Robust bootstrap inference for linear time-varying coefficient models: Some Monte Carlo evidence

Yicong Lin[1,2]
Mingxuan Song[1]

1 Vrije Universiteit Amsterdam

2 Tinbergen Institute

Robust bootstrap inference for linear time-varying coefficient models: Some Monte Carlo evidence

Yicong Lin[*,1,2] and Mingxuan Song[1]

[1]Vrije Universiteit Amsterdam

[2]Tinbergen Institute

August, 2023

Abstract

We propose two robust bootstrap-based simultaneous inference methods for time series models featuring time-varying coefficients and conduct an extensive simulation study to assess their performance. Our exploration covers a wide range of scenarios, encompassing serially correlated, heteroscedastic, endogenous, nonlinear, and nonstationary error processes. Additionally, we consider situations where the regressors exhibit unit roots, thus delving into a nonlinear cointegration framework. We find that the proposed moving block bootstrap and sieve wild bootstrap methods show superior, robust small sample performance, in terms of empirical coverage and length, compared to the sieve bootstrap introduced by Friedrich and Lin (2022) for stationary models. We then revisit two empirical studies: herding effects in the Chinese new energy market and consumption behaviors in the U.S. Our findings strongly support the presence of herding behaviors before 2016, aligning with earlier studies. However, we diverge from previous research by finding no substantial herding evidence between around 2018 and 2021. In the second example, we find a time-varying cointegrating relationship between consumption and income in the U.S.

Keywords: time-varying models, bootstrap inference, simultaneous confidence bands, energy market, nonlinear cointegration.

JEL Classification: C14, C22, C63, Q56

---

[*]Corresponding author: Department of Econometrics and Data Science, Vrije Universiteit Amsterdam, De Boelelaan 1105, 1081 HV, Amsterdam, the Netherlands. E-mail address: yc.lin@vu.nl.

1   Introduction

A wide range of empirical studies in recent years have often revealed the presence of trending patterns and time-varying relationships (see, for instance, Lee and Shaddick, 2007; Zanin and Marra, 2012; Chang et al., 2016; Silvapulle et al., 2017; Hailemariam et al., 2019; Li et al., 2019; Friedrich et al., 2020; Uddin et al., 2020; Wang et al., 2021; Gao et al., 2021; Awaworyi Churchill et al., 2021; Ren et al., 2022; Anand et al., 2023; Friedrich et al., 2023; Sun et al., 2023). To capture evolving relationships over time, it is common to enable coefficients in regression models to evolve deterministically and smoothly across time. Subsequently, these coefficient curves are estimated nonparametrically (Cai, 2007). Both pointwise confidence intervals and simultaneous confidence bands have been established to quantify the uncertainty of estimation in these models. Pointwise intervals, for example, help determine whether a coefficient curve significantly deviates from zero at a specific time point. However, it is also crucial in empirical studies, as mentioned above, to make assertions about the coefficient curve being significantly different from zero over a particular period, in order to draw stronger conclusions about the overall tendency. Friedrich and Lin (2022) propose a residual-based sieve bootstrap (SB) framework specifically for these purposes. They find that when dealing with weakly stationary regressors and errors, using the SB in small samples offers superior performance compared to the asymptotic approach proposed by Zhou and Wu (2010).

Nevertheless, in scenarios involving more general dynamics, such as error processes with nonlinear dependence structures or unit root regressors, the performance of bootstrap methods remains unknown in the existing literature. This gap in understanding could limit the potential applications of bootstrap methods. This paper aims to address this gap by providing the first set of finite sample evidence that shows bootstrap methods can yield robust and accurate results in more complex contexts. More specifically, based on the SB procedure by Friedrich and Lin (2022), we propose two additional bootstrap schemes tailored for time-varying models. These schemes, moving block bootstrap (MBB) initially proposed by Kunsch (1989), and sieve wild bootstrap (SWB) by Cavaliere and Taylor (2008, 2009) and Smeekes and Taylor (2012), have been previously explored for different purposes in time series. Our findings suggest that the MBB and SWB outperform the SB method in terms of empirical coverage and length of intervals/bands. This advantage is particularly noticeable when dealing with nonstationary, nonlinear error processes, or nonlinear cointegration as considered in previous studies such as Park and Hahn (1999), Phillips et al. (2017), and Li et al. (2020). Comparing the two proposed methods, the MBB displays overall greater robustness than the SWB and could be considered a preferred choice in practical applications.

We then revisit two empirical examples. First, we consider herding effects in the Chinese renewable energy market. Previous research mostly relies on time-constant regressions (see Chang et al., 2000, for

instance), or solely constructs pointwise intervals in time-varying models to detect herding behaviors (see Ren and Lucey, 2023). Since simultaneous bands can be used to examine herding effects for any sub-periods, we further construct simultaneous bands using the proposed bootstrap methods. We find evidence of herding behaviors in one sub-period before 2016, supporting Ren and Lucey (2023). However, our results also suggest that the time-varying herding effects may be insignificant in most of the other sub-periods, thereby diverging from the conclusions drawn by Ren and Lucey (2023). The primary reason for this difference lies in the occurrence of false discoveries when focusing solely on pointwise intervals. This underscores the importance of conducting simultaneous inference in practice. Second, we study the relationship between consumption and income in the U.S. using time-varying cointegrating regression models. Our analysis suggests that these variables have nonlinear cointegration without the presence of a drift component, instead of linear cointegration with a time-varying drift, see the model employed in Kapetanios et al. (2020).

The remainder of this paper is structured as follows. Section 2 introduces the model and nonparametric estimation. Section 3 describes the proposed residual-based bootstrap methods. We also introduce the construction of confidence intervals/bands in this section. In Section 4, we conduct an extensive Monte Carlo study. The empirical applications are provided in Section 5. Section 6 concludes. Additional simulation and empirical results are reported in the Online Appendix.

2 The set-up and nonparametric estimation

We consider the following regression models with time-varying coefficients:

$$
y_t = \beta_{0,t} + \sum_{j=1}^{d} \beta_{j,t} x_{j,t} + u_t = \boldsymbol{\beta}_t' \boldsymbol{x}_t + u_t, \qquad t = 1, \ldots, n, \tag{2.1}
$$

where y_t is a dependent variable, $\boldsymbol{x}_t = (1, x_{1,t}, \ldots, x_{d,t})'$ stacks explanatory variables, and $\{u_t\}$ is an error process that possibly has serial dependence and heteroscedasticity. Moreover, $\boldsymbol{\beta}_t :=$ $(\beta_{0,t}, \beta_{1,t}, \ldots, \beta_{d,t})'$ is a $(d+1)$-dimensional vector of coefficients that varies over time, capturing the time-varying relationships. Model (2.1) has been widely adopted in empirical studies (see, .e.g, Park and Park, 2013; Mikayilov et al., 2018; Kapetanios et al., 2020) due to its interpretability and robustness to model misspecification. In line with earlier studies, for instance, Cai (2007), we assume $\boldsymbol{\beta}_t = \boldsymbol{\beta}(t/n)$, where $\boldsymbol{\beta}(\cdot) = (\beta_0(\cdot), \ldots, \beta_d(\cdot))' : [0,1] \to \mathbb{R}^{d+1}$ is a vector of *unknown* smooth functions. The focus of the current paper is to construct bootstrap-based simultaneous confidence bands and pointwise confidence intervals for $\boldsymbol{\beta}(\cdot)$. This focus is particularly on scenarios where errors exhibit nonlinear temporal dependence, and/or the regressors contain unit roots.

There is an urgent need to gain a better understanding of the performance and implementation

of bootstrap methods in the aforementioned scenarios. This understanding is crucial for ensuring robust inference and catering to the prevalence in empirical studies. When the regressors are strictly stationary, the benefits of employing bootstrap methods over the asymptotic approach (see Zhou and Wu, 2010; Karmakar et al., 2022) have been demonstrated in Friedrich and Lin (2022), when errors are linearly dependent. The performance of bootstrap methods for handling nonlinear, and possibly nonstationary, errors is largely unexplored in the existing literature. Moreover, in cases involving regressors with unit roots, leading to a nonlinear cointegrating framework, Li et al. (2020) illustrate the occurrence of an asymptotic degeneracy problem, leading to a degenerate asymptotic distribution. A remedy for this concern involves a nonstandard transformation that is path-dependent. However, practically implementing such an approach requires addressing two key factors: (1) the need for second-order bias correction in cases of endogeneity, and (2) the presence of prior knowledge concerning the cointegration structure of nonstationary regressors. Moreover, simultaneous inference based on large sample asymptotics has not yet been established in this setting. Bootstrap methods offer a potential avenue to deal with these challenges. To the best of our knowledge, no bootstrap method is currently available for this specific setting.

For inference, we shall first estimate the coefficient curves. We employ the local linear estimation, proposed by Cai (2007), as described in the next section. Alternatively, one can formulate $\boldsymbol{\beta}_t$ as a latent process, written in a state-space representation. Subsequently, one can apply methods such as the Kalman filter (Durbin and Koopman, 2012) or score-driven filters (Creal et al., 2013) to estimate this process.

2.1 The local linear estimation

A brief overview of the estimation is provided here; we refer the interested reader to Cai (2007) for more details. Let $\tau_t = t/n$, then $\beta_{j,t} = \beta_j(\tau_t)$, for $j = 0, \ldots, d$. By a first-order Taylor approximation, for $\tau \in [0, 1]$, we have $\boldsymbol{\beta}_t = \boldsymbol{\beta}(\tau_t) \approx \boldsymbol{\beta}(\tau) + \boldsymbol{\beta}^{(1)}(\tau)(\tau_t - \tau)$, where $\boldsymbol{\beta}^{(1)}(\cdot) = \left(\beta_0^{(1)}(\cdot), \ldots, \beta_d^{(1)}(\cdot)\right)'$, and $\beta_j^{(1)}(\cdot)$, $j = 0, \ldots, d$, is the first-order derivative of $\beta_j(\cdot)$. Model (2.1) can then be approximated by:

$$y_t = \boldsymbol{\beta}_t' \boldsymbol{x}_t + u_t \approx \boldsymbol{\beta}(\tau)' \boldsymbol{x}_t + \boldsymbol{\beta}^{(1)}(\tau)'(\tau_t - \tau)\boldsymbol{x}_t + u_t = \boldsymbol{z}_t(\tau)' \boldsymbol{\theta}(\tau) + u_t, \tag{2.2}$$

where $\boldsymbol{z}_t(\tau) = \left(\boldsymbol{x}_t', (\tau_t - \tau)\boldsymbol{x}_t'\right)'$ and $\boldsymbol{\theta}(\tau) = \left(\boldsymbol{\beta}(\tau)', \boldsymbol{\beta}^{(1)}(\tau)'\right)'$. For each $\tau \in [0, 1]$, the local linear estimator of $\boldsymbol{\theta}(\tau)$ can be obtained by solving the following minimization problem:

$$\hat{\boldsymbol{\theta}}(\tau) = \underset{\boldsymbol{\theta}(\tau)}{\operatorname{argmin}} \sum_{t=1}^{n} \left(y_t - \boldsymbol{z}_t(\tau)' \boldsymbol{\theta}(\tau)\right)^2 K\left(\frac{\tau_t - \tau}{h}\right), \tag{2.3}$$

where $K(\cdot)$ is a kernel function and $h \downarrow 0$ is a bandwidth. A closed-form expression of $\hat{\boldsymbol{\theta}}(\tau)$ exists, given by

$$\hat{\boldsymbol{\theta}}(\tau) = \begin{pmatrix} \hat{\boldsymbol{\beta}}(\tau) \\ \hat{\boldsymbol{\beta}}^{(1)}(\tau) \end{pmatrix} = \left(\boldsymbol{Z}(\tau)' \boldsymbol{K}_h(\tau) \boldsymbol{Z}(\tau) \right)^{-1} \boldsymbol{Z}(\tau)' \boldsymbol{K}_h(\tau) \boldsymbol{y}, \tag{2.4}$$

where $\boldsymbol{Z}(\tau) = \left(\boldsymbol{z}_1(\tau), \ldots, \boldsymbol{z}_n(\tau) \right)'$, $\boldsymbol{y} = (y_1, \ldots, y_n)'$, and $\boldsymbol{K}_h(\tau) = \text{diag} \left[K\left(\frac{\tau_1 - \tau}{h}\right), \ldots, K\left(\frac{\tau_n - \tau}{h}\right) \right]$ is a diagonal matrix.

Indeed, the selections of the kernel function $K(\cdot)$ and bandwidth h need to be made prior to estimation. In the nonparametric literature, it is widely recognized that the choice of the kernel function generally has a minimal impact, unlike the bandwidth selection. For this reason, we employ the common Epanechnikov kernel given by $K(x) = 3/4(1-x^2)\mathbb{1}_{\{|x| \leq 1\}}$ throughout the paper, where $\mathbb{1}_{\{\cdot\}}$ is an indicator function. The selection of the bandwidth involves a trade-off between bias and variance. If the bandwidth is excessively large, it results in oversmoothed estimates and substantial bias. Conversely, an overly small bandwidth leads to higher variance. Furthermore, given our emphasis on inference, it is crucial to note that an "optimal" bandwidth for estimation does not always coincide with the one "optimal" for constructing confidence intervals/bands. These intervals/bands typically necessitate capturing local peaks and troughs. In the subsequent sections, we employ local modified-cross-validation (LMCV) as introduced by Friedrich and Lin (2022), which combines local cross-validation from Vieu (1991) and modified cross-validation from Chu and Marron (1991). It is worth highlighting that a data-driven bandwidth offers practical guidance, but ensuring reasonable outcomes requires visual assessment. If necessary, minor adjustments around the automatically determined bandwidth may be applied in practice, see, for instance, Friedrich et al. (2020).

With the estimation method in place, it is time to discuss the bootstrap methods for constructing confidence intervals/bands of $\boldsymbol{\beta}(\cdot)$ next.[1]

3 Bootstrap inference

Our proposed bootstrap procedures are detailed in Section 3.1. Building upon these procedures, we then outline the steps to obtain confidence intervals/bands in Section 3.2.

3.1 Bootstrap procedures

The sieve bootstrap, proposed in Friedrich and Lin (2022), handles serial correlation in error processes by fitting regression residuals with AR(p) models and subsequently resampling the residuals

---

[1]We provide the Python code for implementing local linear estimation and the proposed bootstrap methods in Section 3 for constructing confidence intervals/bands on the website: https://yiconglin.com/research-2/.

from the estimated AR(p) model. The underlying idea is that any purely deterministic, strictly stationary processes with zero mean, regardless of their linear dependence characteristics, may be well approximated by an AR(∞) process in a similar spirit to Wold's representation (Kreiss et al., 2011). While their simulation results indicate a degree of robustness, even when dealing with nonstationary processes (specifically, unconditional heteroscedasticity), a wild component could be added on top of this to introduce additional uncertainty for capturing heteroscedasticity. This suggestion aligns with their Remark 5 and leads to the method called the sieve wild bootstrap (SWB). Recall $\tau_t = t/n$, $t = 1, \ldots, n$. The procedure can be summarized in six steps as follows.

The SWB

SWB1 Estimate Model (2.1) using a larger bandwidth $\tilde{h}$ than the bandwidth h, i.e., $\tilde{h} > h$, used for the original estimate $\hat{\boldsymbol{\beta}}(\cdot)$. Compute the residuals as $\hat{u}_t = y_t - \tilde{\boldsymbol{\beta}}(\tau_t)'\boldsymbol{x}_t$, $t = 1, \ldots, n$, where $\tilde{\boldsymbol{\beta}}(\tau_t)$ is the resulting estimator obtained using $\tilde{h}$ in estimation.

SWB2 Fit an AR(p) model to the residuals $\{\hat{u}_t\}$, then compute the residuals of AR(p) model by $\hat{\varepsilon}_{t,p} = \hat{u}_t - \sum_{j=1}^{p} \hat{\phi}_j \hat{u}_{t-j}$, $t = p + 1, \ldots, n$. Obtain the recentered residuals by $\tilde{\varepsilon}_{t,p} = \hat{\varepsilon}_{t,p} - (n-p)^{-1} \sum_{t=p+1}^{n} \hat{\varepsilon}_{t,p}$. Let $\tilde{\varepsilon}_{t,p} = 0$ for $t = 1, \ldots, p$.

SWB3 Obtain $\varepsilon_t^* = \nu_t^* \tilde{\varepsilon}_{t,p}$, $t = 1, \ldots, n$, where $\nu_t^* \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, 1)$.

SWB4 Generate the bootstrap innovations $u_t^* = \sum_{j=1}^{p} \hat{\phi}_j u_{t-j}^* + \varepsilon_t^*$ and subsequently the bootstrap observations $y_t^* = \tilde{\boldsymbol{\beta}}(\tau_t)'\boldsymbol{x}_t + u_t^*$, where $\tilde{\boldsymbol{\beta}}(\tau_t)$ is obtained in the step SWB1.

SWB5 With the bootstrap data $\{(y_t^*, \boldsymbol{x}_t)\}$, obtain the bootstrap estimator $\hat{\boldsymbol{\beta}}^*(\cdot)$ using the same bandwidth h as used for $\hat{\boldsymbol{\beta}}(\cdot)$.

SWB6 Repeat SWB3 to SWB5 B times. Let

$$\hat{q}_{j,\alpha}(\tau) = \inf \left\{ u \in \mathbb{R} : \ \mathbb{P}^* \left(\hat{\beta}_j^*(\tau) - \tilde{\beta}_j(\tau) \leq u \right) \geq \alpha \right\}, \qquad j = 0, \ldots, d, \tag{3.1}$$

be the 100αth percentile of the B centered bootstrap statistics $\hat{\beta}_j^*(\tau) - \tilde{\beta}_j(\tau)$, where $\mathbb{P}^*$ denotes for the probability measure conditional on the samples.

A few comments are in order. First, an oversmoothed $\tilde{h}$ is required in SWB1 for bootstrap to capture the second-order bias of the local linear estimator originated from the first-order Taylor approximation (2.2). This is known as a vital technical step when bootstrapping nonparametric regression models; the underlying intuition has been discussed in Friedrich and Lin (2022, Remark 4). In practice, the results are generally not sensitive to the selection of $\tilde{h}$. Second, the lag length $p \equiv p(n)$ in SWB2 can be determined using common information criteria such as AIC and BIC. We

address the selection of $\tilde{h}$ and p in Section 4. Finally, similar to other wild bootstrap procedures, we draw ν_t^* from a standard normal distribution in SWB3. Indeed, other distributions with zero mean and unit variance could also be employed.[2]

Another frequently employed bootstrap approach that can potentially accommodate serial dependence and heteroscedasticity is the moving block bootstrap (MBB), initially proposed by Kunsch (1989). For instance, Brüggemann et al. (2016) have demonstrated the asymptotic validity of the MBB within a VAR framework even when error processes are mixing; linear dependence is generally not required. Such findings suggest that the MBB has the potential for broader application in settings with complex dynamics. As such, we propose the following MBB procedure for inference in our time-varying coefficient models.

More specifically, the procedure begins by stacking residuals into fixed-size, overlapping blocks of consecutive observations, followed by resampling the blocks. In this way, the original (linear/nonlinear) dependence structure can be preserved within each block. Similarly, the specific steps can be summarized in six steps, with the first and the last two steps remaining the same as in the SWB procedure. Therefore, we only highlight the different steps below.

The MBB

MBB2 Divide the residuals into $(n - \ell + 1)$ overlapping blocks, $\{B_i, i = 1, \ldots, n - \ell + 1\}$ with $B_i = (\hat{u}_i, \ldots, \hat{u}_{i+\ell-1})$, where $\ell < n$ is a block length.

MBB3 Let $N = \lceil n/\ell \rceil$, where $\lceil \cdot \rceil$ is the ceilling function. Obtain $B_1^*, \ldots, B_N^*$ by drawing randomly with replacement from the blocks $B_1, \ldots, B_{n-\ell+1}$.

MBB4 Generate the bootstrap innovations $u_1^*, \ldots, u_n^*$ by laying out $B_1^*, \ldots, B_N^*$. Specifically, let $(u_1^*, \ldots, u_n^*, u_{n+1}^*, \ldots, u_{N\ell}^*) \equiv (B_1^*, \ldots, B_N^*)$, then discard the last $N\ell - n$ values $u_{n+1}^*, \ldots, u_{N\ell}^*$. Subsequently, compute the bootstrap observations $y_t^* = \tilde{\boldsymbol{\beta}}(\tau_t)' \boldsymbol{x}_t + u_t^*$.

Note that, under appropriate rate assumptions, the residuals in the first steps of both the SWB and the MBB may be written as $\hat{u}_t = u_t + o_p(1)$, where $o_p(1)$ is uniformly over $[0, 1]$. As previously mentioned, the SWB, like the SB, captures the inherent serial dependence of u_t by fitting an AR(p) model to $\hat{u}_t$, whereas the MBB preserves the dependence structure through blocks of length ℓ. Therefore, the tuning parameter $\ell \equiv \ell(n)$ in MBB2 plays a role similar to the lag length p in the SWB, determining the extent of retained temporal dependence. It is worth noting that one could potentially incorporate a wild component alongside the MBB. Our extensive simulations show that the MBB yields satisfactory empirical coverage, and including a wild component would therefore only increase

---

[2]Alternatively, one could generate the bootstrap innovations as $\varepsilon_t^* = \nu_t^* \hat{\varepsilon}_{t,p}$ without recentering $\hat{\varepsilon}_{t,p}$. The conditional mean of ε_t^* remains due to the mean-zero property of ν_t^*. The results are similar without recentering in our investigation.

the empirical length. During our preliminary investigation, we also implemented the autoregressive wild bootstrap originally proposed by Smeekes and Urban (2014). However, our simulations revealed less satisfactory performance, and consequently, we have chosen to omit the results to conserve space.

3.2 Confidence intervals and bands

Using the bootstrap percentiles $\hat{q}_{j,\alpha}(\cdot)$, $j = 0, \ldots, d$, obtained in Section 3.1, we can immediately construct $(1 - \alpha)$-level pointwise confidence intervals of $\beta_j(\cdot)$:

$$I_{j,\alpha}^*(\tau) = \left[\hat{\beta}_j(\tau) - \hat{q}_{j,1-\alpha/2}(\tau), \ \hat{\beta}_j(\tau) - \hat{q}_{j,\alpha/2}(\tau) \right], \qquad \tau \in (0, 1). \tag{3.2}$$

For any given time point τ, pointwise intervals are designed to ensure that the probability of $\beta_j(\tau) \in I_{j,\alpha}^*(\tau)$ is at least $1 - \alpha$. In numerous applications (see Section 5), it is often crucial to study the evolving relationships over a certain time period, say $G \subset [0, 1]$. In this case, the probability of $\beta_j(\tau) \in I_{j,\alpha}^*(\tau)$ for all $\tau \in G$ is (simultaneously) required to be no less than $1 - \alpha$, leading to the so-called simultaneous confidence bands.

Constructing simultaneous bands using large sample asymptotics presents substantial challenges; it involves various nuisance parameters that are hard to estimate (Zhou and Wu, 2010; Karmakar et al., 2022). Moreover, asymptotic simultaneous bands suffer from slow convergent speeds, often at a log-log rate, requiring a large sample size for accurate coverage. Therefore, we opt for the bootstrap-based procedure in Bühlmann (1998), Friedrich et al. (2020), and Friedrich and Lin (2022), which has demonstrated good performance in small samples. It can be formulated as follows:

Step 1 Let $G = \bigcup_{i=1}^m U_i(h)$, where $U_i(h) = [\tau_i - ah, \tau_i + bh]$ with $\tau_i \in (0, 1)$, $a, b \geq 0$. For each $\alpha_p \in [1/B, \alpha]$ and $\tau \in G$, we compute pointwise quantiles $\hat{q}_{j,1-\alpha_p/2}(\tau)$ and $\hat{q}_{j,\alpha_p/2}(\tau)$, $j = 0, \ldots, d$.

Step 2 Choose a new significance level $\hat{\alpha}_s = \hat{\alpha}_s(\alpha)$ given by

$$\hat{\alpha}_s = \operatorname*{argmin}_{\alpha_p \in [1/B,\alpha]} \left| \mathbb{P}^* \left(\hat{q}_{j,\alpha_p/2}(\tau) \leq \hat{\beta}_j^*(\tau) - \tilde{\beta}_j(\tau) \leq \hat{q}_{j,1-\alpha_p/2}(\tau), \ \forall \tau \in G \right) - (1 - \alpha) \right|. \tag{3.3}$$

Step 3 Given $\hat{\alpha}_s$, construct the simultaneous bands:

$$I_{j,\hat{\alpha}_s}^{G*}(\tau) = \left[\hat{\beta}_j(\tau) - \hat{q}_{j,1-\hat{\alpha}_s/2}(\tau), \ \hat{\beta}_j(\tau) - \hat{q}_{j,\hat{\alpha}_s/2}(\tau) \right], \qquad \tau \in G.$$

We briefly discuss the procedure. Note that Step 2 essentially determines a new level $\hat{\alpha}_s$ such that

$$\frac{\# \left\{ \hat{\beta}_j^*(\tau) - \tilde{\beta}_j(\tau) \in \left[\hat{q}_{j,\hat{\alpha}_s/2}(\tau), \hat{q}_{j,1-\hat{\alpha}_s/2}(\tau) \right], \ \forall \tau \in G \right\}}{B} \approx 1 - \alpha,$$

where $\#E$ counts how many times the event E occurs in the bootstrap samples. It often leads to a level $\hat{\alpha}_s$ that is much smaller than α to ensure that the simultaneous coverage is approximately $1 - \alpha$. It resembles the Bonferroni correction in multiple testing problems. Moreover, the relationship between α_p and coverage is straightforward: an increase of α_p leads to lower coverage, meaning a smaller value of $\mathbb{P}^*\big(\hat{q}_{j,\alpha_p/2}(\tau) \leq \hat{\beta}_j^*(\tau) - \tilde{\beta}_j(\tau) \leq \hat{q}_{j,1-\alpha_p/2}(\tau), \ \forall \tau \in G\big)$. This monotonic relationship offers practical convenience. In the implementation, one can start with $\alpha_p = 1/B$. If the resulting coverage is lower than $1 - \alpha$, the procedure halts with $\hat{\alpha}_s = 1/B$. If it is greater than $1 - \alpha$, one can then proceed with $\alpha_p = 2/B, 3/B, \ldots, \alpha$ until finding the largest coverage smaller than $1 - \alpha$. Let's say α_p^0 is the corresponding level. Then set $\hat{\alpha}_s = \alpha_p^0 - 1/B$.

Establishing the asymptotic bootstrap validity of the SWB and the MBB is a non-trivial task in the current setting, particularly when considering flexible dynamic structures. We leave this for future research. We can nevertheless provide extensive simulation evidence to illustrate their performance, as will be presented in the next section.

4   A simulation study

In this section, we conduct a Monte Carlo study to investigate the finite sample performance of the proposed bootstrap procedures under different scenarios. As a benchmark, we also implement the sieve bootstrap (SB) proposed by Friedrich and Lin (2022) for linearly dependent errors. For performance comparison, we adopt their method of calculating empirical (pointwise and simultaneous) coverage and length of confidence bands, see their p. 12.[3] Given our practical interests, we evaluate the empirical simultaneous coverage over three different subsets of $[0, 1]$: (i) $G_{sub} = U_1(h) \bigcup U_4(h)$, (ii) $G = \bigcup_{i=1}^4 U_i(h)$, and (iii) $\{i/n, \ i = 1, \ldots, n\}$ (labeled as FS, meaning full sample), where $U_i(h) = \{(i/5) - h + j/100, \ j = 0, \ldots, \lfloor 200h \rfloor\} \bigcap [0, 1]$. Several decisions need to be made prior to implementation. Firstly, we set $\tilde{h} = 2h^{5/9}$ for the oversmoothing bandwidth as suggested in Friedrich and Lin (2022). Secondly, the lag length p in SWB2 is determined through AIC, with $p \in \big[0, 10 \log(n)\big]$. Lastly, we opt for the common choice of block length $\ell = 1.75n^{1/3}$ in MBB2, see, e.g., Smeekes and Urbain (2014). All results are reported based on 500 replicates and $B = 1,299$ bootstrap draws.

---

[3](i) Empirical pointwise coverage: For $j = 0, \ldots, d$ and for each Monte Carlo iteration, we calculate the percentage of $\beta_j(\tau_t)$ covered by the bootstrap intervals for $t = 1, \ldots, n$. Then, we calculate the average of these percentages over the total of M iterations. (ii) Empirical simultaneous coverage: For $j = 0, \ldots, d$ and for each Monte Carlo iteration, we determine whether the set $\{\beta_j(\tau), \ \tau \in G\}$ is entirely contained within the confidence bands over G. If this condition is met, it is counted as a success for achieving simultaneous coverage over G. The empirical simultaneous coverage is then calculated as the success rate across all M iterations. (iii) Empirical length: For each Monte Carlo iteration, we compute the median length of intervals/bands across the time grid $\{1/n, 2/n, \ldots, n/n\}$. The average of these medians is then computed over M iterations. For GARCH settings, we use the median of the medians of the lengths instead of the average of the medians to mitigate the impact of outliers.

4.1 Data generating processes

Consider the time-varying coefficient model as adopted in Friedrich and Lin (2022):

$$y_t = \beta_1(t/n)x_{1,t} + \beta_2(t/n)x_{2,t} + u_t, \qquad t = 1, \ldots, n,$$
$$\beta_1(\tau) = 1.5 \exp\big(-10(\tau - 0.2)^2\big) + 1.6 \exp\big(-8(\tau - 0.8)^2\big),$$
$$\beta_2(\tau) = -0.5\tau - 0.5 \exp\big(-5(\tau - 0.8)^2\big),$$

where the function $\beta_1(\cdot)$ has two peaks, making simultaneous coverage relatively challenging, and $\beta_2(\cdot)$ gradually decreases. Let $\boldsymbol{x}_t = (x_{1,t}, x_{2,t})'$ follow a bivariate VAR(1) process:

$$\boldsymbol{x}_t = \boldsymbol{A}\boldsymbol{x}_{t-1} + \boldsymbol{\xi}_t, \qquad \boldsymbol{\xi}_t \overset{\text{i.i.d.}}{\sim} \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_2). \tag{4.1}$$

The eigenvalues of the coefficient matrix $\boldsymbol{A}$ control the degree of serial dependence of $\{\boldsymbol{x}_t\}$. Similar to Chang (2004), we generate $\boldsymbol{A}$ as follows. First, create a 2×2 random matrix $\boldsymbol{U}$ with elements drawn from a standard uniform distribution $U(0,1)$. Second, let $\boldsymbol{A} = \boldsymbol{H}\boldsymbol{L}\boldsymbol{H}'$, where $\boldsymbol{L} = \text{diag}(\lambda_1, \lambda_2)$ and $\boldsymbol{H} = \boldsymbol{U}(\boldsymbol{U}'\boldsymbol{U})^{-1/2}$. We consider three cases: $(\lambda_1, \lambda_2) \in \{(0.3, 0.2), (1, 0.2), (1, 1)\}$. The first case $(\lambda_1, \lambda_2) = (0.3, 0.2)$ represents strictly stationary regressors. The second choice $(\lambda_1, \lambda_2) = (1, 0.2)$ signifies a combination of $I(1)$ and $I(0)$ regressors, while the final selection $(\lambda_1, \lambda_2) = (1, 1)$ yields both $I(1)$ regressors.

We consider five different error processes $\{u_t\}$, with the first one being an autoregressive process, which serves as a benchmark.

AR $\quad u_t = \phi u_{t-1} + \varepsilon_t$, where $\varepsilon_t \overset{\text{i.i.d.}}{\sim} \mathcal{N}\big(0, (1 - \phi^2)/2\big)$. Let $\phi = 0.3$.

ENDO $\quad u_t$ follows the same AR process process above. Nevertheless, ε_t is allowed to be correlated with the innovation $\boldsymbol{\xi}_t$ of $\boldsymbol{x}_t$, which implies endogeneity (labeled as ENDO). More specifically, let $\boldsymbol{\eta}_t := (\boldsymbol{\xi}_t', \varepsilon_t)' \overset{\text{i.i.d.}}{\sim} \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma})$, where $\boldsymbol{\Sigma} = \begin{pmatrix} 1 & \rho & \rho^2 \\ \rho & 1 & \rho \\ \rho^2 & \rho & \sigma_\varepsilon^2 \end{pmatrix}$, $\sigma_\varepsilon^2 = (1 - \phi^2)/2$, $\phi = 0.3$, and $\rho \in \{0.3, 0.5\}$.

GARCH $\quad u_t = \sigma_t \nu_t$, where ν_t follows a standard normal distribution, $\sigma_t^2 = (1 - \alpha - \beta) + \alpha u_{t-1}^2 + \beta \sigma_{t-1}^2$, with $(\alpha, \beta) \in \{(0.2, 0.7), (0.3, 0.6)\}$. The case with $(\alpha, \beta) = (0.2, 0.7)$ is labeled as GARCH1, while the other case is labeled as GARCH2.

NL1 $\quad u_t = 1/4 \sum_{j=0}^{500} a(t/n)^j \zeta_{t-j}$ where $a(\tau) = 1/2 - (\tau - 1/2)^2$ and $\zeta_t \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$.

NL2 $\quad u_t = 1/8 \big[\sum_{j=0}^{500} a(t/n)^j \zeta_{t-j}\big] \big[\sum_{j=0}^{500} c(t/n)^j \epsilon_{t-j}\big]$, where $a(\tau)$ is defined in NL1, $c(\tau) = 1/4 + \tau/2$. Moreover, ζ_t and ϵ_t follow standard normal distributions.

Some remarks can be provided as follows. In the second case (ENDO), we introduce endogeneity

through the parameter ρ. A larger value of ρ indicates a stronger level of endogeneity. Note that in cointegration literature, endogeneity is allowed when $\boldsymbol{x}_t$ are $I(1)$ due to the super-consistency properties. However, this introduces second-order bias, in addition to the bias originating from the Taylor approximation in Section 2.1, resulting in extra challenges when conducting inference (Li et al., 2020). Fully modified estimation is often employed to tackle this bias in asymptotic inference (see, e.g., Lin and Reuvers, 2022a,b). Bootstrap methods may offer an automatic way to mitigate the issue. Moreover, in the third scenario (GARCH), we introduce conditional heteroscedasticity. Lastly, we consider two nonstationary scenarios with highly nonlinear (NL) dependence, namely, NL1 and NL2, similar to the settings in Zhou and Wu (2010). To examine the impact of the dynamics present in $\{u_t\}$, we first consider fix the values of bandwidth $h \in \{0.09, 0.12, 0.15\}$. Subsequently, we assess the influence of bandwidth selection using the LMCV approach as mentioned in Section 2.1.

We provide the complete results in the Online Appendix Section A. To save space, we highlight the key findings here and organize the discussions for each simulation setting as described on p. 10.

4.2 Autoregressive errors (AR)

The empirical coverage of 95%-level confidence intervals/bands in the benchmark settings with AR errors can be found in Table 1. Our observations are summarized as follows.

(a) When the regressors are stationary, matching the settings considered in Friedrich and Lin (2022), it is evident that all bootstrap methods yield values of empirical coverage that are close to the nominal coverage of 95%. As found in Friedrich and Lin (2022), we observe that the pointwise coverage of all methods is relatively insensitive to the choice of bandwidth. However, the selection of bandwidth proves to be crucial for simultaneous coverage, with, for instance, $h = 0.12$ outperforming other values. This observation is expected, as the choice of bandwidth plays a critical role in determining whether the local linear estimates can capture local fluctuations. Consequently, this impacts whether the simultaneous bands can uniformly cover the true curves (across specific subsets of the interval $[0, 1]$, such as G_{sub}).

(b) When there is a unit root in regressors, leading to a nonlinear cointegration regression, the MBB has the overall best empirical coverage, followed by the SWB, even while maintaining a shorter empirical length (Table A2, Online Appendix). The SB mostly performs well except for the full sample simultaneous coverage. The SWB consistently achieves higher empirical coverage than the SB, albeit at the cost of a longer empirical length. This difference arises because the SWB introduces additional randomness through a wild component.

Table 1: Empirical coverage 95%-level confidence intervals and bands for $\beta_1(\cdot)$ and $\beta_2(\cdot)$ in the presence of serially correlated errors (AR). The panel labeled as "$I(0)$ $\boldsymbol{x}_t$" corresponds to stationary regressors with $(\lambda_1, \lambda_2) = (0.3, 0.2)$ for Eq. (4.1), whereas "$I(0)/I(1)$ $\boldsymbol{x}_t$" represents a combination of stationary and unit root regressors with $(\lambda_1, \lambda_2) = (1, 0.2)$.

	h	β_1				β_2			
		PW	G_{sub}	G	FS	PW	G_{sub}	G	FS
					$I(0)$ $\boldsymbol{x}_t$				
	0.09	0.941	0.930	0.920	0.896	0.941	0.936	0.932	0.902
SB	0.12	0.940	0.934	0.898	0.926	0.941	0.940	0.918	0.914
	0.15	0.934	0.910	0.892	0.880	0.940	0.914	0.912	0.890
	0.09	0.946	0.964	0.958	0.944	0.947	0.966	0.958	0.942
SWB	0.12	0.945	0.948	0.954	0.944	0.946	0.972	0.972	0.946
	0.15	0.938	0.930	0.920	0.912	0.945	0.956	0.950	0.942
	0.09	0.941	0.934	0.932	0.918	0.943	0.930	0.932	0.912
MBB	0.12	0.941	0.912	0.892	0.934	0.940	0.940	0.918	0.938
	0.15	0.932	0.910	0.898	0.894	0.941	0.944	0.930	0.918
					$I(0)/I(1)$ $\boldsymbol{x}_t$				
	0.09	0.962	0.968	0.972	0.818	0.962	0.976	0.968	0.818
SB	0.12	0.966	0.962	0.968	0.882	0.968	0.962	0.954	0.870
	0.15	0.964	0.926	0.934	0.906	0.971	0.942	0.938	0.906
	0.09	0.965	0.974	0.982	0.872	0.967	0.990	0.994	0.880
SWB	0.12	0.969	0.980	0.974	0.896	0.971	0.978	0.976	0.908
	0.15	0.968	0.964	0.960	0.930	0.975	0.956	0.954	0.922
	0.09	0.966	0.960	0.962	0.954	0.967	0.958	0.962	0.954
MBB	0.12	0.967	0.968	0.958	0.972	0.970	0.956	0.954	0.964
	0.15	0.960	0.944	0.952	0.952	0.972	0.968	0.960	0.966

4.3 Endogenous errors (ENDO)

As mentioned, when regressors are unit root nonstationary, common least square estimators become super-consistent. Endogenous errors, i.e., errors that are correlated with regressors, are therefore allowed in the cointegration literature. We present the results of $\rho = 0.5$ in Table 2. The case of $\rho = 0.3$ yields slightly better results, see Table A4.

(c) When strong endogeneity is present, the MBB consistently achieves the most accurate coverage compared to the other two methods. The empirical full sample simultaneous coverage of the SB clearly deteriorates in the presence of strong endogeneity, with approximately 88% (β_1) and 85% (β_2) for the full sample, in contrast to the MBB's approximately 94% (β_1) and 90% (β_2). However, it is important to highlight that all methods maintain accurate empirical pointwise coverage.

Table 2: Empirical coverage of confidence bands for $\beta_1(\cdot)$ and $\beta_2(\cdot)$ with endogenous errors (ENDO), where $\rho = 0.5$ (level of endogeneity) and $(\lambda_1, \lambda_2) = (1, 1)$.

	h	β_1				β_2			
		PW	G_{sub}	G	FS	PW	G_{sub}	G	FS
	0.09	0.947	0.932	0.954	0.814	0.918	0.876	0.910	0.746
SB	0.12	0.956	0.928	0.950	0.880	0.934	0.874	0.898	0.838
	0.15	0.954	0.898	0.898	0.882	0.935	0.870	0.876	0.848
	0.09	0.953	0.962	0.980	0.884	0.924	0.912	0.962	0.864
SWB	0.12	0.961	0.964	0.970	0.906	0.939	0.906	0.922	0.888
	0.15	0.959	0.928	0.934	0.890	0.939	0.904	0.912	0.888
	0.09	0.956	0.936	0.958	0.922	0.935	0.890	0.902	0.862
MBB	0.12	0.963	0.940	0.958	0.944	0.946	0.898	0.914	0.900
	0.15	0.957	0.910	0.914	0.928	0.943	0.888	0.890	0.902

Table 3: Empirical coverage of confidence bands for $\beta_1(\cdot)$ and $\beta_2(\cdot)$ with GARCH errors (GARCH1) and $(\lambda_1, \lambda_2) = (1, 0.2)$ (nonlinear cointegration).

	h	β_1				β_2			
		PW	G_{sub}	G	FS	PW	G_{sub}	G	FS
	0.09	0.953	0.936	0.920	0.742	0.947	0.938	0.924	0.734
SB	0.12	0.956	0.928	0.914	0.826	0.955	0.922	0.910	0.800
	0.15	0.954	0.888	0.882	0.834	0.957	0.888	0.884	0.824
	0.09	0.956	0.960	0.964	0.812	0.952	0.952	0.944	0.806
SWB	0.12	0.959	0.944	0.938	0.858	0.959	0.942	0.940	0.832
	0.15	0.959	0.916	0.928	0.890	0.961	0.912	0.910	0.856
	0.09	0.963	0.946	0.938	0.938	0.958	0.932	0.938	0.918
MBB	0.12	0.962	0.940	0.928	0.950	0.961	0.930	0.932	0.950
	0.15	0.955	0.924	0.934	0.932	0.963	0.924	0.928	0.924

4.4 Conditional heteroscedastic errors (GARCH)

We present the results for the GARCH1 setting in Table 3, where $\boldsymbol{x}_t$ includes a unit root, i.e., $(\lambda_1, \lambda_2) = (1, 0.2)$. Similar findings are obtained in other settings, as evidenced in Tables A5 to A8 in the Online Appendix.

(d) As before, the MBB outperforms the other two methods in terms of coverage and has a slightly shorter length than SWB bands in most cases. The lower empirical simultaneous coverage of the SB is expected because the SB is not specifically designed for heteroscedastic data.

Table 4: Empirical coverage and length of 95%-level intervals and bands for $\beta_1(\cdot)$ and $\beta_2(\cdot)$ with nonstationary, nonlinear errors (NL1) and $(\lambda_1, \lambda_2) = (1, 0.2)$ (nonlinear cointegration).

	h	β_1				β_2			
		PW	G_{sub}	G	FS	PW	G_{sub}	G	FS
		Empirical Coverage							
	0.09	0.979	0.990	0.990	0.836	0.982	0.996	0.994	0.822
SB	0.12	0.979	0.976	0.970	0.860	0.986	0.988	0.988	0.856
	0.15	0.971	0.926	0.924	0.890	0.987	0.936	0.942	0.896
	0.09	0.980	0.994	0.990	0.844	0.983	0.998	0.998	0.852
SWB	0.12	0.980	0.988	0.990	0.884	0.987	0.994	0.998	0.878
	0.15	0.974	0.938	0.940	0.904	0.989	0.950	0.948	0.910
	0.09	0.990	0.990	0.992	0.992	0.989	0.990	0.986	0.978
MBB	0.12	0.986	0.980	0.978	0.988	0.989	0.990	0.986	0.992
	0.15	0.971	0.938	0.938	0.952	0.987	0.990	0.984	0.988
		Empirical Length							
	0.09	0.289	0.411	0.441	0.456	0.276	0.390	0.436	0.425
SB	0.12	0.274	0.358	0.398	0.441	0.274	0.376	0.402	0.438
	0.15	0.268	0.386	0.384	0.408	0.271	0.379	0.386	0.411
	0.09	0.302	0.450	0.482	0.533	0.292	0.430	0.474	0.492
SWB	0.12	0.291	0.395	0.447	0.482	0.295	0.420	0.445	0.483
	0.15	0.284	0.428	0.425	0.444	0.283	0.414	0.434	0.455
	0.09	0.271	0.380	0.409	0.418	0.283	0.393	0.435	0.450
MBB	0.12	0.254	0.323	0.360	0.398	0.270	0.372	0.385	0.427
	0.15	0.237	0.343	0.338	0.360	0.254	0.362	0.361	0.385

4.5 Nonlinear, nonstationary errors (NL)

Recall that our particular focus lies in scenarios where there might be nonlinear cointegration and highly nonlinear errors. The outcomes for the NL1 setting, where $\boldsymbol{x}_t$ contains a unit root, are presented in Table A12, and similar observations hold true for the NL2 setting and when $\boldsymbol{x}_t$ is stationary (Tables A9 - A12, Online Appendix).

(e) All methods yield a slightly conservative pointwise coverage. This is due to the super-consistency of local linear estimators in nonlinear cointegration regressions. The empirical pointwise coverage is accurate for all methods with a smaller bandwidth of $h = 0.09$ when the regressors are stationary, see Tables A9 and A11.

(f) When dealing with nonstationary regressors, it is noteworthy that the SB exhibits the lowest empirical simultaneous coverage for the full sample construction, with the SWB following closely behind. On the other hand, the MBB demonstrates over-coverage, but it also has the shortest empirical length. Consequently, the over-coverage observed in the MBB is not a severe concern.

Table 5: Empirical coverage of 95%-level intervals and bands for $\beta_1(\cdot)$ and $\beta_2(\cdot)$ when the bandwidth is selected by LMCV(ℓ), i.e., leaving $(2\ell + 1)$ out with $\ell = 0, 2, 4, 6$, as well as their averages (labeled as AVG), see Friedrich and Lin (2022) for further details. Here, we have NL1 errors and stationary regressors, i.e., $(\lambda_1, \lambda_2) = (0.3, 0.2)$.

			β_1				β_2		
		PW	G_{sub}	G	FS	PW	G_{sub}	G	FS
	LMCV0	0.950	0.950	0.952	0.916	0.957	0.970	0.970	0.908
	LMCV2	0.943	0.952	0.952	0.922	0.958	0.978	0.974	0.932
SWB	LMCV4	0.936	0.946	0.932	0.916	0.958	0.970	0.958	0.920
	LMCV6	0.935	0.928	0.932	0.914	0.958	0.978	0.964	0.938
	AVG	0.948	0.956	0.962	0.932	0.959	0.974	0.962	0.932
	LMCV0	0.949	0.932	0.922	0.902	0.954	0.954	0.942	0.920
	LMCV2	0.938	0.926	0.898	0.896	0.954	0.960	0.938	0.930
MBB	LMCV4	0.932	0.896	0.886	0.886	0.954	0.952	0.920	0.930
	LMCV6	0.929	0.880	0.858	0.894	0.955	0.954	0.922	0.930
	AVG	0.943	0.934	0.918	0.940	0.956	0.952	0.918	0.934

Interestingly, in cases where the regressors are stationary, we note that the MBB displays greater sensitivity to bandwidth selection compared to the other methods (Tables A9 and A11).

Based on our Monte Carlo results, we can now draw several key observations. First, the SWB method exhibits robustness in capturing error processes that feature characteristics like serial correlation, heteroscedasticity, or a combination of both. However, it is worth noting that the SWB method tends to produce confidence intervals/bands with relatively larger empirical lengths. In contrast to the SWB method, the SB method produces bands with shorter empirical lengths. However, it may not perform as accurately when confronted with heteroscedasticity. The MBB method proves to be a versatile choice when dealing with error processes that exhibit complex and flexible structures, such as high nonlinearity and endogeneity. Additionally, when working with $I(1)$ regressors, the MBB method seems to be the superior option, offering good empirical coverage while maintaining relatively shorter lengths.

4.6 Impacts of data-driven bandwidth

In a final investigation, we assess the performance of the bootstrap methods when the bandwidth is selected using an automatic procedure, specifically LMCV as described in Section 2.1. Our earlier results have clearly demonstrated that empirical simultaneous coverage is highly sensitive to the choice of bandwidth for all the methods employed in this study. For our implementation, we adopt the same set of tuning parameters as provided in Section 4.1 of Friedrich and Lin (2022). We consider "leave $(2\ell + 1)$ out", where $\ell \in \{0, 2, 4, 6\}$ and their averages (labeled as "AVG"). The complete set of results is presented in Tables A13 - A15 in the Online Appendix; the results for highly nonlinear NL1

errors with stationary regressors are presented in Table 5.

    (g) Both bootstrap methods perform well when using the AVG across different dynamics of the error process. However, upon closer examination, we observe that the MBB is more sensitive to the tuning parameter ℓ, which determines how many observations to be left out according to Friedrich and Lin (2022), in contrast to the SWB.

Given this observation, the AVG shall be employed in our empirical applications in the next section.

5   Empirical applications

In this section, we revisit two empirical examples using the proposed bootstrap methods. The first example is about the presence of herding behaviors in the Chinese renewable energy market, while the second investigates time variations in consumption behavior.

5.1   Herding effects in Chinese renewable energy market

Herding effects in stock markets pertain to the phenomenon where investors tend to follow the actions and judgments of others rather than making independent decisions based on their own information and analysis. It is widely recognized that the herding phenomenon can lead to stocks being improperly priced, resulting in market inefficiencies. As highlighted in Ren and Lucey (2023), China, being one of the largest producers of various forms of renewable energy and the leading global investor in this sector, possesses the potential to play a significant role in worldwide renewable energy investments. Nonetheless, realizing this potential requires a focus on stabilizing the financial market and improving market efficiency. Thus, identifying potential herding behaviors in the stock market is important.

    To offer insight into the modeling approach for capturing herding effects, we begin by introducing a basic linear regression model as proposed in Chang et al. (2000):

$$\text{CSAD}_{m,t} = \gamma_0 + \gamma_1 R_{m,t} + \gamma_2 |R_{m,t}| + \gamma_3 R_{m,t}^2 + \varepsilon_t, \qquad t = 1, \ldots, n. \tag{5.1}$$

Here, $\text{CSAD}_{m,t} = N^{-1} \sum_{i=1}^{N} |R_{i,t} - R_{m,t}|$ represents the cross-sectional absolute deviation of returns (CSAD) that measures the dispersion; γ_0 is an intercept; $R_{i,t}$ is the stock return of individual stock i at time t; $R_{m,t}$ corresponds to the average market return across N stocks at time t. In a rational stock market, one would expect a linear and positive relationship between dispersion and market returns. On the other hand, during periods of substantial price fluctuations, irrational investors are more likely to display herding behaviors. Consequently, the expected changes in dispersion might appear less pronounced than anticipated or could potentially even go the opposite. This can lead to a

negative and nonlinear relationship between $\text{CSAD}_{m,t}$ and $R_{m,t}$. For this reason, $R^2_{m,t}$ is included in the model. A negative value of the coefficient γ_3, associated $R^2_{m,t}$, could be interpreted as indicative of local herding behavior.

Numerous existing studies indicate that market conditions, inherently subject to temporal variations, exert varying degrees of influence on herding behavior, see Ren and Lucey (2023) and references therein. Hence, the parameters in Model (5.1) might be subject to time variation. To identify time-varying herding effects across different periods, Ren and Lucey (2023) propose the following model:

$$\text{CSAD}_{m,t} = \gamma_0(t/n) + \gamma_1(t/n)R_{m,t} + \gamma_2(t/n)|R_{m,t}| + \gamma_3(t/n)R^2_{m,t} + \gamma_4(t/n)\text{CSAD}_{m,t-1} + \varepsilon_t. \quad (5.2)$$

The inclusion of the first lag of $\text{CSAD}_{m,t}$ aims to reduce the level of serial dependence in the error process ε_t. This adjustment is made since their bootstrap method (a wild bootstrap) lacks the capability to account for serial correlation in errors, which sets it apart from our proposed procedures. For the purpose of direct comparison, we adopt their model (5.2). However, to ensure the robustness of our findings, we also present results in the Online Appendix (Section B.2) by omitting the lagged term. Importantly, the conclusions and results remain qualitatively the same across both cases.

To account for situations where herding effects might be influenced by other markets during periods without within-industry evidence of herding, the next model includes an additional explanatory variable, $R_{int,t}$, representing the return of another market of interest that could affect herding effects:

$$\text{CSAD}_{m,t} = \gamma_0(t/n) + \gamma_1(t/n)R_{m,t} + \gamma_2(t/n)|R_{m,t}| +$$
$$\gamma_3(t/n)R^2_{m,t} + \gamma_4(t/n)R^2_{int,t} + \gamma_5(t/n)\text{CSAD}_{m,t-1} + \varepsilon_t. \quad (5.3)$$

Negative values of $\gamma_4(t/n)$ can be interpreted as an indication of herding influenced by the market of interest when there is no within-industry herding.

The data for $\text{CSAD}_{m,t}$ and $R_{m,t}$ used in this paper were obtained from Ren and Lucey (2023). These values were calculated using China Securities Index Co., Ltd. (CSI) New Energy Index data, covering the period from January 2015 to April 2022, resulting in a total of $n = 1,778$ observations. The CSI New Energy Index comprises a selection of the leading 80 securities involved in different segments of the renewable energy sector in China. This index serves as a benchmark for evaluating the overall performance of the Chinese renewable energy industry. In addition, we use the CSI 300 Index price as $R_{int,t}$. The CSI 300 Index tracks the performance of the top 300 stocks traded on the Shanghai Stock Exchange and the Shenzhen Stock Exchange, serving as a barometer for the overall Chinese stock market. Furthermore, employing the dynamic programming algorithm for estimating
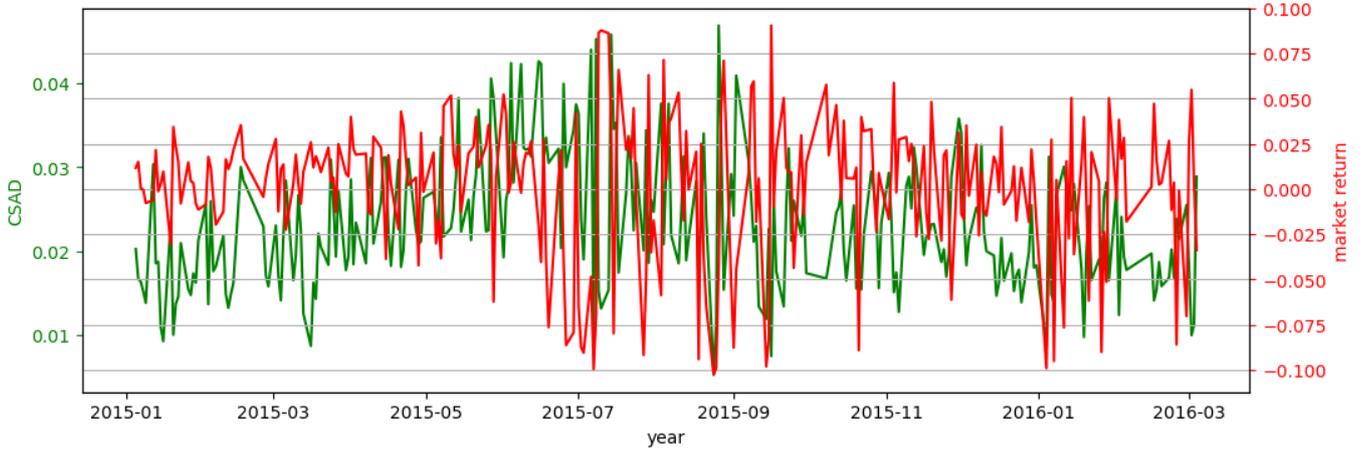
Figure 1: CSAD and the market return in G1.
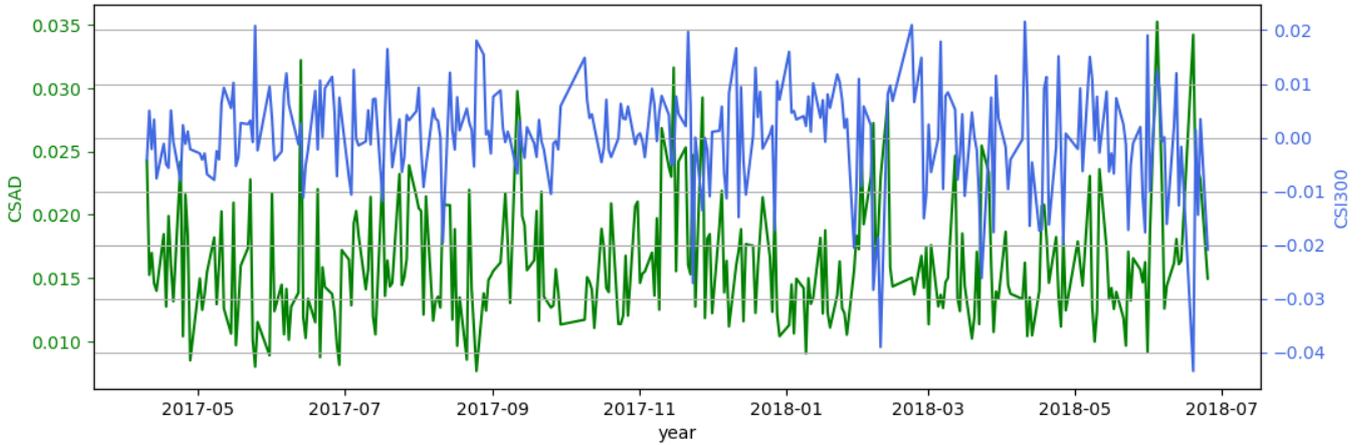
Figure 2: CSAD and CSI300 return in G3.

multiple break dates developed by Bai and Perron (2003), Ren and Lucey (2023) found five break dates within the dataset. The entire time period can be partitioned by these five break dates into six consecutive sub-periods denoted as G1 to G6 in chronological order. These partitions are indicated by the dark-grey vertical lines in Figure 3 below. For instance, the period preceding the first break date (G1) spans until approximately March 2016.

To gain further insights, we plot the CSAD alongside the market return during the G1 period in Figure 1 and with the return of CSI300 during G3 in Figure 2. In Figure 1, the movement of the two variables appears to be in opposite directions during the G1 period. In Figure 2, a negative correlation possibly exists between CSI300 and CSAD in G3. These plots suggest that within-industry herding may exist in G1 while herding influenced by the whole market may be present in G3.

5.1.1 Empirical results of herding behavior

We first discuss the results of Model (5.2). Pointwise intervals and simultaneous bands of $\gamma_3(t/n)$ are constructed using the proposed methods in Section 3. These simultaneous bands, as elaborated upon in Section 3.2, are established using both the full sample and two sub-intervals. To be precise, the full
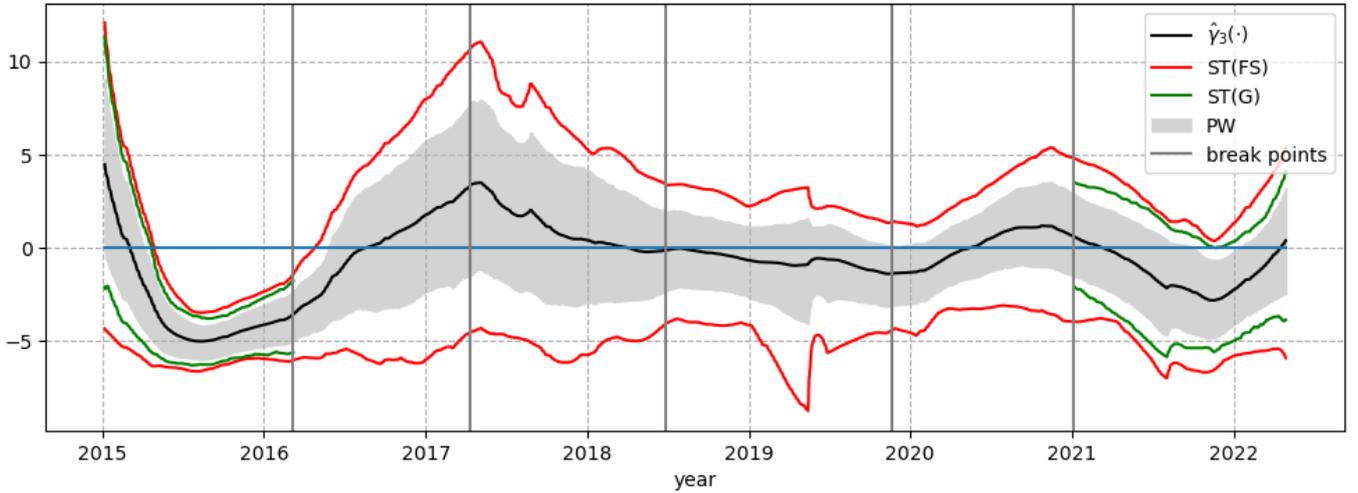
Figure 3: Estimated curve of $\gamma_3(\cdot)$ (black line) and resulting 95%-level confidence intervals (grey, labeled as PW) and bands using the MBB for Model (5.2). Simultaneous bands labeled as ST(FS) are constructed based on the full sample, while ST(G) refers to either sub-interval G1 or G6. The vertical lines in dark grey represent the estimated break dates provided in Ren and Lucey (2023).

sample consists of all $1,778$ observations, while we also focus on two separate sub-intervals: periods G1 (January 2015 to March 2016) and, separately, G6 (starting from May 2021). These sub-intervals are of particular interest as Ren and Lucey (2023) have found significant herding behavior during these periods. However, the evidence so far has been derived from bootstrapping pointwise intervals in time-varying coefficient models, which may not provide statistically valid conclusions about variation over any of the time periods (that obviously contain infinitely many points). This underlines the importance of using simultaneous bands. Foreshadowing our results (e.g., Figures 3 and 5), pointwise intervals can appear significantly narrower compared to simultaneous bands, potentially resulting in false discoveries.

Considering the robust performance of the AVG approach for bandwidth selection demonstrated in the simulations, we employ it to determine the values of the optimal bandwidth. As a result, we obtain bandwidth values of 0.06 and 0.0975 for Models (5.2) and (5.3), respectively. As suggested by various empirical studies such as Friedrich et al. (2020), it is necessary to visually inspect whether data-driven bandwidths lead to reasonable results. Since a bandwidth of 0.06 produces somewhat wiggly estimates and might lead to overfitting, we take 0.0975 for both models, which yields satisfactory results.

Our findings strongly suggest the presence of herding behavior during period G1, while such behavior does not exhibit statistical significance in G6 (Figure 3). More specifically, we observe that the estimated curve $\hat{\gamma}_3(\cdot)$ remains mostly negative during both periods. The 95%-level simultaneous confidence bands using the MBB, derived from both the full sample and the sub-intervals, corroborate the presence of significant herding behavior in G1. This supports Ren and Lucey (2023). The authors argue that the introduction of the Paris Agreement may have stimulated herding behavior in G1, given its direct impact on attracting investors' attention to emerging new energy stocks. However, during
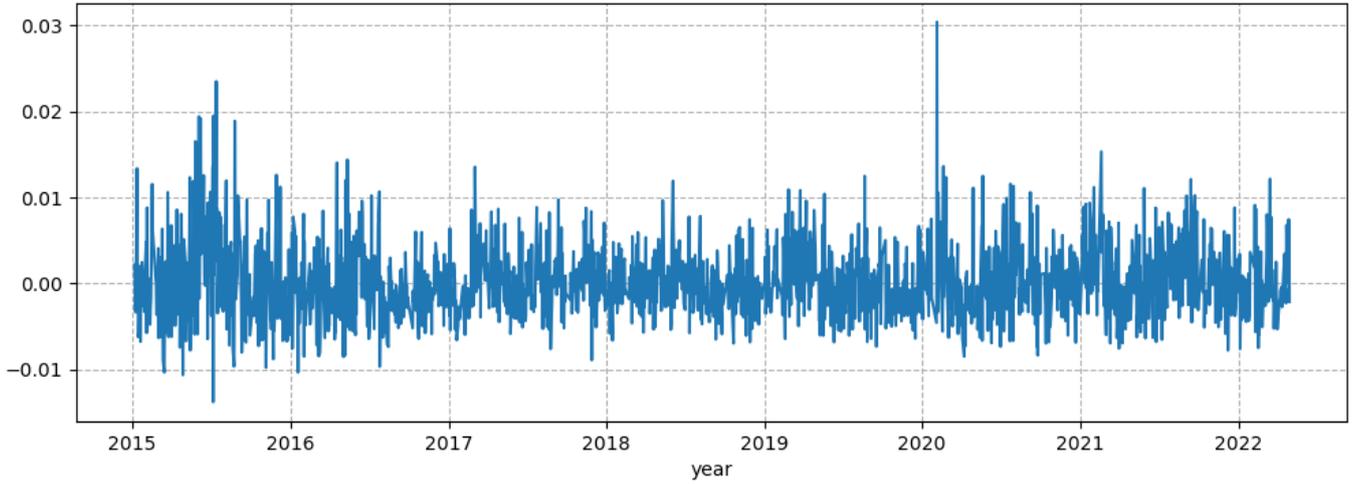
Figure 4: The residuals obtained from Model (5.2) show the existence of both serial correlation and heteroscedasticity.
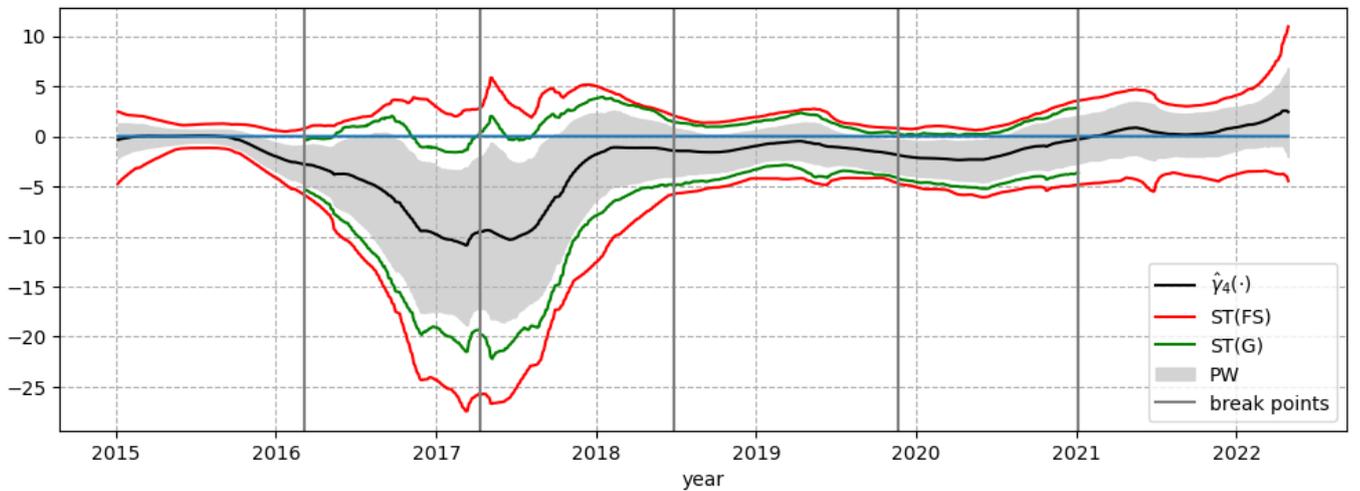
Figure 5: Estimated curve of $\gamma_4(\cdot)$ (black line) and resulting 95%-level confidence intervals (grey, labeled as PW) and bands using the MBB for Model (5.3). Simultaneous bands labeled as ST(FS) are constructed based on the full sample, while ST(G) refers to one of the sub-intervals from G2 to G5. The vertical lines in dark grey represent the estimated break dates provided in Ren and Lucey (2023).

the G6 period, the simultaneous bands contain zero within the bands, indicating that the estimates in G6 are not statistically significant in contrast to the findings in their study. Therefore, herding effects may not be present during this period, possibly due to the impact of COVID-19. During this period, the temporary energy shortage led investors to recognize the importance of renewable energy. This is not surprising, intuitively speaking, since our bootstrap methods show greater robustness to the dynamics of error processes. Despite the inclusion of a lagged term in Models (5.2) and (5.3), there may still exist a level of temporal dependence in errors (Figure 4), potentially resulting in larger type-I errors. Regarding the remaining four sub-periods, simultaneous full sample bands do not provide indications of herding within the industry. These results align with the conclusions presented in Ren and Lucey (2023).

It is worth recalling that Model (5.3) is employed to explore whether herding effects are influenced

by other markets. Therefore, for Model (5.3), we follow the same procedures for $\hat{\gamma}_4(\cdot)$ as in the previous model, but our focus shifts to the sub-periods G2 to G5 due to the absence of significant herding effects observed in the prior model. The estimated curve and the resulting intervals/bands based on the MBB are shown in Figure 5. The simultaneous bands based on one of the sub-intervals from G2 to G5 (green line) reveal significant evidence of herding behavior influenced by other industries during the periods of G2 and G3, in alignment with the findings presented in Ren and Lucey (2023).[4] However, the authors also identify noteworthy evidence around the break date approximately in November 2019. Zooming in, pointwise intervals do suggest significantly negative values at local time points, while the simultaneous bands suggest that this tendency quickly diminishes, leading to overall insignificance. Indeed, examining intervals and bands jointly enables us to draw more powerful conclusions.

In the Online Appendix, specifically in Section B.1, we provide the SWB intervals and bands as part of a robustness check. It is worth noting that the conclusions drawn earlier remain unchanged.

5.2 Case of time-varying consumption behavior

In this section, we examine the classic example of the cointegrating relationship between consumption, income, and interest rates. We adopt a quarterly dataset of aggregate U.S. data obtained from the Federal Reserve Economic Data (FRED). The dataset covers the period from the first quarter of 1959 to the first quarter of 2023, comprising $n = 258$ observations. It includes natural logarithm-transformed personal consumption expenditures ($\log C_t$), disposable personal income ($\log I_t$), also in natural logarithm, and the interest rate (R_t), expressed as a percentage.[5] We plot the data series in Figure 6. It is clear that $\log C_t$ and $\log I_t$ exhibit trending patterns and co-movements. Standard augmented Dickey-Fuller tests imply both variables are trend-nonstationary. Furthermore, R_t is $I(1)$.

Similar to the settings in Park and Hahn (1999) and Li et al. (2020), we begin with the following base model, which incorporates a time-varying cointegrating relationship:

$$\log C_t = \beta_1(t/n) \log I_t + \beta_2(t/n) R_t + \varepsilon_t, \qquad t = 1, \ldots, n. \tag{M1}$$

With a data-driven bandwidth of 0.07, the full sample simultaneous bands indicate substantial time variations in $\beta_1(\cdot)$ during the periods from around 1969 to 1982 (marked by high inflation in the U.S.)

---

[4]In fact, the Chinese stock market experienced a prolonged bubble collapse starting in June 2015, reaching its lowest point around February 2016.

[5]The data is available at:

- Personal consumption expenditures: https://fred.stlouisfed.org/series/PCE;
- Disposable personal income: https://fred.stlouisfed.org/series/DSPI;
- Interest rates (federal funds effective rate): https://fred.stlouisfed.org/series/FEDFUNDS.
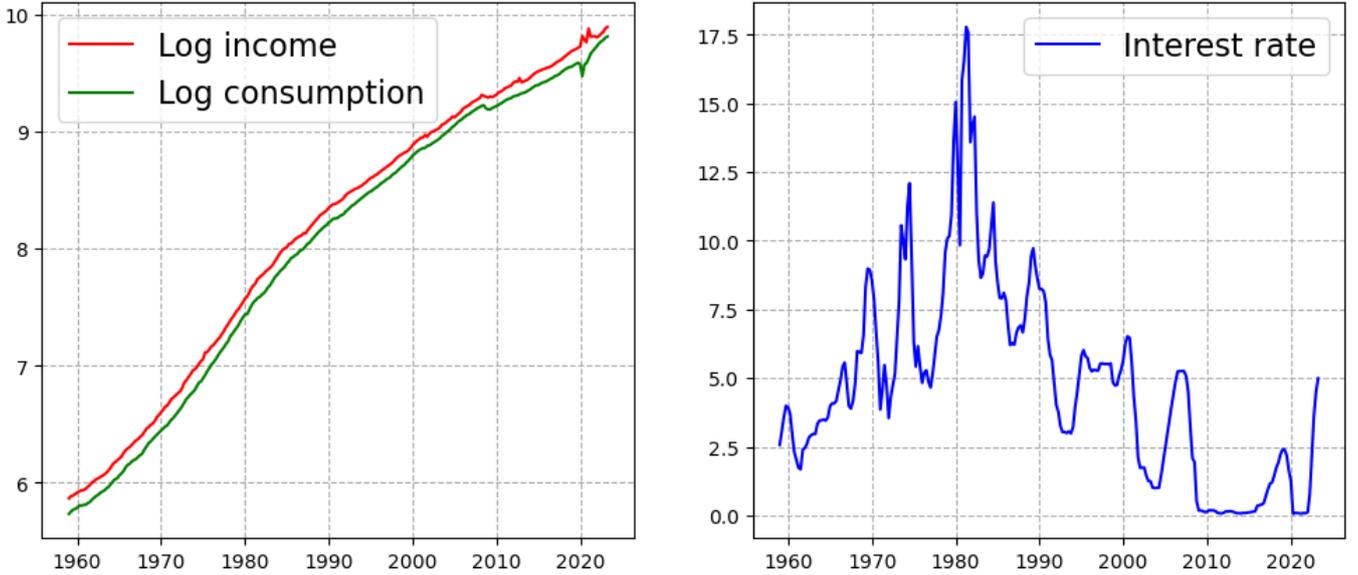
**Figure 6:** Aggregate US data for log consumption, log income, and interest rate from 1959 to 2023.
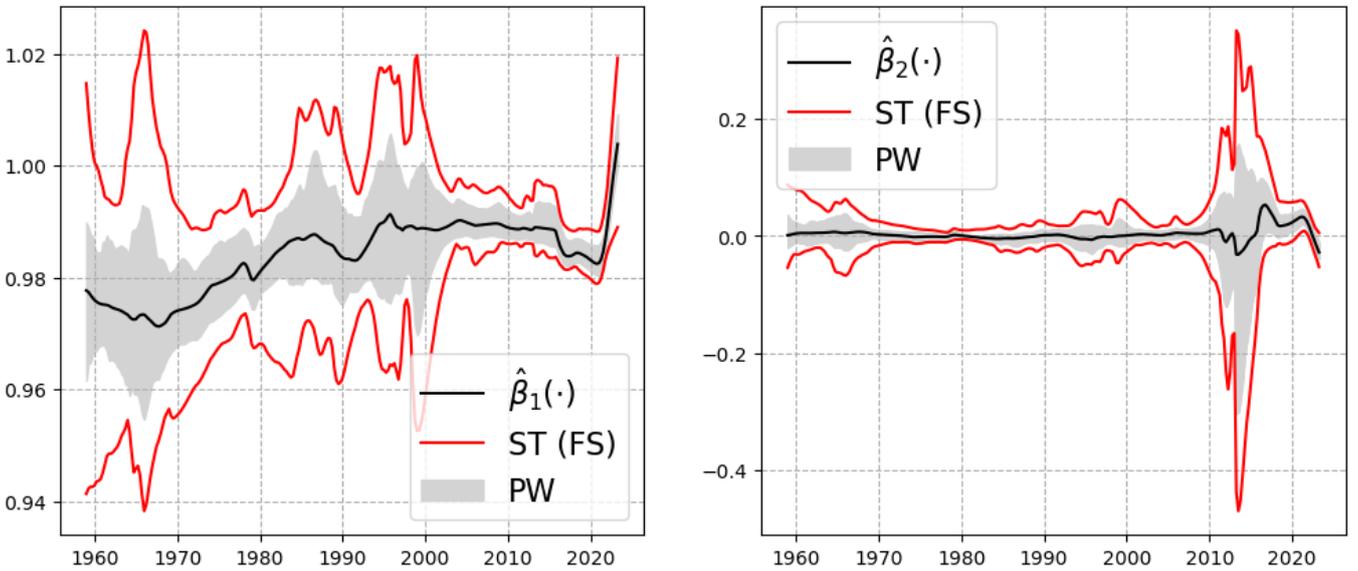
Figure 7: Estimated curves (black line) and resulting 95%-level confidence intervals (grey, labeled as PW) and bands (red) using the MBB for Model (M1).

and from the end of 2001 (the 9/11 attack) onwards, as depicted in Figure 7. Furthermore, there has been a substantial surge since the onset of the COVID-19 pandemic, and this tendency shows a growth pattern potentially exceeding 1. Conversely, we observe that $\beta_2(\cdot)$ is not significantly different from zero across the sample periods.

Considering the absence of evidence that $\beta_2(\cdot)$ significantly deviates from zero, we are led to further explore the following two models:

$$\log\big(C_t/I_t\big) = \beta_0(t/n) + \varepsilon_t,, \tag{M2}$$

$$\log C_t = \beta_1(t/n)\log I_t + \varepsilon_t. \tag{M3}$$
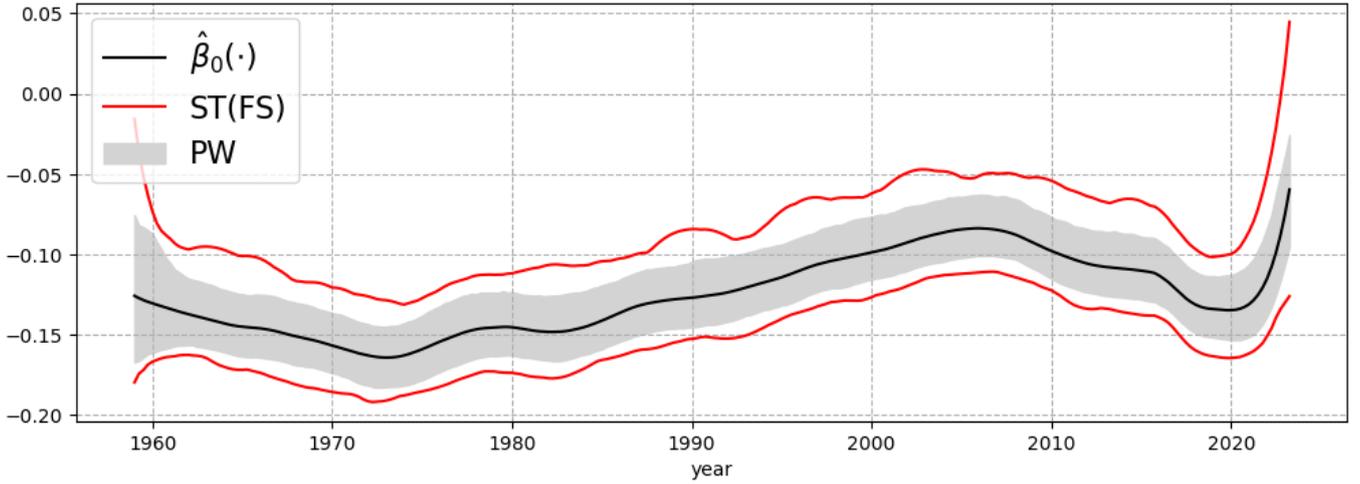
Figure 8: Estimated curve of $\beta_0(\cdot)$ (black line) and resulting 95%-level confidence intervals (grey, labeled as PW) and bands (red) using the MBB for Model (M2).

We first discuss Model (M2). In this model, we use $\beta_0(\cdot)$ to account for potential left-over dynamics when subtracting $\log I_t$ from both sides of Model (M1). Given that we do not impose any specific functional forms on $\beta_0(\cdot)$, this model offers greater flexibility in comparison to common trend break models. Traditional trend break models usually permit only a single break, primarily due to the difficulties associated with estimating and inferring multiple trend breaks (see, e.g., Perron and Zhu, 2005; Harvey and Leybourne, 2014, 2015; Beutner et al., 2023). Moreover, Model (M2) has been employed in Kapetanios et al. (2020) to investigate the so-called "great ratios", which propose that a set of ratios, such as consumption-output and income-output, either remain constant or converge towards a stable value, suggesting a long-run equilibrium relationship. Previous studies examining "great ratios" report mixed evidence regarding the existence of such ratios (see, e.g., Sarantis and Stewart, 1999; Attfield and Temple, 2010; Chudik et al., 2023). Interestingly, Kapetanios et al. (2020) find strong evidence of cointegration by incorporating a time-varying drift. The LHS of Model (M2), which represents the consumption-income ratio, also belongs to this set of great ratios. Displayed in Figure 8, our estimated curve of $\beta_0(\cdot)$ in Model (M2) using the bandwidth of 0.07 aligns with the results presented in Kapetanios et al. (2020). Specifically, it reveals that the time-varying drift is indeed statistically significant across the entire sample period.

On the other hand, Model (M3) is a natural consequence of Model (M1). Note that Model (M3) can be equivalently written as:

$$\log\left(C_t/I_t\right) = \left(\beta_1(t/n) - 1\right)\log I_t + \varepsilon_t.$$

Given that $\log I_t$ is dominated by a mildly nonlinear, deterministic trend (Figure 6), Model (M3) may be well approximated by (M2). In other words, we can write $\left(\beta_1(t/n) - 1\right)\log I_t \approx \beta_0(t/n)$. To verify this conjecture, we show plots of the residuals and fits for both models in Figure 9. Indeed, the
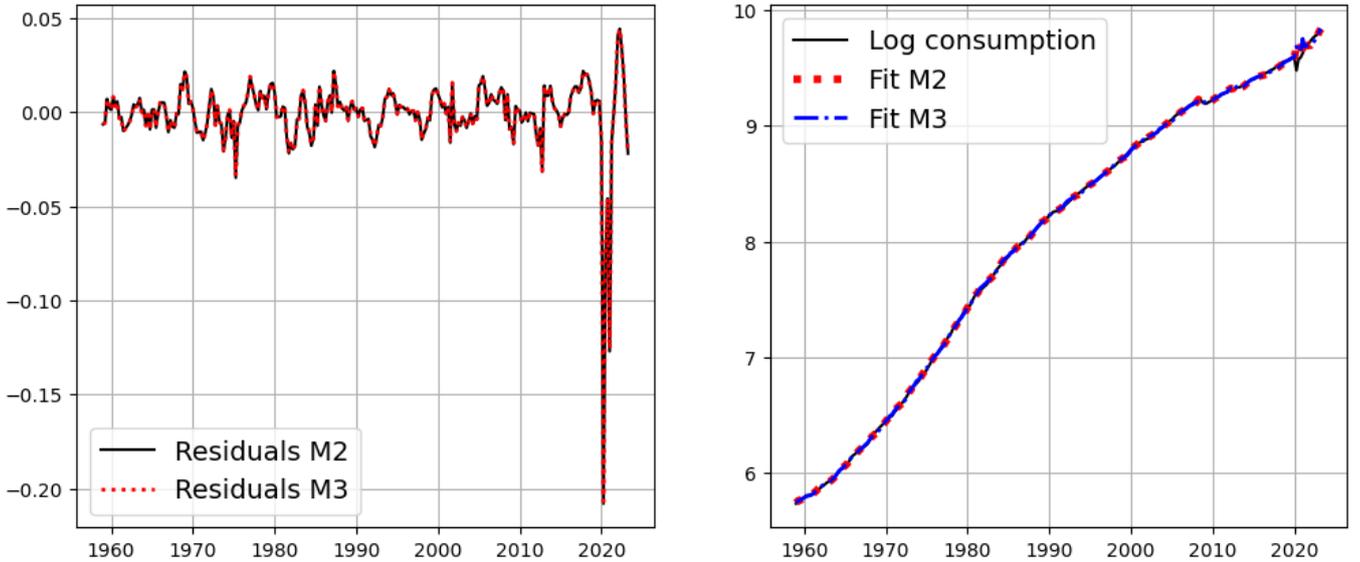
Figure 9: Residuals and fits for Models (M2) and (M3).

residuals and fits display an expected similarity. Additionally, it is worth noting that the residuals exhibit serial correlation but remain mostly stationary, with the exception of the recent sudden shock. This suggests that log consumption and log income are potentially nonlinearly cointegrated without a drift component, instead of linearly cointegrated with a time-varying drift.

6   Conclusions

In this paper, we have studied linear time series models with time-varying parameters. We proposed two bootstrap methods, namely the moving block bootstrap (MBB) and the sieve wild bootstrap (SWB), for conducting robust inference. Our extensive simulation study encompassed scenarios involving serial correlation, heteroscedasticity, endogeneity, nonlinear dependence, nonstationarity in error processes, as well as time-varying cointegrating relationships. The results of our study indicate that both proposed methods outperform the previously suggested sieve bootstrap inference method by Friedrich and Lin (2022) for stationary models. Specifically, the MBB exhibits overall accurate empirical coverage while maintaining a shorter empirical length in comparison to the SWB. Our results provide initial insights into the potential extensions and performance of bootstrap methods for these widely adopted models.

We then applied the proposed methods in two empirical studies. In the first study, we uncovered statistically significant herding effects in the Chinese renewable energy market before 2018. These effects could potentially be attributed to other markets between approximately March 2016 and early 2018, supporting the findings of Ren and Lucey (2023), as confirmed by our simultaneous confidence bands. However, we did not observe significant effects between 2018 and 2021, which diverges from the results reported by Ren and Lucey (2023). In the second application, we explored

the cointegrating relationship between consumption and income in the U.S. Our analysis revealed significant time variations in this relationship. Interestingly, we discovered that both a model with a time-varying drift component (previously adopted in Kapetanios et al., 2020) and a time-varying coefficient model without a drift yielded similar outcomes. This suggests that consumption and income could potentially exhibit nonlinear cointegration without a drift component, instead of being linearly cointegrated with a time-varying drift.

As shown in our study, bootstrap methods can be readily applied without the need for complicated estimation of nuisance parameters, such as second-order bias terms and long-run covariance matrices, or data transformations as proposed in Li et al. (2020). Despite these simplifications, it robustly delivers accurate inference results. Looking ahead, a crucial avenue for future research involves establishing asymptotic bootstrap consistency, particularly in settings with complex dynamics, such as nonlinear cointegration with endogenous regressors.

References

Anand, B., S. Paul, and A. R. Nair (2023). Time-varying effects of oil price shocks on financial stress: Evidence from India. *Energy Economics*, 106703.

Attfield, C. and J. R. Temple (2010). Balanced growth and the great ratios: New evidence for the US and UK. *Journal of Macroeconomics 32*(4), 937–956.

Awaworyi Churchill, S., K. Ivanovski, and M. E. Munyanyi (2021). Income inequality and renewable energy consumption: Time-varying non-parametric evidence. *Journal of Cleaner Production 296*, 126306.

Bai, J. and P. Perron (2003). Computation and analysis of multiple structural change models. *Journal of Applied Econometrics 18*(1), 1–22.

Beutner, E., Y. Lin, and S. Smeekes (2023). GLS estimation and confidence sets for the date of a single break in models with trends. *Econometric Reviews 42*(2), 195–219.

Brüggemann, R., C. Jentsch, and C. Trenkler (2016). Inference in VARs with conditional heteroskedasticity of unknown form. *Journal of Econometrics 191*(1), 69–85.

Bühlmann, P. (1998). Sieve bootstrap for smoothing in nonstationary time series. *Annals of Statistics 26*(1), 48–83.

Cai, Z. (2007). Trending time-varying coefficient time series models with serially correlated errors. *Journal of Econometrics 136*(1), 163–188.

Cavaliere, G. and A. M. R. Taylor (2008). Bootstrap unit root tests for time series with nonstationary volatility. *Econometric Theory 24*(1), 43–71.

Cavaliere, G. and A. M. R. Taylor (2009). Heteroskedastic time series with a unit root. *Econometric Theory 25*(5), 1228–1276.

Chang, E. C., J. W. Cheng, and A. Khorana (2000). An examination of herd behavior in equity markets: An international perspective. *Journal of Banking & Finance 24*(10), 1651–1679.

Chang, Y. (2004). Bootstrap unit root tests in panels with cross-sectional dependency. *Journal of Econometrics 120*(2), 263–293.

Chang, Y., Y. Choi, C. S. Kim, J. I. Miller, and J. Y. Park (2016). Disentangling temporal patterns in elasticities: A functional coefficient panel analysis of electricity demand. *Energy Economics 60*, 232–243.

Chu, C.-K. and J. S. Marron (1991). Comparison of two bandwidth selectors with dependent errors. *Annals of Statistics 19*(4), 1906–1918.

Chudik, A., M. H. Pesaran, and R. P. Smith (2023). Revisiting the great ratios hypothesis. *Oxford Bulletin of Economics and Statistics*, 0305–9049.

Creal, D., S. J. Koopman, and A. Lucas (2013). Generalized autoregressive score models with applications. *Journal of Applied Econometrics 28*(5), 777–795.

Durbin, J. and S. J. Koopman (2012). *Time Series Analysis by State Space Methods* (2nd ed.). Oxford University Press (Oxford).

Friedrich, M., E. Beutner, H. Reuvers, S. Smeekes, J.-P. Urbain, W. Bader, B. Franco, B. Lejeune, and E. Mahieu (2020). A statistical analysis of time trends in atmospheric ethane. *Climatic Change 162*, 105–125.

Friedrich, M. and Y. Lin (2022). Sieve bootstrap inference for linear time-varying coefficient models. *Journal of Econometrics*. In press.

Friedrich, M., Y. Lin, P. Ramdaras, S. Telg, and B. van der Sluis (2023). Time-varying effects of housing attributes and economic environment on housing prices. Tinbergen Institute Discussion Paper. TI 23-039/III.

Friedrich, M., S. Smeekes, and J.-P. Urbain (2020). Autoregressive wild bootstrap inference for nonparametric trends. *Journal of Econometrics 214*(1), 81–109.

Gao, J., B. Peng, and R. Smyth (2021). On income and price elasticities for energy demand: A panel data study. *Energy Economics 96*, 105168.

Hailemariam, A., R. Smyth, and X. Zhang (2019). Oil prices and economic policy uncertainty: Evidence from a nonparametric panel data model. *Energy Economics 83*, 40–51.

Harvey, D. I. and S. J. Leybourne (2014). Break date estimation for models with deterministic structural change. *Oxford Bulletin of Economics and Statistics 76*, 623–642.

Harvey, D. I. and S. J. Leybourne (2015). Confidence sets for the date of a break in level and trend when the order of integration is unknown. *Journal of Econometrics 184*, 262–279.

Kapetanios, G., S. Millard, K. Petrova, and S. Price (2020). Time-varying cointegration with an application to the UK great ratios. *Economics Letters 193*, 109213.

Karmakar, S., S. Richter, and W. B. Wu (2022). Simultaneous inference for time-varying models. *Journal of Econometrics 227*(2), 408–428.

Kreiss, J.-P., E. Paparoditis, and D. N. Politis (2011). On the range of validity of the autoregressive sieve bootstrap. *Annals of Statistics 39*(4), 2103 – 2130.

Kunsch, H. R. (1989). The jackknife and the bootstrap for general stationary observations. *Annals of Statistics*, 1217–1241.

Lee, D. and G. Shaddick (2007). Time-varying coefficient models for the analysis of air pollution and health outcome data. *Biometrics 63*(4), 1253–1261.

Li, D., P. C. Phillips, and J. Gao (2020). Kernel-based inference in time-varying coefficient cointegrating regression. *Journal of Econometrics 215*(2), 607–632.

Li, X., Z. Zhao, et al. (2019). A time varying approach to the stock return–inflation puzzle. *Journal of the Royal Statistical Society Series C 68*(5), 1509–1528.

Lin, Y. and H. Reuvers (2022a). Cointegrating polynomial regressions with power law trends: Environmental kuznets curve or omitted time effects? Tinbergen Institute Discussion Paper. TI 2022-092/III.

Lin, Y. and H. Reuvers (2022b). Fully modified estimation in cointegrating polynomial regressions: Extensions and Monte Carlo comparison. Tinbergen Institute Discussion Paper. 2022-093/III.

Mikayilov, J. I., F. J. Hasanov, and M. Galeotti (2018). Decoupling of CO2 emissions and GDP: A time-varying cointegration approach. *Ecological Indicators 95*, 615–628.

Park, C. and S. Park (2013). Exchange rate predictability and a monetary model with time-varying cointegration coefficients. *Journal of International Money and Finance 37*, 394–410.

Park, J. Y. and S. B. Hahn (1999). Cointegrating regressions with time varying coefficients. *Econometric Theory 15*(5), 664–703.

Perron, P. and X. Zhu (2005). Structural breaks with deterministic and stochastic trends. *Journal of Econometrics 129*, 65–119.

Phillips, P. C., D. Li, and J. Gao (2017). Estimating smooth structural change in cointegration models. *Journal of Econometrics 196*(1), 180–195.

Ren, B. and B. Lucey (2023). Herding in the chinese renewable energy market: Evidence from a bootstrapping time-varying coefficient autoregressive model. *Energy Economics 119*, 106526.

Ren, X., Z. Tong, X. Sun, and C. Yan (2022). Dynamic impacts of energy consumption on economic growth in China: Evidence from a non-parametric panel data model. *Energy Economics 107*, 105855.

Sarantis, N. and C. Stewart (1999). Is the consumption-income ratio stationary? Evidence from panel unit root tests. *Economics Letters 64*(3), 309–314.

Silvapulle, P., R. Smyth, X. Zhang, and J.-P. Fenech (2017). Nonparametric panel data model for crude oil and stock market prices in net oil importing countries. *Energy Economics 67*, 255–267.

Smeekes, S. and A. M. R. Taylor (2012). Bootstrap union tests for unit roots in the presence of nonstationary volatility. *Econometric Theory 28*(2), 422–456.

Smeekes, S. and J.-P. Urbain (2014). A multivariate invariance principle for modified wild bootstrap methods with an application to unit root testing. Maastricht University Research Papers, RM/14/008.

Sun, X., S. Xiao, X. Ren, and B. Xu (2023). Time-varying impact of information and communication technology on carbon emissions. *Energy Economics 118*, 106492.

Uddin, M. M., V. Mishra, and R. Smyth (2020). Income inequality and CO2 emissions in the G7, 1870-2014: Evidence from non-parametric modelling. *Energy Economics 88*, 104780.

Vieu, P. (1991). Nonparametric regression: optimal local bandwidth choice. *Journal of the Royal Statistical Society: Series B (Methodological) 53*(2), 453–464.

Wang, B., J. Wei, X. Tan, and B. Su (2021). The sectorally heterogeneous and time-varying price elasticities of energy demand in china. *Energy Economics 102*, 105486.

Zanin, L. and G. Marra (2012). Rolling regression versus time-varying coefficient modelling: An empirical investigation of the Okun's law in some euro area countries. *Bulletin of Economic Research 64*(1), 91–108.

Zhou, Z. and W. B. Wu (2010). Simultaneous inference of linear models with time varying coefficients. *Journal of the Royal Statistical Society: Series B (Statistical Methodology) 72*(4), 513–531.

Online Appendix to:

Improved bootstrap inference for linear time-varying coefficient models: Some Monte Carlo evidence

Yicong Lin[1,2] and Mingxuan Song[1]

[1] *Vrije Universiteit Amsterdam*

[2] *Tinbergen Institute*

Contents

A   Additional simulation results

Table A1: Empirical coverage and length of 95%-level intervals and bands for $\beta_1(\cdot)$ and $\beta_2(\cdot)$ with $n = 200$ and homoskedastic errors (AR) and $I(0)$ $\boldsymbol{x}_t$.

	h	β_1 PW	G_{sub}	G	FS	β_2 PW	G_{sub}	G	FS
					Empirical Coverage				
SB	0.09	0.941	0.930	0.920	0.896	0.941	0.936	0.932	0.902
	0.12	0.940	0.934	0.898	0.926	0.941	0.940	0.918	0.914
	0.15	0.934	0.910	0.892	0.880	0.940	0.914	0.912	0.890
SWB	0.09	0.946	0.964	0.958	0.944	0.947	0.966	0.958	0.942
	0.12	0.945	0.948	0.954	0.944	0.946	0.972	0.972	0.946
	0.15	0.938	0.930	0.920	0.912	0.945	0.956	0.950	0.942
MBB	0.09	0.941	0.934	0.932	0.918	0.943	0.930	0.932	0.912
	0.12	0.941	0.912	0.892	0.934	0.940	0.940	0.918	0.938
	0.15	0.932	0.910	0.898	0.894	0.941	0.944	0.930	0.918
					Empirical Length				
SB	0.09	0.555	0.772	0.842	0.905	0.560	0.773	0.848	0.906
	0.12	0.479	0.641	0.692	0.771	0.480	0.668	0.695	0.780
	0.15	0.425	0.606	0.621	0.633	0.430	0.610	0.626	0.639
SWB	0.09	0.567	0.836	0.898	1.051	0.571	0.844	0.911	1.064
	0.12	0.486	0.691	0.762	0.813	0.490	0.725	0.766	0.812
	0.15	0.435	0.652	0.676	0.690	0.437	0.658	0.674	0.691
MBB	0.09	0.552	0.758	0.834	0.876	0.554	0.759	0.843	0.880
	0.12	0.476	0.631	0.677	0.759	0.476	0.652	0.679	0.762
	0.15	0.422	0.598	0.614	0.634	0.425	0.599	0.615	0.634

Table A2: Empirical coverage and length of 95%-level intervals and bands for $\beta_1(\cdot)$ and $\beta_2(\cdot)$ with $n = 200$ and homoskedastic errors (AR) and $I(0)/I(1)$ $\boldsymbol{x}_t$.

| | h | | β_1 | | | | β_2 | | |
		PW	G_{sub}	G	FS	PW	G_{sub}	G	FS
					Empirical Coverage				
	0.09	0.962	0.968	0.972	0.818	0.962	0.976	0.968	0.818
SB	0.12	0.966	0.962	0.968	0.882	0.968	0.962	0.954	0.870
	0.15	0.964	0.926	0.934	0.906	0.971	0.942	0.938	0.906
	0.09	0.965	0.974	0.982	0.872	0.967	0.990	0.994	0.880
SWB	0.12	0.969	0.980	0.974	0.896	0.971	0.978	0.976	0.908
	0.15	0.968	0.964	0.960	0.930	0.975	0.956	0.954	0.922
	0.09	0.966	0.960	0.962	0.954	0.967	0.958	0.962	0.954
MBB	0.12	0.967	0.968	0.958	0.972	0.970	0.956	0.954	0.964
	0.15	0.960	0.944	0.952	0.952	0.972	0.968	0.960	0.966
					Empirical Length				
	0.09	0.504	0.711	0.769	0.809	0.469	0.667	0.719	0.721
SB	0.12	0.463	0.598	0.667	0.738	0.427	0.592	0.618	0.686
	0.15	0.426	0.606	0.617	0.641	0.397	0.554	0.568	0.593
	0.09	0.524	0.779	0.841	0.954	0.483	0.729	0.767	0.850
SWB	0.12	0.479	0.655	0.746	0.794	0.438	0.655	0.687	0.740
	0.15	0.436	0.654	0.672	0.689	0.422	0.618	0.644	0.658
	0.09	0.496	0.682	0.737	0.781	0.470	0.657	0.723	0.742
MBB	0.12	0.440	0.574	0.622	0.694	0.416	0.575	0.588	0.659
	0.15	0.400	0.562	0.570	0.594	0.380	0.539	0.546	0.568

Table A3: Empirical coverage and length of 95%-level intervals and bands for $\beta_1(\cdot)$ and $\beta_2(\cdot)$ with $n = 200$ and homoskedastic errors (ENDO) $\rho = 0.3$ and $I(1)$ $\boldsymbol{x}_t$.

	h	β_1				β_2			
		PW	G_{sub}	G	FS	PW	G_{sub}	G	FS
					Empirical Coverage				
	0.09	0.973	0.974	0.984	0.860	0.961	0.946	0.968	0.840
SB	0.12	0.977	0.966	0.972	0.918	0.968	0.944	0.962	0.902
	0.15	0.971	0.942	0.944	0.920	0.965	0.936	0.938	0.912
	0.09	0.976	0.986	0.990	0.900	0.964	0.980	0.992	0.928
SWB	0.12	0.979	0.978	0.978	0.942	0.971	0.972	0.978	0.948
	0.15	0.975	0.956	0.956	0.936	0.968	0.948	0.954	0.934
	0.19	0.980	0.966	0.978	0.968	0.970	0.944	0.960	0.946
MBB	0.12	0.980	0.970	0.974	0.970	0.973	0.960	0.968	0.960
	0.15	0.969	0.962	0.968	0.970	0.968	0.940	0.944	0.954
					Empirical Length				
	0.09	0.508	0.723	0.807	0.821	0.542	0.779	0.868	0.877
SB	0.12	0.473	0.591	0.724	0.766	0.506	0.725	0.787	0.824
	0.15	0.449	0.646	0.662	0.713	0.485	0.700	0.717	0.781
	0.09	0.523	0.790	0.897	0.948	0.555	0.845	0.962	1.013
SWB	0.12	0.488	0.639	0.772	0.855	0.519	0.781	0.832	0.928
	0.15	0.466	0.697	0.715	0.763	0.496	0.749	0.770	0.823
	0.09	0.543	0.755	0.820	0.841	0.622	0.859	0.938	0.963
MBB	0.12	0.495	0.590	0.736	0.766	0.577	0.795	0.859	0.895
	0.15	0.457	0.649	0.665	0.706	0.534	0.753	0.776	0.824

Table A4: Empirical coverage and length of 95%-level intervals and bands for $\beta_1(\cdot)$ and $\beta_2(\cdot)$ with $n = 200$ and homoskedastic errors (ENDO) $\rho = 0.5$ and $I(1)$ $\boldsymbol{x}_t$.

	h	β_1				β_2			
		PW	G_{sub}	G	FS	PW	G_{sub}	G	FS
					Empirical Coverage				
SB	0.09	0.947	0.932	0.954	0.814	0.918	0.876	0.910	0.746
	0.12	0.956	0.928	0.950	0.880	0.934	0.874	0.898	0.838
	0.15	0.954	0.898	0.898	0.882	0.935	0.870	0.876	0.848
SWB	0.09	0.953	0.962	0.980	0.884	0.924	0.912	0.962	0.864
	0.12	0.961	0.964	0.970	0.906	0.939	0.906	0.922	0.888
	0.15	0.959	0.928	0.934	0.890	0.939	0.904	0.912	0.888
MBB	0.09	0.956	0.936	0.958	0.922	0.935	0.890	0.902	0.862
	0.12	0.963	0.94	0.958	0.944	0.946	0.898	0.914	0.900
	0.15	0.957	0.910	0.914	0.928	0.943	0.888	0.890	0.902
					Empirical Length				
SB	0.09	0.553	0.788	0.880	0.894	0.585	0.843	0.936	0.947
	0.12	0.511	0.638	0.788	0.827	0.542	0.774	0.843	0.881
	0.15	0.484	0.695	0.714	0.770	0.515	0.745	0.767	0.828
SWB	0.09	0.569	0.859	0.975	1.033	0.597	0.911	1.039	1.091
	0.12	0.527	0.694	0.840	0.932	0.555	0.835	0.886	0.992
	0.15	0.501	0.748	0.771	0.818	0.528	0.793	0.817	0.870
MBB	0.09	0.598	0.826	0.901	0.923	0.672	0.930	1.013	1.041
	0.12	0.546	0.647	0.811	0.842	0.622	0.854	0.923	0.960
	0.15	0.505	0.709	0.728	0.777	0.575	0.809	0.835	0.886

Table A5: Empirical coverage and length of 95%-level intervals and bands for $\beta_1(\cdot)$ and $\beta_2(\cdot)$ with $n = 200$ and GARCH errors (GARCH1) and $I(0)$ $\boldsymbol{x}_t$.

	h	β_1				β_2			
		PW	G_{sub}	G	FS	PW	G_{sub}	G	FS
					Empirical Coverage				
SB	0.09	0.936	0.890	0.870	0.784	0.943	0.908	0.896	0.814
	0.12	0.937	0.878	0.872	0.828	0.944	0.902	0.882	0.858
	0.15	0.936	0.864	0.864	0.820	0.948	0.894	0.892	0.858
SWB	0.09	0.941	0.92	0.900	0.892	0.947	0.952	0.934	0.884
	0.12	0.940	0.908	0.916	0.860	0.948	0.934	0.932	0.884
	0.15	0.939	0.900	0.908	0.866	0.950	0.924	0.924	0.896
MBB	0.09	0.941	0.908	0.892	0.888	0.949	0.916	0.906	0.902
	0.12	0.941	0.900	0.902	0.928	0.951	0.918	0.898	0.924
	0.15	0.939	0.892	0.894	0.876	0.950	0.918	0.910	0.898
					Empirical Length				
SB	0.09	0.723	1.026	1.116	1.190	0.727	1.026	1.111	1.192
	0.12	0.622	0.838	0.905	1.012	0.617	0.873	0.896	1.002
	0.15	0.555	0.796	0.815	0.834	0.550	0.787	0.804	0.825
SWB	0.09	0.739	1.120	1.198	1.412	0.743	1.118	1.199	1.409
	0.12	0.634	0.912	1.004	1.070	0.628	0.947	0.994	1.063
	0.15	0.561	0.860	0.881	0.912	0.557	0.849	0.881	0.902
MBB	0.09	0.736	1.057	1.159	1.250	0.730	1.054	1.150	1.246
	0.12	0.631	0.862	0.933	1.059	0.623	0.897	0.928	1.045
	0.15	0.560	0.817	0.834	0.861	0.556	0.809	0.835	0.853

Table A6: Empirical coverage and length of 95%-level intervals and bands for $\beta_1(\cdot)$ and $\beta_2(\cdot)$ with $n = 200$ and GARCH errors (GARCH1) and $I(0)/I(1)$ $\boldsymbol{x}_t$.

	h	β_1				β_2			
		PW	G_{sub}	G	FS	PW	G_{sub}	G	FS
		Empirical Coverage							
	0.09	0.953	0.936	0.920	0.742	0.947	0.938	0.924	0.734
SB	0.12	0.956	0.928	0.914	0.826	0.955	0.922	0.910	0.800
	0.15	0.954	0.888	0.882	0.834	0.957	0.888	0.884	0.824
	0.09	0.956	0.960	0.964	0.812	0.952	0.952	0.944	0.806
SWB	0.12	0.959	0.944	0.938	0.858	0.959	0.942	0.940	0.832
	0.15	0.959	0.916	0.928	0.890	0.961	0.912	0.910	0.856
	0.09	0.963	0.946	0.938	0.938	0.958	0.932	0.938	0.918
MBB	0.12	0.962	0.940	0.928	0.950	0.961	0.930	0.932	0.950
	0.15	0.955	0.924	0.934	0.932	0.963	0.924	0.928	0.924
		Empirical Length							
	0.09	0.605	0.848	0.919	0.967	0.588	0.816	0.904	0.888
SB	0.12	0.549	0.705	0.796	0.882	0.514	0.713	0.737	0.832
	0.15	0.511	0.725	0.743	0.766	0.465	0.665	0.686	0.705
	0.09	0.624	0.916	0.995	1.135	0.603	0.912	0.966	1.084
SWB	0.12	0.573	0.781	0.897	0.948	0.522	0.783	0.840	0.888
	0.15	0.525	0.793	0.815	0.837	0.479	0.724	0.748	0.781
	0.09	0.608	0.850	0.938	0.977	0.596	0.863	0.916	1.012
MBB	0.12	0.524	0.686	0.761	0.865	0.512	0.739	0.763	0.873
	0.15	0.470	0.679	0.693	0.716	0.464	0.662	0.678	0.690

Table A7: Empirical coverage and length of 95%-level intervals and bands for $\beta_1(\cdot)$ and $\beta_2(\cdot)$ with $n = 200$ and GARCH errors (GARCH2) and $I(0)$ $\boldsymbol{x}_t$.

	h	β_1				β_2			
		PW	G_{sub}	G	FS	PW	G_{sub}	G	FS
		Empirical Coverage							
	0.09	0.932	0.872	0.83	0.758	0.935	0.890	0.856	0.770
SB	0.12	0.930	0.870	0.812	0.800	0.937	0.906	0.852	0.830
	0.15	0.925	0.836	0.802	0.762	0.937	0.878	0.862	0.818
	0.09	0.936	0.92	0.884	0.880	0.938	0.932	0.886	0.878
SWB	0.12	0.932	0.912	0.884	0.836	0.942	0.934	0.902	0.848
	0.15	0.929	0.892	0.860	0.850	0.941	0.926	0.914	0.890
	0.09	0.938	0.906	0.868	0.864	0.941	0.928	0.888	0.866
MBB	0.12	0.934	0.908	0.872	0.886	0.944	0.912	0.890	0.892
	0.05	0.930	0.886	0.860	0.844	0.942	0.916	0.900	0.884
		Empirical Length							
	0.09	0.695	0.980	1.060	1.157	0.694	0.983	1.060	1.146
SB	0.12	0.595	0.805	0.873	0.965	0.597	0.836	0.875	0.969
	0.15	0.530	0.762	0.780	0.797	0.529	0.751	0.777	0.792
	0.09	0.709	1.071	1.172	1.380	0.704	1.069	1.168	1.367
SWB	0.12	0.603	0.873	0.971	1.019	0.606	0.922	0.969	1.025
	0.15	0.539	0.826	0.851	0.880	0.535	0.817	0.842	0.876
	0.09	0.704	1.041	1.127	1.236	0.701	1.029	1.117	1.215
MBB	0.12	0.602	0.841	0.920	1.024	0.599	0.875	0.911	1.021
	0.15	0.535	0.791	0.812	0.841	0.532	0.789	0.803	0.824

Table A8: Empirical coverage and length of 95%-level intervals and bands for $\beta_1(\cdot)$ and $\beta_2(\cdot)$ with $n = 200$ and GARCH errors (GARCH2) and $I(0)/I(1)$ $\boldsymbol{x}_t$.

	h	β_1				β_2			
		PW	G_{sub}	G	FS	PW	G_{sub}	G	FS
		Empirical Coverage							
	0.09	0.957	0.942	0.934	0.772	0.951	0.936	0.912	0.740
SB	0.12	0.961	0.920	0.906	0.806	0.956	0.938	0.918	0.812
	0.15	0.961	0.882	0.886	0.832	0.960	0.906	0.902	0.840
	0.09	0.960	0.962	0.942	0.836	0.955	0.946	0.934	0.800
SB	0.12	0.963	0.944	0.948	0.844	0.959	0.956	0.954	0.850
	0.15	0.963	0.924	0.924	0.874	0.964	0.936	0.938	0.882
	0.09	0.965	0.946	0.940	0.930	0.959	0.930	0.930	0.922
MBB	0.12	0.964	0.946	0.934	0.942	0.962	0.948	0.936	0.956
	0.15	0.957	0.922	0.928	0.918	0.963	0.944	0.944	0.942
		Empirical Length							
	0.09	0.587	0.846	0.911	0.945	0.556	0.798	0.863	0.864
SB	0.12	0.531	0.695	0.780	0.861	0.498	0.711	0.721	0.809
	0.15	0.492	0.699	0.718	0.735	0.454	0.655	0.665	0.692
	0.09	0.617	0.922	0.990	1.123	0.569	0.881	0.951	0.999
SWB	0.12	0.555	0.757	0.873	0.919	0.515	0.781	0.823	0.879
	0.15	0.512	0.768	0.793	0.816	0.474	0.719	0.735	0.767
	0.09	0.581	0.837	0.906	0.984	0.574	0.826	0.915	0.958
MBB	0.12	0.500	0.685	0.750	0.849	0.496	0.718	0.741	0.835
	0.15	0.441	0.661	0.661	0.680	0.444	0.663	0.655	0.680

Table A9: Empirical coverage and length of 95%-level intervals and bands for $\beta_1(\cdot)$ and $\beta_2(\cdot)$ with $n = 200$ and (NL1) errors and $I(0)$ $\boldsymbol{x}_t$.

	h	β_1				β_2			
		PW	G_{sub}	G	FS	PW	G_{sub}	G	FS
		Empirical Coverage							
	0.09	0.954	0.946	0.942	0.922	0.955	0.966	0.942	0.934
SB	0.12	0.943	0.904	0.890	0.922	0.956	0.960	0.948	0.950
	0.15	0.911	0.838	0.822	0.806	0.955	0.956	0.940	0.928
	0.09	0.958	0.964	0.960	0.958	0.960	0.984	0.974	0.970
SWB	0.12	0.947	0.960	0.944	0.946	0.961	0.976	0.980	0.966
	0.15	0.916	0.888	0.882	0.878	0.960	0.964	0.966	0.954
	0.09	0.953	0.930	0.934	0.950	0.957	0.954	0.948	0.958
MBB	0.12	0.939	0.878	0.858	0.912	0.956	0.950	0.940	0.970
	0.15	0.904	0.810	0.790	0.798	0.953	0.952	0.940	0.946
		Empirical Length							
	0.09	0.238	0.333	0.360	0.386	0.237	0.332	0.361	0.385
SB	0.12	0.207	0.279	0.299	0.336	0.206	0.288	0.298	0.334
	0.15	0.187	0.264	0.271	0.277	0.186	0.265	0.270	0.276
	0.09	0.244	0.364	0.390	0.455	0.243	0.362	0.390	0.453
SWB	0.12	0.212	0.300	0.332	0.350	0.211	0.314	0.331	0.349
	0.15	0.190	0.286	0.296	0.304	0.189	0.286	0.294	0.301
	0.09	0.236	0.325	0.355	0.374	0.235	0.325	0.356	0.374
MBB	0.12	0.204	0.271	0.293	0.327	0.203	0.282	0.290	0.325
	0.15	0.183	0.258	0.264	0.270	0.182	0.258	0.263	0.272

Table A10: Empirical coverage and length of 95%-level intervals and bands for $\beta_1(\cdot)$ and $\beta_2(\cdot)$ with $n = 200$ and (NL1) errors and $I(0)/I(1)$ x_t.

	h	β_1 PW	G_{sub}	G	FS	β_2 PW	G_{sub}	G	FS
					Empirical Coverage				
SB	0.09	0.979	0.990	0.990	0.836	0.982	0.996	0.994	0.822
	0.12	0.979	0.976	0.970	0.860	0.986	0.988	0.988	0.856
	0.15	0.971	0.926	0.924	0.890	0.987	0.936	0.942	0.896
SWB	0.09	0.980	0.994	0.990	0.844	0.983	0.998	0.998	0.852
	0.12	0.980	0.988	0.990	0.884	0.987	0.994	0.998	0.878
	0.15	0.974	0.938	0.940	0.904	0.989	0.950	0.948	0.910
MBB	0.09	0.990	0.990	0.992	0.992	0.989	0.990	0.986	0.978
	0.12	0.986	0.980	0.978	0.988	0.989	0.990	0.986	0.992
	0.15	0.971	0.938	0.938	0.952	0.987	0.990	0.984	0.988
					Empirical Length				
SB	0.09	0.289	0.411	0.441	0.456	0.276	0.390	0.436	0.425
	0.12	0.274	0.358	0.398	0.441	0.274	0.376	0.402	0.438
	0.15	0.268	0.386	0.384	0.408	0.271	0.379	0.386	0.411
SWB	0.09	0.302	0.450	0.482	0.533	0.292	0.430	0.474	0.492
	0.12	0.291	0.395	0.447	0.482	0.295	0.420	0.445	0.483
	0.15	0.284	0.428	0.425	0.444	0.283	0.414	0.434	0.455
MBB	0.09	0.271	0.380	0.409	0.418	0.283	0.393	0.435	0.450
	0.12	0.254	0.323	0.360	0.398	0.270	0.372	0.385	0.427
	0.15	0.237	0.343	0.338	0.360	0.254	0.362	0.361	0.385

Table A11: Empirical coverage and length of 95%-level intervals and bands for $\beta_1(\cdot)$ and $\beta_2(\cdot)$ with $n = 200$ and (NL2) errors and $I(0)$ $\boldsymbol{x}_t$.

| | h | \multicolumn{4}{c}{β_1} | \multicolumn{4}{c}{β_2} |
		PW	G_{sub}	G	FS	PW	G_{sub}	G	FS
Empirical Coverage
	0.09	0.967	0.952	0.954	0.936	0.972	0.968	0.958	0.928
SB	0.12	0.949	0.904	0.906	0.920	0.974	0.966	0.960	0.946
	0.15	0.884	0.748	0.762	0.748	0.975	0.968	0.966	0.940
	0.09	0.969	0.974	0.980	0.974	0.974	0.984	0.980	0.950
SWB	0.12	0.953	0.942	0.946	0.942	0.976	0.988	0.984	0.958
	0.15	0.890	0.824	0.836	0.844	0.976	0.980	0.974	0.962
	0.09	0.966	0.950	0.958	0.954	0.973	0.968	0.964	0.958
MBB	0.12	0.944	0.892	0.900	0.940	0.973	0.972	0.968	0.970
	0.15	0.867	0.696	0.710	0.708	0.973	0.956	0.958	0.952
Empirical Length
	0.09	0.157	0.223	0.242	0.264	0.156	0.222	0.240	0.263
SB	0.12	0.137	0.187	0.204	0.226	0.136	0.195	0.202	0.225
	0.15	0.124	0.179	0.184	0.188	0.123	0.178	0.182	0.186
	0.09	0.160	0.248	0.270	0.321	0.159	0.243	0.266	0.318
SWB	0.12	0.140	0.206	0.231	0.239	0.138	0.211	0.227	0.238
	0.15	0.126	0.196	0.201	0.211	0.125	0.195	0.201	0.208
	0.09	0.155	0.224	0.242	0.262	0.154	0.221	0.240	0.260
MBB	0.12	0.135	0.185	0.200	0.227	0.133	0.191	0.198	0.225
	0.15	0.121	0.176	0.181	0.185	0.120	0.175	0.179	0.183

Table A12: Empirical coverage and length of 95%-level intervals and bands for $\beta_1(\cdot)$ and $\beta_2(\cdot)$ with $n = 200$ and (NL2) errors and $I(0)/I(1)$ $\boldsymbol{x}_t$.

	h	β_1 PW	G_{sub}	G	FS	β_2 PW	G_{sub}	G	FS
					Empirical Coverage				
SB	0.09	0.981	0.986	0.982	0.852	0.987	0.990	0.998	0.870
	0.12	0.983	0.972	0.976	0.882	0.992	0.996	0.996	0.866
	0.15	0.971	0.918	0.920	0.898	0.993	0.976	0.982	0.940
SWB	0.09	0.983	0.988	0.988	0.886	0.988	0.995	0.998	0.892
	0.12	0.984	0.986	0.984	0.902	0.992	0.998	0.998	0.896
	0.15	0.974	0.944	0.940	0.914	0.994	0.988	0.988	0.954
MBB	0.09	0.994	0.994	0.996	0.994	0.996	0.996	0.996	0.992
	0.12	0.991	0.988	0.990	0.992	0.995	0.998	0.998	0.998
	0.15	0.975	0.948	0.954	0.956	0.993	0.986	0.988	0.990
					Empirical Length				
SB	0.09	0.234	0.331	0.360	0.370	0.231	0.324	0.366	0.362
	0.12	0.235	0.295	0.343	0.374	0.235	0.332	0.353	0.379
	0.15	0.233	0.331	0.343	0.352	0.239	0.331	0.347	0.363
SWB	0.09	0.245	0.372	0.396	0.439	0.243	0.364	0.408	0.431
	0.12	0.248	0.330	0.386	0.426	0.245	0.370	0.392	0.422
	0.15	0.247	0.370	0.376	0.399	0.253	0.371	0.395	0.408
MBB	0.09	0.230	0.317	0.337	0.360	0.231	0.324	0.357	0.369
	0.12	0.221	0.263	0.312	0.342	0.220	0.307	0.315	0.348
	0.15	0.209	0.296	0.292	0.316	0.208	0.286	0.296	0.316

Table A13: Empirical coverage of 95%-level intervals and bands for $\beta_1(\cdot)$ and $\beta_2(\cdot)$ when the bandwidth is selected by LMCV(ℓ), i.e., leaving $(2\ell + 1)$ out with $\ell = 0, 2, 4, 6$, as well as their averages (labeled as AVG), see Friedrich and Lin (2022). Here, we have AR errors and $I(0)$ $\boldsymbol{x}_t$.

	β_1				β_2			
	PW	G_{sub}	G	FS	PW	G_{sub}	G	FS
				SWB				
LMCV0	0.940	0.930	0.940	0.912	0.941	0.948	0.942	0.916
LMCV2	0.942	0.936	0.946	0.920	0.942	0.950	0.946	0.912
LMCV4	0.938	0.942	0.948	0.930	0.941	0.942	0.944	0.934
LMCV6	0.940	0.934	0.924	0.914	0.941	0.938	0.940	0.926
AVG	0.942	0.936	0.934	0.924	0.941	0.942	0.934	0.928
				MBB				
LMCV0	0.937	0.906	0.916	0.898	0.936	0.902	0.890	0.876
LMCV2	0.937	0.920	0.918	0.896	0.939	0.890	0.888	0.892
LMCV4	0.936	0.896	0.908	0.904	0.939	0.902	0.904	0.896
LMCV6	0.935	0.904	0.892	0.902	0.938	0.914	0.902	0.902
AVG	0.939	0.922	0.904	0.898	0.938	0.900	0.890	0.894

Table A14: Empirical coverage of 95%-level intervals and bands for $\beta_1(\cdot)$ and $\beta_2(\cdot)$ when the bandwidth is selected by LMCV(ℓ), see Figure A13 for further details. Here, we have AR errors and $I(0)/I(1)$ $\boldsymbol{x}_t$.

	β_1				β_2			
	PW	G_{sub}	G	FS	PW	G_{sub}	G	FS
				SWB				
LMCV0	0.961	0.970	0.974	0.854	0.959	0.982	0.978	0.868
LMCV2	0.963	0.980	0.986	0.896	0.963	0.976	0.976	0.884
LMCV4	0.963	0.964	0.970	0.892	0.966	0.976	0.978	0.902
LMCV6	0.964	0.970	0.968	0.904	0.969	0.974	0.982	0.928
AVG	0.964	0.976	0.974	0.888	0.965	0.978	0.992	0.900
				MBB				
LMCV0	0.969	0.950	0.966	0.950	0.964	0.946	0.960	0.920
LMCV2	0.971	0.970	0.968	0.970	0.967	0.954	0.964	0.936
LMCV4	0.971	0.970	0.974	0.972	0.969	0.946	0.958	0.956
LMCV6	0.970	0.964	0.962	0.958	0.969	0.946	0.954	0.960
AVG	0.972	0.966	0.968	0.968	0.970	0.948	0.962	0.948

Table A15: Empirical coverage of 95%-level intervals and bands for $\beta_1(\cdot)$ and $\beta_2(\cdot)$ when the bandwidth is selected by LMCV(ℓ), see Figure A13 for further details. Here, we have GARCH1 errors and $I(0)$ $\boldsymbol{x}_t$.

	β_1				β_2			
	PW	G_{sub}	G	FS	PW	G_{sub}	G	FS
				SWB				
LMCV0	0.929	0.894	0.864	0.844	0.942	0.920	0.898	0.862
LMCV2	0.930	0.884	0.882	0.842	0.943	0.914	0.908	0.880
LMCV4	0.929	0.914	0.892	0.860	0.945	0.934	0.924	0.904
LMCV6	0.927	0.898	0.888	0.858	0.947	0.940	0.930	0.912
AVG	0.929	0.906	0.898	0.866	0.944	0.916	0.910	0.896
				MBB				
LMCV0	0.933	0.868	0.852	0.826	0.938	0.908	0.866	0.848
LMCV2	0.934	0.886	0.880	0.868	0.940	0.908	0.878	0.858
LMCV4	0.933	0.876	0.874	0.878	0.940	0.922	0.898	0.870
LMCV6	0.931	0.888	0.870	0.866	0.940	0.906	0.874	0.858
AVG	0.935	0.886	0.874	0.860	0.940	0.916	0.892	0.880

B   Additional empirical results

B.1   The SWB confidence intervals and bands for Models (5.2) and (5.3)
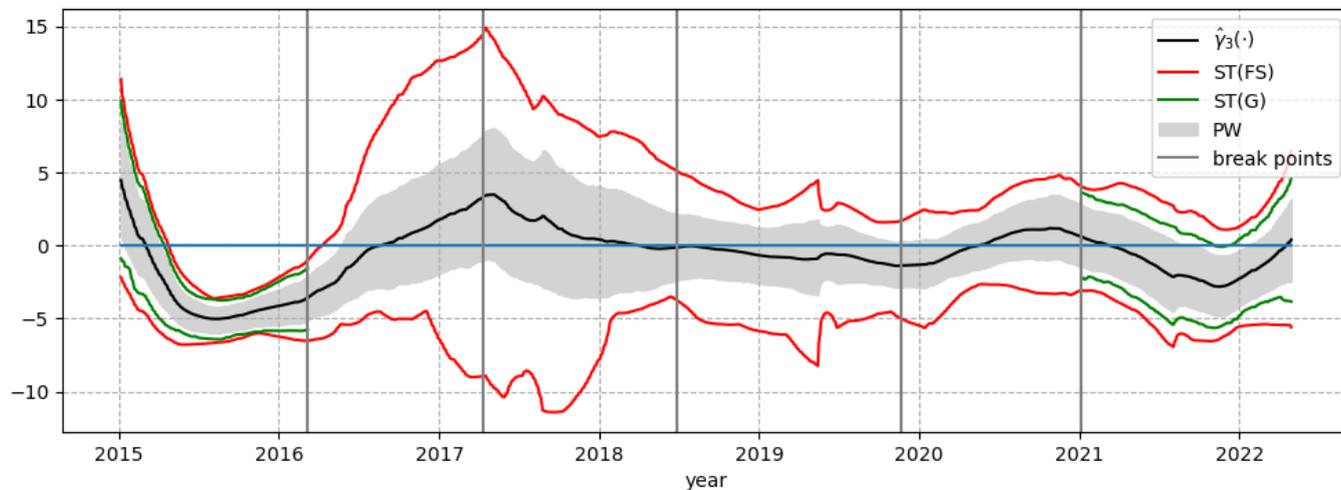

Figure 10: Estimated curve of $\gamma_3(\cdot)$ (black line) and resulting 95%-level confidence intervals (grey, labeled as PW) and bands using the SWB for Model (5.2), see Figure 3 for further information.
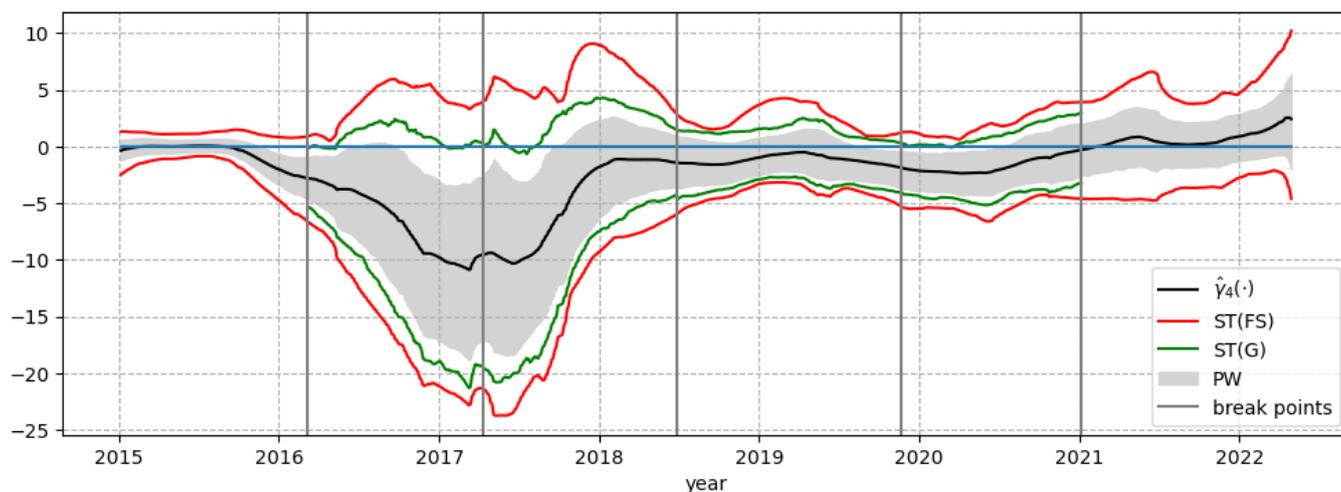

Figure 11: Estimated curve of $\gamma_4(\cdot)$ (black line) and resulting 95%-level confidence intervals (grey, labeled as PW) and bands using the SWB for Model (5.3), see Figure 5 for further information..

B.2 Models without lag terms

Our bootstrap methods allow error processes that are serially dependent. As a robustness check, we consider the lag-free counterparts of Models (5.2) and (5.3):

$$\text{CSAD}_{m,t} = \gamma_0(t/n) + \gamma_1(t/n)R_{m,t} + \gamma_2(t/n)|R_{m,t}| + \gamma_3(t/n)R_{m,t}^2 + \varepsilon_t, \tag{B.1}$$

$$\text{CSAD}_{m,t} = \gamma_0(t/n) + \gamma_1(t/n)R_{m,t} + \gamma_2(t/n)|R_{m,t}| + \gamma_3(t/n)R_{m,t}^2 + \gamma_4(t/n)R_{int,t}^2 + \varepsilon_t. \tag{B.2}$$

B.2.1 The MBB confidence intervals and bands for Models (B.1) and (B.2)
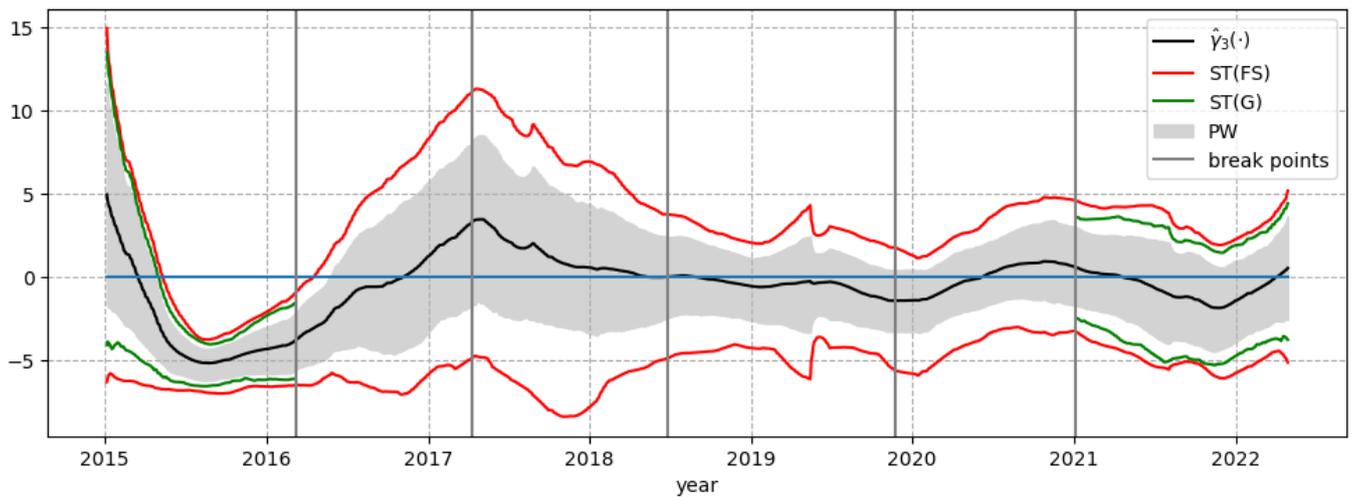
Figure 12: Estimated curve of $\gamma_3(\cdot)$ (black line) and resulting 95%-level confidence intervals (grey, labeled as PW) and bands using the MBB for Model (B.1), see Figure 3 for further information.
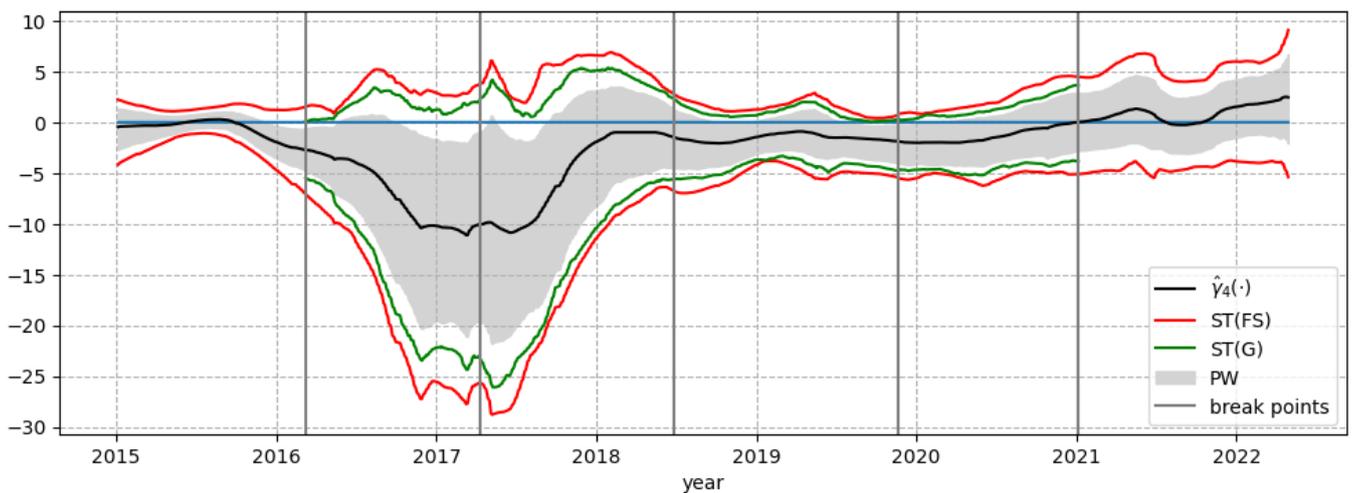

Figure 13: Estimated curve of $\gamma_4(\cdot)$ (black line) and resulting 95%-level confidence intervals (grey, labeled as PW) and bands using the MBB for Model (B.2), see Figure 5 for further information.

## B.2.2 The SWB confidence intervals and bands for Models (B.1) and (B.2)
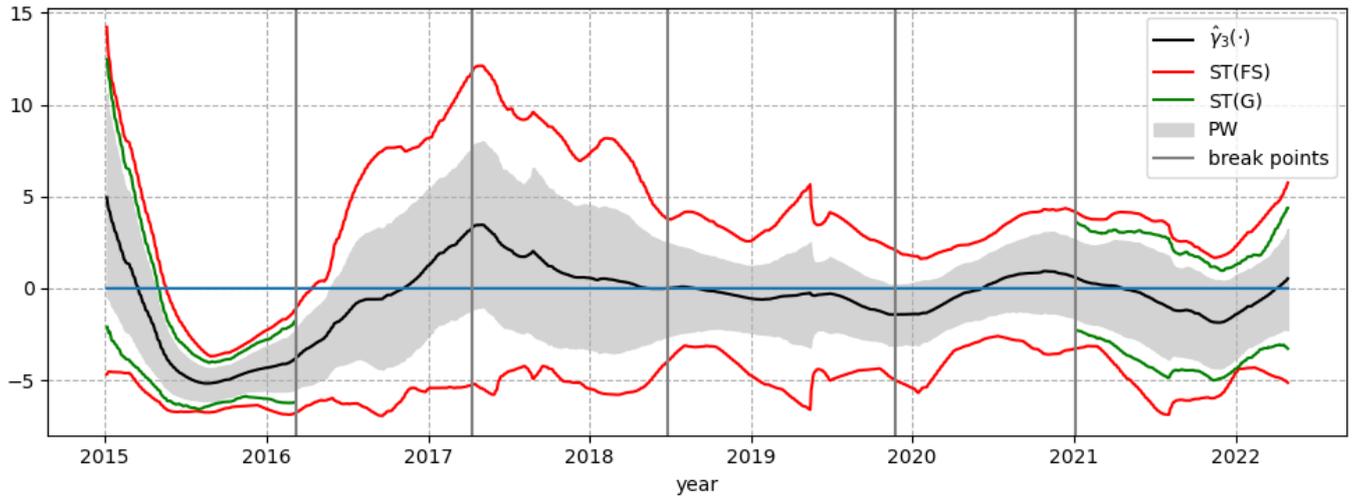
Figure 14: Estimated curve of $\gamma_3(\cdot)$ (black line) and resulting 95%-level confidence intervals (grey, labeled as PW) and bands using the SWB for Model (B.1), see Figure 3 for further information.
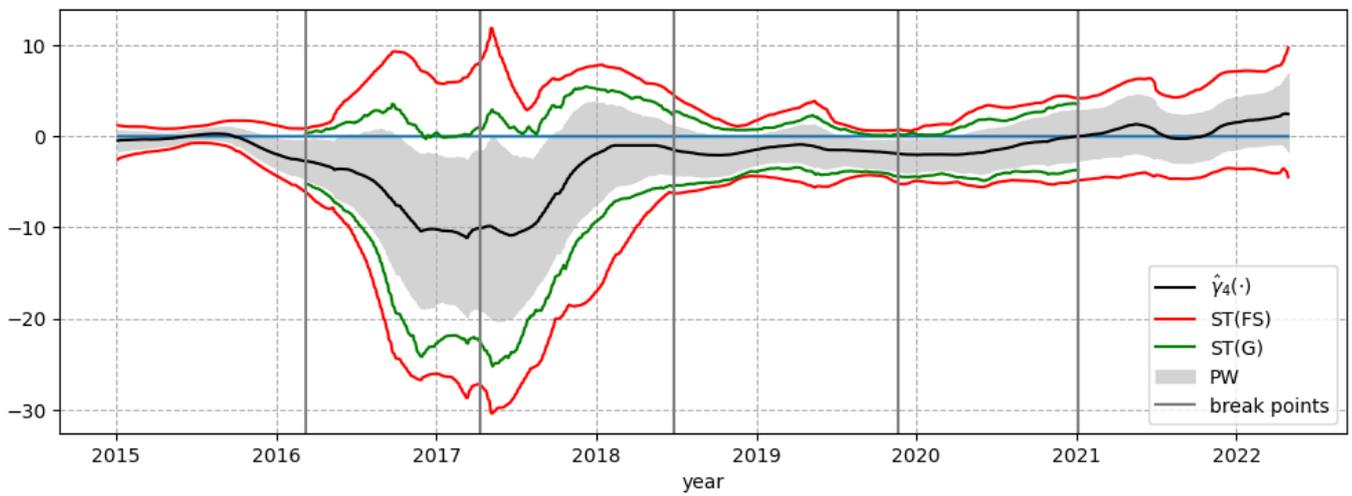

Figure 15: Estimated curve of $\gamma_4(\cdot)$ (black line) and resulting 95%-level confidence intervals (grey, labeled as PW) and bands using the SWB for Model (B.2), see Figure 5 for further information.