| Statistica Sinica Preprint No: SS-2022-0397 | |
|---|---|
| **Title** | Localizing Multivariate CAViaR |
| **Manuscript ID** | SS-2022-0397 |
| **URL** | http://www.stat.sinica.edu.tw/statistica/ |
| **DOI** | 10.5705/ss.202022.0397 |


| | |
|---|---|
| **Complete List of Authors** | Xiu Xu,<br>Yegor Klochkov,<br>Li Chen and<br>Wolfgang Karl Härdle |
| **Corresponding Authors** | Li Chen |
| **E-mails** | lichen812@xmu.edu.cn |

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0397



# LOCALIZING MULTIVARIATE CAVIAR

Xiu Xu [1] Yegor Klochkov [2] Li Chen [3] Wolfgang Karl Härdle[4,5]

[1] *Finance and Economics College, Jimei University, China*

[2] *ByteDance, United Kingdom*

[3] *Wang Yanan Institute for Studies in Economics, Xiamen University, China*

[4] *School of Business and Economics, Humboldt-Universität zu Berlin, Germany*

[5] *IDA Institute Digital Assets, Bucharest University of Economic Studies, Romania*

*Abstract:* Risk transmission among financial markets and their participants is time-evolving, especially for extreme risk scenarios. Possibly sudden time variation of such risk structures asks for quantitative techniques that can cope with such situations. Here we present a novel localized multivariate CAViaR-type model to respond to the challenge of time-varying risk contagion. For this purpose, we construct a test for parameter homogeneity with totally data-driven critical values. We prove that these critical values lead to the required confidence level. Based on this test, we propose an estimation procedure that adapts to a possible time-variation of the parameter. A comprehensive simulation study supports the effectiveness of our approach in detecting structural changes in multivariate CAViaR. Finally, when applying for the US and German financial markets, we can trace out the dynamic tail risk spillovers and find that the US market appears to play a dominant role in risk transmissions, especially in volatile market periods.

*Key words and phrases:* conditional quantile autoregression, local parametric approach, change point detection, multiplier bootstrap.

Corresponding author: Li Chen (lichen812@xmu.edu.cn).

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0397

## 1. Introduction

Financial risk dependence and the mechanism of risk spillover among international equity markets have attracted increasing attentions among theorists, empirical researchers, and practitioners. A risk contagion is generated through dependence between extreme negative shocks across financial markets. It is well-known that large downside market movements occurring in one country would unavoidably have substantial effects on other international equity markets. There now exists a wide-spread consensus in the empirical literature that the dependence between the returns of financial assets is non-Gaussian with asymmetric marginals, nonlinear features, and shows massive time-variation (DÍnnocenzo et al.; 2024; Gong et al.; 2021; Okimoto; 2008).

In order to address some of these properties Engle and Manganelli (2004) propose a conditional autoregressive value at risk (CAViaR) model to specify the evolution of conditional quantile over time for univariate time series. Later, White et al. (2015) (denoted as WKM) built up a multivariate framework for multiple time series as well as various quantile levels. It can be considered as a vector autoregressive extension to quantile models with the underlying Value at Risk (VaR) processes not only being autocorrelated but also cross-sectionally intertwined. When applying multivariate CAViaR to financial institutions, it presents valuable results in capturing the sensitivity of financial entities to institutional specific and market-wide shocks of the system. However it fails to cope with time-variation or local stationary regimes. We focus on this essential feature and propose an extension towards a localized multivariate CAViaR framework that allows us to estimate and forecast the dynamics of financial risk dependence.

The majority of existing literature uses volatility as a risk measure and investigate the volatility risk contagions (Engle; 2002; Shahzad et al.; 2021). Although volatility is a fundamental instrument to measure risk shifts, it has been commonly criticized as only

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0397

capturing the properties of second moments of the return time series and ignoring extreme market events structure (Hong et al.; 2009; Han et al.; 2016). Besides, the volatility risk measure is symmetric and equally values the gains and losses, which contradicts the fact that investors tend to be more sensitive to the negative returns and especially for significant downside risk, e.g., financial crisis. Therefore, a volatility risk measure is not enough to evaluate the financial risk interdependence. On the contrary, VaR is commonly utilized to measure asymmetric risk, i.e., to evaluate the loss given a predetermined probability of extreme events. Although not a perfect (non-coherent) risk measure, it has been accepted as a standard for financial regulations, e.g., a criterion by the Basel committee on banking supervision (Tsay and Chen; 2018).

One commonly observes the unstable and time-varying interdependence of financial risk and tail contagion in empirical studies (Li et al.; 2023; Mehlitz and Auer; 2021; Baele and Inghelbrecht; 2010). The risk contagion is caused by dependence between relatively extreme negative shocks across international financial markets. A constant parametric model over a long-run time series is certainly at limit to portray elements of non-stationarity. Gerlach et al. (2011) propose a time-varying quantile model using a bayesian approach for univariate time series, and specify the model dynamics up to finite parameters. This research focuses on time-varying parameters in dynamic multivariate quantile processes to capture the dynamic interdependence and time-varying tail spillovers. We propose a framework to localize multivariate autoregressive conditional quantiles by exploiting a local parametric approach, denoted as LMCR (Local Multivariate CAViaR) model for simplicity. The advantages of our approach are at least twofold: (1) we investigate the extreme tail risk spillovers among financial time series and (2) we examine the dynamic interdependence pattern of tail risk contagion in a time-varying context.

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0397

Our LMCR model is based on the local parametric approach (LPA), which utilizes a parametric model over an adaptively chosen interval of homogeneity. The so-called interval of homogeneity is that one can not reject the null hypothesis that the model parameters are constant within this interval. The essential idea of LPA is to find the longest homogeneous interval that guarantees a relatively small modeling bias with sequential backward-looking examination, see, e.g., Spokoiny (1998, 2009). The significant advantage of this approach lies in achieving a balance between the modeling bias and parameter variability, and has been widely exploited in time-varying model analysis, see, e.g., Chen et al. (2010); Chen and Niu (2014); Niu et al. (2017); Xu et al. (2018); Zbonáková et al. (2018). Recent advances in multiplier bootstrap (MBS) allow us to construct data-driven critical values for homogeneity tests based on change point detection, see Suvorikova and Spokoiny (2017); Spokoiny and Zhilova (2015). The MBS only relies on the model equations and is free from any specific distribution assumptions, which can favourably eliminate the model misspecification issue. In our research, we extend LPA to multivariate quantile regression and develop LMCR model with multiplier boostrap technique. Due to the local homogeneous interval detection, we study the finite sample properties of WKM rather than asymptotic ones in section 2. In particular, we establish a bahadur-type expansion based on uniform exponential inequality, see Lemma 1, and compare it with the multiplier bootstrap counterpart.

Our approach appears to be well suited to capture shifting asymmetric dependence among different markets. It is worth mentioning that earlier research investigates the co-movements of substantial changes by utilizing copula-based methods, see Chen et al. (2022); Lee and Lee (2022). We emphasize that rather than relying on a fixed specification of a copula, we localize parametric modeling of risk dependence via a multivariate CAViaR model. A simulation study under various parameter change scenarios demonstrates the

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0397

performance of LMCR. Besides, when applying LMCR to tail risk analysis of the US and German market index, we find that at 1% quantile level, the typical identified interval lengths in daily time series include 140 days on average. At the higher quantile level (5%), the selected interval lengths range roughly between 160-230 days. This is an important message given the current historical simulation based risk measures are at 250 days. These findings might therefore change the regulatory risk measurement tools. The model also presents appealing merits in tracing the dynamics of tail risk spillover. We find that the US market appears to play a dominant role in risk transmissions of shocks to the German market, especially in volatile market periods.

The rest of this paper is structured as follows. In Section 2, we derive new high-probability bounds and a Bahadur-type expansion for the multivariate CAViaR model. The results serve as a non-asymptotic extension of WKM. In Section 3, we propose a homogeneity test based on a sequential change point detection, in which the critical values are evaluated via the novel multiplier bootstrap technique. We also present the corresponding theoretical results. In Section 4, we propose the adaptive approach for WKM under the possible time variation estimation, and summarize the implementation procedure for our LMCR model. In Section 5 and Section 6, a simulation study and an empirical application examine the performance of our approach. Section 7 concludes the paper. The proofs of the theorems are relegated to the Supplementary Material.

## 2. Model

We consider time-varying multivariate quantile regression in time series analysis. Define the data set $\mathcal{Y} = \{\mathbf{Y}_t : t = 1, \ldots, T\}$, with each $\mathbf{Y}_t$ being a $n \times 1$ column, and $Y_{it}$ is referred to $i$th component of this vector. Denote the natural filtration $\mathcal{F}_t = \sigma\{\mathbf{Y}_1, \ldots, \mathbf{Y}_t\}$ and we wish to estimate the quantiles of $Y_{it}$ conditioned on $\mathcal{F}_{t-1}$ at any given moment

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0397

$t = 1, \ldots, T$.

Similar to multivariate CAViaR (WKM), our LMCR model also assumes that conditional quantiles $q_{it}(\boldsymbol{\theta}, \mathcal{Y}) = \inf\{y : \mathsf{P}(Y_{it} \leq y \mid \mathcal{F}_{t-1}) \geq \tau\}$ follow an autoregressive process as,

$$q_{it}(\boldsymbol{\theta}, \mathcal{Y}) = \Psi_t^\top \gamma_i + \sum_{k=1}^{q} \sum_{j=1}^{n} \boldsymbol{\beta}_{ijk} q_{jt-k}(\boldsymbol{\theta}, \mathcal{Y}), \tag{2.1}$$

where $\mathcal{F}_{t-1}$–measurable $\Psi_t \in \mathbb{R}^d$ denote predictors available at time $t$, which typically include lagged values of times series $\mathbf{Y}_t$. We assume that the parameter set $\boldsymbol{\theta} = \left((\gamma_i)_{i=1}^n, (\boldsymbol{\beta}_{ijk})_{i,j,k=1}^{n,n,q}\right) \in \mathbb{R}^{nd+n^2q}$ is finite in the above parametric model. For the sake of simplicity, omitting the dependence on the time series outputs, we write $q_{it}(\boldsymbol{\theta}) = q_{it}(\boldsymbol{\theta}, \mathcal{Y})$ in what follows below. Besides, we denote $\|\cdot\|$ as the Euclidian norm of a vector or the spectral norm of a matrix, and denote $|\cdot|$ as the absolute value except that $|\mathcal{I}|$ means the number of observations within the sample interval $\mathcal{I}$ in the following.

For any sample interval $\mathcal{I} = [a, b]$ with the integers $a, b \in \{1, 2, \ldots, T\}$ and $a < b$, i.e. the interval $[a, b]$ denotes the interval of integers from $a$ to $b$, we write

$$(Y_{it}, \Psi_t)_{t \in \mathcal{I}} \sim \text{LMCR}(\boldsymbol{\theta}), \tag{2.2}$$

if the equation (2.1) is fulfilled on this interval with parameter $\boldsymbol{\theta}$, which is the localized multivariate CAViaR model, denoted as LMCR. This model characterizes a local tail behavior of the observed time series. Note that it may not hold in the whole historical sample due to the potential structural change, whereas we assume that there is a local data interval within which the equation (2.1) is satisfied with constant parameters.

The parameter set $\boldsymbol{\theta}$ in equation (2.1) can be estimated via the quasi-maximum likelihood estimation (QMLE), which is mathematically equivalent to minimizing the loss function typical for quantile regression. For a given quantile level $\tau \in (0, 1)$, we denote the check function $\rho_\tau(u) = u(\tau - \mathbf{1}_{[u \leq 0]})$ where $\mathbf{1}_{[\cdot]}$ is the indicator function, and

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0397

take

$$\ell_t(\boldsymbol{\theta}) = -\sum_{i=1}^{n} \rho_\tau\{Y_{it} - q_{it}(\boldsymbol{\theta})\}, \tag{2.3}$$

as the quasi log-probability of the observation $t$. The log-likelihood based on the sample interval $\mathcal{I}$ for a fixed quantile level $\tau$ is written as

$$L_{\mathcal{I}}(\boldsymbol{\theta}) = \sum_{t\in\mathcal{I}} \ell_t(\boldsymbol{\theta}) \tag{2.4}$$

and the estimator based on this set of observations is

$$\widetilde{\boldsymbol{\theta}}_{\mathcal{I}} = \arg\max_{\boldsymbol{\theta}\in\Theta_0} L_{\mathcal{I}}(\boldsymbol{\theta}), \tag{2.5}$$

where the parameter set $\Theta_0$ is a certain bounded subset from $\mathbb{R}^{n^2+np}$.

Note that the major contribution of LMCR is to analyze the time-varying parameter characteristics in multivariate quantile regressions. A constant parametric model over long-run time series is certainly at limit to portray the features of non-stationarity, which widely exist in empirical financial time series. We endeavor to capture the potential structural change through monitoring an "optimal" longest sample interval, i.e. homogeneous interval, at each time point $t$. The so-called homogeneous interval is the largest sample interval within which one can not reject the null hypothesis that the model parameters are constant under a certain significance level. For more details we refer to Section 3.

In this context, we are interested in the local properties of the response variables within a certain finite sub-interval $\mathcal{I} \subseteq [1, T]$, see Equation (2.2). Hence we need to derive the non-asymptotic bounds concerning the corresponding estimators with exponentially high probabilities in the following. Specifically, we will show that under certain conditions, for any positive value $\mathtt{x} > 0$, it holds with probability at least $1 - 6e^{-\mathtt{x}}$ that

$$\|\tilde{\boldsymbol{\theta}}_{\mathcal{I}} - \boldsymbol{\theta}^*\| \lesssim C_1\sqrt{\frac{\mathtt{x} + \log|\mathcal{I}|}{|\mathcal{I}|}}, \tag{2.6}$$

where $\boldsymbol{\theta}^*$ is the true parameter set defined in the following assumption 1, and $C_1$ is a positive constant. This type of bounds are requisite in LMCR model since they allow to

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0397



take union bounds over a growing amount of intervals $\mathcal{I}$, which is crucial in Section 3 when dealing with the local change point test. We emphasize that in the analysis, we do not take into account the dependence on $n$ and $p$ and treat them as constants.

Before we discuss the finite sample properties, some assumptions are introduced in the following part. In order to obtain the bounds with exponentially high probabilities in Equation (2.6), it requires an almost sure boundedness of the functions $q_{it}(\boldsymbol{\theta})$ as well as for their derivatives up to the second order.

## 2.1 Assumptions

We say that the LMCR model is "homogeneous" on an interval $\mathcal{I}$ if it satisfies the following assumption.

**Assumption 1.** We assume that there exists the "true" parameter set $\boldsymbol{\theta}^* \in \Theta_0$ such that $q_{it}^* = q_{it}(\boldsymbol{\theta}^*, \mathcal{Y})$ for each $i = 1, \ldots, n$ and $t \in \mathcal{I}$, and we label that the model is homogeneous within the time interval $\mathcal{I} \subseteq [1, T]$.

Assumption 1 delivers the correct specification and ensures the identification of the model. By assuming the existence of the unknown true parameter within a certain sample interval (the so-called homogeneous interval), it ensures the parameter set $\boldsymbol{\theta}^*$ optimizes the estimation objective function locally. Obviously, Assumption 1 implies that $\boldsymbol{\theta}^* = \arg\max \mathsf{E}\ell_t(\boldsymbol{\theta})$ for each $t \in \mathcal{I}$, then we can obtain that $\boldsymbol{\theta}^* = \arg\max \mathsf{E}L_{\mathcal{I}}(\boldsymbol{\theta})$, which falls into the general framework of maximum likelihood estimators (White; 1996; Spokoiny; 2017), and is prepared for the following local change point test for homogeneity in Section 3. The next assumption controls the values and derivatives of the quantile regression functions.

**Assumption 2.** For any $\epsilon > 0$ there exists $\delta = \delta(\epsilon) > 0$ such that whenever $\|\boldsymbol{\theta} - \boldsymbol{\theta}^*\| \geq \epsilon$, it holds that

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0397



$$\mathsf{P}\left(\max_{i\leq n}|q_{it}(\boldsymbol{\theta})-q_{it}(\boldsymbol{\theta}^*)|\geq\delta\right)\geq\delta,\qquad\text{for all } t\in\mathcal{I}. \tag{2.7}$$

Recall that $\|\cdot\|$ denotes the Euclidian norm of a vector and the spectral norm of a matrix. In addition, we denote $f_{it}(\cdot)$ as the density of the innovations $\varepsilon_{it}=Y_{it}-q_{it}^*$ conditional on $\mathcal{F}_{t-1}$. A bound is commonly required for this density function in quantile regressions, see Koenker and Machado (1999); Koenker and Xiao (2006). For instance, the value of this density at point 0 appears in the median asymptotic distribution, which is a classical result in statistics. The next assumption ensures the desired boundedness of the density function $f_{it}(\cdot)$. To avoid technical difficulties, we assume that a uniform bound exists everywhere.

**Assumption 3.** (i) For $s=0,1,2$ there are constants $D_s>0$ such that for each $i,t$ and for each $\boldsymbol{\theta}\in\Theta_0$ it holds pointwise $|q_{it}(\boldsymbol{\theta})|\leq D_0$, $\|\nabla q_{it}(\boldsymbol{\theta})\|\leq D_1$ and $\|\nabla^2 q_{it}(\boldsymbol{\theta})\|\leq D_2$. (ii) the conditional density function of innovations $\varepsilon_{it}$ are bounded from above, i.e. $f_{it}(x)\leq\bar{f}$ for each $i,t$ and $x\in\mathbb{R}$. (iii) Additionally, the conditional density of innovations satisfies $f_{it}(x)\geq\underline{f}$ for $|t|\leq\delta_0$, and $\delta_0$ is a positive constant.

Assumption 3 (i) ensures the quantile function as well as its derivatives up to second order are bounded. Assumption 3 (ii) is the standard bounded conditional density assumption in the literature of quantile regression. Assumption 3 (iii) ensures the density is bounded from below, which means we have even a small testing interval enough observations. Note that Assumption 3 is akin to Assumption 5 of WKM, but we replace the bounds in mean with almost sure boundness. This is a payment for finite sample exponential bounds that follow below.

The set of assumptions has to be different since the essence of our paper is to localize the WKM model. This requires finer and possibly stronger appearing assumptions since WKM are only dealing with a single-likelihood optimization problem. In contrast, we need to handle the fluctuations of the maximum of likelihood ratios over time-varying

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0397



intervals. There are potential structural breaks for the response variable, which is the main focus of our localizing Multivariate CAViaR model.

Furthermore, in extending the setting to dependent situations, we impose the following assumptions. Recall the definition of the mixing coefficients, see Chanda (1974); Bradley (2005). For any sub $\sigma$-fields $\mathcal{A}_1, \mathcal{A}_2$ of same probability space $(\Omega, \mathcal{F}, \mathsf{P})$, we define

$$\alpha(\mathcal{A}_1, \mathcal{A}_2) = \sup_{A \in \mathcal{A}_1, B \in \mathcal{A}_2} |\mathsf{P}(A \cap B) - \mathsf{P}(A)\mathsf{P}(B)|,$$

$$\beta(\mathcal{A}_1, \mathcal{A}_2) = \sup_{(A_i) \subset \mathcal{A}_1, (B_i) \subset \mathcal{A}_2} \sum_{i,j} |\mathsf{P}(A_i \cap B_j) - \mathsf{P}(A_i)\mathsf{P}(B_j)|,$$

where in the latter the supremum is taken over all finite partitions $(A_i) \subset \mathcal{A}_1$ and $(B_j) \subset \mathcal{A}_2$ of $\Omega$. Then the coefficients

$$a_k((X_t)) = \sup_t \alpha(\sigma(X_1, \ldots, X_t), \sigma(X_{t+k}, \ldots, X_T)),$$

$$b_k((X_t)) = \sup_t \beta(\sigma(X_1, \ldots, X_t), \sigma(X_{t+k}, \ldots, X_T))$$

denote $\alpha$– and $\beta$–mixing coefficients of the process $(X_t)_{t \leq T}$, respectively.

**Assumption 4.** (i) Suppose that the sequence of vectors $(q_{\cdot t}(\boldsymbol{\theta}), \nabla q_{\cdot t}(\boldsymbol{\theta}))$ is $\alpha$–mixing with $\alpha(m) \leq \exp(-\gamma m)$ for some constants $\gamma > 0$; (ii) The sequence of vectors $\nabla q_{\cdot t}(\boldsymbol{\theta}^*)$ is $\beta$–mixing with coefficients $\beta(m) \leq m^{-\delta'}$ for some constants $\delta' > 1$; (iii) for each $i = 1, \ldots, n$, the innovations $(\varepsilon_{it})_{t \in \mathcal{I}}$ are independent and satisfy $\mathsf{P}(\varepsilon_{it} < 0) = \tau$.

Assumption 4 helps us in controlling the stochastics of random sums of the involved test elements and it is deeply rooted in the strong results from mixing theory. Assumption 4 (i) is designed for the whole parameter set and it is weaker than (ii) which controls the dependencies at the true parameters, while $\beta$-mixing implies $\alpha$-mixing. We use different conditions under different reference that we need as precise as possible. The last part (iii) is natural and standard in quantile regressions which fixes the quantile level. Assumption

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0397



4 is required to prove Lemma 1 and Theorem 1, and it basically ensures the sequence is a $\beta$-mixing process. Since the GARCH process is $\beta$-mixing, it makes sense that we generate GARCH processes in the simulation at Section 5, which fulfill all these assumptions.

Before imposing the last assumption, we introduce some additional notations. The *score* function of the likelihood in Equation (2.4) is written as

$$\nabla L_{\mathcal{I}}(\boldsymbol{\theta}) = \sum_{t\in\mathcal{I}} \nabla \ell_t(\boldsymbol{\theta}) = \sum_{t\in\mathcal{I}} \mathbf{g}_t(\boldsymbol{\theta}),$$

where we denote $\mathbf{g}_t(\boldsymbol{\theta}) = \nabla \ell_t(\boldsymbol{\theta})$. Using the definition of log-likelihood in Equation (2.3), we can explicitly write $\mathbf{g}_t(\boldsymbol{\theta}) = \sum_i \nabla q_{it}(\boldsymbol{\theta})\psi_\tau(Y_{it} - q_{it}(\boldsymbol{\theta}))$, where $\psi_\tau(u) = \tau - \mathbf{1}_{[u\leq 0]}$ is the directional derivative of $\rho_\tau(u)$. The assumption below touches the information matrix as well as the variance of the score and is required to obtain the boundedness of the estimator in finite sample.

**Assumption 5.** The vector $(\mathbf{q}_t^*, \nabla \mathbf{q}_t(\boldsymbol{\theta}^*), \boldsymbol{\varepsilon}_t)$, $t \in \mathcal{I}$ is a stationary process. Additionally, the matrices

$$Q = \sum_{i=1}^{n} \mathsf{E}\left\{ f_{it}(0)\nabla q_{it}(\boldsymbol{\theta}^*)[\nabla q_{it}(\boldsymbol{\theta}^*)]^\top \right\}, \qquad V = \mathsf{Var}\{\mathbf{g}_t(\boldsymbol{\theta}^*)\}$$

are strictly positive definite and symmetric.

## 2.2 Consistency of the estimator

Here we present the consistency properties of the estimator $\widetilde{\boldsymbol{\theta}}_{\mathcal{I}}$, as the length of the interval $|\mathcal{I}|$ tends to infinity. Note that instead of showing the standard asymptotic results such as convergence in probability or square mean, we provide bounds with exponentially large probabilities. This allows us to consider the growing amount of intervals simultaneously and it is crucial for the sequential interval test in Section 3.

Denote $\lambda_t(\boldsymbol{\theta}) = \mathsf{E}\mathbf{g}_t(\boldsymbol{\theta})$. Recall that one of the main tools in providing convergence and asymptotic normality of $M$-estimators is to derive uniform deviation bounds for the

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0397



score function, see e.g., White (1996); Spokoiny (2017). The following lemma provides a uniform deviation bound with an exponential probability.

**Lemma 1.** *Suppose that Assumptions 3 and 4 hold on an interval $\mathcal{I}$. Then, for any $\mathtt{x} > 0$,*

$$\sup_{\|\boldsymbol{\theta}-\boldsymbol{\theta}^*\|\le r} \frac{1}{|\mathcal{I}|^{1/2}} \left\| \sum_{t\in\mathcal{I}} \mathbf{g}_t(\boldsymbol{\theta}) - \boldsymbol{\lambda}_t(\boldsymbol{\theta}) - \mathbf{g}_t(\boldsymbol{\theta}^*) + \boldsymbol{\lambda}_t(\boldsymbol{\theta}^*) \right\| \le \diamondsuit(|\mathcal{I}|, r, \mathtt{x}),$$

*with probability at least $1 - e^{-\mathtt{x}}$, where*

$$\diamondsuit(T, r, \mathtt{x}) = C_0 \left\{ r\sqrt{\mathtt{x}} + r^{1/2}\sqrt{\mathtt{x} + \log T} + T^{-1/2}(\log T)^2(r\mathtt{x} + \mathtt{x} + \log T) \right\}$$

*and $C_0$ is a positive constant that does not depend on $T, r, \mathtt{x}$, and $r$ is a positive constant.*

**Remark 1.** In the above error term $\diamondsuit(T, r, \mathtt{x})$, the second item with $r^{1/2}$ comes from the fact that $\mathbf{g}_t(\boldsymbol{\theta})$ contains non-differentiable generalized errors, $\psi_\tau(Y_{it} - q_{it}(\boldsymbol{\theta})) = \mathbf{1}[Y_{it} \ge q_{it}(\boldsymbol{\theta})] - \tau$, which are Bernoulli random variables and can not be handled by chaining-type argument. This is different from the case of smooth score, see Spokoiny (2017).

Given the above results we can further bound the score function uniformly over the whole parameter set. This allows us to have the following consistency results. One can find the detailed proof in online appendix.

**Proposition 1.** *Suppose that Assumptions 1–5 hold on the interval $\mathcal{I}$. It holds with probability $\ge 1 - 6e^{-\mathtt{x}}$,*

$$\|\widetilde{\boldsymbol{\theta}}_{\mathcal{I}} - \boldsymbol{\theta}^*\| \le C_1 \sqrt{\frac{\mathtt{x} + \log|\mathcal{I}|}{|\mathcal{I}|}},$$

*where $C_1$ is a positive constant and does not depend on $|\mathcal{I}|$ and $\mathtt{x}$.*

## 2.3 Local quadratic expansion

The next step is to provide the analogous asymptotic normality properties of the estimator $\widetilde{\boldsymbol{\theta}}_{\mathcal{I}}$ by utilizing a local Fisher expansion. The main tool explores a linear approximation of

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0397

the gradient of the likelihood function, which can be achieved by means of Proposition 1, see the detailed proof in the online appendix. The following results can serve as a non-asymptotic adaptation of the classical asymptotic normality by central limit theorem.

**Proposition 2.** *Suppose that Assumptions 1–5 hold on the interval $\mathcal{I}$. Then, for any $0 < \mathtt{x} \leq |\mathcal{I}|$, it holds with probability at least $1 - 3e^{-\mathtt{x}}$,*

$$\begin{aligned} \left\| \sqrt{|\mathcal{I}|} Q(\widetilde{\boldsymbol{\theta}}_{\mathcal{I}} - \boldsymbol{\theta}^*) - \boldsymbol{\xi}_{\mathcal{I}} \right\| \leq & \bar{C} \frac{(\mathtt{x} + \log|\mathcal{I}|)^{3/4}}{|\mathcal{I}|^{1/4}}, \\ \left| L(\widetilde{\boldsymbol{\theta}}_{\mathcal{I}}) - L(\boldsymbol{\theta}^*) - \|\boldsymbol{\xi}_{\mathcal{I}}\|^2/2 \right| \leq & \bar{\bar{C}} \frac{(\mathtt{x} + \log|\mathcal{I}|)^{3/4}}{|\mathcal{I}|^{1/4}}, \end{aligned} \tag{2.8}$$

*where $\boldsymbol{\xi}_{\mathcal{I}} = \frac{1}{\sqrt{|\mathcal{I}|}} \sum_{t \in \mathcal{I}} Q^{-1} \mathbf{g}_t(\boldsymbol{\theta}^*)$, and $\bar{C}, \bar{\bar{C}}$ are some positive constants which do not depend on $|\mathcal{I}|$ and $\mathtt{x}$.*

**Remark 2.** The above inequalities (2.8) serve as a non-asymptotic version of the central limit theorem (CLT) for the estimators as Theorem 2 in WKM. This follows from the fact that the sequence $(Q^{-1}\mathbf{g}_t(\boldsymbol{\theta}^*))_{t \leq T}$ satisfies CLT as a martingale difference sequence, see also Theorem 5.24 in White (2014). Specifically, with probability close to one (for example, by letting $\mathtt{x} = c \log|\mathcal{I}|$ with a large $c > 1$), the distance between $\sqrt{|\mathcal{I}|} Q(\widetilde{\boldsymbol{\theta}}_{\mathcal{I}} - \boldsymbol{\theta}^*)$ and $\boldsymbol{\xi}_{\mathcal{I}}$ can be bounded by a quantity that is $O((\log|\mathcal{I}|)^{3/4}/|\mathcal{I}|^{1/4})$, which can be relatively small by choosing a large interval length $|\mathcal{I}|$.

## 3. Homogeneity testing via local change point detection

The essential element of the LMCR technique is to find an "optimal" interval $\mathcal{I}$ at each time point $t$. Generally there is a trade-off between modelling bias and variability, i.e., within a longer (shorter) sample interval the parameter variability is relatively small (large) while the model bias is potentially large (small). Here we present the established standard on how to find the longest sample interval within a certain acceptable modelling bias, i.e. the homogeneous interval.

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0397



### 3.1 Local change point detection

For a given interval $\mathcal{I} = [a, b] \subseteq [1, T]$ with $a, b \in \{1, 2, \ldots, T\}$ and $a < b$, we test whether there is a structural change in the parametric model (2.1). A natural alternative is that there exists a break point $s \in (a, b)$ such that there are different parameters on sub-interval $A_s = [a, s]$ and sub-interval $B_s = [s + 1, b]$. Hence one tests the null hypothesis

$$\mathbf{H_0}(\mathcal{I}): \quad (Y_{it}, \Psi_t)_{t\in\mathcal{I}} \sim \text{LMCR}(\boldsymbol{\theta}^*_{\mathcal{I}}),\ \boldsymbol{\theta}^*_{\mathcal{I}} \in \Theta_0,$$

against the alternative hypothesis

$$\begin{aligned}\mathbf{H_1}(\mathcal{I}): \quad & (Y_{it}, \Psi_t)_{t\in A_s} \sim \text{LMCR}(\boldsymbol{\theta}^*_{A_s}),\\ & (Y_{it}, \Psi_t)_{t\in B_s} \sim \text{LMCR}(\boldsymbol{\theta}^*_{B_s}) \text{ with } \boldsymbol{\theta}^*_{A_s} \neq \boldsymbol{\theta}^*_{B_s}.\end{aligned}$$

This is a natural idea and has been put forward in numerous papers, e.g. Härdle et al. (2022). In order to construct the test statistics, we consider a set of candidates for a break point $\mathcal{S}(\mathcal{I}) \subset (a, b)$ and for each such candidate $s \in \mathcal{S}(\mathcal{I})$ the test statistics is written as,

$$T_{\mathcal{I},s} = L_{A_{\mathcal{I},s}}(\widetilde{\boldsymbol{\theta}}_{A_{\mathcal{I},s}}) + L_{B_{\mathcal{I},s}}(\widetilde{\boldsymbol{\theta}}_{B_{\mathcal{I},s}}) - L_{\mathcal{I}}(\widetilde{\boldsymbol{\theta}}_{\mathcal{I}}), \tag{3.1}$$

where $A_{\mathcal{I},s} = [a, s]$ represents observations to the left from break point and $B_{\mathcal{I},s} = [s+1, b]$ are the observations to the right from break point candidate $s \in \mathcal{I}$. The existence of break point among the candidates is detected using

$$T_{\mathcal{I}} = \max_{s\in\mathcal{S}(\mathcal{I})} T_{\mathcal{I},s}. \tag{3.2}$$

Given a certain confidence level $\alpha$, we need to construct a critical value $\mathfrak{z}_{\mathcal{I},\alpha}$ such that under the null hypothesis it holds

$$\mathsf{P}\left(T_{\mathcal{I}} > \mathfrak{z}_{\mathcal{I},\alpha}\right) = \alpha, \tag{3.3}$$

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0397



which stands for the false alarm rate. Determining there critical values is a crucial task to be discussed later.

Note that Spokoiny (2009) has initiated this local change point detection approach and guaranteed the theoretical foundations, see also Suvorikova and Spokoiny (2017). The asymptotic distribution of the test statistics $T_{\mathcal{I}}$ in (3.2) will consolidate the theoretical foundations of the judgment of breakpoints. A derivation of asymptotic properties, requiring (unrealistically) increasing test intervals, is challenging but doable. There are two ways to address this issue: First one could rely on Andrews (1993) to describe the limiting distribution where each sub-interval length goes to infinity. Second, one could employ the multiplier bootstrap technique, also used in Jirak (2015); Suvorikova and Spokoiny (2017), where critical values for uniform confidence bands are derived. Both ways would go beyond the scope and message of this paper though. Meanwhile, in our local change point detection, we perceive the change point within a finite sample interval, which implies the sample size can not go to infinity and the candidate locations $\mathcal{I}$ are fixed. Hence, the distribution of $T_{\mathcal{I}}$ in (3.2) is not of standard form and depends on the chosen likelihood and the data, see Spokoiny and Zhilova (2015). In addition, in order to obtain the asymptotic distribution, we need to know in advance the distribution of the change point candidate set $\mathcal{S}(\mathcal{I})$ relative to the sample size (Andrews; 1993).

Previous research use the so-called *propagation approach* to construct critical values (Xu et al.; 2018; Niu et al.; 2017). This technique is based on simulated test statistics under a predetermined data distribution assumption. For instance, the latter paper calculates the critical values via a skewed normal distribution. However, in practice the *true* distribution is unfortunately unknown, hence a predetermined model is possibly misspecified. Rather than relying on a prescribed data distribution assumption, we construct critical values $\mathfrak{z}_{\mathcal{I},\alpha}(\mathcal{Y})$ in a completely data-driven way. With the corresponding data

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0397



interval for testing, we exploit the multiplier bootstrap technique to construct the critical values, which is introduced in the next section.

### 3.2 Multiplier bootstrap

In this section we present more details about the multiplier bootstrap (*MBS*) used to calculate the critical values needed in (3.3). Due to the unavailability of the asymptotic distribution of the breakpoint test statistic $T_{\mathcal{I}}$ in (3.2), we apply *MBS* and introduce the corresponding bootstrap test statistics. Theorem 1 justifies that the distribution of statistics in the bootstrap world can mimic the unknown distribution of the original statistics $T_{\mathcal{I}}$. *MBS* is a very flexible tool to replicate the distribution in a data-driven way, and we can choose the critical values for the test without any knowledge of the asymptotic or pivotal distribution.

Its idea is to simulate the unknown distribution of the original log-likelihood by introducing *MBS* with each item reweighted,

$$L_{\mathcal{I}}^{\circ}(\boldsymbol{\theta}) = \sum_{t \in \mathcal{I}} w_t \ell_t(\boldsymbol{\theta}),$$

where $(w_t)_{t \leq T}$ is a given random sequence of i.i.d. weights independent of the sample. For sake of simplicity, we additionally assume that they have sub-Gaussian tails.

**Assumption 6.** The weights $(w_t)_{t \leq T}$ are independent with $\mathsf{E}(w_t) = 1$ and $\mathsf{Var}(w_t) = 1$. Besides there is a positive constant $C_w$ such that for each $t$ it holds $\mathsf{E} \exp\{(w_t/C_w)^2\} \leq 2$.

Denote the corresponding bootstrap estimator

$$\widetilde{\boldsymbol{\theta}}_{\mathcal{I}}^{\circ} = \arg\max L_{\mathcal{I}}^{\circ}(\boldsymbol{\theta}),$$

while the expectation of bootstrap log-likelihood with respect to the simulated weights is obviously maximized by the original estimator,

$$\widetilde{\boldsymbol{\theta}}_{\mathcal{I}} = \arg\max \mathsf{E}^{\circ} L_{\mathcal{I}}^{\circ}(\boldsymbol{\theta}) = \arg\max L_{\mathcal{I}}(\boldsymbol{\theta}),$$

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0397



where $\mathsf{E}^\circ[\cdot] = \mathsf{E}[\cdot \mid \mathcal{Y}]$ denotes the expectation in the "bootstrap world". Spokoiny and Zhilova (2015) show that with a high probability the distribution of simulated likelihood ratio $L^\circ_{\mathcal{I}}(\widetilde{\boldsymbol{\theta}}^\circ_{\mathcal{I}}) - L^\circ_{\mathcal{I}}(\widetilde{\boldsymbol{\theta}}_{\mathcal{I}})$ in the "bootstrap world" mimics the distribution of original likelihood ratio $L_{\mathcal{I}}(\widetilde{\boldsymbol{\theta}}_{\mathcal{I}}) - L_{\mathcal{I}}(\boldsymbol{\theta}^*)$ up to some errors that decrease with a growing sample.

**Proposition 3.** *Suppose that Assumptions 1–5 and 6 hold on the interval $\mathcal{I}$. Then, there is $T_0 > 0$ such that $T \geq T_0$ and $\mathtt{x} \leq T$ on the probability of at least $1 - e^{-\mathtt{x}}$, it holds with probability at least $1 - e^{-\mathtt{x}}$ conditioned on the data, that*

$$\begin{aligned} \left\| \sqrt{|\mathcal{I}|} Q(\widetilde{\boldsymbol{\theta}}^\circ_{\mathcal{I}} - \widetilde{\boldsymbol{\theta}}_{\mathcal{I}}) - \boldsymbol{\xi}^\circ_{\mathcal{I}} \right\| &\leq C' \frac{(\mathtt{x} + \log T)^{3/4}}{T^{1/4}}, \\ \left| L^\circ_{\mathcal{I}}(\widetilde{\boldsymbol{\theta}}^\circ_{\mathcal{I}}) - L^\circ_{\mathcal{I}}(\widetilde{\boldsymbol{\theta}}_{\mathcal{I}}) - \|\boldsymbol{\xi}^\circ_{\mathcal{I}}\|^2/2 \right| &\leq C'' \frac{(\mathtt{x} + \log T)^{3/4}}{T^{1/4}}, \end{aligned}$$

*where $\boldsymbol{\xi}^\circ_{\mathcal{I}} = \frac{1}{\sqrt{T}} \sum_{t \in \mathcal{I}} w_t Q^{-1} \mathbf{g}_t(\boldsymbol{\theta}^*)$ and $C', C''$ are positive constants which do not depend on $T$ and $\mathtt{x}$.*

The multiplier bootstrap technique has been applied to change point detection in time series, see Suvorikova and Spokoiny (2017); Avanesov and Buzun (2018). The bootstrap test for a local change point $s$ on the interval $\mathcal{I}$ is introduced as,

$$\begin{aligned} T^\circ_{\mathcal{I},s} =& L^\circ_{A_s}(\widetilde{\boldsymbol{\theta}}^\circ_{A_s}) + L^\circ_{B_s}(\widetilde{\boldsymbol{\theta}}^\circ_{B_s}) - \sup\{L^\circ_{A_s}(\boldsymbol{\theta}) + L^\circ_{B_s}(\boldsymbol{\theta} + \widetilde{\boldsymbol{\theta}}_{B_s} - \widetilde{\boldsymbol{\theta}}_{A_s})\}, \qquad (3.4) \\ T^\circ_{\mathcal{I}} =& \max_{s \in \mathcal{S}(\mathcal{I})} T^\circ_{\mathcal{I},s}. \end{aligned}$$

Note that the shift $\widetilde{\boldsymbol{\theta}}_{B_s} - \widetilde{\boldsymbol{\theta}}_{A_s}$ in Equation (3.4) is devoted to compensate the biases of estimators $\widetilde{\boldsymbol{\theta}}^\circ_{A_s}$ and $\widetilde{\boldsymbol{\theta}}^\circ_{B_s}$ in the bootstrap world, which is not required in the original test in Equation (3.1). The bootstrap bias originates from the fact that the resampled datasets are drawn from the empirical distribution of the observed data, which might not be an accurate representation of the true population distribution. This bias exists noticeable for small sample size, or when the sample is skewed, see Hall (1986); Efron and Tibshirani (1994); Abadie and Imbens (2008); Hesterberg (2011). The error correction

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0397



has been throughly discussed as early as in Härdle and Mammen (1993) in which the proposed wild bootstrap is equivalent to MBS here, see also Hall (1994); DiCiccio and Efron (1996); Davison and Hinkley (1997); MacKinnon (2006).

The above test can further be used to simulate the critical values, since its distribution conditioned on the data mimics the distribution of original test $T_{\mathcal{I}}$ in Equation (3.2) with a high probability, as the following theorem states.

**Theorem 1.** *Suppose that Assumptions 2-5 and 6 hold on an interval* $\mathcal{I} \subseteq [1, T]$. *Suppose that the set of break points satisfies for some constants* $\alpha_0 > 0$,

$$\max_{s \in \mathcal{S}(\mathcal{I})} (|A_{\mathcal{I},s}|, |B_{\mathcal{I},s}|) \geq \alpha_0 |\mathcal{I}|. \tag{3.5}$$

*Then, there are positive constants* $C, c > 0$ *that do not depend on* $|\mathcal{I}|$, *such that it holds with probability at least* $1 - 1/|\mathcal{I}|$,

$$\sup_{z \in \mathbb{R}} |\mathsf{P}\left(T_{\mathcal{I}} > z\right) - \mathsf{P}^{\circ}(T^{\circ}_{\mathcal{I}} > z)| \lesssim C|\mathcal{I}|^{-c}.$$

This theorem justifies that the distribution of the bootstrap statistics $T^{\circ}_{\mathcal{I}}$ mimics the unknown distribution of the original statistics $T_{\mathcal{I}}$. Hence one can construct critical values from the bootstrap statistics:

$$z^{\circ}_{\mathcal{I}}(\alpha) = z^{\circ}_{\mathcal{I}}(\alpha; \mathbf{Y}) = \inf\{z \ : \ \mathsf{P}^{\circ}(T^{\circ}_{\mathcal{I}} > z) \leq \alpha\}. \tag{3.6}$$

The resulting values are fully data-dependent and can be estimated via Monte-Carlo simulations with arbitrary precision (see Sections 5 for details). Given the theorem above, we can use these data-dependent critical values for the original test in Equation (3.2) on the same data interval.

**Corollary 1.** *Under the assumptions of Theorem 1, we have*

$$|\mathsf{P}\{T_{\mathcal{I}} > z^{\circ}_{\mathcal{I}}(\alpha)\} - \alpha| \leq C|\mathcal{I}|^{-c},$$

*where some positive constants* $C, c > 0$ *do not depend on the interval length.*

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0397

## 4. Localizing Multivariate CAViaR

A real data time series can not be globally fitted by one single parametric model with constant parameters. In this context we assume that at each time point $t \in \{1, \ldots, T\}$, there exists a historical sample interval $[t-m, t]$, over which the data process follows a constant parametric model, in our case equation (2.1). The essential idea of LMCR is to identify the longest sample interval, i.e. the interval of homogeneity, in which a stable parametric model can be achieved. For the sake of computation simplicity, the interval of homogeneity will be selected among a set of nested interval candidates through a sequential detesting procedure. Finally, the parameter vector at every time point $t$ is estimated using the adaptively selected homogeneous data interval.

### Interval Selection

The practical way of selecting this homogeneous interval is as follows. To alleviate the computational burden, choose $(K+1)$ nested intervals with length $n_k = |\mathcal{I}_k|$, $k = 0, \ldots, K$, i.e., $\mathcal{I}_0 \subset \mathcal{I}_1 \subset \cdots \subset \mathcal{I}_K$. Interval lengths are usually taken to be geometrically increasing with $n_k = \lceil n_0 c^k \rceil$, where a constant $c > 1$ is slightly greater than one, so that in the worst case one only neglects a small proportion of unknown homogeneous intervals. We assume that the initial interval $\mathcal{I}_0$ is small enough, so that the model parameters are constant within this interval.

### Local Change Point Detection Test

Our target is to detect the longest homogeneous interval by conducting a sequential testing procedure based on Section 3. At each time $t$, we examine the homogeneity of the parameter over interval $\mathcal{I}_k$ against the alternative of homogeneity over interval $\mathcal{I}_{k-1}$ for every $k = 1, \ldots, K$ sequentially. By our assumption, $\mathcal{I}_0$ is homogeneous. The resulting

interval of homogeneity would then be the last before the first rejected one. Therefore, for each $k = 1, \ldots, K$ we examine a set of breaking points within $\mathcal{S}_k = \mathcal{I}_k \setminus \mathcal{I}_{k-1}$ outside of the interval that we already tested. The algorithm at step $k$ is visualized in Figure 1.

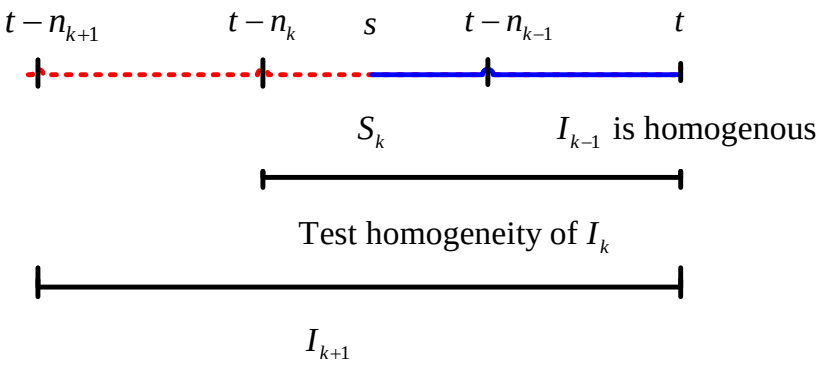


Figure 1: Sequential testing for parameter homogeneity in interval $\mathcal{I}_k$ with length $n_k$ ending at fixed time point $t$.

The hypotheses of the test at step $k$ read as

$H_0$ : parameter homogeneity of $\mathcal{I}_k$ vs $H_1 : \exists$ change point within $\mathcal{S}_k = \mathcal{I}_k \setminus \mathcal{I}_{k-1}$.

The test statistics, i.e. the statistics in equation (3.1), is

$$T_{\mathcal{I}_k,s} = L_{A_{\mathcal{I}_k,s}}(\widetilde{\boldsymbol{\theta}}_{A_{\mathcal{I}_k,s}}) + L_{B_{\mathcal{I}_k,s}}(\widetilde{\boldsymbol{\theta}}_{B_{\mathcal{I}_k,s}}) - L_{\mathcal{I}_{k+1}}(\widetilde{\boldsymbol{\theta}}_{\mathcal{I}_{k+1}}), \tag{4.1}$$

where $A_{\mathcal{I}_k,s} = [t-n_{k+1}, s]$ and $B_{\mathcal{I}_k,s} = [s+1, t]$ are sub-intervals of $\mathcal{I}_{k+1}$. Since the change point position is unknown, we test every point $s \in \mathcal{S}_k = [t - n_k, t - n_{k-1}]$.

According to the homogeneous testing procedure in Section 3, we reject the $k$th interval, if

$$\max_{s \in \mathcal{S}_k} T_{\mathcal{I}_k,s} > \mathfrak{z}^{\circ}_{\mathcal{I}_k}(\alpha), \tag{4.2}$$

where $\mathfrak{z}^{\circ}_{\mathcal{I}_k}(\alpha)$ is generated through multiplier bootstrap in equation (3.6).

Note that if the model is homogeneous on a historical interval $[t - n^*, t]$, then due to Corollary 1 we will accept homogeneity of each interval $\mathcal{I}_k = [t - n_k, t]$ with $n_k \leq n^*$ with

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0397

a high probability. If an interval $\mathcal{I}_k$ remains homogeneous, then the estimator $\widetilde{\boldsymbol{\theta}}_{\mathcal{I}_k}$ has small bias, while the variance decreases with growing number of observations according to Proposition 2. Hence the candidate with least variance corresponds to the largest chosen interval of homogeneity, and the final estimator reads as

$$\widehat{\boldsymbol{\theta}} = \widetilde{\boldsymbol{\theta}}_{\mathcal{I}_{\widehat{k}}}, \qquad \widehat{k} = \max\{k \ : \ \mathcal{I}_k \text{ is not rejected against } \mathcal{I}_{k-1}\}. \tag{4.3}$$

**Critical Values**

The critical value defines the level of significance for the aforementioned test statistic (4.1). In classical hypothesis testing, critical values are selected to ensure a prescribed test level, the probability of rejecting the null hypothesis under null hypothesis is true (type I error). In the considered framework, we similarly control the loss of this 'false alarm' of detecting a non-existing change point. Based on Theorem 1 in section 3.2, we can mimic the distribution of the test statistic (4.1) using the corresponding one with multiplier bootstrap in (3.4). We can use the critical values in the bootstrap world given a significance level for the test statistic on the same data interval by Equation (3.6).

**Summary of LCMR Approach**

Before we numerically analyze the proposed procedure in the next two sections, we summarize the LCMR scheme:

1. Select intervals $\mathcal{I}_k$, $\mathcal{S}_k$, $A_{k,s}$ and $B_{k,s}$, and at each time point $t$ compute the test statistics $T_{\mathcal{I}_k}$ for the step $k = 1, \ldots, K$, see equation (4.1).

2. Testing procedure - select the set of critical values given a tuning parameter $\alpha$ , see section 3.2.

3. Interval of homogeneity is considered as the interval $I_{\widehat{k}}$ for which the null has been

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0397

first rejected at step $\widehat{k}+1$; $\widehat{k} = \max_{k \le K} \{k : T_{\mathcal{I}_\ell} \le \mathfrak{z}^\circ_{\mathcal{I}_\ell}(\alpha), \ell \le k\}$.

4. Adaptive estimation - the adaptively estimated parameter vector at the interval of homogeneity $\widehat{\boldsymbol{\theta}} = \widetilde{\boldsymbol{\theta}}_{\mathcal{I}_{\widehat{k}}}$.

## 5. Simulation

In this section we study the effectiveness of our adaptive approach in detecting the structure breaks in numerical analysis. Following the setup of WKM and the simulation study in Gerlach et al. (2011) and Hong et al. (2009), we generate the data time series using a two-variate GARCH process:

$$\begin{aligned} \sigma_{1t} &= \widetilde{\beta}_{11}\sigma_{1t-1} + \widetilde{\beta}_{12}\sigma_{2t-1} + \widetilde{\gamma}_{11}|y_{1t-1}| + \widetilde{\gamma}_{12}|y_{2t-1}| + \widetilde{c}_1 \\ \sigma_{2t} &= \widetilde{\beta}_{21}\sigma_{1t-1} + \widetilde{\beta}_{22}\sigma_{2t-1} + \widetilde{\gamma}_{21}|y_{1t-1}| + \widetilde{\gamma}_{22}|y_{2t-1}| + \widetilde{c}_2 \\ Y_{it} &= \sigma_{it}\varepsilon_{it}, \qquad \varepsilon_{it} \sim \mathrm{N}(0,1) \text{ i.i.d.} \qquad i = 1,2 \end{aligned} \tag{5.1}$$

Denote the parameter set $\widetilde{\theta} = (\widetilde{\beta}_{ij}, \widetilde{\gamma}_{ij}, \widetilde{c}_i)$ where $i, j = 1, 2$.

Note that at a given quantile level $\tau$, the quantile process $q_{it}(\tau) = \mathrm{Quant}_\tau(Y_{it} \mid \mathcal{F}_{t-1})$ satisfies $q_{it}(\tau) = \Phi^{-1}(\tau)\sigma_{it}$, where $\Phi^{-1}(\tau)$ is the quantile function of the standard normal distribution. Therefore, we have the following equations

$$\begin{aligned} q_{1t}(\tau) &= \beta_{11}q_{1t-1}(\tau) + \beta_{12}q_{2t-1}(\tau) + \gamma_{11}|y_{1t-1}| + \gamma_{12}|y_{2t-1}| + c_1 \\ q_{2t}(\tau) &= \beta_{21}q_{1t-1}(\tau) + \beta_{22}q_{2t-1}(\tau) + \gamma_{21}|y_{1t-1}| + \gamma_{22}|y_{2t-1}| + c_2, \end{aligned} \tag{5.2}$$

where the parameter set $\theta_\tau = (\beta_{ij}, \gamma_{ij}, c_i)$ consists of ten coefficients $\beta_{ij} = \widetilde{\beta}_{ij}$, $\gamma_{ij} = \Phi^{-1}(\tau)\widetilde{\gamma}_{ij}$, and $c_i = \Phi^{-1}(\tau)\widetilde{c}_i$ for $i, j = 1, 2$.

In the simulation we generate the time series $(Y_{it})_{t=1}^{500}$ with the initial variances $\sigma_{i1} = 1$

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0397

and two parameter sets

$$\boldsymbol{\theta}_{left} = (0.5, 0, 0, 0.5, 0, 0.2, 0.2, 0, 0.5, 0.5),$$

$$\boldsymbol{\theta}_{right} = (-0.5, 0, 0, 0.5, 0, 0.2, 0.2, 0, 0.5, 0.5),$$

so that before the break $t \leq s = 250$ the time series satisfy the GARCH process (5.1) with the parameter set $\boldsymbol{\theta}_{left}$ and after the break with the parameter set $\boldsymbol{\theta}_{right}$. For each time point we test a nested sequence of intervals $I_0 \subset I_1 \subset \cdots \subset I_K$ with lengths $n_k = \lceil c^k |I_0| \rceil$. Following the Interval Selection part in Section 4, we take $K = 9$, $|I_0| = 60$ and $c = 1.2$. The considered lengths of intervals are therefore,

$$\{60, 72, 87, 104, 125, 150, 180, 215, 258\}. \tag{5.3}$$

Figure 2 presents the results of the detected homogenous interval length. Figures 3 and 4 show the estimated conditional quantiles $\widehat{q}_{it}$ based on the observations available at a point $t-1$ with the corresponding selected homogeneity intervals. One can observe that the proposed LMCR model favourably forecasts the swift of both time series. Figure 3 displays a natural lag situation, since we assume the smallest homogeneous test internal as given in (5.3) has a length of 60 observations. Hence the forecasting of reaching time for the switch (change point) is a bit lagged. It means even at point t there is a change point, since the smallest homogeneous interval is 60, one can not cut the sample at change point t and naturally there is a bit lag with the forecasting of change point based on the detecting procedure. These results are also consistent with the selected homogeneous intervals in Figure 2.

Note that the optimization problem (2.5) is computationally involved. We deal with a highly non-concave target function that may even have various local maxima. Indeed, the quantile functions (2.1) are polynomials of a multivariate parameter, with the total degree growing up to the number of observations. Notice also that the equation (2.1) is a

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0397

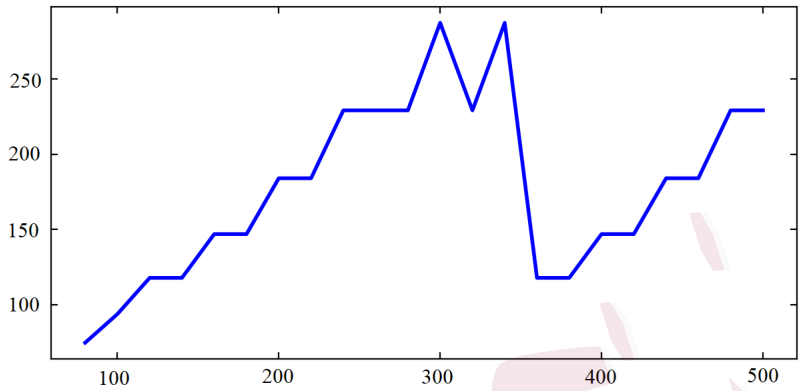


Figure 2: Selected length of homogeneous intervals for timepoints 80 to 500.

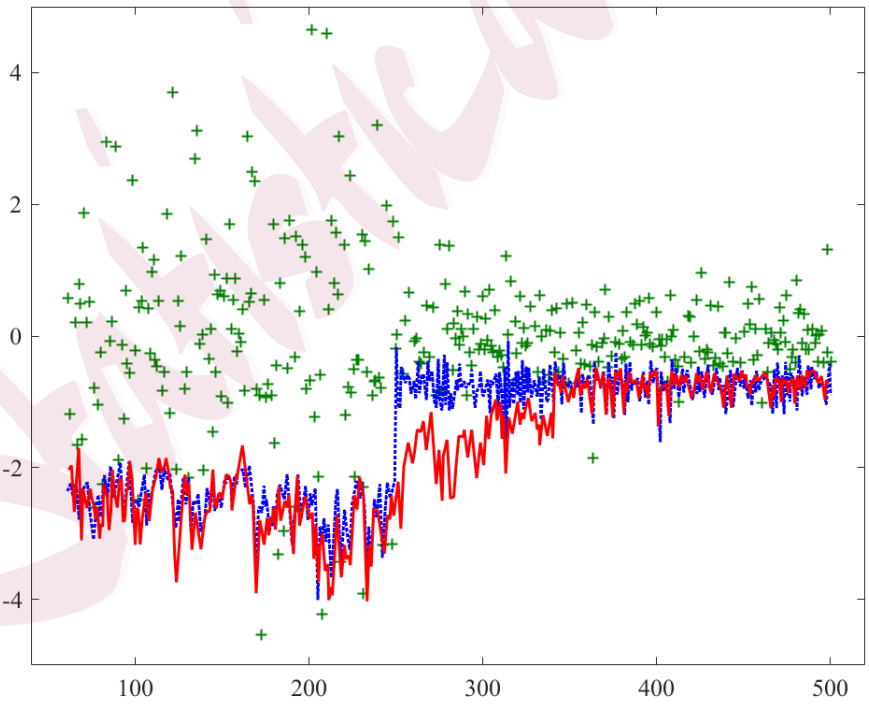


Figure 3: LMCR's predicted quantile one step ahead (red solid line), actual quantile (blue dashed line) and the original simulated time series (green plus points) for $i = 1$ in (5.2).

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0397

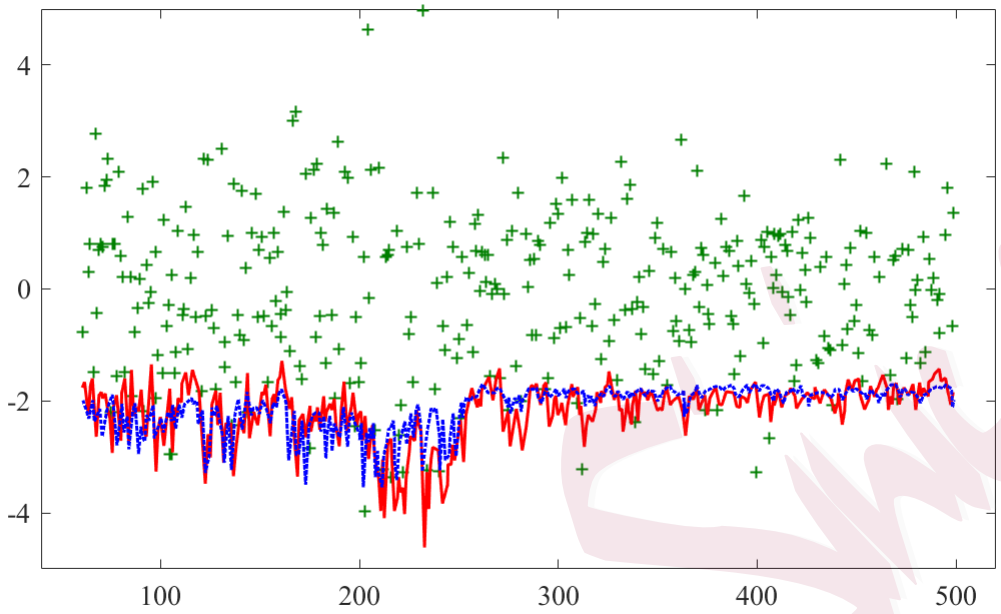


Figure 4: LMCR's predicted quantile one step ahead (red solid line), actual quantile (blue dashed line) and the original simulated time series (green plus points) for $i = 2$ in (5.2).

simple recurrent neural network with a linear activation function and one can use software developed specifically for fitting neural networks. In the simulations, we use python's `Keras` package, which exploits gradient descent, and the procedure is well optimized.

## 6. Application

### 6.1 Data and Parameter Dynamics

We consider two stock markets, namely, the S&P 500 and DAX series. Daily index returns are obtained from Datastream and our data cover the period from 3 January 2005 to 29 December 2017, in total 3390 trading days. The daily returns evolve similarly across the selected markets and all present relatively large variations during the financial crisis period from 2008–2010, see Figure S1 in online appendix S8. Although the return time series exhibit nearly zero-mean with slightly pronounced skewness values, all present

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0397



comparatively high kurtosis, see Table T1 that collects the summary statistics in online appendix S8.

We utilize the bivariate model (5.2) to study the selected (daily) stock market indices. Indeed it is better to study higher time series dimensions. However, more variates lead to more parameters need to be estimated, which tremendously affects the model estimation accuracy and stability in practice. For instance, under the three dimensions data scenario, the estimated parameters in the model alike (5.2) is 21 (one 3 by 3 parameter matrix for the lag quantile items, plus another one 3 by 3 parameter matrix for the lag return items, and 3 intercepts for each time series) with only one lag order for each variables, compared with 10 parameters in bivariate scenario. Hence the parameters to be estimated are more than two times, which will terrifically reduce the estimation efficiency. Particularly, the proposed LMCR model is to detect the longest sample intervals under which the model parameters are constant using a sequential test. Using this finite sub-interval, more parameters underlie larger challenges to obtain the global optimal parameter set in computation. Following the application set in White et al. (2015) with WKM model, which utilize the bivariate model to analyze 230 financial time series through each pair of variables between the market index and each of the financial institutions, the LMCR application still study the bivariate situation with model (5.2).

We firstly consider different interval lengths (e.g., 60 and 500 observations) and analyze the corresponding time-variation estimates via two quantile levels, namely $\tau = 0.01$ and $\tau = 0.05$. Due to page limit, the results are moved to the online appendix, see Figures S2 and S3 in online Appendix S8. One may observe a relatively large variability of the estimated parameters while fitting the model over short data intervals and vice versa.

Key empirical results from the presented fixed rolling window exercise can be summarized as follows: (a) there exists a trade-off between the modeling bias and parameter

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0397



variability across different estimation setups, (b) the characteristics of time series of estimated parameter values as well as the estimation quality results demand the application of an adaptive method that successfully accommodates time-varying parameters, (c) data intervals covering 60 to 500 observations may provide a good balance between the bias and variability. Motivated by these findings, we now turn to LMCR.

We exactly follow the steps as described in Section 4 to implement LMCR in the application. In line with the aforementioned empirical results, we select $(K+1) = 13$ intervals, starting with 60 observations (three months) and ending with 500 observations (two trading years), i.e., we consider the set

$$\{60, 75, 94, 118, 148, 185, 231, 289, 361, 451, 500\}$$

with the coefficient $c = 1.25$ in accordance with the literature. In addition, we assume the model parameters are constant within the initial interval $I_0 = 60$.

Meanwhile, we use the initial two-year time series, i.e. from 3 January 2005 to 30 December 2006, as the training sample to simulate the critical values. We exactly follow the procedure described in Section 3.2 to operate the simulation. We set two cases of the tuning parameter: the conservative case $\alpha = 0.8$ and the modest case $\alpha = 0.9$ to choose the critical values. We present the empirical results in the next section.

## 6.2 Results

LMCR accommodates and reacts to structural changes. From the fixed rolling window exercise in subsection 6.1 one observes time-varying parameter characteristics while facing the trade-off between parameter variability and the modelling bias. How to account for the effects of potential market changes on the tail risk based on the intervals of homogeneity? In the application, we employ LMCR to estimate the tail risk exposure as well as to analyze the cross-sectional spillover effects between the two selected stock

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0397



markets. Using the time series of the adaptively selected interval length, one can trace out the dynamic tail risk spillovers and identify the distinct roles in risk transmissions.

**A. Homogeneous Intervals**

The interval of homogeneity in tail quantile dynamics is obtained here by the LMCR framework for the time series of DAX and S&P 500 returns. Using the sequential local change point detection test, the optimal interval length is considered at two quantile levels, namely, $\tau = 0.01$ and $\tau = 0.05$, see Figures 5 and 6. All figures present the estimated lengths of the interval of homogeneity in trading days using the selected stock market indices from 1 January 2007 to 29 December 2017. The upper panel depicts the conservative risk case $\alpha = 0.8$, whereas the lower panel denotes the modest risk case $\alpha = 0.9$.

In a similar way, the intervals of homogeneity are slightly shorter in the conservative risk case $\alpha = 0.8$, as compared to the modest risk case $\alpha = 0.9$. The average daily selected optimal interval length supports this, see, e.g., Table 1. The results are presented for the selected quantile levels at the conservative and modest risk cases, $\alpha = 0.8$ and $\alpha = 0.9$, respectively. In general the average lengths of selected intervals range between 7-10 months of daily observations across different markets. At quantile levels $\tau = 0.05$, the intervals of homogeneity are slightly larger than the intervals at $\tau = 0.01$.

**B. One-Step-Ahead Forecasts of Tail Risk Exposure**

Based on LMCR, one can directly estimate dynamic tail risk exposure. The tail risk at smaller quantile level is relatively lower than risk at higher levels, see Figure 7. Here the estimated quantile risk exposure for the two stock market indices from 1 January 2007 to 29 December 2017 is displayed for two quantile levels. The left panel represents the

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0397



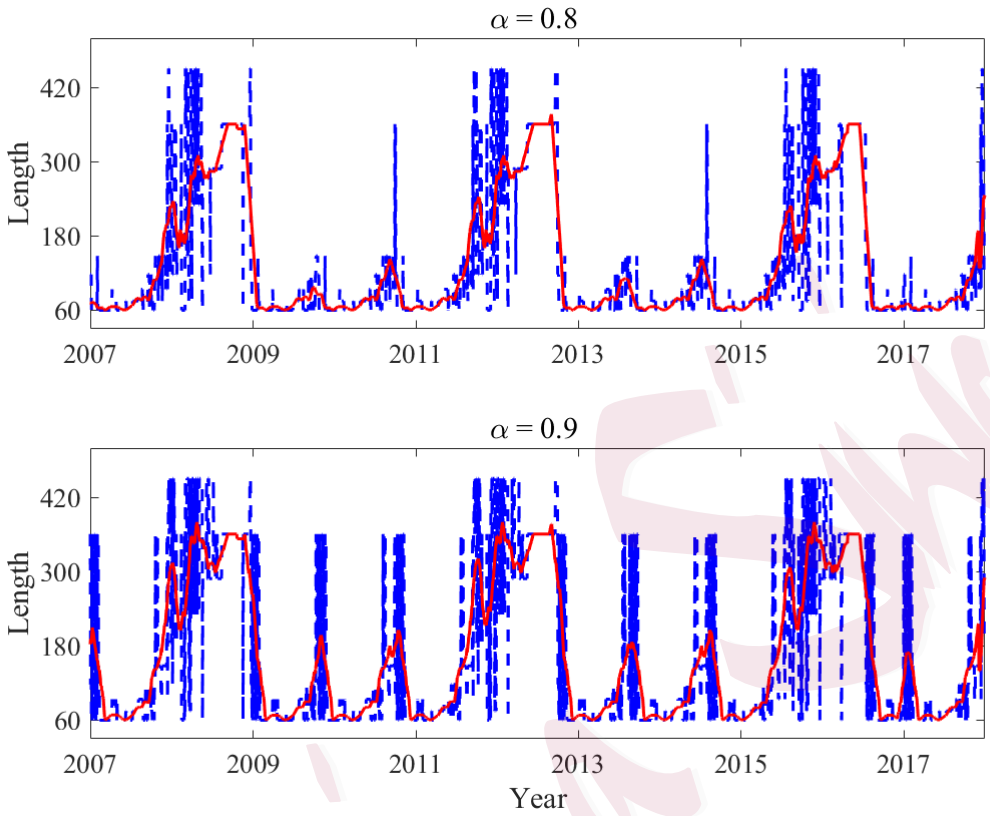


Figure 5: Estimated length of the interval of homogeneity for the selected stock markets for the conservative (upper panel, $\alpha = 0.8$) and the modest (lower panel, $\alpha = 0.9$) risk cases, for quantile level $\tau = 0.01$. The red solid line denotes one-month smoothed values.

| | $\alpha = 0.8$ | $\alpha = 0.9$ |
|---|---|---|
| $\tau = 0.05$ | 159 | 231 |
| $\tau = 0.01$ | 143 | 171 |

Table 1: Mean value of the adaptively selected intervals. Note: the average number of trading days of the adaptive interval length is provided for the DAX and S&P 500 market indices at quantile levels, $\tau = 0.05$ and $\tau = 0.01$, and the conservative ($\alpha = 0.80$) and the modest ($\alpha = 0.90$) risk case.

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0397



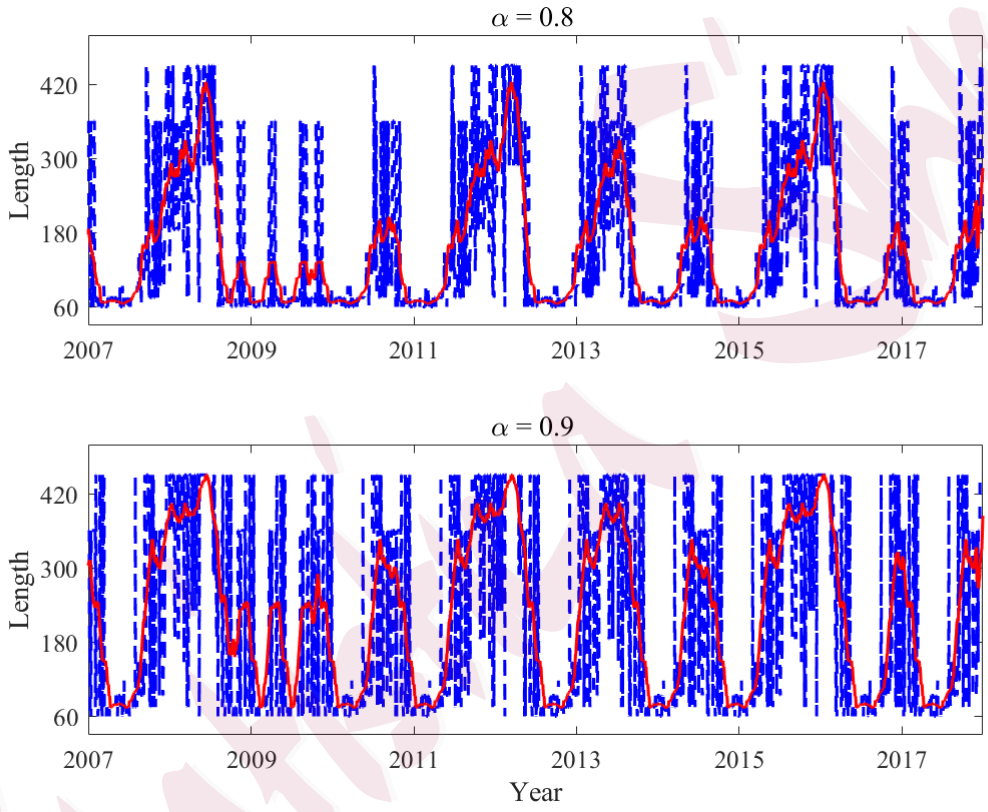


Figure 6: Estimated length of the interval of homogeneity for the selected stock markets for the conservative (upper panel, $\alpha = 0.8$) and the modest (lower panel, $\alpha = 0.9$) risk cases, for quantile level $\tau = 0.05$. The red solid line denotes one-month smoothed values.

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0397



conservative risk case $\alpha = 0.8$ results, whereas the right panel considers the modest risk case $\alpha = 0.9$. The latter leads on average to slightly lower variability, as compared to the conservative risk case which results in marginally shorter homogeneity intervals.

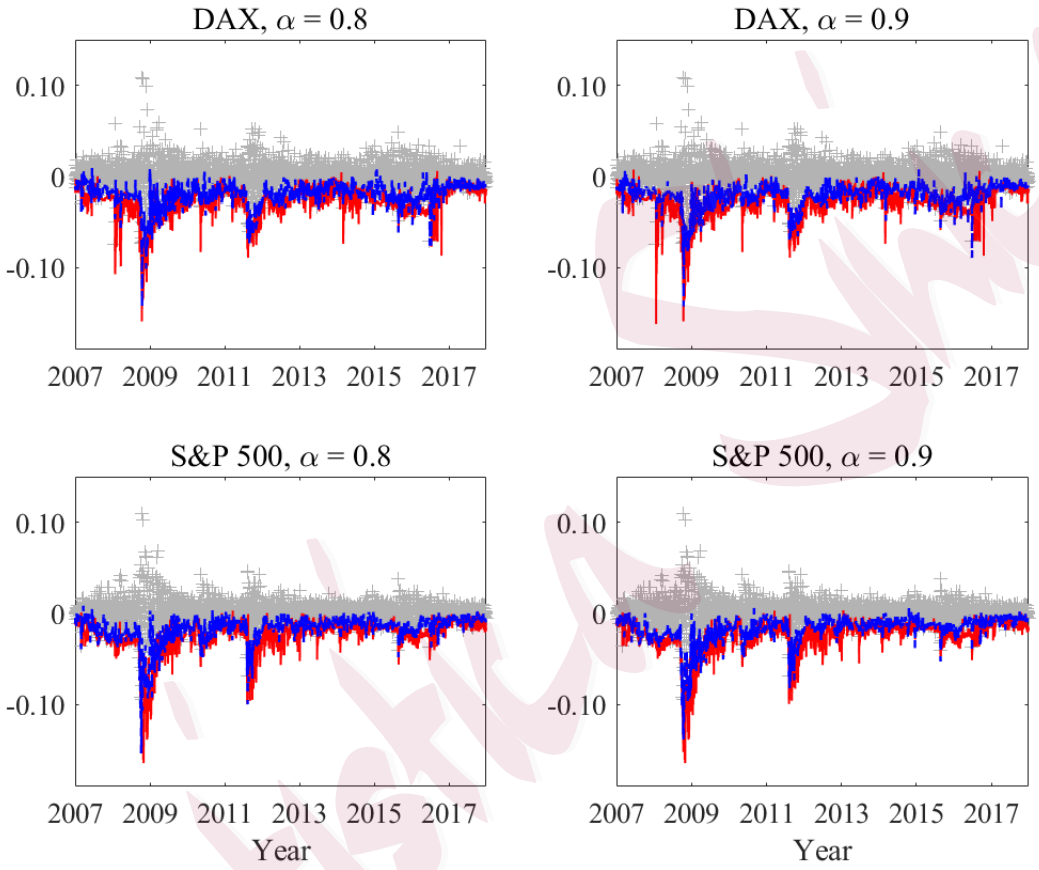


Figure 7: One-step ahead forecasts of quantile risk exposure at level $\tau = 0.05$ (blue dashed line) and $\tau = 0.01$ (red solid line) for return time series of DAX and S&P 500 indices (grey plus points) from 1 January 2007 to 29 December 2017. The left panel shows results of the conservative risk case $\alpha = 0.8$ and the right panel depicts results of the modest risk case $\alpha = 0.9$.

We also provide the out-of-sample forecasting performance of LMCR, see Table 2. We report the results for two quantile levels $\tau = 0.01$ and $\tau = 0.05$, as well as the results for WKM estimation with rolling window 250 observations, listed as WKM rolling. The

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0397



performance is evaluated using the number of VaR exceedances. The return time series is transformed into a time series of indicator functions which take value one if the return exceeds the VaR and zero otherwise. For the estimation at quantile level $\tau = 0.01$ ($\tau = 0.05$), one would expect returns exceeding the VaR 1% (5%) on average. The table reveals that the prediction of LMCR are relatively more precise, in which the values are more closed to the targets 1% (5%) compared to WKM with rolling window estimation, either for conservative risk case $\alpha = 0.8$ or modest risk case $\alpha = 0.9$ for both time series DAX and S&P 500 at the two selected quantile levels.

| | LMCR | | | | WKM rolling | |
|---|---|---|---|---|---|---|
| | $\alpha = 0.8$ | | $\alpha = 0.9$ | | | |
| | DAX | S&P 500 | DAX | S&P 500 | DAX | S&P 500 |
| $\tau = 0.01$ | 1.25% | 1.18% | 1.22% | 1.04% | 1.39% | 1.71% |
| $\tau = 0.05$ | 5.33% | 5.47% | 5.08% | 5.44% | 5.79% | 5.45% |

Table 2: Out-of-sample predictive performance using the number of VaR exceedances. WKM rolling denotes WKM estimation with rolling window 250 observations.

**C. Time-Varying Coefficient Estimates**

The transitions among the financial markets are revealed by the cross-sectional coefficients (Adams et al.; 2014). The dynamics of the two coefficients, $\beta_{12}$ and $\beta_{21}$, are representations of spillover effects. Figures 8 and 9 plot their dynamics from S&P 500 to DAX, $\beta_{12}$ and the ones from DAX to S&P 500, $\beta_{21}$. The upper (lower) panel represent the case of quantile level $\tau = 0.01$ ($\tau = 0.05$). The blue lines show results of the conservative risk case $\alpha = 0.8$ and the red lines depict results of the modest risk case $\alpha = 0.9$.

Moreover, it shows that the cross-sectional coefficient $\beta_{12}$ presents larger and more

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0397

volatile dynamics compared with the coefficient $\beta_{21}$ for both quantile levels $\tau = 0.01$ and $\tau = 0.05$. The shifting of the risk spillovers from US market to German market tend to be more intensive, especially during the unstable market period, e.g. the 2008 financial crisis period and the 2012 European sovereign debt crisis. Hence, compared with the spillovers from DAX to S&P 500, the US market appears to play dominate role in risk transmissions of shocks to DAX indice, especially in volatile time.

Besides, the time series in Figures 8 and 9 at different quantile levels are obviously quite different. The reason is twofold, first the lower quantile level ($\tau = 0.01$) are showing higher variation and second the $\alpha$ sensitivity as described via Figures 6 and 5 yields shorter intervals for the conservative case. This is clearly visible in the upper panel of Figure 8, where in the financial crisis the blue (conservative case) are indicating heavier tail spillovers. The visible difference between 0.01 and 0.05 quantile levels are driven by the tail behaviour of the data and in our opinion can not be attributed to a theoretical feature since the tail behaviour varies over time.

## 7. Conclusion

The cross-sectional tail risk dependence among financial markets is time-varying, and LMCR is constructed to cope with this challenge in evaluating the risk contagion. A local adaptive approach assumes that at any given point of time, there is a historical interval of observations over which the time series follows a parametric model. By utilizing a local change point detection procedure, one can sequentially determine the interval of homogeneity over which the time series behavior can be approximated described by a fixed parameter. LMCR estimates the tail risk transmission by relying on the longest detected interval of homogeneity.

A comprehensive simulation study supports the effectiveness of our approach in detecting structural changes in multivariate tail risk estimation. When setting the quantile

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0397

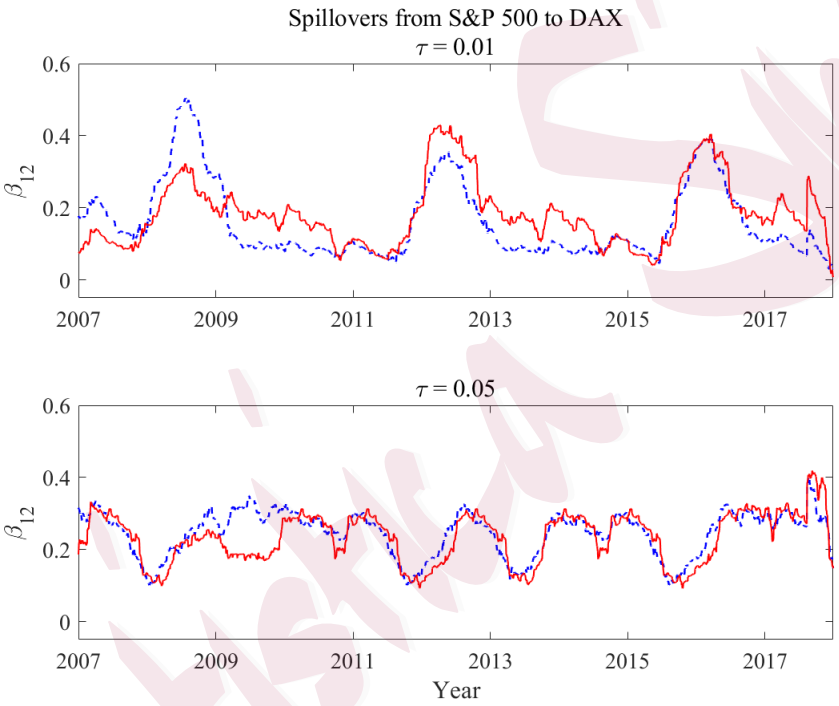


Figure 8: Time-varying coefficients $\beta_{12}$ at quantile level $\tau = 0.01$ (upper panel) and $\tau = 0.05$ (lower panel) between DAX and S&P 500. The blue dashed lines show results of the conservative risk case $\alpha = 0.8$ and the red solid lines depict results of the modest risk case $\alpha = 0.9$.

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0397

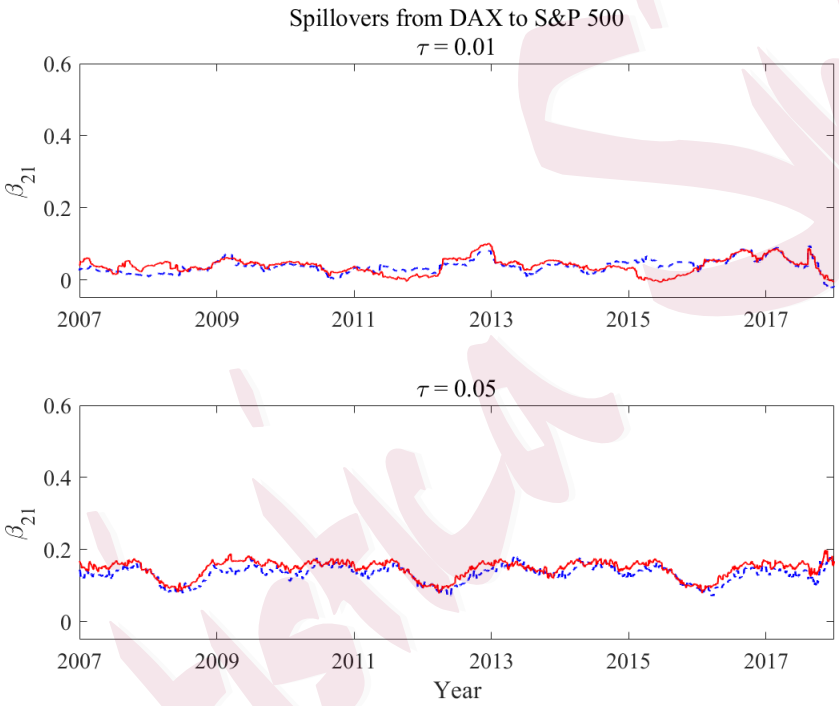


Figure 9: Time-varying coefficients $\beta_{21}$ at quantile level $\tau = 0.01$ (upper panel) and $\tau = 0.05$ (lower panel) between DAX and S&P 500 . The blue dashed lines show results of the conservative risk case $\alpha = 0.8$ and the red solid lines depict results of the modest risk case $\alpha = 0.9$.

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0397



levels at $\tau = 0.05$ and $\tau = 0.01$ in an application of stock market indices DAX and S&P 500, the dynamic tail risk measures are successfully obtained. Besides, the developed approach permits a delineation of the shifting tail risk spillover effects. We find that the US market tends to play a prominent role in risk transmissions of shocks to the German market, especially in volatile times.

## Supplementary Material

The online Supplementary Material includes the appendix for proofs.

## Acknowledgments

Xiu Xu acknowledges the support of Natural Science Foundation of China (NSFC, Grant Numbers: 72273095, 71803140). Li Chen gratefully acknowledges the support of the National Natural Science Foundation of China (NSFC, Grant Numbers: 72103173, 72033008, and 72233002). This paper was supported through "IDA Institute of Digital Assets", CF166/15.11.2022, contract number 760046/23.05.2023, financed under the Romania's National Recovery and Resilience Plan, Apel nr. PNRR-III-C9-2022-I8. We gratefully acknowledge the support of the Marie Sklodowska-Curie Actions under the European Union's Horizon Europe research and innovation program for the Industrial Doctoral Network on Digital Finance, acronym: DIGITAL, Project No. 101119635.

## References

Abadie, A. and Imbens, G. W. (2008). On the failure of the bootstrap for matching estimators, *Econometrica* **76**(6): 1537–1557.

Adams, Z., Füss, R. and Gropp, R. (2014). Spillover effects among financial institu-

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0397



tions: A state-dependent sensitivity value-at-risk approach, *Journal of Financial and Quantitative Analysis* **49**(3): 575–598.

Andrews, D. W. (1993). Tests for parameter instability and structural change with unknown change point, *Econometrica* **61**(4): 821–856.

Avanesov, V. and Buzun, N. (2018). Change-point detection in high-dimensional covariance structure, *Electronic Journal of Statistics* **12**: 3254–3294.

Baele, L. and Inghelbrecht, K. (2010). Time-varying integration, interdependence and contagion, *Journal of International Money and Finance* **29**(5): 791–818.

Bradley, R. C. (2005). Basic Properties of Strong Mixing Conditions. A Survey and Some Open Questions, *Probability Surveys* **2**: 107–144.

Chanda, K. C. (1974). Strong mixing properties of linear stochastic processes, *Journal of Applied Probability* **11**(2): 401–408.

Chen, X., Xiao, Z. and Wang, B. (2022). Copula-based time series with filtered nonstationarity, *Journal of Econometrics* **228**(1): 127–155.

Chen, Y., Härdle, W. K. and Pigorsch, U. (2010). Localized realized volatility modeling, *Journal of the American Statistical Association* **105**(492): 1376–1393.

Chen, Y. and Niu, L. (2014). Adaptive dynamic Nelson–Siegel term structure model with applications, *Journal of Econometrics* **180**(1): 98–115.

Davison, A. C. and Hinkley, D. V. (1997). *Bootstrap methods and their application*, number 1, Cambridge university press.

DiCiccio, T. J. and Efron, B. (1996). Bootstrap confidence intervals, *Statistical Science* **11**(3): 189–228.

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0397



DÍnnocenzo, E., Lucas, A., Schwaab, B. and Zhang, X. (2024). Modeling extreme events: time-varying extreme tail shape, *Journal of Business & Economic Statistics* **42**(3): 903–917.

Efron, B. and Tibshirani, R. J. (1994). *An introduction to the bootstrap*, Chapman and Hall/CRC.

Engle, R. (2002). Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models, *Journal of Business & Economic Statistics* **20**(3): 339–350.

Engle, R. F. and Manganelli, S. (2004). CAViaR: Conditional autoregressive value at risk by regression quantiles, *Journal of Business & Economic Statistics* **22**(4): 367–381.

Gerlach, R. H., Chen, C. W. and Chan, N. Y. (2011). Bayesian time-varying quantile forecasting for value-at-risk in financial markets, *Journal of Business & Economic Statistics* **29**(4): 481–492.

Gong, X., Liu, Y. and Wang, X. (2021). Dynamic volatility spillovers across oil and natural gas futures markets based on a time-varying spillover method, *International Review of Financial Analysis* **76**: 101790.

Hall, P. (1986). On the bootstrap and confidence intervals, *The Annals of Statistics* pp. 1431–1452.

Hall, P. (1994). Methodology and theory for the bootstrap, *Handbook of econometrics* **4**: 2341–2381.

Han, H., Linton, O., Oka, T. and Whang, Y.-J. (2016). The cross-quantilogram: measuring quantile dependence and testing directional predictability between time series, *Journal of Econometrics* **193**(1): 251–270.

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0397



Härdle, W. K., Huang, C. and Khowaja, K. (2022). Uniform inference for generalized random forests, *Available at SSRN: http://dx.doi.org/10.2139/ssrn.4079006* .

Härdle, W. and Mammen, E. (1993). Comparing nonparametric versus parametric regression fits, *The Annals of Statistics* pp. 1926–1947.

Hesterberg, T. (2011). Bootstrap, *Wiley Interdisciplinary Reviews: Computational Statistics* **3**(6): 497–526.

Hong, Y., Liu, Y. and Wang, S. (2009). Granger causality in risk and detection of extreme risk spillover between financial markets, *Journal of Econometrics* **150**(2): 271–287.

Jirak, M. (2015). Uniform change point tests in high dimension, *The Annals of Statistics* **43**(6): 2451–2483.

Koenker, R. and Machado, J. A. (1999). Goodness of fit and related inference processes for quantile regression, *Journal of the American Statistical Association* **94**(448): 1296–1310.

Koenker, R. and Xiao, Z. (2006). Quantile autoregression, *Journal of the American Statistical Association* **101**(475): 980–990.

Lee, H.-T. and Lee, C.-C. (2022). A regime-switching real-time copula garch model for optimal futures hedging, *International Review of Financial Analysis* **84**: 102395. https://doi.org/10.1016/j.irfa.2022.102395.

Li, H., Huang, X. and Guo, L. (2023). Extreme risk dependence and time-varying spillover between crude oil, commodity market and inflation in china, *Energy Economics* **127**: 107090. https://doi.org/10.1016/j.eneco.2023.107090.

MacKinnon, J. G. (2006). Bootstrap methods in econometrics, *Economic Record* **82**: S2–S18.

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0397



Mehlitz, J. S. and Auer, B. R. (2021). Time-varying dynamics of expected shortfall in commodity futures markets, *Journal of Futures Markets* **41**(6): 895–925.

Niu, L., Xu, X. and Chen, Y. (2017). An adaptive approach to forecasting three key macroeconomic variables for transitional china, *Economic Modelling* **66**: 201–213.

Okimoto, T. (2008). New evidence of asymmetric dependence structures in international equity markets, *Journal of Financial and Quantitative Analysis* **43**(3): 787–815.

Shahzad, S. J. H., Naeem, M. A., Peng, Z. and Bouri, E. (2021). Asymmetric volatility spillover among chinese sectors during covid-19, *International Review of Financial Analysis* **75**: 101754. https://doi.org/10.1016/j.irfa.2021.101754.

Spokoiny, V. (1998). Estimation of a function with discontinuities via local polynomial fit with an adaptive window choice, *The Annals of Statistics* **26**(4): 1356–1378.

Spokoiny, V. (2009). Multiscale local change point detection with applications to value-at-risk, *The Annals of Statistics* **37**(3): 1405–1436.

Spokoiny, V. (2017). Penalized maximum likelihood estimation and effective dimension, *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, Vol. 53, Institut Henri Poincaré, pp. 389–429.

Spokoiny, V. and Zhilova, M. (2015). Bootstrap confidence sets under model misspecification, *The Annals of Statistics* **43**(6): 2653–2675.

Suvorikova, A. and Spokoiny, V. (2017). Multiscale change point detection, *Theory of Probability & Its Applications* **61**(4): 665–691.

Tsay, R. S. and Chen, R. (2018). *Nonlinear time series analysis*, Vol. 891, John Wiley & Sons.

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0397



White, H. (1996). *Estimation, inference and specification analysis*, number 22, Cambridge university press.

White, H. (2014). *Asymptotic theory for econometricians*, Academic press.

White, H., Kim, T.-H. and Manganelli, S. (2015). VAR for VaR: Measuring tail dependence using multivariate regression quantiles, *Journal of Econometrics* **187**(1): 169–188.

Xu, X., Mihoci, A. and Härdle, W. K. (2018). lCARE-localizing Conditional Autoregressive Expectiles, *Journal of Empirical Finance* **48**: 198–220.

Zbonáková, L., Li, X. and Härdle, W. K. (2018). Penalized adaptive forecasting with large information sets and structural changes, *Available at SSRN: http://dx.doi.org/10.2139/ssrn.3237444* .