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 of Science in Quantitative Economics and Finance

Author:
Yevhen Perehuda

Supervisor:
Assistant Professor
Benjamin Holcblat

Reviewer:
Associate Professor
Julien Pénasse

August 2023

Abstract

Over recent decades, the Cumulative Sum (CUSUM) test has emerged as a standard diagnostic tool for time-series structural break evaluation. At its core, the CUSUM test monitors the cumulative sum of residuals and identifies the parameter shift if the sum surpasses the critical value at some point. However, there is limited research on CUSUM's role in regression models with the ARCH effect — a mechanism capturing volatility autocorrelation. In this research, I conduct an in-depth simulation study of residual-based CUSUM tests, focusing on structural breaks in such X-ARCH models in different scenarios. The methods I used in the study lean on the general theory by Grønneberg & Holcblat (2019*a*), refined for the ARCH context. Results show that this approach effectively identifies changes in both the mean and X covariates. The Monte-Carlo simulations consistently demonstrate stable empirical size for different X and ARCH parameters, which highlights the reliability of the residual CUSUM test. For practical application, I further validated the approach by analyzing the yen/dollar exchange rate.

Contents

1 Introduction

Time-series data, due to its temporal nature, often exhibits changing patterns and structural shifts. Detecting such changes has principal 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.

ARCH is a statistical concept used to describe time series data where the variance of the current error term or volatility is a function of the previous time periods' error terms. In simpler terms, it captures the phenomenon where periods of high volatility are likely to be followed by high volatility and low volatility by low volatility. Recognizing and accounting for ARCH patterns is crucial because many financial time series, like stock returns, exhibit such volatility clustering. Ignoring ARCH effects can lead to incorrect model specifications, leading to unreliable forecasts and misguided policy decisions.

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

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

where ϵ_t is a sequence of length n of independent and identically distributed (IID) random variables with zero-mean, σ_t denotes time-varying conditional volatility, and $X_{T,t-1}$ defines time-dependent covariates that can rely on sample size T. Subsequently, $y_{T,t}$ becomes T-dependent too.

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

Pioneering work of Brown et al. (1975) has set the stage for detecting change-points in time-series. They introduced CUSUM method, which is a statistical method designed to detect shifts in the mean or other parameters of the time-series. It operates by continuously accumulating the differences between observed and predicted values. When this cumulative sum exceeds a certain threshold, it signals a potential change or shift in the process, meaning that the applied model has a misspecification.

The overlap of CUSUM tests with ARCH models has been minimally addressed in research. Though Kim et al. (2000), Lee et al. (2004) and Kulperger & Yu (2005) initiated discussions on the topic, there is a distinct lack of extensive studies on mean change tests for regression models with ARCH errors. Many traditional findings related to the CUSUM test are contingent on errors that are IID or operate under the martingale difference principle (referenced by works like Krämer et al. (1988), Ploberger & Krämer (1992)). This is despite the acknowledged need to account for error autocorrelation in certain time-series data. Notably, Grønneberg & Holcblat (2019*a*) introduced a comprehensive set of theoretical tools and flexible guidelines to create new residual-based CUSUM tests. These tests are constructed with intentionally broad assumptions to ensure their wide applicability. Importantly, they are not solely tied to IID errors but can also handle more complex error structures, including ARCH and generalized autoregressive conditional heteroskedasticity (GARCH) effects.

It is crucial to acknowledge the constraints of the suggested method. Although the examined CUSUM test is versatile, its efficacy in identifying shifts in parameters that influence the second moment in ARCH models might be limited. Nonetheless, its ability to detect changes in the first moment remains notably robust.

This thesis is organised as follows: Section 2 offers a brief overview 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 my model. Section 4 showcases my Monte-Carlo simulation study, elucidating the test's empirical properties. In Section 5, I conduct real data analysis, emphasizing my test's practical applications. Finally, Section 6 concludes the study, summarizing my findings and pointing towards potential areas for future research. The appendices provide additional details, with Appendix A offering a comprehensive list of abbreviations and notations used throughout the thesis, Appendix B presents additional tables with simulation results, and Appendix C provides critical values for limiting distribution of the CUSUM test used in Section 4.2

2 Literature review

The foundational paper of Page (1955) introduced the problem of testing the change-point in time-series. Page's work became 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. The alternative hypothesis is in case of changing at some unknown point. Change-point estimation finds diverse applications, which can include monitoring of intensive-care patients (Fried & Imhoff 2004), climatic time-series (Gallagher et al. 2013)and detecting financial bubbles (Phillips et al. 2011, 2015).

Due to De Prado (2018), structural break tests can be broadly categorized into 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). The CUSUM test is a prevalent choice among these methods as a diagnostic tool due to its simplicity in understanding and straightforward implementation.

CUSUM tests can be divided into estimation-, score- 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 count either recursive (Brown et al. 1975, Krämer et al. 1988, Otto & Breitung 2023) or standard residuals(Ploberger & Krämer 1992, 1996, Grønneberg & Holcblat 2019*a*, Lee et al. 2004). The latter ones 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, 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. This includes 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) model (Glosten et al. 1993, Zakoian 1994), which also considers the leverage effect. Additionally, the introduction of the IGARCH (Integrated GARCH) model (Engle & Bollerslev 1986) recognizes the persistence in volatility by constraining the sum of ARCH and GARCH parameters to be one. Also, the book of Francq & Zakoian (2019) is considered as one of the main references for GARCH models, where they provide a comprehensive study of the family of GARCH models and their application. Such continual advancements reflect the industry's effort to model and predict financial volatilities with increasing precision.

In the domain of the intersection of CUSUM tests with ARCH and GARCH models, Inclan & Tiao (1994) were pioneers in proposing the CUSUM test, which was primarily devised to detect variance shifts within IID samples. Subsequently, Kim et al. (2000) adapted the CUSUM test to the GARCH(1,1) model. They believed that changing the GARCH parameters would affect the variance. While their theoretical grounding was sound, in practice, the CUSUM test exhibited significant size distortions and reduced power. Addressing these limitations, Lee et al. (2004) 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 is 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. 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-GARCH models, and the research by Oh & Lee (2018*a*,*b*) delved into the intricacies of ARMA-GARCH. Yet, to the best of my knowledge, no explicit tests have been formulated for a regression model with covariates and ARCH errors (X-ARCH model). Importantly, the CUSUM tests discussed in the current and previous paragraphs targeted structural breaks affecting only the second moment, because these tests operated over squares of residuals. Regarding the first moment - mean, there are no studies that would investigate it in the ARCH setup.

The work of Grønneberg & Holcblat (2019*a*) introduced a notable advancement by establishing the limit behaviour of the partial-sum process of ARMAX residuals and introducing a universal algorithm for pivotal CUSUM test statistics. Crucially, their proposed statistic can detect mean shifts in regressions across various error types, including autocorrelated ones. This innovation has a lot of implications, particularly for a vast array of models under the ARMAX-GARCH umbrella. This work is exceptionally relevant in the context of the X-ARCH model, addressing a previously unexplored area in the literature.

3 Residual-based CUSUM test

Before conducting simulations, it's essential to have a robust theoretical understanding. The success of a simulation depends on precise statistics and understanding the residual processes, which requires an in-depth exploration of the relevant theory.

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

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

Our main focus is on the following standard ($\mathfrak{U}_T(s)$) and average-corrected ($\mathfrak{V}_T(s)$) 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 μ and λ s.t. (a) $\sqrt{T}(\hat{\mu} - \mu) = O_P(1)$ and (b) $\sqrt{T}(\hat{\lambda} - \lambda) = O_P(1)$*

In my study, I make no specific assumptions about the estimators of the X-ARCH parameters other than they are $O_P(T^{-\frac{1}{2}})$ away from their targets (population values). Such an assumption is essential only for parameters that affect the X part. Hence, if the parameters are correctly identified, the results are generally valid regardless of the estimation method chosen.

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 important for proof of the presence of specific power series in B applied to the error term u_t in terms of almost sure convergence, which is evident from Lemma 7(i) by Grønneberg & Holcblat (2019*b*). Furthermore, Assumption 2(a) paves the way for the application of the Phillips–Solo device methodology (Phillips & Solo 1992). It is worth noting that Assumption 2(a) is relatively mild, primarily because it necessitates the existence of only the first moment. This requirement stands in contrast to the work by Lee et al. (2004), Deng & Perron (2008), Oh & Lee (2018*a*), which demands the existence of the fourth moment, or by Ploberger & Krämer (1992) with the second moment.

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. This assumption's generality is one of the main breakthrough parts, given the inherent heteroscedastic nature of numerous financial and economic time-series and the challenges in negating autocorrelation in errors. Many models often produce errors that deviate from being just IID or martingale differences, this assumption highlights the broad applicability and relevance of Assumption 2(b) in diverse scenarios. Assumption 2(b), together with the Phillips–Solo device, ensures the asymptotic vanishing of the partial-sum average process of the 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 Lemma 1, we find the gap between error and residuals appearing asymptotically and after we get rid of nuisance terms with the help of additional assumptions.

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

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

Part (a) of Theorem 1 provides insights into the asymptotic difference between the scaled residual partial-sum process and its error counterpart. The subsequent part, Theorem 1(b) and (c), together with satisfied Assumption 3(c), allow to mask the covariate nuisance term, while analysing the partial-sum process of *average-corrected residuals*. For derivation, the Grønneberg & Holcblat (2019*b*) also used the Phillips–Solo device. Meanwhile, Theorem 1(c) presents the most straightforward approach for obtaining a pivotal statistic for residual-based CUSUM tests, particularly when the covariates do not have a zero-mean. However, in order to derive a pivotal statistic in conventional scenarios, it is imperative that σ_u should be consistently estimated.

Proposition 1. *Suppose an 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 I focus on the univariate scenario and use, consequently, variance and standard deviation in place of a covariance matrix. Proposition 1, combined with Theorem 1(b) and (c), provides pivot CUSUM statistics when used in conjunction with 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° *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) and MacNeill (1978)), while condition (a) simply requires that standard empirical deviation approach the actual deviation. 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, referring to (Smirnov 1948). Sometimes in the literature, such limiting distribution is called as "Kolmogorov" distribution.

Remark 1. Through my analysis, I identified a constraint in the aforementioned residual-based CUSUM test. Despite the mild assumptions, making the test a versatile toolbox for detecting changes, it struggles to detect alterations in parameters that affect only the second moment.

4 Simulation study

In empirical research, especially when developing or introducing a new statistical test, it is essential to validate its properties under controlled scenarios. Simulation studies offer a structured environment to understand the behaviour and characteristics of a proposed test. In this section, I embark on an extensive simulation exercise to empirically validate the effectiveness of the CUSUM test in the context of regression models with ARCH errors.

Here, I investigated two critical aspects of any statistical test: its size and 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 my exploration, I implemented these tests at the significance levels of $\alpha = 0.01$ and 0.05. The Monte-Carlo simulations 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. I used sample sizes of $T = 500$, 1000, and 1500 for each test. The value of ϵ_t was taken from the standard normal distribution $N(0, 1)$. In addition, it should be highlighted the fraction a of the sample, which is influenced by structural change, is chosen from the list $(0.1, 0.3, 0.5, 0.7, 0.9)$. Such a way of changing the breakpoint fraction was inspired by Ploberger & Krämer (1992).

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 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 critical values.
4. Repeat steps 1-3 multiple times (e.g., 1000 iterations) to compute empirical size and power (measured in percentages).

Here, I use Python 3.10 running on Windows 11 and the packages "sklearn" and "arch". The code is publicly available through GitHub[1].

In the following subsections, I will comprehensively assess the effectiveness of the CUSUM test within the X-ARCH framework by investigating three distinct scenarios:

[1] `https://github.com/Eugen17/X-ARCH-CUSUM-test/tree/main`

(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 measure the test's efficiency in identifying additional structural break which is not known; and (3) seasonal dummy model, 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 error

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 quite similar to 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 : θ 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.

Table 1 shows the basic results, which include simulation for initial $\theta = (0, 0.2, 0.2)$. To ensure a comprehensive evaluation, I 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-B6.

My simulation results show the robustness of the CUSUM test, revealing no size distortions given Tables 2, B2, B4 and B6. An intriguing observation, consistent across simulations, is the diminished sensitivity of the CUSUM tests if the change-point is located near the series' boundaries. This characteristic is inherent to CUSUM tests and warrants consideration when drawing conclusions from the results.

Further, as I introduced more pronounced shifts in the mean, $\Delta\mu = \mu' - \mu$, the power of the test improved, especially evident in larger sample sizes. This phenomenon underscores the test's ability to accurately detect pronounced mean shifts even with 1% significance. Tables B1, B3, B5 show that the test stays stable even with increasing ARCH parameters ω and α, leading just to slightly reduced efficiency. However, for smaller mean shifts, especially when the structural break is proximate to the series' extremities, it is recommended to take caution.

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	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	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	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 powers of the CUSUM test in the X–ARCH(1) model (constant mean and (μ, ω, α) = (0.0, 0.2, 0.2)) with the shift in mean $\Delta\mu$.

T	Level of significance	
	1%	**5%**
500	0.81	3.84
1000	0.84	4.26
1500	0.91	4.28

Table 2: Empirical sizes of the CUSUM test in the X–ARCH(1) model (constant mean and (μ, ω, α) = (0.0, 0.2, 0.2)) with the shift in mean $\Delta\mu$.

4.2 Change-point covariate model with IID and ARCH errors

Examples 4 and 6, by Grønneberg & Holcblat (2019*a*), introduced a novel CUSUM test tailored for a model with a change-point covariate. Specifically, this model employed a dummy variable $I\{t \leqslant \mathfrak{p}T\}$, representing a change-point at the initial $\mathfrak{p}$-th fraction of the sample. The authors originally emphasized that the statistic possesses a limiting distribution, as shown in Equation 9 and Table C12, but under the premise of a zero-mean IID shock.

4.2.1 Model with IID error

First, I examine the model previously discussed with only IID errors. My testing primarily focus on the shift in μ. However, rather than altering the location of the change as in the previous subsection, I adjust the parameter $\mathfrak{p}$ within the covariate.

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

where u_t is IID standard normal random variable.

Since the indicator covariate does not meet Assumption 3(c), average-corrected residuals are not effective in eliminating nuisance terms. Therefore, only Theorem 1 (a) is applicable. The CUSUM test, initially introduced by Grønneberg & Holcblat (2019*a*), is defined as:

$$\hat{\sigma}_u^{-1}\sqrt{T} \sup_{s\in[0,1]} \left|\hat{\mathfrak{U}}_T(s) - [\min(s,\mathfrak{p})/\mathfrak{p}]\hat{\mathfrak{U}}_T(1)\right| \xrightarrow{\mathscr{L}} \sup_{s\in[0,1]} \mid B(s) - [\min(s,\mathfrak{p})/\min(1,\mathfrak{p})]B(1) \mid, \tag{9}$$

where B is a Brownian motion. For this test, I found the critical values via simulation, which are listed in Table C12.

The following hypotheses were considered for the testing:

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

vs.

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

This test poses the question: given knowledge of the first change-point occurring at the $\mathfrak{p}$-th fraction, can we detect a second change-point in the middle with a change of $\Delta\mu$?

The empirical sizes and powers of the CUSUM test in the simple regression model (only with IID errors) with the shift in mean $\Delta\mu$ are presented in Table 3 and B7. The test, while possessing power, is a bit less proficient at detecting shifts compared to the model in Section 4.1. Concurrently, it is noteworthy from Table B7 that the test does not show any discrepancies in size.

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	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	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	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 3: Empirical powers of the CUSUM test in the regression model (IID shock) and with the shift in mean $\Delta\mu$.

4.2.2 Model with ARCH error

Through my additional experimental analysis, I found that the same limiting distribution can be observed even for a known zero-mean regression model, but that incorporates an ARCH error. This insight expands the applicability of the CUSUM test described. The model is then revised to:

$$\begin{aligned} y_t &= \lambda I\{t \leqslant \mathfrak{p}T\} + u_t, \\ u_t &= \sigma_t \epsilon_t, \quad \sigma_t^2 = \omega + \alpha u_{t-1}^2. \end{aligned} \tag{10}$$

The following hypotheses were considered for the testing:

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

H_1 : θ change to $\theta' = (\mu', \lambda, \omega, \alpha)$ 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	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	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	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 4: Empirical powers of the CUSUM test in the X–ARCH(1) model (indicator regressor and $(\mu, \lambda, \omega, \alpha)$ = (0.0, 2, 0.2, 0.2)) with the shift in mean $\Delta\mu$.

I discovered that the model with ARCH using the CUSUM test requires a significantly reduced sample size for comparable accuracy to the model with IID samples. For example, it more effectively identifies the change-point in mean when $\Delta\mu = 0.1$ and sample size $T = 500$ compared to the IID scenario. Table B8 illustrates that the test size converges to the intended significance level.

Additionally, I explored whether the test identifies a shift in the parameter λ at the midpoint of the known change-point section ($\mathfrak{p}/2$). This would imply an extra structural break, which is similar to an already known one but of a shorter duration.

The second conducted simulation considered the testing:

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

vs.

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

Table B9 indicates that there are no size violations. Due to Table 5, the test demonstrates high power starting from $\Delta\lambda = 0.5$ and $\mathfrak{p} = 0.5$, with positive correlations evident with $\mathfrak{p}$, and that is intuitive. Imagine, if the known change-point is at $\mathfrak{p} = 0.3$, it means that the unknown change-point has a fraction of just $0.15T$. This limited sample size makes identifying unanticipated shifts in CUSUM statistics challenging. Therefore, for such changes in λ, the enhanced sensitivity of the new CUSUM becomes evident for larger $\mathfrak{p}$.

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	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	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	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 5: Empirical powers of the CUSUM test in the X–ARCH(1) model (indicator regressor and $(\mu, \lambda, \omega, \alpha) = (0.0, 2, 0.2, 0.2)$) with the shift in parameter $\Delta\lambda$.

4.3 Seasonal dummy model with ARCH error

Seasonal patterns frequently appear in real-world time-series. A common approach to capture this seasonality is to incorporate seasonal dummy variables:

$$\begin{aligned} y_t &= \mu + \sum_{k=1}^{d} \lambda_k I\{t \equiv k (\bmod d)\} + u_t, \\ u_t &= \sigma_t \epsilon_t, \quad \sigma_t^2 = \omega + \alpha u_{t-1}^2, \end{aligned} \tag{11}$$

where d is the number of seasonal dummies and λ_k – parameters for them. It is evident that the inclusion of seasonal dummies introduces additional terms in the expansion of the residual's partial sum. However, these terms become negligible when considering average-corrected residuals, due to Proposition 2 by Grønneberg & Holcblat (2019*a*), because such covariates satisfy Assumptions 3(a)-(c). For my simulation study, I chose $d = 3$ and initial $\boldsymbol{\lambda} = (1, -1, 2)$.

4.3.1 Mean change

In this section, I conduct Monte-Carlo simulations to examine a shift in mean, similar to the approach taken in Section 4.1, but for varying breakpoint fractions.

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

Tables 6 and B10 shows a pattern similar to Tables 1 and 2. For all setups, the sizes align closely with the nominal level. The test also produces reasonably good powers for $\Delta\lambda$ starting at 0.5 regardless of the sample size. Overall, my findings strongly support the validity of the CUSUM for identifying shifts in the mean.

4.3.2 Parameter change

Besides investigating shifts in mean, I also delved into the ability to detect changes in one of the seasonal parameters, particularly λ_1.

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	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	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	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 7: Empirical powers of the CUSUM test in the X–ARCH(1) model (seasonal dummies, (μ, ω, α) = (0.0, 0.2, 0.2) and λ = (1, -1, 2)) with the shift in the parameter $\Delta\lambda_1$.

The results from Tables 7 and B11 suggest that in such scenarios, the test exhibits some kind of size distortion and considerably reduced power compared to earlier configurations. The decreased effectiveness can potentially be attributed to the increased number of covariates, making estimation more challenging and leading to a violation of Assumption 1(a). As for the lowered power (e.g., for $\Delta\lambda_1 = 0.5$ case), the reason can be the modification of just one seasonal covariate. This shift in the covariate now affects not the entire fraction $(1-a)T$ of the time-series but only a segment, precisely $(1-a)T/3$. Naturally, the more localized alteration is less prominent. Nevertheless, the test maintains good power for larger sample sizes and larger values of the shift.

5 Real data analysis

In my analysis of real data, I examined the yen/dollar exchange rate returns from January 5, 1998, to January 27, 2003, as shown in Figure 1. This data mirrored that of Lee et al. (2004).

Figure 1: Exchange rate YEN/USD from Jan 5, 1998 to Jan 27, 2003.

Upon applying the Ljung–Box and LM–ARCH tests, it became clear that the model incorporating ARCH(1) effectively captures the dynamics of this series. For both tests, the p-values are extremely close to zero, significantly lower than typical significance thresholds (e.g. 0.05). As a result, we reject the null hypothesis of no autocorrelation in residuals, supporting the hypothesis of ARCH effect presence.

For the fitting, I employed the model outlined in Equation 1 with a constant mean. The founded change-point looks more interpretable than in (Lee et al. 2004). The residual-based CUSUM test yielded a value of 8.05, as visible in Figure 2, on 7 October 1998. The value of 8.05 surpasses the theoretical critical value of 1.63 at the 0.01 significance level, suggesting a significant shift in the mean. Notably, this date falls towards the end of 1998. This specific time frame might be significant due to its association with the tail end of the Asian financial crisis.

Moreover, the Japanese yen experienced its most significant two-day surge on 6-8 October since it began freely floating in February 1973, after the dissolution of the Bretton Woods agreement, as noted by Cooper & Talbot (1999). The shift in the yen's exchange rate was primarily linked to the unwinding of the so-called "yen carry trade" and intensified market reactions during the Asian Financial Crisis. Due to 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.

The Fall of 1998 also stands out due to its new monetary and economic policies. Around this period, Japan witnessed the introduction of several groundbreaking regulatory measures aiming for stability, as noted by Ghosh (2000). These events could have played a role in the observed shift:

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

Drilling down further, the pre-change segment, spanning from 5 January to 7 October 1998, conformed to the model:

$$\begin{aligned} Y_t &= 0.0074 + u_t, \\ u_t &= \sigma_t \epsilon_t, \quad \sigma_t^2 = 0.000000058 + 0.2u_{t-1}^2. \end{aligned} \tag{12}$$

In contrast, the post-change phase is best described by the model:

$$\begin{aligned} Y_t &= 0.0086 + u_t, \\ u_t &= \sigma_t \epsilon_t, \quad \sigma_t^2 = 0.00000018 + 0.2u_{t-1}^2. \end{aligned} \tag{13}$$

Visual inspection of the Figure 1 also supports the presence of a shift in mean around the determined date.

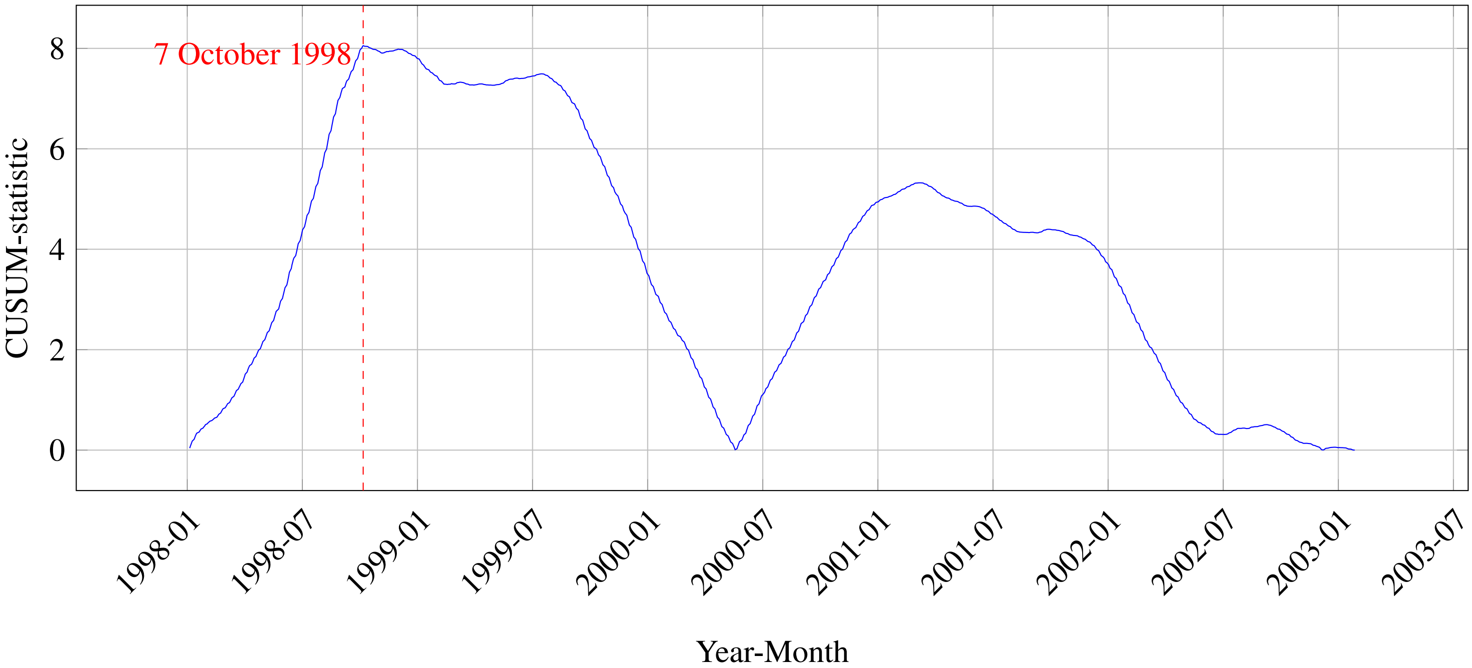

Figure 2: CUSUM-statistics from Jan 5, 1998 to Jan 27, 2003.

6 Conclusion

In this thesis, I carried out a simulation study exploring the intersection of CUSUM tests with X-ARCH models. I launched a Monte-Carlo simulation focused on the residual-based CUSUM test, with an emphasis on pinpointing structural breaks in the first moment of X-ARCH models. Through this, I was able to successfully determine the limiting asymptotic distribution.

I have made some alterations to the general assumptions presented by Grønneberg & Holcblat (2019*a*,*b*), to fit the ARCH context. The versatility of the test arises from its ability to encompass a variety of error structures, going beyond the typical IID errors. Under Assumptions 1-3, I outlined the conditions that guarantee the test statistic's convergence to the supremum of the standard Brownian Bridge. Even if Assumption 3(c) might not always be met, Theorem 1 continues to provide a foundation for the CUSUM test, as demonstrated in Section 4.2.2.

The test, I examined, is adaptable due to its general assumptions and the use of standard residuals. This makes its application more straightforward since standard residuals are readily available, unlike the recursive ones. My simulation study confirms the effectiveness of the investigated residual-based test, demonstrating its consistent capability in detecting shifts in both the mean and X covariates. The test showed significant power, particularly for changes of 0.5 or greater, regardless of the significance level and sample size. Importantly, the Monte-Carlo simulations consistently reflected consistent empirical sizes across varied ARCH parameters, highlighting the reliability of the CUSUM test.

For a practical application, I assessed the test using the yen/dollar exchange rate data spanning 1998 to 2003, and preliminary inspected it for the existence of the ARCH effect. Consequently, the test detected a structural shift precisely during key events in the Japanese economy at the end of 1998. Additionally, the correctness of the identified structural breaks can be straightforwardly validated through visual inspection.

In essence, I contend that the studied residual-based CUSUM test serves as a functional tool for change-point identifying in X-ARCH models across diverse scenarios. As a next step, we envision broadening the simulation study to more complex models, particularly focusing on different X-GARCH models.

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.

Engle, R. F. & Bollerslev, T. (1986), 'Modelling the persistence of conditional variances', *Econometric reviews* **5**(1), 1–50.

Francq, C. & Zakoian, J.-M. (2019), *GARCH models: structure, statistical inference and financial applications*, John Wiley & Sons.

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.

Glosten, L. R., Jagannathan, R. & Runkle, D. E. (1993), 'On the relation between the expected value and the volatility of the nominal excess return on stocks', *The journal of finance* **48**(5), 1779–1801.

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.

Kulperger, R. & Yu, H. (2005), 'High moment partial sum processes of residuals in garch models and their applications'.

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.

MacNeill, I. B. (1978), 'Properties of sequences of partial sums of polynomial regression residuals with applications to tests for change of regression at unknown times', *The Annals of Statistics* pp. 422–433.

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.

Zakoian, J.-M. (1994), ‘Threshold heteroskedastic models’, *Journal of Economic Dynamics and control* **18**(5), 931–955.

A Abbreviations and notations

This appendix provides a comprehensive list of abbreviations and notations used throughout the document.

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)
s.t.	such that
w.p.a.	with probability approaching one
IID	Independent and Identically Distributed

Notations	Meaning
sup	supremum
$\mathbb{I}\{\}$	Indicator function
mod	modulus operation
λ	covariates parameter
μ	mean or constant
σ	standard deviation
ϵ	ARCH innovation, strong white noise process
ω	intercept in the ARCH model
α	ARCH coefficient
T	number of observations (sample size)

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	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	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	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 powers of the CUSUM test in the X–ARCH(1) model (constant mean and (μ, ω, α) = (0.0, 0.2, 0.4)) with the shift in mean $\Delta\mu$.

T	Level of significance	
	1%	**5%**
500	0.88	4.42
1000	0.88	4.48
1500	0.92	4.56

Table B2: Empirical sizes of the CUSUM test in the X–ARCH(1) model (constant mean and (μ, ω, α) = (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	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	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	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 B3: Empirical powers of the CUSUM test in the X–ARCH(1) model (constant mean and (μ, ω, α) = (0.0, 0.4, 0.2)) with the shift in mean $\Delta\mu$.

T	Level of significance	
	1%	**5%**
500	0.82	4.08
1000	0.8	4.34
1500	0.94	4.5

Table B4: Empirical sizes of the CUSUM test in the X–ARCH(1) model (constant mean and (μ, ω, α) = (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	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	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	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 B5: Empirical powers of the CUSUM test in the X–ARCH(1) model (constant mean and (μ, ω, α) = (0.0, 0.4, 0.4)) with the shift in mean $\Delta\mu$.

T	Level of significance	
	1%	**5%**
500	0.82	4.08
1000	0.8	4.34
1500	0.94	4.5

Table B6: Empirical sizes of the CUSUM test in the X–ARCH(1) model (constant mean and (μ, ω, α) = (0.0, 0.4, 0.4)) with the shift in mean $\Delta\mu$.

T	Level of significance	
	1%	**5%**
500	0.84	4.8
1000	0.86	4.86
1500	1.28	5.32

Table B7: Empirical sizes of the CUSUM test in the regression model (IID shock) and with the shift in mean $\Delta\mu$.

T	Level of significance	
	1%	**5%**
500	1.08	4.82
1000	0.94	4.9
1500	0.78	4.82

Table B8: Empirical sizes of the CUSUM test in the X–ARCH(1) model (indicator regressor and $(\mu, \lambda, \omega, \alpha)$ = (0.0, 2, 0.2, 0.2)) with the shift in mean $\Delta\mu$.

T	Level of significance	
	1%	**5%**
500	1.18	4.84
1000	1.04	4.58
1500	0.94	4.98

Table B9: Empirical sizes of the CUSUM test in the X–ARCH(1) model (indicator regressor and $(\mu, \lambda, \omega, \alpha)$ = (0.0, 2, 0.2, 0.2)) with the shift in parameter $\Delta\lambda$.

T	Level of significance	
	1%	**5%**
500	0.88	4.82
1000	0.94	4.46
1500	0.74	4.76

Table B10: Empirical sizes of the CUSUM test in the X–ARCH(1) model (seasonal dummies, (μ, ω, α) = (0.0, 0.2, 0.2) and λ = (1, -1, 2)) with the shift in mean $\Delta\mu$.

T	Level of significance	
	1%	**5%**
500	0.48	3.9
1000	0.64	3.84
1500	0.74	3.66

Table B11: Empirical sizes of the CUSUM test in the X–ARCH(1) model (seasonal dummies, (μ, ω, α) = (0.0, 0.2, 0.2) and λ = (1, -1, 2)) with the shift in the parameter $\Delta\lambda_1$.

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 C12: CUSUM test's critical values for different values of p for X-ARCH(1) model from Equation 9.