbox for detecting changes, it falls short in detecting alterations in parameters that affect the second moment. While it's proficient at identifying changes impacting the first moment and can accommodate a variety of error structures, including ARCH-like errors, it is not capable when targeting parameters exclusive to volatility dynamics in ARCH models.

# 4 Simulation study

In empirical research, especially when developing or introducing a new statistical test, it is crucial to validate its properties under controlled scenarios. Given the ARCH model's aptness in capturing volatility dynamics, it is crucial to understand how the CUSUM test performs in the presence of ARCH effects, especially when the series is devoid of any systematic mean changes. Simulation studies offer a structured environment to understand the behavior and characteristics of a proposed test. In this section, we embark on an extensive simulation exercise to empirically validate the effectiveness of the CUSUM test in the context of regression models with ARCH errors.

The primary motivation behind these simulations is to investigate two critical aspects of any statistical test: its size and its power. The size of a test refers to the probability of incorrectly rejecting a true null hypothesis, which, in the context of structural breaks, would signify the test's propensity to detect a break when none exists. The power of a test relates to its capability to correctly reject a false null hypothesis. In this setting, it indicates the test's ability to accurately identify an existent break.

In our exploration, we implemented these tests at the significance levels of $\alpha = 0.01$ and $0.05$. The Monte-Carlo simulation included 1000 iterations for each setup of parameters independent of which model was tested. The empirical size and power were determined based on the frequency of null hypothesis rejection. We used sample sizes of $T = 500$, 1000, and 1500 for each test. The value of $\epsilon_t$ was taken from the standard normal distribution $N(0, 1)$.

Each simulation will involve generating synthetic time-series data based on the X-ARCH model. The parameters will be systematically varied to mimic different real-world scenarios and to understand the sensitivity and robustness of the CUSUM test.

For each simulation setup, the following steps will be followed:

1. Generate time-series data based on the X-ARCH model with specific parameters.

2. Apply the CUSUM test due to Theorem 1 to this data.

3. Record the outcome of the test and compare it to the known truth.

4. Repeat the above steps multiple times (e.g., 1000 iterations) to compute empirical size and power.

Here, we use Python 3.10 running on Windows 11 and the packages "sklearn" and "arch".

In our study, we will comprehensively assess the effectiveness of the CUSUM test within the X-ARCH framework by investigating three distinct scenarios: (1) Constant Mean, where the data assumes a consistent mean throughout the series, enabling us to evaluate the test's power and size in the simplest case; (2) Change-point type covariate model, introducing a known structural break in the series to gauge the test's efficiency in identifying additional structural break which is not known; and (3) Seasonal, where the data exhibits periodic patterns, which is typical for many time-series, challenging the test's robustness against potential seasonal fluctuations that could obscure genuine structural breaks. Each scenario offers unique insights into the CUSUM test's performance under varying conditions.

## 4.1 Constant mean model with ARCH

The constant mean model, as the name implies, assumes that the data series maintains a steady mean throughout its duration. This assumption simplifies the underlying structure and allows us to focus on the ARCH effects. So, the corresponding equation looks almost like Equation 1:

$$\begin{aligned} y_t &= \mu + u_t, \\ u_t &= \sigma_t \epsilon_t, \quad \sigma_t^2 = \omega + \alpha u_{t-1}^2. \end{aligned} \tag{7}$$

The following hypotheses were considered for the testing:

$H_0$ : The true parameter $\theta = (\mu, \omega, \alpha)$ does not change over time-series $y_1, \ldots, y_n$. vs.

$H_1$ : $\theta$ change to $\theta' = (\mu', \omega, \alpha)$ at $[n \cdot a]$, where $0 < 1 - a < 1$ is the fraction of time-series with the structural break.

To ensure a comprehensive evaluation, we additionally tested the model across different parameter sets: $\theta = [0, 0.2, 0.4), (0, 0.4, 0.4)$, and $(0, 0.4, 0.2)$. The corresponding results for each of these parameter configurations are documented in Tables B1, B2and B3 respectively.

Our simulations, as depicted across these tables, attest to the robustness of our CUSUM tests, revealing minimal size distortions. An intriguing observation, consistent across simulations, is the diminished sensitivity of the CUSUM tests as the change point nears the series' boundaries. This characteristic is inherent to CUSUM tests and warrants consideration when drawing conclusions from the results.

Further, as we introduced more pronounced shifts in the mean, $\Delta\mu = \mu' - \mu$, the power of the test enhanced, especially evident in larger sample sizes. This phenomenon underscores the test's ability to accurately detect pronounced mean shifts. However, for subtler mean shifts, especially when the structural break is proximate to the series' extremities, the test's power dwindles, leading to reduced detection efficacy.

| T | $\Delta\mu$ | Significance 1% | | | | | Significance 5% | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Break point fraction *a* | | | | | Break point fraction *a* | | | | |
| | | *0.1* | *0.3* | *0.5* | *0.7* | *0.9* | *0.1* | *0.3* | *0.5* | *0.7* | *0.9* |
| **500** | *size($\Delta\mu = 0$)*[1] | 0.8 | | | | | 3.84 | | | | |
| | 0.05 | 1.2 | 2 | 4.2 | 3.7 | 0.9 | 5.8 | 9.3 | 15 | 11.1 | 4 |
| | 0.1 | 1.6 | 16.8 | 27.2 | 15.8 | 1.6 | 7.2 | 37.1 | 48.6 | 34 | 7.6 |
| | 0.5 | 93.1 | 100 | 100 | 100 | 95.5 | 99.6 | 100 | 100 | 100 | 99.2 |
| | 1.0 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| **1000** | *size* | 0.8 | | | | | 4.26 | | | | |
| | 0.05 | 1 | 6.3 | 11.9 | 6.6 | 1.3 | 5.3 | 19.3 | 28.3 | 19.7 | 6.8 |
| | 0.1 | 3.5 | 45.3 | 60.7 | 41.6 | 3.1 | 11.3 | 68.7 | 79.7 | 65.8 | 13 |
| | 0.5 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| | 1.0 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| **1500** | *size* | 0.9 | | | | | 4.28 | | | | |
| | 0.05 | 1.4 | 11 | 20.9 | 10 | 2 | 6.2 | 25 | 41.6 | 30.5 | 7.1 |
| | 0.1 | 4.8 | 65.1 | 82.7 | 66.2 | 6.6 | 18 | 84.7 | 93.3 | 85.8 | 18.3 |
| | 0.5 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| | 1.0 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |

Table 1: Empirical sizes and powers of the CUSUM test in the X–ARCH(1) model (with constant mean and $(\mu, \omega, \alpha) = (0.0, 0.2, 0.2)$) with the shift in mean $\Delta\mu$

## 4.2 Change-point covariate model with ARCH effect

In order to check the performance of the new CUSUM test by (Grønneberg & Holcblat 2019*a*):

### 4.2.1 Model with IID error

, that is, a change point at the first pth fraction of the sample. T

$$y_t = \lambda I\{t \leqslant \mathfrak{p}T\} + u_t, \tag{8}$$

where $u_t$ is IID standard normal random variable.

$$\begin{aligned} \hat{u}_t &= Y_t - \hat{\lambda} I\{t \leqslant \mathfrak{p}T\} \\ &= (\lambda - \hat{\lambda}) I\{t \leqslant \mathfrak{p}T\} + u_t \end{aligned} \tag{9}$$

[1]Here and further we denote as *size* case when there is no parameter change in time-series (i.e. $\Delta\mu = 0$ or $\Delta\lambda = 0$).

$$\sqrt{T}\hat{\mathfrak{U}}_T(s) = \sqrt{T}(\lambda-\hat{\lambda})\frac{1}{T}\sum_{t=1}^{\lfloor Ts\rfloor} I\{t \leqslant \mathfrak{p}T\} + \sqrt{T}\mathfrak{U}_T(s)$$
$$= \min(s,\mathfrak{p})\sqrt{T}(\lambda-\hat{\lambda}) + \sqrt{T}\mathfrak{U}_T(s) + o_P(1) \tag{10}$$

$$\begin{aligned}\sqrt{T}\hat{\mathfrak{U}}_T(s) - [\min(s,\mathfrak{p})/\mathfrak{p}]\sqrt{T}\hat{\mathfrak{U}}_T(1) &= \min(s,\mathfrak{p})\sqrt{T}(\lambda-\hat{\lambda}) + \sqrt{T}\mathfrak{U}_T(s) \\ &- [\min(s,\mathfrak{p})/\min(1,\mathfrak{p})]\left[\min(1,\mathfrak{p})\sqrt{T}(\lambda-\hat{\lambda}) + \sqrt{T}\mathfrak{U}_T(1)\right] + o_P(1) \\ &= \sqrt{T}\mathfrak{U}_T(s) - [\min(s,\mathfrak{p})/\min(1,\mathfrak{p})]\sqrt{T}\mathfrak{U}_T(1) + o_P(1)\end{aligned} \tag{11}$$

$$\hat{\sigma}_u^{-1}\sqrt{T}\sup_{s\in[0,1]}\left|\hat{\mathfrak{U}}_T(s) - [\min(s,\mathfrak{p})/\mathfrak{p}]\hat{\mathfrak{U}}_T(1)\right| \xrightarrow{\mathscr{L}} \sup_{s\in[0,1]} \mid B(s) - [\min(s,\mathfrak{p})/\min(1,\mathfrak{p})]B(1) \mid, \tag{12}$$

where $B$ is a Brownian motion. This CUSUM test, firstly described by (Grønneberg & Holcblat 2019*a*) has critical values founded via simulation listed in Table C4.

| **T** | $\Delta\mu$ | **Significance 1%** | | | | | **Significance 5%** | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | **Value of $\mathfrak{p}$** | | | | | **Value of $\mathfrak{p}$** | | | | |
| | | *0.1* | *0.3* | *0.5* | *0.7* | *0.9* | *0.1* | *0.3* | *0.5* | *0.7* | *0.9* |
| **500** | *size* | 0.84 | | | | | 4.8 | | | | |
| | 0.05 | 1.3 | 2.2 | 3.8 | 2.3 | 1 | 7.5 | 10 | 11.8 | 7.3 | 5.4 |
| | 0.1 | 7.6 | 10.2 | 13.2 | 4.7 | 2.2 | 19.9 | 23.8 | 31.7 | 13.3 | 9 |
| | 0.5 | 100 | 100 | 100 | 99.8 | 78.2 | 100 | 100 | 100 | 100 | 94.4 |
| | 1.0 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| **1000** | *size* | 0.86 | | | | | 4.86 | | | | |
| | 0.05 | 3 | 4.8 | 5.9 | 2.8 | 1.5 | 9.5 | 13 | 17.6 | 9 | 7 |
| | 0.1 | 15.7 | 19.8 | 33.8 | 9.5 | 3.5 | 32.5 | 39.6 | 57.2 | 25.1 | 13.1 |
| | 0.5 | 100 | 100 | 100 | 100 | 99.8 | 100 | 100 | 100 | 100 | 100 |
| | 1.0 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| **1500** | *size* | 1.28 | | | | | 5.32 | | | | |
| | 0.05 | 5.1 | 7.6 | 9.8 | 3.5 | 1.5 | 14.2 | 19.2 | 23.9 | 12.6 | 8.1 |
| | 0.1 | 23.6 | 38.8 | 51.9 | 16 | 4.6 | 43.2 | 60.5 | 73.1 | 35.6 | 17.7 |
| | 0.5 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| | 1.0 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |

Table 2: Empirical sizes and powers of the CUSUM test in the regression model with IID shock and with the shift in mean $\Delta\mu$

### 4.2.2 Model with ARCH effect

$$y_t = \lambda I\{t \leqslant \mathfrak{p}T\} + u_t,$$
$$u_t = \sigma_t \epsilon_t, \quad \sigma_t^2 = 0.2 + 0.2u_{t-1}^2. \tag{13}$$

The following hypothesis were considered for the testing:

$H_0$ : The true parameter $\theta = (\lambda, \mu, \omega, \alpha)$ does not change over time-series $y_1, \ldots, y_n$.

vs.

$H_1$ : $\theta$ change to $\theta' = (\lambda, \mu', \omega, \alpha)$ at $[T/2]$.

| **T** | $\Delta\mu$ | **Significance 1%** | | | | | **Significance 5%** | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | **Value of $\mathfrak{p}$** | | | | | **Value of $\mathfrak{p}$** | | | | |
| | | *0.1* | *0.3* | *0.5* | *0.7* | *0.9* | *0.1* | *0.3* | *0.5* | *0.7* | *0.9* |
| **500** | *size* | 1.08 | | | | | 4.82 | | | | |
| | 0.05 | 6.5 | 9.7 | 14.3 | 4.1 | 2.3 | 17.8 | 23.3 | 34.1 | 12.2 | 10.5 |
| | 0.1 | 38.5 | 49.9 | 66.8 | 28.1 | 8.9 | 59 | 70.7 | 84.5 | 51 | 22.3 |
| | 0.5 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| | 1.0 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| **1000** | *size* | 0.94 | | | | | 4.9 | | | | |
| | 0.05 | 12.9 | 22.2 | 37.1 | 11.1 | 3.7 | 30.4 | 41.4 | 58.7 | 24.8 | 14.9 |
| | 0.1 | 74 | 85 | 96 | 63.6 | 20.4 | 87.9 | 95.2 | 99.5 | 82.9 | 43.7 |
| | 0.5 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| | 1.0 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| **1500** | *size* | 0.78 | | | | | 4.82 | | | | |
| | 0.05 | 25.7 | 36.9 | 57.7 | 18.4 | 5.4 | 45 | 59.6 | 77.2 | 39.4 | 16.7 |
| | 0.1 | 90.5 | 96.8 | 99.8 | 87 | 38.9 | 96.8 | 99.4 | 100 | 95.4 | 66 |
| | 0.5 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| | 1.0 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |

Table 3: Empirical sizes and powers of the CUSUM test in the X–ARCH(1) model (indicator regressor and $(\lambda, \mu, \omega, \alpha)$ = (2, 0, 0.2, 0.2)) with the shift in mean $\Delta\mu$

The second conducted simulation considered the testing:

$H_0$ : The true parameter $\theta = (\lambda, \mu, \omega, \alpha)$ does not change over time-series $y_1, \ldots, y_n$.

vs.

$H_1$ : $\theta$ change to $\theta' = (\lambda', \mu, \omega, \alpha)$ at $[\mathfrak{p}/2]$.

| T | $\Delta\lambda$ | **Significance 1%** | | | | | **Significance 5%** | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | **Value of $\mathfrak{p}$** | | | | | **Value of $\mathfrak{p}$** | | | | |
| | | *0.1* | *0.3* | *0.5* | *0.7* | *0.9* | *0.1* | *0.3* | *0.5* | *0.7* | *0.9* |
| **500** | *size* | 1.18 | | | | | 4.84 | | | | |
| | 0.05 | 1.5 | 0.8 | 1.2 | 2 | 3.6 | 4.4 | 4.9 | 5.8 | 8.2 | 13.3 |
| | 0.1 | 0.5 | 0.6 | 1.1 | 4.5 | 20.9 | 4.6 | 5.8 | 6.1 | 16.6 | 43.4 |
| | 0.5 | 1.4 | 7.6 | 93 | 100 | 100 | 4.6 | 34 | 99.3 | 100 | 100 |
| | 1.0 | 0.8 | 94.1 | 100 | 100 | 100 | 4.4 | 99.5 | 100 | 100 | 100 |
| **1000** | *size* | 1.04 | | | | | 4.58 | | | | |
| | 0.05 | 0.7 | 1.3 | 1.3 | 3.2 | 9.9 | 5 | 5.7 | 6.1 | 11.8 | 23.6 |
| | 0.1 | 1 | 1.4 | 2.1 | 15.6 | 44.9 | 4.6 | 4.9 | 10.9 | 36 | 67.1 |
| | 0.5 | 0.4 | 49 | 100 | 100 | 100 | 4.3 | 84.2 | 100 | 100 | 100 |
| | 1.0 | 1.7 | 100 | 100 | 100 | 100 | 13.4 | 100 | 100 | 100 | 100 |
| **1500** | *size* | 0.94 | | | | | 4.98 | | | | |
| | 0.05 | 1.7 | 1.1 | 1.3 | 4.4 | 14 | 6.1 | 4.9 | 6.5 | 16.7 | 31.7 |
| | 0.1 | 1 | 1 | 3.6 | 29.5 | 72.1 | 5.1 | 4.4 | 17.7 | 55.8 | 88.1 |
| | 0.5 | 1.1 | 85.4 | 100 | 100 | 100 | 5 | 98.4 | 100 | 100 | 100 |
| | 1.0 | 5.5 | 100 | 100 | 100 | 100 | 28.9 | 100 | 100 | 100 | 100 |

Table 4: Empirical sizes and powers of the CUSUM test in the X–ARCH(1) model (indicator regressor and $(\lambda, \mu, \omega, \alpha)$ = (2, 0, 0.2, 0.2)) with the shift in parameter $\Delta\lambda$

## 4.3 Seasonal dummies model with ARCH

$$\begin{aligned} y_t &= \mu + \sum_{k=1}^{d} \lambda_k I\{t \equiv k \pmod d\} + u_t, \\ u_t &= \sigma_t \epsilon_t, \quad \sigma_t^2 = \omega + \alpha u_{t-1}^2. \end{aligned} \tag{14}$$

### 4.3.1 Mean change

The following hypotheses were considered for the testing:

$H_0$ : The true parameter $\theta = (\mu, \boldsymbol{\lambda}, \omega, \alpha)$ does not change over time-series $y_1, \ldots, y_n$.

vs.

$H_1$ : $\theta$ change to $\theta' = (\mu', \boldsymbol{\lambda}, \omega, \alpha)$ at $[n \cdot a]$, where $0 < 1 - a < 1$ is the fraction of time-series with the structural break.

| T | $\Delta\mu$ | **Significance 1%** | | | | | **Significance 5%** | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | **Break point fraction *a*** | | | | | **Break point fraction *a*** | | | | |
| | | *0.1* | *0.3* | *0.5* | *0.7* | *0.9* | *0.1* | *0.3* | *0.5* | *0.7* | *0.9* |
| **500** | *size* | 0.9 | | | | | 4.3 | | | | |
| | 0.05 | 1.3 | 2.9 | 4.7 | 3.3 | 0.7 | 6.1 | 9.9 | 15.2 | 12.1 | 4.8 |
| | 0.1 | 2 | 18.3 | 27.6 | 17.8 | 1.1 | 7.7 | 37.5 | 48.7 | 38.4 | 6.5 |
| | 0.5 | 95.1 | 100 | 100 | 100 | 95.1 | 99.6 | 100 | 100 | 100 | 99.3 |
| | 1.0 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| **1000** | *size* | 0.88 | | | | | 4.66 | | | | |
| | 0.05 | 1.2 | 7.3 | 11.5 | 6.5 | 1.9 | 6.2 | 19.3 | 28.7 | 20.1 | 7.6 |
| | 0.1 | 2.9 | 43.4 | 59.9 | 42.8 | 4.6 | 12 | 68.9 | 80.9 | 67.6 | 13.6 |
| | 0.5 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| | 1.0 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| **1500** | *size* | 1.04 | | | | | 4.92 | | | | |
| | 0.05 | 2.2 | 10.2 | 18.7 | 13.1 | 2.2 | 7 | 27.7 | 38.5 | 28.6 | 8.7 |
| | 0.1 | 5.4 | 68.4 | 84.8 | 65.6 | 5.9 | 20.1 | 85.4 | 94.4 | 86 | 19.7 |
| | 0.5 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| | 1.0 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |

Table 5: Empirical sizes and powers of the CUSUM test in the X–ARCH(1) model (with seasonal dummies, $(\mu, \lambda, \omega, \alpha)$ = (0.0, 0.2, 0.2) and $\lambda$ = (1, -1, 2)) with the shift in mean $\Delta\mu$

### 4.3.2 Parameter change

The following hypotheses were considered for the testing:

$H_0$ : The true parameter $\theta = (\mu, \boldsymbol{\lambda}, \omega, \alpha)$ does not change over time-series $y_1, \ldots, y_n$.

vs.

$H_1$ : $\theta$ change to $\theta' = (\mu, \boldsymbol{\lambda}', \omega, \alpha)$ at $[n \cdot a]$, where $0 < 1 - a < 1$ is the fraction of time-series with the structural break.

| **T** | $\Delta\lambda_1$ | **Significance 1%** | | | | | **Significance 5%** | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | **Break point fraction *a*** | | | | | **Break point fraction *a*** | | | | |
| | | *0.1* | *0.3* | *0.5* | *0.7* | *0.9* | *0.1* | *0.3* | *0.5* | *0.7* | *0.9* |
| **500** | *size* | 0.48 | | | | | 3.9 | | | | |
| | 0.05 | 0.8 | 0.4 | 1.2 | 1.1 | 0.6 | 4.3 | 4.7 | 5.4 | 5.2 | 5.4 |
| | 0.1 | 0.6 | 1.6 | 1.6 | 1.2 | 0.7 | 4 | 4.7 | 7.5 | 5.5 | 3.3 |
| | 0.5 | 2.4 | 30 | 41.7 | 28.2 | 1.8 | 10.3 | 52.3 | 66.9 | 53.1 | 10.6 |
| | 1.0 | 13.2 | 96.7 | 98.6 | 93.9 | 13 | 40.8 | 99.6 | 99.8 | 98.8 | 33.4 |
| **1000** | *size* | 0.64 | | | | | 3.84 | | | | |
| | 0.05 | 0.5 | 0.9 | 1.5 | 1.3 | 0.6 | 3.4 | 5 | 6.8 | 4.6 | 4 |
| | 0.1 | 1.4 | 2.4 | 4 | 1.6 | 1.8 | 4.7 | 8.4 | 12.7 | 7.1 | 5.7 |
| | 0.5 | 4.9 | 67.4 | 83.5 | 66.6 | 5.5 | 18.6 | 86.4 | 94.4 | 85.2 | 18.8 |
| | 1.0 | 43.4 | 100 | 100 | 99.9 | 40.9 | 76.9 | 100 | 100 | 100 | 75.5 |
| **1500** | *size* | 0.74 | | | | | 3.66 | | | | |
| | 0.05 | 0.7 | 1.5 | 1.3 | 2.1 | 0.7 | 4.9 | 5.1 | 6.6 | 7.1 | 3.6 |
| | 0.1 | 1.2 | 2.1 | 3.6 | 2.7 | 0.9 | 4.9 | 9.8 | 11.5 | 9.8 | 4.4 |
| | 0.5 | 9.2 | 88.2 | 96.4 | 86.7 | 8.4 | 29 | 96.8 | 99.2 | 96.1 | 28.1 |
| | 1.0 | 79 | 100 | 100 | 100 | 76.5 | 96.4 | 100 | 100 | 100 | 95.8 |

Table 6: Empirical sizes and powers of the CUSUM test in the X–ARCH(1) model (with seasonal dummies, $(\mu, \lambda, \omega, \alpha)$ = (0.0, 0.2, 0.2) and $\lambda$ = (1, -1, 2)) with the shift in parameter $\Delta\lambda_1$

## 5 Real data analysis

For our real data analysis, we examined the yen/dollar exchange rate returns spanning from January 5, 1998, to January 27, 2003 (Figure 1). This data mirrored that of (Lee et al. 2003), employing the model depicted in Equation 1 with a constant mean. But in our case, the founded change points can be considered as more interpretable.


Figure 1: Plot of the exchange rate YEN/USD from Jan 5, 1998 to Jan 27, 2003.

Upon applying the Ljung–Box and LM–ARCH tests, we discerned that the model incorporating ARCH(1) effectively captures the dynamics of this series. For both tests, the p-values are extremely close to zero, much smaller than common significance levels (e.g., 0.05). This means we reject the null hypothesis of no autocorrelation in residuals, supporting the hypothesis of ARCH existence.

The residual-based parameter change test yielded a value of 7.51, as showcased in Figure 2, at two specific points: 7 October 1998 and 8 December 1998. This value eclipses the theoretical critical benchmark of 1.63 at the 0.01 level, indicating a significant change point in the mean. Intriguingly, both dates align at the end of 1998. This period's pertinence might be attributed to the decline of the Asian financial crisis.

Moreover, the Japanese yen experienced its most significant two-day surge on 6-8 October since it started its free float in February 1973, following the fallout from the Bretton Woods agreement (Cooper & Talbot 1999). The move in the exchange rate yen was largely attributed to the unwinding of the so-called "yen carry trade" and intensified market reactions during the Asian Financial Crisis. Given Japan's low-interest rates, global investors heavily engaged in the yen carry trade, borrowing in yen to invest in higher-yielding currencies. This pressure intensified in October 1998, when Wall Street institutions rescued Long-Term Capital Management, leading to the liquidation of its risk bets in its global portfolio.

For latter date (Dec 8, 1998) is also particularly interesting due to its timing around several unprecedented key events in Japan during this time which had stabilising goal (Ghosh 2000). Consequently, those might have influenced the shift:

- October 1998: Financial reconstruction bills were passed.
- November 1998: The Long-Term Credit Bank of Japan was nationalized.
- November also saw a massive 23.9 trillion-yen economic package announced by the government.
- December marked the nationalization of the Nippon Credit Bank.

Drilling down further, the pre-change segment, spanning from late January to October 1998, adheres to the model:

$$\begin{aligned} &Y_t = 0.0074 + u_t, \\ &u_t = \sigma_t \epsilon_t, \quad \sigma_t^2 = 0.000000058 + 0.2u_{t-1}^2. \end{aligned} \tag{15}$$

Contrastingly, the post-change phase is best encapsulated by the model:

$$\begin{aligned} &Y_t = 0.0086 + u_t, \\ &u_t = \sigma_t \epsilon_t, \quad \sigma_t^2 = 0.00000018 + 0.2u_{t-1}^2. \end{aligned} \tag{16}$$

In case if we consider the first period is until December 1998, the model looks like this:

$$\begin{aligned} &Y_t = 0.0076 + u_t, \\ &u_t = \sigma_t \epsilon_t, \quad \sigma_t^2 = 0.00000016 + 0.2u_{t-1}^2. \end{aligned} \tag{17}$$

where after the break point it changes to:

$$\begin{aligned} &Y_t = +u_t, \\ &u_t = \sigma_t \epsilon_t, \quad \sigma_t^2 = 0.00000018 + 0.2u_{t-1}^2. \end{aligned} \tag{18}$$

Visual check on Figs. 1 also supports the existence of a change in mean.

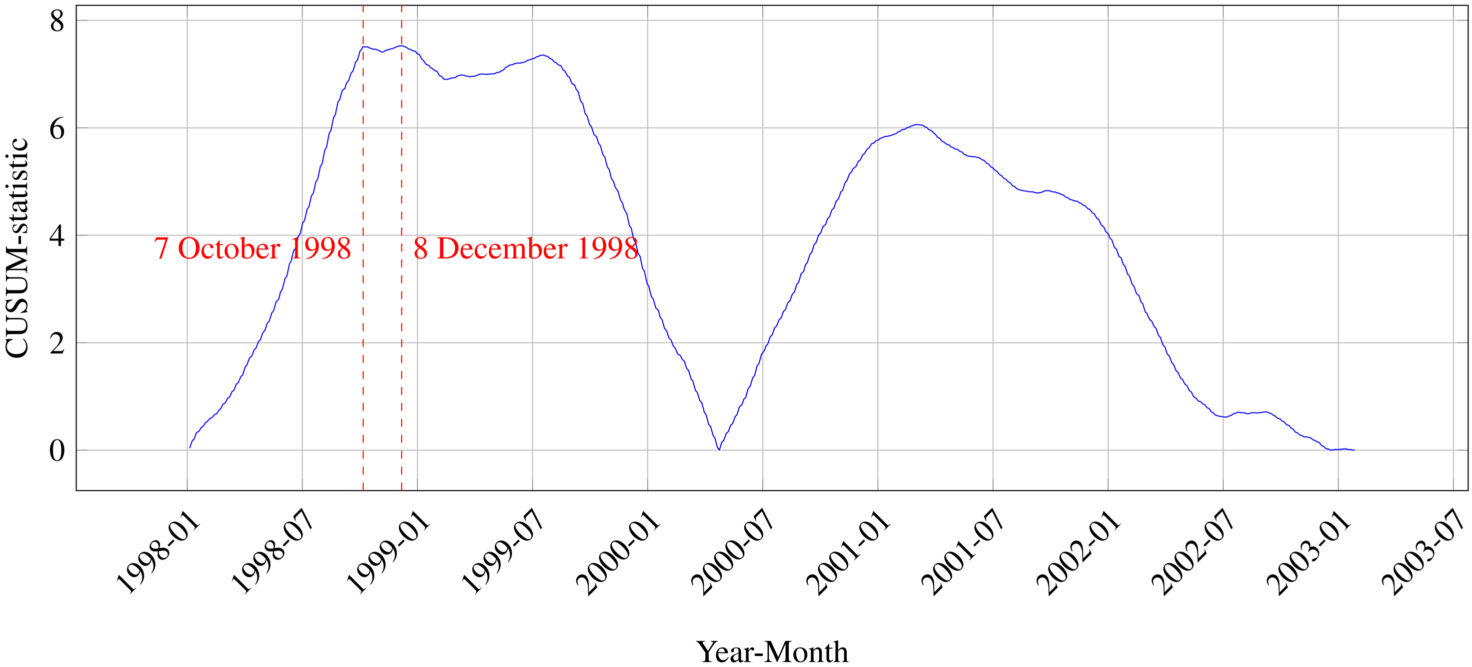


Figure 2: Plot of the CUSUM-statistics from Jan 5, 1998 to Jan 27, 2003.

# 6 Conclusion

has no severe size distortions in most cases

# References


Andreou, E. & Ghysels, E. (2002), 'Detecting multiple breaks in financial market volatility dynamics', *Journal of Applied Econometrics* **17**(5), 579–600.

Bai, J. & Perron, P. (1998), 'Estimating and testing linear models with multiple structural changes', *Econometrica* pp. 47–78.

Berkes, I., Horváth, L. & Kokoszka, P. (2004), 'Testing for parameter constancy in garch (p, q) models', *Statistics & probability letters* **70**(4), 263–273.

Billingsley, P. (1968), *Convergence of probability measures*, John Wiley & Sons.

Bollerslev, T. (1986), 'Generalized autoregressive conditional heteroskedasticity', *Journal of econometrics* **31**(3), 307–327.

Brown, R. L., Durbin, J. & Evans, J. M. (1975), 'Techniques for testing the constancy of regression relationships over time', *Journal of the Royal Statistical Society Series B: Statistical Methodology* **37**(2), 149–163.

Cooper, N. & Talbot, J. (1999), 'The yen/dollar exchange rate in 1998: Views from options markets', *Bank of England Quarterly Bulletin* **37**, 68–77.

Davidson, J. (1994), *Stochastic limit theory: An introduction for econometricians*, OUP Oxford.

De Prado, M. L. (2018), *Advances in financial machine learning*, John Wiley & Sons.

Deng, A. & Perron, P. (2008), 'The limit distribution of the cusum of squares test under general mixing conditions', *Econometric Theory* **24**(3), 809–822.

Engle, R. F. (1982), 'Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation', *Econometrica: Journal of the econometric society* pp. 987–1007.

Fried, R. & Imhoff, M. (2004), 'On the online detection of monotonic trends in time series', *Biometrical Journal: Journal of Mathematical Methods in Biosciences* **46**(1), 90–102.

Gallagher, C., Lund, R. & Robbins, M. (2013), 'Changepoint detection in climate time series with long-term trends', *Journal of Climate* **26**(14), 4994–5006.

Ghosh, B. N. (2000), *Global financial crises and reforms: cases and caveats*, Vol. 27, Routledge.

Grønneberg, S. & Holcblat, B. (2019*a*), 'On partial-sum processes of armax residuals', *The Annals of Statistics* **47**(6), 3216–3243.

Grønneberg, S. & Holcblat, B. (2019*b*), 'Supplement to "on partial-sum processes of armax residuals."'.

Inclan, C. & Tiao, G. C. (1994), 'Use of cumulative sums of squares for retrospective detection of changes of variance', *Journal of the American Statistical Association* **89**(427), 913–923.

Kim, S., Cho, S. & Lee, S. (2000), 'On the cusum test for parameter changes in garch (1, 1) models', *Communications in Statistics-Theory and Methods* **29**(2), 445–462.

Krämer, W., Ploberger, W. & Alt, R. (1988), 'Testing for structural change in dynamic models', *Econometrica: Journal of the Econometric Society* pp. 1355–1369.

Lee, S., Ha, J., Na, O. & Na, S. (2003), 'The cusum test for parameter change in time series models', *Scandinavian Journal of Statistics* **30**(4), 781–796.

Lee, S. & Lee, J. (2014), Residual based cusum test for parameter change in ar-garch models, *in* 'Modeling Dependence in Econometrics: Selected Papers of the Seventh International Conference of the Thailand Econometric Society, Faculty of Economics, Chiang Mai University, Thailand, January 8-10, 2014', Springer, pp. 101–111.

Lee, S., Tokutsu, Y. & Maekawa, K. (2004), 'The cusum test for parameter change in regression models with arch errors', *Journal of the Japan Statistical Society* **34**(2), 173–188.

Nelson, D. B. (1991), 'Conditional heteroskedasticity in asset returns: A new approach', *Econometrica: Journal of the econometric society* pp. 347–370.

Oh, H. & Lee, S. (2018*a*), 'On change point test for arma–garch models: Bootstrap approach', *Journal of the Korean Statistical Society* **47**(2), 139–149.

Oh, H. & Lee, S. (2018*b*), 'On score vector-and residual-based cusum tests in arma–garch models', *Statistical Methods & Applications* **27**, 385–406.

Otto, S. & Breitung, J. (2023), 'Backward cusum for testing and monitoring structural change with an application to covid-19 pandemic data', *Econometric Theory* **39**(4), 659–692.

Page, E. (1955), 'A test for a change in a parameter occurring at an unknown point', *Biometrika* **42**(3/4), 523–527.

Phillips, P. C., Shi, S. & Yu, J. (2015), ‘Testing for multiple bubbles: Historical episodes of exuberance and collapse in the s&p 500’, *International economic review* **56**(4), 1043–1078.

Phillips, P. C. & Solo, V. (1992), ‘Asymptotics for linear processes’, *The Annals of Statistics* pp. 971–1001.

Phillips, P. C., Wu, Y. & Yu, J. (2011), ‘Explosive behavior in the 1990s nasdaq: When did exuberance escalate asset values?’, *International economic review* **52**(1), 201–226.

Ploberger, W. & Krämer, W. (1992), ‘The cusum test with ols residuals’, *Econometrica: Journal of the Econometric Society* pp. 271–285.

Ploberger, W. & Krämer, W. (1996), ‘A trend-resistant test for structural change based on ols residuals’, *Journal of Econometrics* **70**(1), 175–185.

Smirnov, N. (1948), ‘Table for estimating the goodness of fit of empirical distributions’, *The annals of mathematical statistics* **19**(2), 279–281.

# A Abbreviations and notations

This appendix provides a comprehensive list of abbreviations and notations used throughout the document. This reference is intended to assist readers in understanding and interpreting the terms and symbols used.

| Abbreviation | Meaning |
|---|---|
| ARCH | Autoregressive Conditional Heteroskedasticity |
| X-ARCH | Regression model with covariates and ARCH errors |
| CUSUM | Cumulative Sum |
| GARCH | Generalized Autoregressive Conditional Heteroskedasticity |
| AR | Autoregressive Model |
| ARMA | Autoregressive Moving Average |
| ARMAX | Autoregressive Moving Average with exogenous inputs (covariates) |
| OLS | Ordinary least-squares |
| EGARCH | Exponential GARCH |
| IGARCH | Integrated GARCH |
| TGARCH | Threshold GARCH |
| i.e. | id est (that is) |
| e.g. | exempli gratia (for example) |
| IID | Independent and Identically Distributed |

| Notations | Meaning |
|---|---|
| sup | supremum |
| $\mathbb{I}\{\}$ | Indicator function |
| mod | modulus operation |
| $\lambda$ | covariates parameter |
| $\mu$ | mean or constant |
| $\sigma$ | standard deviation |
| $\epsilon$ | ARCH innovation, strong white noise process |
| $\omega$ | intercept in the ARCH model |
| $\alpha$ | ARCH coefficient |
| T | number of observations |

# B Additional tables and figures

| T | $\Delta\mu$ | Significance 1% | | | | | Significance 5% | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Break point fraction | | | | | Break point fraction | | | | |
| | | *0.1* | *0.3* | *0.5* | *0.7* | *0.9* | *0.1* | *0.3* | *0.5* | *0.7* | *0.9* |
| **500** | *size* | 0.88 | | | | | 4.42 | | | | |
| | 0.05 | 0.8 | 2.4 | 2.9 | 1.6 | 1.1 | 4.4 | 8.2 | 10.3 | 9.9 | 5.4 |
| | 0.1 | 1.4 | 11.7 | 20.2 | 9.1 | 1.5 | 7 | 29.8 | 38.4 | 24.8 | 7.8 |
| | 0.5 | 79.7 | 100 | 100 | 100 | 81.3 | 96.6 | 100 | 100 | 100 | 96.8 |
| | 1.0 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| **1000** | *size* | 0.88 | | | | | 4.48 | | | | |
| | 0.05 | 1 | 5.7 | 7.7 | 4.6 | 1.4 | 6.1 | 17.4 | 19.9 | 16.1 | 7.4 |
| | 0.1 | 2.3 | 31.4 | 46.5 | 30.5 | 2 | 9.8 | 53.6 | 69.8 | 57.1 | 9.5 |
| | 0.5 | 99.9 | 100 | 100 | 100 | 99.5 | 100 | 100 | 100 | 100 | 100 |
| | 1.0 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| **1500** | *size* | 0.88 | | | | | 4.56 | | | | |
| | 0.05 | 0.8 | 9.3 | 15.8 | 7.8 | 1.4 | 7.1 | 24.8 | 29.7 | 22.8 | 6.1 |
| | 0.1 | 4 | 50 | 69.4 | 47.7 | 2.7 | 14.6 | 73.9 | 85.8 | 73.8 | 12.9 |
| | 0.5 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| | 1.0 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |

Table B1: Empirical sizes and powers of the CUSUM test in the X–ARCH(1) model (with constant mean and $(\mu, \omega, \alpha)$ = (0.0, 0.2, 0.4)) with the shift in mean $\Delta\mu$

| T | $\Delta\mu$ | **Significance 1%** | | | | | **Significance 5%** | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | **Break point fraction** | | | | | **Break point fraction** | | | | |
| | | *0.1* | *0.3* | *0.5* | *0.7* | *0.9* | *0.1* | *0.3* | *0.5* | *0.7* | *0.9* |
| **500** | *size* | 0.82 | | | | | 4.08 | | | | |
| | 0.05 | 0.5 | 1.7 | 3.3 | 1.3 | 1 | 3.3 | 6.5 | 8.8 | 6.9 | 4.9 |
| | 0.1 | 0.9 | 6.2 | 12.1 | 5.9 | 1.5 | 6 | 19.2 | 28.6 | 19.7 | 5.9 |
| | 0.5 | 49 | 100 | 100 | 100 | 46.8 | 78.5 | 100 | 100 | 100 | 79.3 |
| | 1.0 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| **1000** | *size* | 0.8 | | | | | 4.34 | | | | |
| | 0.05 | 0.7 | 3.1 | 5.3 | 3 | 0.9 | 3.7 | 10.8 | 15.9 | 11.7 | 4.8 |
| | 0.1 | 2.4 | 19.6 | 27.8 | 17.4 | 1.4 | 9.2 | 39.2 | 50.8 | 35.8 | 7.8 |
| | 0.5 | 95.3 | 100 | 100 | 100 | 96.2 | 99.6 | 100 | 100 | 100 | 99.5 |
| | 1.0 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| **1500** | *size* | 0.94 | | | | | 4.5 | | | | |
| | 0.05 | 1 | 5.8 | 9.4 | 5.6 | 1 | 5.2 | 17 | 22.3 | 15.6 | 5.9 |
| | 0.1 | 1.9 | 27.9 | 44.2 | 29.3 | 3 | 11.1 | 53.8 | 68.8 | 54.4 | 9.8 |
| | 0.5 | 99.9 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| | 1.0 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |

Table B2: Empirical sizes and powers of the CUSUM test in the X–ARCH(1) model (with constant mean and $(\mu, \omega, \alpha)$ = (0.0, 0.4, 0.2)) with the shift in mean $\Delta\mu$

| T | $\Delta\mu$ | **Significance 1%** | | | | | **Significance 5%** | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | **Break point fraction** | | | | | **Break point fraction** | | | | |
| | | *0.1* | *0.3* | *0.5* | *0.7* | *0.9* | *0.1* | *0.3* | *0.5* | *0.7* | *0.9* |
| **500** | *size* | 1.06 | | | | | 4.58 | | | | |
| | 0.05 | 1 | 1.5 | 2.8 | 1.5 | 1 | 5 | 5.5 | 7.7 | 5.8 | 4.5 |
| | 0.1 | 1 | 4.6 | 6.8 | 4.8 | 1.4 | 4.6 | 14.3 | 20 | 16.2 | 5.3 |
| | 0.5 | 27.9 | 99.8 | 100 | 99.8 | 30.9 | 63.3 | 100 | 100 | 99.9 | 63.2 |
| | 1.0 | 99.8 | 100 | 100 | 100 | 99.9 | 100 | 100 | 100 | 100 | 100 |
| **1000** | *size* | 0.84 | | | | | 4.16 | | | | |
| | 0.05 | 0.9 | 3 | 5 | 3 | 1.8 | 5.1 | 8.3 | 12.9 | 9.3 | 5.4 |
| | 0.1 | 1.1 | 10.3 | 21.4 | 11.6 | 1.4 | 7.9 | 26.1 | 39.9 | 28.5 | 6.6 |
| | 0.5 | 81.7 | 100 | 100 | 100 | 81.4 | 96.6 | 100 | 100 | 100 | 95.9 |
| | 1.0 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| **1500** | *size* | 0.94 | | | | | 4.3 | | | | |
| | 0.05 | 1.9 | 4.6 | 6 | 3.7 | 0.8 | 5.9 | 13 | 17.3 | 12.8 | 5.5 |
| | 0.1 | 2.1 | 21.2 | 30.3 | 20.8 | 2.8 | 9.2 | 44.4 | 54.5 | 41.7 | 9.5 |
| | 0.5 | 98.1 | 100 | 100 | 100 | 98.5 | 100 | 100 | 100 | 100 | 99.9 |
| | 1.0 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |

Table B3: Empirical sizes and powers of the CUSUM test in the X–ARCH(1) model (with constant mean and $(\mu, \omega, \alpha)$ = (0.0, 0.4, 0.4)) with the shift in mean $\Delta\mu$

# C Simulation of the critical values for change-point covariate model with ARCH effect

| p | 90% Quantile | 95% Quantile | 99% Quantile |
|---|---|---|---|
| 0.1 | 1.8535 | 2.1220 | 2.6515 |
| 0.3 | 1.6491 | 1.8827 | 2.3526 |
| 0.5 | 1.4601 | 1.6502 | 2.0346 |
| 0.7 | 1.3075 | 1.4593 | 1.7685 |
| 0.9 | 1.2186 | 1.3490 | 1.6307 |

Table C4: Critical values for different values of $p$ for X-ARCH(1) model from Equation 12