Metrika (2022) 85:771–808
https://doi.org/10.1007/s00184-021-00844-z

Estimating a gradual parameter change in an AR(1)-process

Marie Hušková[1] · Zuzana Prášková[1] · Josef G. Steinebach[2]

Received: 16 July 2020 / Accepted: 30 September 2021 / Published online: 20 October 2021

Abstract
We discuss the estimation of a change-point t_0 at which the parameter of a (non-stationary) AR(1)-process possibly changes in a gradual way. Making use of the observations $X_1, \ldots, X_n$, we shall study the least squares estimator $\widehat{t_0}$ for t_0, which is obtained by minimizing the sum of squares of residuals with respect to the given parameters. As a first result it can be shown that, under certain regularity and moment assumptions, $\widehat{t_0}/n$ is a consistent estimator for τ_0, where $t_0 = \lfloor n\tau_0 \rfloor$, with $0 < \tau_0 < 1$, i.e., $\widehat{t_0}/n \stackrel{P}{\to} \tau_0$ $(n \to \infty)$. Based on the rates obtained in the proof of the consistency result, a first, but rough, convergence rate statement can immediately be given. Under somewhat stronger assumptions, a precise rate can be derived via the asymptotic normality of our estimator. Some results from a small simulation study are included to give an idea of the finite sample behaviour of the proposed estimator.

Keywords AR(1)-process · Gradual change · Change-point estimator · Consistency · Convergence rate · Asymptotic normality

Mathematics Subject Classification 62F10 · 62F12 · 62M10

✉ Josef G. Steinebach
jost@math.uni-koeln.de

Marie Hušková
huskova@karlin.mff.cuni.cz

Zuzana Prášková
praskova@karlin.mff.cuni.cz

1 Department of Probability and Mathematical Statistics, Faculty of Mathematics and Physics, Charles University, Sokolovská 83, CZ, 18675 Praha 8, Czech Republic

2 Department of Mathematics and Computer Science, University of Cologne, Weyertal, 86–90, 50931 Cologne, Germany

1 Introduction and statistical framework

In this work we study the estimation of a change-point at which the parameter of a (non-stationary) AR(1)-process possibly changes in a gradual way. More precisely, we observe a time series $X_1, \ldots, X_n$ possessing the structure

$$X_t = (\beta_0 + \beta_1 g(t, t_0))X_{t-1} + e_t \quad (t = 1, 2, \ldots), \quad \text{with} \quad X_0 = e_0, \tag{1.1}$$

where $\{e_t\}_{t=0,1,\ldots}$ is a sequence of independent, identically distributed (i.i.d.) random variables with $Ee_0 = 0$, $0 < Ee_0^2 = \sigma^2$, $Ee_0^4 < \infty$, β_0, β_1 are unknown parameters satisfying

$$|\beta_0| < 1, \quad \beta_1 = \beta_{1,n} \to 0, \quad |\beta_1|\sqrt{n} \to \infty \quad (n \to \infty), \tag{1.2}$$

and $g(\cdot, t_0) = g_n(\cdot, t_0)$ is a (known) real function such that

$$g_n(t, t_0) = 0 \quad (t \le t_0) \quad \text{and} \quad g_n(t, t_0) > 0 \quad (t > t_0) \tag{1.3}$$

(see more detailed assumptions below). That is, we assume that the parameter β_0 of the (stationary) AR(1)-process changes *gradually* at an unknown time-point $t_0 = t_{0,n} = \lfloor n\tau_0 \rfloor$, with $0 < \tau_0 < 1$, $\lfloor \cdot \rfloor$ denoting the integer part, and our aim is to provide an estimator for t_0 making use of the observations $X_1, \ldots, X_n$ and under certain assumptions on the function $g_n(\cdot, t_0)$ to be specified below.

Remark 1 (a) As in earlier works, we only study the case of a gradual change under "local alternatives" here, i.e., under $\beta_{1,n} \to 0$, but $\beta_{1,n}\sqrt{n} \to \infty$ as $n \to \infty$ (cf., e.g., Dümbgen (1991), Jarušková (1998a), Hušková (1998a), Hušková (2001), or Hušková and Steinebach (2002)).

(b) In case of the gradual change function g_n being unknown, it would be sufficient to have an estimating function $\widehat{g}_n$ (say), which approximates g_n at a certain rate. For a more detailed discussion we refer to Remarks 5 and 7 below.

Note that, if $g(\cdot, t_0)$ is a bounded function, then

$$b := \sup_{t \ge 1} |\beta_0 + \beta_1 g(t, t_0)| < 1 \tag{1.4}$$

for sufficiently large n and, by a repeated application of (1.1),

$$X_t = e_t + \sum_{j=1}^{t} e_{t-j} \prod_{i=0}^{j-1} (\beta_0 + \beta_1 g(t-i, t_0)) \quad (t = 1, 2, \ldots). \tag{1.5}$$

Before we turn to formulating our main results, we give a brief account of related works, particularly concerning the detection of *gradual* changes in various dependent data sets. Most of the earlier papers on the change analysis in autoregressive processes deal with *abrupt* changes, either in the mean or in the autoregressive parameters, respectively in the variance of the error process. Picard (1985) proposed a procedure for testing changes in the covariance structure of an AR(p)-process based on a likelihood ratio approach and obtained the asymptotic distribution of the likelihood estimators of the change parameters. Gaussian-type likelihood ratio procedures for testing abrupt changes in autoregressive models were studied in Davis et al. (1995), where also the limit distribution of the test statistic was established. Gombay (2008) used efficient score vectors to develop statistics that are able to test changes in any of the parameters of a Gaussian AR(p)-model separately, or in any collection of them, and she also studied large sample properties of the change-point estimator.

Other results were based on partial sums of residual processes, see, e.g., Horváth (1993) for testing or Bai (1994) for proving consistency of the change-point estimator. Hušková et al. (2007) used an approach based on partial sums of weighted residuals and obtained asymptotic distributions for various max-type test statistics together with proving the consistency of the change-point estimator in an AR(p)-model. Moreover, bootstrap versions of the proposed tests were studied in Hušková et al. (2008).

A quasi-maximum likelihood method was used in Bai (2000) to analyze vector autoregressive models (VAR) possessing multiple structural changes. The approach developed in Davis et al. (1995) has been extended to VAR models by Dvořák (2015), who dealt with asymptotic tests for an abrupt change in the autoregressive parameters and the variance structure under various assumptions on the correlations of the errors. Kirch et al. (2015) extended the class of max-type change-point statistics considered in Hušková et al. (2007) to the VAR case and epidemic change alternatives and developed a new approach taking possible misspecification of the model into account.

Slama and Saggou (2017) considered a Bayesian analysis of a possible change in the parameters of an AR(p)-model and developed a test, which can detect a change in any of the parameters separately. Moreover, the posterior density of the change-point is given by using a Gibbs sampler. Many references on the change-point analysis in time series can also be found in a survey paper by Aue and Horváth (2013).

Concerning *gradual* changes in autoregression, Salazar (1982) studied a model similar to (1.1) from a Bayesian point of view. Under the assumption of normality of the error process and with some joint prior distribution of the change-point t_0 and other parameters under consideration, he obtained a joint posterior distribution from which the marginal distribution of the change-point could be obtained via numerical integration. A similar, though not identical problem was solved by Venkatesan and Arumugam (2007), who considered an AR(p)-model with a gradual switch in the parameters over a finite interval. Here again, computation of the posterior distribution of the change-point requires the use of numerical integration.

He et al. (2008) derived a parameter constancy test in a stationary vector autoregressive model against the hypothesis that the parameters of the model change smoothly over time. Though model (1.1) could be considered a special case of the model studied in He et al. (2008), the authors treat other type of smooth functions and do not consider any estimator of the breaking point.

Our approach below is motivated by the previous work by Hušková, see, e.g., Hušková (1998a, b, 1999, 2001), Jarušková (1998a, b, 1999, 2001, 2002, 2003), or by Hušková and Steinebach (2000), Hušková and Steinebach (2002), respectively by Albin and Jarušková (2003). In the above cited papers a gradual-type change in the mean of a location model is considered and asymptotic tests for detecting the change together with limit properties of the estimator of the change-point are developed for various types of smoothly changing parameters. More specifically, the mentioned model can be written in the form

$$Y_t = \mu + \delta_n\, g\big((t - t_0)/n\big) + \epsilon_t, \quad t = 1, \ldots, n, \tag{1.6}$$

where μ, δ_n, t_0 are unknown parameters, $\epsilon_1, \ldots, \epsilon_n$ are i.i.d. errors, with zero mean and finite moments of order $2 + \Delta$, $\Delta > 0$, and the function g satisfies the assumption

$$g(x) = 0, \quad (x \leq 0), \quad g(x) > 0 \quad (x > 0),$$

together with other assumptions specified for the formulated problems.

Döring and Jensen (2015), and also Döring (2015a, b), extended the methodology proposed by Hušková (1999) to regression models with independently distributed random regressors. Wang (2007) studied the same location model as Hušková (1999) with errors that exhibit long memory dependence, and Slabý (2001) considered a test based on ranks. Hlávka and Hušková (2017), motivated by gender differences observed in a real data set, proposed a two-sample gradual change test that leads to more precise results than the application of a procedure based on the standard two-sample t-test. Račkauskas and Tamulis (2013) studied epidemic changes in a location model, in which the transition between regimes is gradual.

Several authors studied smooth changes in other contexts. For example, Aue and Steinebach (2002) discuss an extension of Hušková's (1999) approach to certain statistical models, which cover more general classes of stochastic processes satisfying an invariance principle (see also Kirch and Steinebach (2006), Steinebach (2000), Steinebach and Timmermann (2011) or Timmermann (2014), Timmermann (2015)).

Vogt and Dette (2015) developed a nonparametric method to estimate a smooth change-point in a locally stationary framework and established the rate of convergence of the change-point estimator. Their procedure allows to deal with a wide variety of stochastic characteristics including the mean, covariances and higher moments.

Hoffmann et al. (2018) and Hoffmann and Dette (2019) discuss statistical inference for the detection and the localization of gradual changes in the jump characteristic of a discretely observed Itô semimartingale.

Quessy (2019) proposed a general class of consistent test statistics for the detection of gradual changes in copulas and developed their large-sample properties.

Now, let us turn to our problem. We shall study the least squares estimator $\widehat{t_0}$ for t_0, which is obtained by minimizing

$$S(b_0, b_1, t_*) = \sum_{t=1}^{n} [X_t - (b_0 + b_1 g(t, t_*)) X_{t-1}]^2$$

with respect to (w.r.t.) $b_0, b_1 \in \mathbb{R}$, $t_* = 0, 1, \ldots, \lfloor n(1-\delta)\rfloor$, $\delta > 0$ arbitrarily small, i.e.,

$$S(\widehat{b_0}, \widehat{b_1}, \widehat{t_0}) = \min_{b_0, b_1, t_*} S(b_0, b_1, t_*) = \min_{t_*} \min_{b_0, b_1} S(b_0, b_1, t_*). \tag{1.7}$$

Remark 2 The technical condition $t_* \leq \lfloor n(1-\delta)\rfloor$, with $\delta > 0$ fixed, could be weakened to allow for $\delta = \delta_n \to 0$ $(n \to \infty)$ at a certain rate, which, however, would depend on the parameter $\beta_1 = \beta_{1,n}$ from (1.2) and the function $g = g_n$ from (1.3) as well. Since $\beta_{1,n}$ is unknown, one should choose $\delta > 0$ fixed, but small, for practical use.

Via partial derivatives, it is not difficult to show that, for fixed t_*,

$$\widehat{b_0}(t_*) = \frac{\sum_{t=1}^n X_t X_{t-1}}{\sum_{t=1}^n X_{t-1}^2} - \widehat{b_1}(t_*) \frac{\sum_{t=1}^n g(t, t_*) X_{t-1}^2}{\sum_{t=1}^n X_{t-1}^2} \qquad \text{and} \tag{1.8}$$

$$\widehat{b_1}(t_*) = \frac{\sum_{t=1}^n X_t X_{t-1} g(t, t_*) - \frac{\sum_{j=1}^n X_j X_{j-1}}{\sum_{j=1}^n X_{j-1}^2} \sum_{t=1}^n X_{t-1}^2 g(t, t_*)}{\sum_{t=1}^n X_{t-1}^2 g^2(t, t_*) - \frac{\left(\sum_{t=1}^n g(t,t_*) X_{t-1}^2\right)^2}{\sum_{t=1}^n X_{t-1}^2}}. \tag{1.9}$$

On plugging this into (1.7), we obtain

$$\begin{aligned} S(\widehat{b_0}, \widehat{b_1}, \widehat{t_0}) = \min_{t_*} \Bigg[& \sum_{t=1}^n \left(X_t - \frac{\sum_{j=1}^n X_j X_{j-1}}{\sum_{j=1}^n X_{j-1}^2} X_{t-1} \right)^2 \\ & - \widehat{b_1}^2(t_*) \sum_{t=1}^n X_{t-1}^2 \left(g(t, t_*) - \frac{\sum_{j=1}^n X_{j-1}^2 g(j, t_*)}{\sum_{j=1}^n X_{j-1}^2} \right)^2 \Bigg]. \end{aligned} \tag{1.10}$$

Since the first term in (1.10) does not depend on t_*, a combination of (1.7)–(1.10) eventually results in

$$\widehat{t_0} = \arg\max_{t_*} \frac{\left[\sum_{t=1}^{n} X_t X_{t-1} g(t,t_*) - \frac{\sum_{j=1}^{n} X_j X_{j-1}}{\sum_{j=1}^{n} X_{j-1}^2} \sum_{t=1}^{n} X_{t-1}^2 g(t,t_*)\right]^2}{\sum_{t=1}^{n} X_{t-1}^2 g^2(t,t_*) - \frac{\left(\sum_{t=1}^{n} g(t,t_*) X_{t-1}^2\right)^2}{\sum_{t=1}^{n} X_{t-1}^2}}. \tag{1.11}$$

Remark 3 Note that

$$\max_{t_*} \frac{\left|\sum_{t=1}^{n} X_t X_{t-1} g(t,t_*) - \frac{\sum_{j=1}^{n} X_j X_{j-1}}{\sum_{j=1}^{n} X_{j-1}^2} \sum_{t=1}^{n} X_{t-1}^2 g(t,t_*)\right|}{\left(\sum_{t=1}^{n} X_{t-1}^2 g^2(t,t_*) - \frac{\left(\sum_{t=1}^{n} g(t,t_*) X_{t-1}^2\right)^2}{\sum_{t=1}^{n} X_{t-1}^2}\right)^{1/2}} \tag{1.12}$$

can be used as a test statistic for testing “no change” versus “there is a change”, even if the true function g is unknown, just some integral has to be nonzero (see, e.g., Hušková and Steinebach (2002)). In practice, before starting to estimate $t_0 = \lfloor n\tau_0 \rfloor$, one should first carry out such a test for the existence of a change-point τ_0, with $0 < \tau_0 < 1$.

For our theoretical studies of $\widehat{t_0}$ below, it will be convenient to make use of the model equation (1.1) and rewrite (1.11), after a multiplication with $1/n$, as

$$\begin{aligned}\widehat{t_0} = \arg\max_{t_*} & \frac{\left[\beta_1 \left(\frac{1}{n}\sum_{t=1}^{n} g(t,t_0) g(t,t_*) X_{t-1}^2 - \frac{\frac{1}{n}\sum_{j=1}^{n} g(j,t_0) X_{j-1}^2 \frac{1}{n}\sum_{j=1}^{n} g(j,t_*) X_{j-1}^2}{\frac{1}{n}\sum_{j=1}^{n} X_{j-1}^2}\right)\right.}{\frac{1}{n}\sum_{t=1}^{n} g^2(t,t_*) X_{t-1}^2 - \frac{\left(\frac{1}{n}\sum_{j=1}^{n} g(j,t_*) X_{j-1}^2\right)^2}{\frac{1}{n}\sum_{j=1}^{n} X_{j-1}^2}} \\ & \frac{\left. + \frac{1}{n}\sum_{t=1}^{n} e_t X_{t-1} g(t,t_*) - \frac{\frac{1}{n}\sum_{j=1}^{n} e_j X_{j-1} \frac{1}{n}\sum_{j=1}^{n} g(j,t_*) X_{j-1}^2}{\frac{1}{n}\sum_{j=1}^{n} X_{j-1}^2}\right]^2}{\frac{1}{n}\sum_{t=1}^{n} g^2(t,t_*) X_{t-1}^2 - \frac{\left(\frac{1}{n}\sum_{j=1}^{n} g(j,t_*) X_{j-1}^2\right)^2}{\frac{1}{n}\sum_{j=1}^{n} X_{j-1}^2}}.\end{aligned} \tag{1.13}$$

For later asymptotics it may also be convenient to express $\widehat{t_0}$ as

$$\widehat{t_0} = \arg\max_{t_*} \frac{\left[\beta_1 \frac{1}{n}\sum_{t=1}^{n} \widetilde{g}_n(t,t_0) \widetilde{g}_n(t,t_*) X_{t-1}^2 + \frac{1}{n}\sum_{t=1}^{n} e_t X_{t-1} \widetilde{g}_n(t,t_*)\right]^2}{\frac{1}{n}\sum_{t=1}^{n} \widetilde{g}_n^2(t,t_*) X_{t-1}^2}, \tag{1.14}$$

where

$$\widetilde{g}_n(t,t_*) = g(t,t_*) - \frac{\sum_{j=1}^{n} g(j,t_*) X_{j-1}^2}{\sum_{j=1}^{n} X_{j-1}^2}. \tag{1.15}$$

The paper is organized as follows. Based on the required assumptions, which are collected first, Sect. 2 presents the main results of our work. As a first statement it can be shown in Theorem 1 that $\widehat{t_0}/n$ is a consistent estimator for τ_0, where $t_0 = \lfloor n\tau_0 \rfloor$, with $0 < \tau_0 < 1$, i.e., $\widehat{t_0}/n \xrightarrow{P} \tau_0$ $(n \to \infty)$. Based on the rates obtained in the proof of Theorem 1, a rough convergence rate estimate can immediately be given (see Theorem 2). Under somewhat stronger assumptions, a precise rate can then be derived in Theorem 3 by showing that our estimator has an asymptotically normal limit distribution. In Sect. 3, some results from a small simulation study are included to give an idea of the finite sample behaviour of the proposed estimator. Section 4 collects some auxiliary results which are used in the proofs of the main theorems. The latter are finally given in Sect. 5.

2 Assumptions and main results

For our asymptotic results we assume the gradual change function $g(\cdot, t_*)$ to satisfy the following assumptions:

(A.1) For every $t_* = 0, 1, \ldots, n-1$, the function $g(\cdot, t_*)$ is of the form

$$g(t, t_*) = g_0\Big(\frac{t - t_*}{n}\Big), \qquad t = 0, 1, \ldots, n,$$

where $g_0 : (-\infty, 1] \to \mathbb{R}$ is a real function satisfying:

(A.2) It holds that

$$g_0(x) = 0 \qquad (x \le 0) \qquad \text{and} \qquad g_0(x) > 0 \qquad (0 < x \le 1).$$

(A.3) The function $g_0 : (-\infty, 1] \to \mathbb{R}$ is bounded and Lipschitz continuous, i.e.

$$|g_0(x)| \le D_1 \qquad \text{and} \qquad |g_0(x) - g_0(y)| \le D_2|x - y|, \qquad x, y \le 1,$$

with some positive constants D_1 and D_2.

(A.4) It holds that

$$|g_0(x) - g_0(y) - (x - y)g_0'(y)| \le D_3|x - y|^{1+\Delta}, \quad 0 < x, y < 1,$$
$$|g_0(x) - x g_{0+}'(0)| \le D_3|x|^{1+\Delta}, \quad 0 \le x < 1,$$
$$\Big(\int_0^1 \widetilde{g}_0(x - \tau_0)\, \widetilde{g}_0'(x - \tau_0) dx\Big)^2 < \int_0^1 \widetilde{g}_0^2(x - \tau_0) dx \int_0^1 \widetilde{g}_0'^2(x - \tau_0) dx,$$
$$\text{with} \quad \widetilde{g}_0(x) = g_0(x) - \int_0^1 g_0(y - \tau_0) dy,$$

where $g_{0+}'(0)$ denotes the right derivative at 0, $g_0'(\cdot)$ denotes the derivative, assumed to be bounded and Riemann integrable, $1/2 < \Delta \le 1$, and D_3 is a positive constant.

Remark 4 (a) The case of a negative change, i.e., $g_0(x) < 0$ for $0 < x \le 1$, can be reduced to the positive one by just reparametrizing $\bar{\beta}_1 := -\beta_1$ and $\bar{g}_0(\cdot) = -g_0(\cdot)$.

(b) The function g_0, for example, could be such that $g_0(x) = \pm x_+^{\kappa}$ $(x \leq 1)$, where x_+ denotes the positive part of x and $\kappa \geq 1$ is a fixed exponent.

In our first result it will be shown that $\widehat{t_0}/n$ is a consistent estimator for τ_0, where $t_0 = \lfloor n\tau_0 \rfloor$, with $0 < \tau_0 < 1$, i.e., $\widehat{t_0}/n \xrightarrow{P} \tau_0$ $(n \to \infty)$.

Theorem 1 *Let Assumptions* (A.1)–(A.3) *be satisfied. Then, under the model* (1.1) *and the corresponding conditions formulated above, the estimator* $\widehat{t_0}$ *from* (1.11) *is consistent, i.e.,*

$$\frac{\widehat{t_0}}{n} \xrightarrow{P} \tau_0 \qquad (n \to \infty). \tag{2.1}$$

Remark 5 In case of an unknown change function g, it will be obvious from the proof of Theorem 1 that (2.1) still holds, if g in (1.11) is replaced by an estimator $\widehat{g}_n$ at a rate $o_P(\beta_1)$, more precisely, if there is an estimating function $\widehat{g}_0 = \widehat{g}_{0,n}$ such that, as $n \to \infty$,

$$\max_{x \in [0,1]} \left|\widehat{g}_{0,n}(x) - g_0(x)\right| = o_P(\beta_1). \tag{2.2}$$

In this case, $\widehat{g}_n(t, t_*) = \widehat{g}_0((t - t_*)/n)$ can be used in (1.11) resp. (1.13) instead of $g(t, t_*) = g_0((t - t_*)/n)$ and, in view of the rate $o_P(\beta_1)$, the convergence in (2.1) will be retained.

If, for example, $g_0(x) = x_+^{\kappa}$, with some $\kappa \geq 1$, it would be sufficient to have an estimator $\widehat{\kappa}_n$ such that $|\widehat{\kappa}_n - \kappa| = o_P(\beta_1)$, e.g., $|\widehat{\kappa}_n - \kappa| = O_P(1/\sqrt{n})$ as $n \to \infty$. Such estimates have been obtained in other settings (cf., e.g., Döring and Jensen (2015) for a regression model). In our time series setting, it is an open question and has to be left for future research.

Another possible model to deal with would be the case $\beta_1 g_0(x) = \beta_1 x_+ + \beta_2 x_+^2$, with unknown parameters β_1, β_2. Here, least squares estimation means to minimize

$$\widetilde{S}(b_0, b_1, b_2, t_*) = \sum_{t=1}^{n} \left[X_t - \left(b_0 + b_1\left(\frac{t - t_*}{n}\right)_+ + b_2\left(\frac{t - t_*}{n}\right)_+^2\right) X_{t-1}\right]^2$$

w.r.t. $b_0, b_1, b_2 \in \mathbb{R}$, $t_* = 0, 1, \ldots, \lfloor n(1 - \delta)\rfloor$, $\delta > 0$, and then to modify the corresponding steps in the proofs. For the sake of conciseness, this may also be left for further investigations.

The proofs of Theorem 1 and Remark 5 are postponed to Sect. 5.

Remark 6 It would also be quite straightforward to get a consistent estimator of β_1, i.e., $\widehat{b_1}(\widehat{t_0})$, together with some limiting properties. For the sake of conciseness, we want to omit details here.

Also, if the function $g(\cdot)$ is only known up to a multiplicative constant, then the resulting estimator is still consistent, but the limit distribution below, however, would depend on this multiplicative constant, which is unknown.

On checking the proof of Theorem 1 more carefully, a rough rate of consistency for our estimator $\widehat{t_0}$ can be obtained as follows.

Theorem 2 *Under the conditions of Theorem 1, assume that the limit function $f(\tau_*)$ in (5.4) is twice continuously differentiable in a small neighbourhood of τ_0, with $f''(\tau_*) > D$ for some $D > 0$. Then, with $\widehat{t_0} = \lfloor n\widehat{\tau_0} \rfloor$, for every sequence $\{\varepsilon_n\}$ with $\varepsilon_n \to 0$,*

$$|\widehat{\tau_0} - \tau_0| = O_P\big(|\beta_1|^{1/2}\big) + o_P\left(\frac{1}{|\beta_1|^{1/2}\,\varepsilon_n\, n^{1/4}}\right) \quad (n \to \infty). \tag{2.3}$$

Remark 7 If (2.2) in Remark 5 is replaced by

$$\max_{x\in[0,1]} \big|\widehat{g}_{0,n}(x) - g_0(x)\big| = O_P\Big(\frac{1}{\sqrt{n}}\Big), \tag{2.4}$$

the approximation rate $O_P\big(|\beta_1|^{1/2}\big) + o_P\big(1/(|\beta_1|^{1/2}\,\varepsilon_n\, n^{1/4})\big)$ in Theorem 2 can be retained.

The proofs of Theorem 2 and Remark 7 are also postponed to Sect. 5.

Remark 8 If, for example, $\beta_1 = n^{-\alpha}$, with $0 < \alpha < 1/2$, then ε_n could be chosen as $(\log n)^{-p}$, with $p > 0$, so that one would have the polynomial consistency rate

$$|\widehat{\tau_0} - \tau_0| = \begin{cases} O_P\big(\frac{1}{n^{\alpha/2}}\big), & \text{if} \quad 0 < \alpha < 1/4, \\ o_P\big(\frac{\log^p n}{n^{1/4-\alpha/2}}\big), & \text{if} \quad 1/4 \le \alpha < 1/2, \end{cases} \qquad (n \to \infty). \tag{2.5}$$

Next it will be shown that the estimator $\widehat{t_0}$ of t_0 (or equivalently $\widehat{\tau_0}$ of τ_0) has an asymptotically normal limit distribution.

Theorem 3 *Let Assumptions (A.1)–(A.4) be satisfied. Then, as $n \to \infty$,*

$$\beta_1\sqrt{n}\,\widetilde{H}\frac{\sigma^2}{1-\beta_0^2}\frac{\widehat{t_0} - t_0}{n} \xrightarrow{d} \mathcal{N}\Big(0, \frac{\sigma^4}{1-\beta_0^2}\widetilde{H}\Big), \tag{2.6}$$

or, in a standardized form,

$$\frac{\beta_1}{\sqrt{1-\beta_0^2}}\sqrt{\widetilde{H}}\sqrt{n}\,\frac{\widehat{t_0} - t_0}{n} \xrightarrow{d} \mathcal{N}(0, 1), \tag{2.7}$$

equivalently

$$\frac{\beta_1}{\sqrt{1-\beta_0^2}}\sqrt{\widetilde{H}}\sqrt{n}\,(\widehat{\tau_0} - \tau_0) \xrightarrow{d} \mathcal{N}(0, 1), \tag{2.8}$$

where

$$\widetilde{H} = \int_0^1 \widetilde{g}_0'^2(x-\tau_0)dx - \frac{\left(\int_0^1 \widetilde{g}_0(x-\tau_0)\widetilde{g}_0'(x-\tau_0)dx\right)^2}{\int_0^1 \widetilde{g}_0^2(x-\tau_0)dx}. \tag{2.9}$$

Remark 9 Note that the limit distribution in Theorem 3 does not depend on σ^2.

For the proof of Theorem 3 see also Sect. 5.

3 Some simulations

In this section, before we turn to the proofs of Theorems 1–3, we first present some results from a small simulation study. We simulated observations of the time series (1.1) with the function $g_0(x) = x_+, x \le 1$, for various combinations of β_0 and β_1 and for various change-points t_0. The errors e_t were considered to be i.i.d. with a standard normal distribution. The first 50 simulated values were deleted to start computations with stationary observations X_t for $t = 1, \ldots, t_0$. We simulated either $n = 500$, 1000 or 5000 observations of (1.1). For each realization of $\{X_t, t = 1, \ldots, n\}$, we estimated the change-point $\widehat{t_0}$ according to (1.11) with the given function g_0 and for t_* running from 0 to $\lfloor n(1-\delta)\rfloor$, where $\delta > 0$ denotes the proportion of observations, which were excluded. For each combination we used 10 000 simulation runs.

The change-point was chosen to be either $t_0 = n/4$, $n/2$ or $3/4n$, $\tau_0 = t_0/n$, and $\delta = 0.05$. The parameters β_0 and β_1 should satisfy the asymptotic relation (1.2), i.e., $|\beta_0| < 1$, $\beta_1 \to 0$, and $|\beta_1|\sqrt{n} \to \infty$ as $n \to \infty$. For β_1 we used small multiples of $1/\sqrt{\log\log n}$ such that the asymptotic condition (1.2) and also the condition $|\beta_0 + \beta_1 g_0((t-t_0)/n)| < 1$ for $t = 1, \ldots, \lfloor n(1-\delta)\rfloor$ were satisfied. Though in fact β_1 depends on n, we used the same value always for all considered sample sizes n, to make the presentation of the results more transparent.

In Tables 1, 2, 3, 4, the estimator $\widehat{\tau_0}$ computed as $\widehat{\tau_0} = \widehat{t_0}/n$, the 90% confidence interval for τ_0 computed from the Monte Carlo percentiles of $\sqrt{n}(\widehat{\tau_0} - \tau_0)$, and estimators of β_0, β_1 computed from (1.8) and (1.9) with $t_* = \widehat{t_0}$ are presented. The empirical standard deviations of the corresponding point estimators are given in parentheses.

It can be seen that the point estimators of τ_0 (or t_0, respectively) and of β_0 are systematically slightly underestimated, but otherwise behave quite well for all considered variants. The estimators of β_1 are more volatile, but they converge to the true value with growing number of observations. To study the behaviour of the estimator of the change-point in finite samples, we also displayed histograms and Q-Q plots of the standardized statistic in (2.8). Some results are presented in Figs. 1, 2, 3, 4, 5, 6. The simulations show that the behaviour of the estimator $\widehat{t_0}$ depends both on the size of the sample and on the parameters β_0 and β_1. The assumption $|\beta_0 + \beta_1 g_0((t-t_0)/n)| < 1$ guarantees the stability of the time series (1.1) even after the local change, expressed by the time varying part of the autoregressive coefficient. Due to the local character of the change the convergence to the normal distribution is slow and it can only be

Table 1 Estimators of τ_0, β_0, β_1; $t_0 = n/2$

n	β_0	β_1	τ_0	$\hat{\tau}_0$	CI-90	$\hat{\beta}_0$	$\hat{\beta}_1$
500	0	1.8	0.5	0.4834	(0.3454; 0.6954)	−0.0101	1.8084
				(0.1108)		(0.0752)	(0.4584)
1000				0.4924	(0.4004; 0.6129)	−0.0044	1.7934
				(0.0680)		(0.0485)	0.2764
5000				0.4987	(0.4592; 0.5423)	−0.0004	1.7956
				(0.0253)		(0.0205)	(0.1116)
500	0.3	1.2	0.5	0.4699	(0.2729; 0.8539)	0.2821	1.2632
				(0.1617)		(0.0769)	(0.6576)
1000				0.4849	(0.3529; 0.6934)	0.2916	1.2139
				(0.1073)		(0.0507)	(0.2828)
5000				0.4974	(0.4426; 0.5626)	0.2988	1.1984
				(0.0367)		(0.0200)	(0.1056)
500	0.5	0.8	0.5	0.4676	(0.1856; 0.9496)	0.4809	0.9647
				(0.2085)		(0.0701)	(0.9136)
1000				0.4743	(0.2863; 0.8113)	0.4886	0.8559
				(0.1533)		(0.0490)	(0.8657)
5000				0.4958	(0.4172; 0.5906)	0.4982	0.8035
				(0.0538)		(0.0187)	(0.1031)

expected for sufficiently large samples. In Figs. 1, 2, 3, 4, 5, 6 convergence to normality is demonstrated for respective sample sizes $n = 500$ (left panels) and $n = 5000$ (right panels). The role of the parameter β_1, which represents the rate of the change of the autoregressive coefficient, can be seen on comparing Figs. 3 and 4. In Fig. 3, the value of $\beta_0 + \beta_1 g_0((t - t_0)/n)$ changes from 0 to 0.9, while in Fig. 4 it varies from 0.5 to 0.9 for the same values of $t = t_0 + 1, \dots, n$. The gradual change in the first case is much faster than in the second one, where the change is slow and the increments to the autoregressive coefficient are very small. In this case, the graph of the function on the right-hand side of (1.11) can be very flat and its global maximum can be incorrectly detected or it is detected either very soon or very late. This explains the large values of the side column in the histogram and the large skewness of the test statistic, especially in the left part of Fig. 4. Also the position of t_0, and thus the length of the stationary part of the time series under consideration, plays a role. In general, we observe better results, when the change occurs in the middle of the observed time intervals. For smaller sample sizes the kurtosis of the standardized statistics is larger and the finite sample distribution has heavier tails than the normal distribution, but it improves with growing sample size.

Table 2 Estimators of τ_0, β_0, β_1; $t_0 = n/4$

n	β_0	β_1	τ_0	$\hat{\tau}_0$	CI-90	$\hat{\beta}_0$	$\hat{\beta}_1$
500	0	1.2	0.25	0.2398	(0.0438; 0.4398)	− 0.0157	1.2201
				(0.1220)		(0.0972)	0.2178
1000				0.2383	(0.0983; 0.4563)	− 0.0120	1.2037
				(0.0993)		(0.0770)	(0.1457)
5000				0.2472	(0.1876; 0.3177)	− 0.0024	1.1993
				(0.0402)		(0.0312)	(0.0612)
500	0.3	0.9	0.25	0.2329	(− 0.0091; 0.4729)	0.2797	0.9215
				(0.1522)		(0.0949)	(0.2184)
1000				0.2325	(0.0685; 0.4745)	0.2847	0.9032
				(0.1175)		(0.0734)	(0.1222)
5000				0.2445	(0.1735; 0.3364)	0.2962	0.8980
				(0.0522)		(0.0318)	(0.0500)
500	0.5	0.6	0.25	0.2472	(− 0.1008; 0.4872)	0.4837	0.6680
				(0.1894)		(0.0801)	(0.5534)
1000				0.2374	(0.0004; 0.4824)	0.4868	0.6218
				(0.1494)		(0.0624)	(0.1413)
5000				0.2387	(0.1402; 0.3955)	0.4941	0.6002
				(0.0765)		(0.0312)	(0.0492)

Table 3 Estimators of τ_0, β_0, β_1; $t_0 = 3n/4$

n	β_0	β_1	τ_0	$\hat{\tau}_0$	CI-90	$\hat{\beta}_0$	$\hat{\beta}_1$
500	0	3.6	0.75	0.7346	(0.6466; 0.8806)	− 0.0050	3.6483
				(0.0826)		(0.0543)	(1.5051)
1000				0.7443	(0.6843; 0.8273)	− 0.0018	3.5945
				(0.0463)		(0.0378)	(0.8508)
5000				0.7486	(0.7234; 0.7786)	− 0.0004	3.5824
				(0.0169)		(0.0167)	(0.3220)
500	0.3	2.2	0.75	0.7122	(0.5682; 1.0682)	0.2875	2.5911
				(0.1579)		(0.0582)	(2.0569)
1000				0.7333	(0.6323; 0.9073)	0.2947	2.3065
				(0.0989)		(0.0385)	(1.1574)
5000				0.7476	(0.7026; 0.8019)	0.2991	2.0645
				(0.0309)		(0.0161)	(0.3533)
500	0.5	1.6	0.75	0.6958	(0.5258; 1.2448)	0.4853	1.9656
				(0.1961)		(0.0548)	(1.9813)
1000				0.7229	(0.6019; 0.9759)	0.4928	1.7661
				(0.1289)		(0.0366)	(1.0926)
5000				0.7473	(0.6963; 0.8101)	0.4990	1.6135
				(0.0360)		(0.0146)	(0.3064)

Table 4 Estimators of τ_0, β_0, β_1; $t_0 = n/2$ (negative values of β_0)

n	β_0	β_1	τ_0	$\hat{\tau}_0$	CI-90	$\hat{\beta}_0$	$\hat{\beta}_1$
500	−0.8	3.4	0.5	0.4949	(0.4389; 0.5669)	− 0.7953	3.3398
				(0.0388)		(0.0400)	(0.3545)
1 000				0.4980	(0.4590; 0.5430)	− 0.7953	3.3398
				(0.0255)		(0.0277)	(0.2407)
5 000				0.4994	(0.4832; 0.5171)	− 0.7995	3.3908
				(0.0104)		(0.0121)	(0.1010)
500	−0.5	2.5	0.5	0.4926	(0.3454; 0.6954)	− 0.5012	2.4940
				(0.0687)		(0.0576)	(0.4363)
1 000				0.4973	(0.4303; 0.5758)	− 0.4996	2.4964
				(0.0442)		(0.0402)	(0.2967)
5 000				0.4997	(0.4719; 0.5281)	− 0.4998	2.4994
				(0.0172)		(0.0175)	(0.1226)
500	−0.3	2.4	0.5	0.4904	(0.3924; 0.6244)	− 0.3030	2.3728
				(0.0743)		(0.0652)	(0.4055)
1 000				0.4958	(0.4288; 0.5788)	− 0.3020	2.3869
				(0.0461)		(0.0438)	(0.2679)
5 000				0.4990	(0.4702; 0.5314)	− 0.3001	2.3938
				(0.0185)		(0.0195)	(0.1118)

4 Some auxiliary results

In this section, we collect a series of auxiliary results, which are used in the proofs of our Theorems 1–3. In the sequel, C denotes a generic positive constant, independent of t, t_* and n, which may vary from case to case.

Lemma 1 *Under the assumptions of Theorem 1, as $n \to \infty$,*

$$\max_{t=1,\ldots,n} EX_{t-1}^2 = O(1) \quad \textit{and} \quad \max_{t=1,\ldots,n} EX_{t-1}^4 = O(1), \tag{4.1}$$

$$\Big|\frac{1}{n}\sum_{t=1}^{n} EX_{t-1}^2 - \frac{\sigma^2}{1-\beta_0^2}\Big| = O\big(|\beta_1|\big), \tag{4.2}$$

$$\Big|\frac{1}{n}\sum_{t=1}^{n} g(t,t_0)EX_{t-1}^2 - \frac{\sigma^2}{1-\beta_0^2}\int_0^1 g_0(x-\tau_0)dx\Big| = O\big(|\beta_1|\big). \tag{4.3}$$

Proof (a) In view of (1.5) and the independence and moment assumptions on $\{e_t\}$,

$$EX_{t-1}^2 = \sigma^2 \sum_{j=0}^{t-1} c_j^2 \quad \text{with} \quad c_0 = 1,\; c_j = \prod_{i=0}^{j-1}\big(\beta_0 + \beta_1 g(t-1-i, t_0)\big),\; j = 1, \ldots, t-1.$$

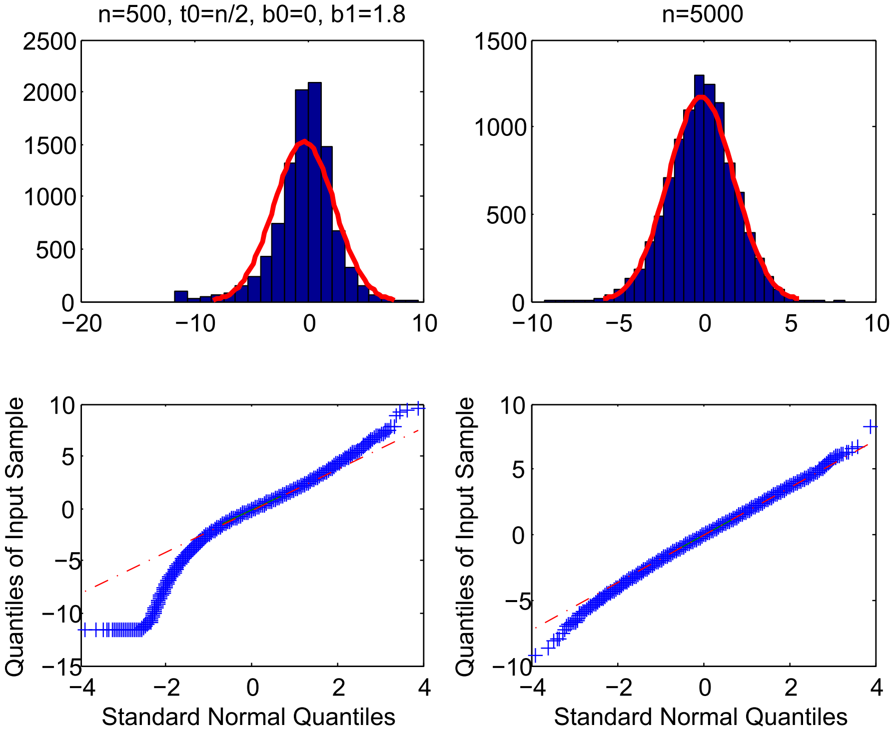

Fig. 1 Histograms and Q-Q plots of $\sqrt{n}(\widehat{\tau}_0 - \tau_0)$, left: $n = 500$, right: $n = 5\,000$, $t_0 = n/2$, $\beta_0 = 0$, $\beta_1 = 1.8$

Now, according to (1.2), for sufficiently large n,

$$\begin{aligned}&\big(\beta_0 + \beta_1 g(t-1-i, t_0)\big)^2\\&\quad = \beta_0^2 + \beta_1\big(2\beta_0 g(t-1-i, t_0) + \beta_1 g^2(t-1-i, t_0)\big) \le \beta_0^2 + |\beta_1|\, C =: q_1,\end{aligned}$$

where $0 \le q_1 < 1$. So, $c_j^2 \le q_1^j$ and

$$EX_{t-1}^2 \le \sigma^2 \frac{1-q_1^t}{1-q_1} \le \sigma^2 \frac{1}{1-q_1} = O(1).$$

Again due to the moment and independence assumptions on $\{e_t\}$, a similar estimation yields

$$EX_{t-1}^4 = Ee_0^4 \sum_{j=0}^{t-1} c_j^4 + 3\sigma^4 \sum_{j \ne k} c_j^2 c_k^2 \le Ee_0^4 \sum_{j=0}^{t-1} c_j^4 + 3\sigma^4 \Big(\sum_{j=0}^{t-1} c_j^2\Big)^2 = O(1),$$

which proves (4.1).

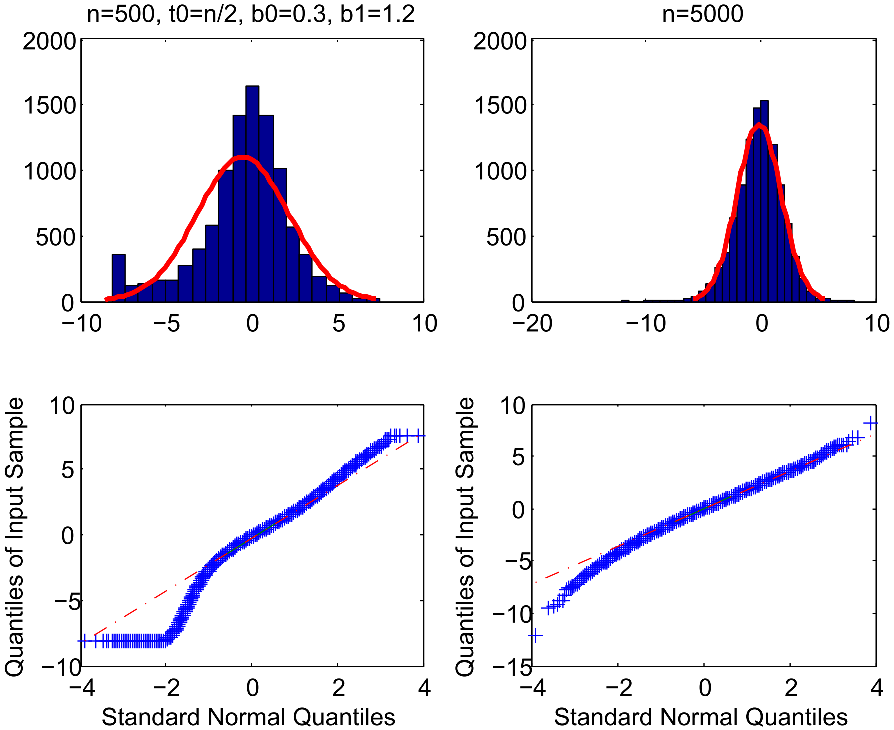

Fig. 2 Histograms and Q-Q plots of $\sqrt{n}(\widehat{\tau}_0 - \tau_0)$, left: $n = 500$, right: $n = 5\,000$, $t_0 = n/2$, $\beta_0 = 0.3$, $\beta_1 = 1.2$

(b) Moreover, since $\beta_1 \to 0$ as $n \to \infty$,

$$EX_{t-1}^2 - \frac{\sigma^2}{1-\beta_0^2} \le \sigma^2\Big(\frac{1}{1-\beta_0^2-|\beta_1|C} - \frac{1}{1-\beta_0^2}\Big)$$
$$= \sigma^2\frac{|\beta_1|C}{(1-\beta_0^2-|\beta_1|C)(1-\beta_0^2)} \le C|\beta_1|.$$

Analogously,

$$EX_{t-1}^2 \ge \sigma^2\frac{1-q_2^t}{1-q_2},$$

with $0 \le q_2 := \beta_0^2 - |\beta_1|C < 1$, and, as $n \to \infty$,

$$EX_{t-1}^2 - \frac{\sigma^2}{1-\beta_0^2} \ge \sigma^2\Big(\frac{1}{1-\beta_0^2+|\beta_1|C} - \frac{1}{1-\beta_0^2}\Big) - \sigma^2\frac{q_2^t}{1-q_2} \ge -C\big(|\beta_1| + q_2^t\big).$$

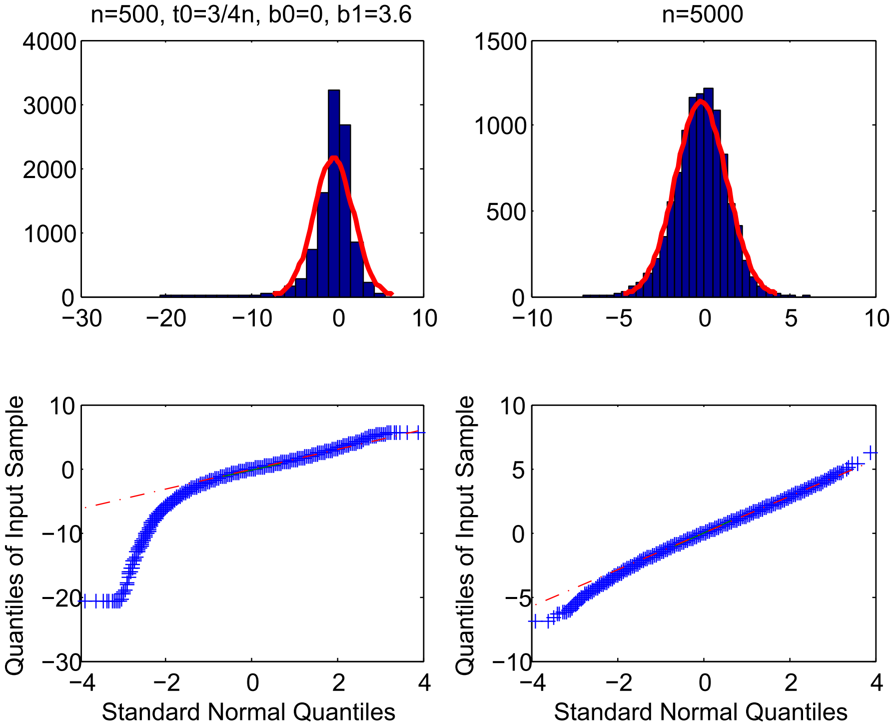

Fig. 3 Histograms and Q-Q plots of $\sqrt{n}(\widehat{\tau}_0 - \tau_0)$, left: $n = 500$, right: $n = 5\,000$, $t_0 = 3n/4$, $\beta_0 = 0$, $\beta_1 = 3.6$

Hence,

$$\left|EX_{t-1}^2 - \frac{\sigma^2}{1-\beta_0^2}\right| \leq C\big(|\beta_1| + q_2^t\big), \tag{4.4}$$

which suffices to prove (4.2), since $\sum_{t=1}^n q_2^t \leq C$ and $\beta_1 \gg 1/n$ as $n \to \infty$.

(c) In view of (4.4) and the assumptions on β_1 and g, as $n \to \infty$,

$$\begin{aligned}
&\left|\frac{1}{n}\sum_{t=1}^n g(t,t_0)EX_{t-1}^2 - \frac{\sigma^2}{1-\beta_0^2}\int_0^1 g_0(x-\tau_0)dx\right| \\
&\leq \frac{1}{n}\sum_{t=1}^n \left|g(t,t_0)\right|\left|EX_{t-1}^2 - \frac{\sigma^2}{1-\beta_0^2}\right| + \frac{\sigma^2}{1-\beta_0^2}\left|\frac{1}{n}\sum_{t=1}^n g(t,t_0) - \int_0^1 g_0(x-\tau_0)dx\right| \\
&\leq C\Big(|\beta_1| + \frac{1}{n}\sum_{t=1}^n q_2^t + \sum_{t=1}^n \int_{\frac{t-1}{n}}^{\frac{t}{n}} \Big|g_0\Big(\frac{t-t_0}{n}\Big) - g_0(x-\tau_0)\Big|dx\Big) \\
&\leq C\Big(|\beta_1| + \frac{1}{n} + \sum_{t=1}^n \int_{\frac{t-1}{n}}^{\frac{t}{n}} \Big|\frac{t}{n} - x - \frac{t_0}{n} + \tau_0\Big|dx\Big) \leq C\Big(|\beta_1| + \frac{1}{n}\Big) = O\big(|\beta_1|\big),
\end{aligned}$$

which completes the proof. □

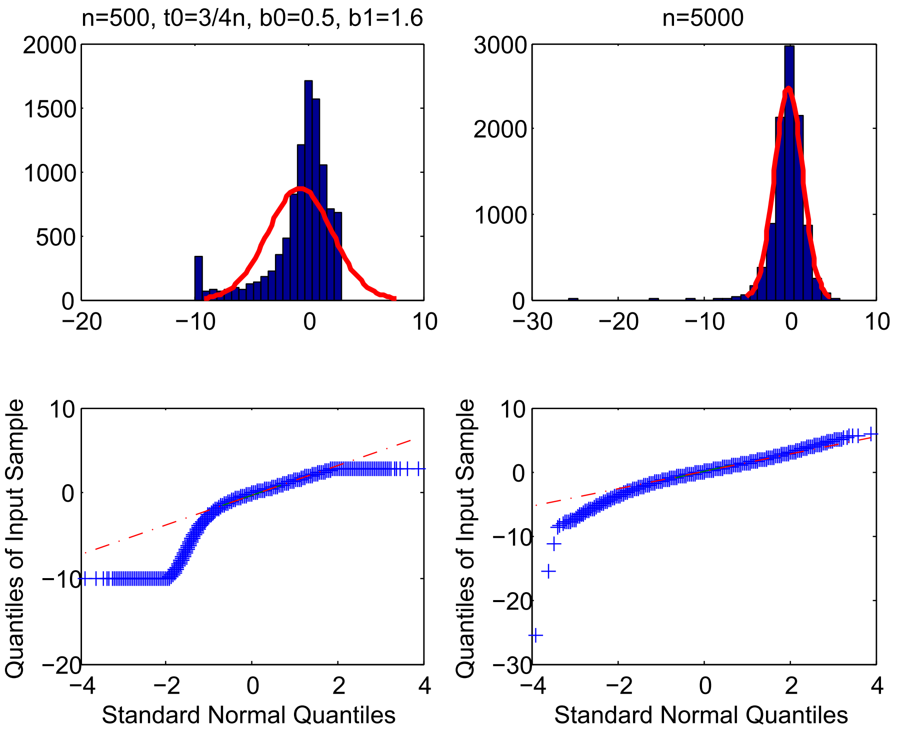

Fig. 4 Histograms and Q-Q plots of $\sqrt{n}(\widehat{\tau}_0 - \tau_0)$, left: $n = 500$, right: $n = 5\,000$, $t_0 = 3n/4$, $\beta_0 = 0.5$, $\beta_1 = 1.6$

Lemma 2 *Under the assumptions of Theorem 1, as $n \to \infty$, with $t_* = \lfloor n\tau_* \rfloor$,*

$$\max_{t_*} \left| \frac{1}{n} \sum_{t=1}^{n} g(t, t_0) g(t, t_*) E X_{t-1}^2 - \frac{\sigma^2}{1-\beta_0^2} \int_0^1 g_0(x - \tau_0) g_0(x - \tau_*) dx \right| = O\big(|\beta_1|\big), \tag{4.5}$$

$$\max_{t_*} \left| \frac{1}{n} \sum_{t=1}^{n} g(t, t_*) E X_{t-1}^2 - \frac{\sigma^2}{1-\beta_0^2} \int_0^1 g_0(x - \tau_*) dx \right| = O\big(|\beta_1|\big), \tag{4.6}$$

and

$$\max_{t_*} \left| \frac{1}{n} \sum_{t=1}^{n} g^2(t, t_*) E X_{t-1}^2 - \frac{\sigma^2}{1-\beta_0^2} \int_0^1 g_0^2(x - \tau_*) dx \right| = O\big(|\beta_1|\big). \tag{4.7}$$

Proof The proof of (4.5) – (4.7) is similar to that of (4.3), so that details can be omitted. Note that the functions $g_0(\cdot - \tau_0) g_0(\cdot - \tau_*)$, $g_0(\cdot - \tau_*)$ and $g_0^2(\cdot - \tau_*)$ are also bounded and Lipschitz continuous, uniformly in $0 \le \tau_* \le 1$. □

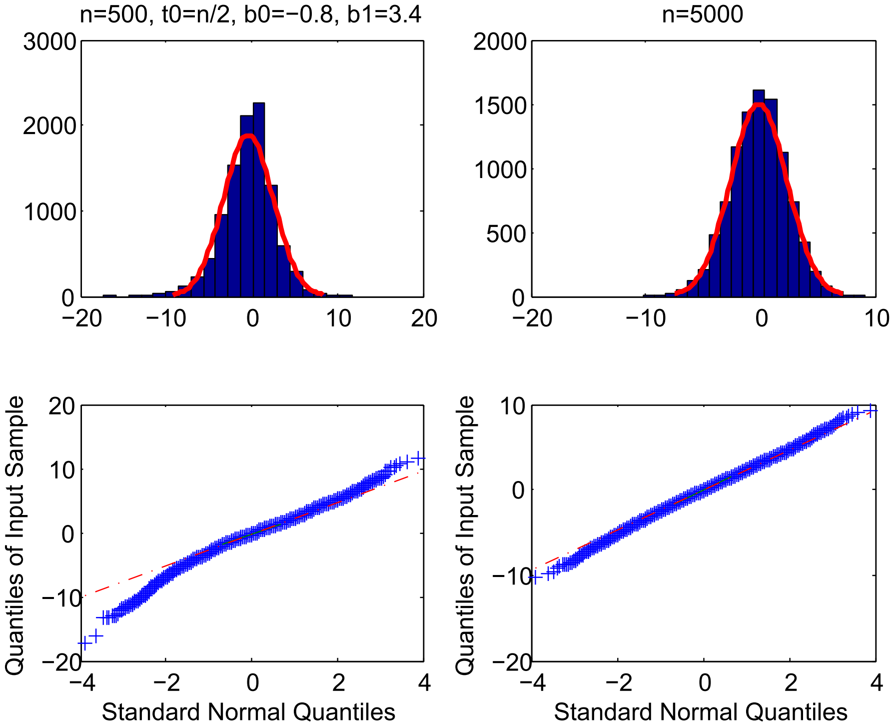

Fig. 5 Histograms and Q-Q plots of $\sqrt{n}(\widehat{\tau}_0 - \tau_0)$, left: $n = 500$, right: $n = 5\,000$, $t_0 = n/2$, $\beta_0 = -0.8$, $\beta_1 = 3.4$

Remark 10 It is obvious from the proofs of Lemmas 1 and 2 that, if the function g_0 is just continuous (instead of Lipschitz continuous), assertions (4.3)–(4.7) still hold true with $t_* = \lfloor n\tau_* \rfloor$, $\tau_* \in [0, 1]$ fixed (instead of $\max_{t_*}$), but with $O\big(|\beta_1|\big)$ being replaced by $o(1)$.

Lemma 3 *Under the assumptions of Theorem 1, as $n \to \infty$, with $t_* = \lfloor n\tau_* \rfloor$,*

$$\Big|\sum_{j=1}^{n} e_j X_{j-1}\Big| = O_P\big(\sqrt{n}\big), \tag{4.8}$$

$$\Big|\sum_{t=1}^{n} e_t X_{t-1} g(t, t_*)\Big| = O_P\big(\sqrt{n}\big), \quad \textit{for every fixed} \;\; \tau_* \in [0, 1], \tag{4.9}$$

and

$$\frac{\varepsilon_n}{\sqrt{n}} \max_{t_*} \Big|\sum_{t=1}^{n} e_t X_{t-1} g(t, t_*)\Big| = o_P(1), \quad \textit{for every sequence } \{\varepsilon_n\}_{n=1,2,\ldots} \textit{ with } \varepsilon_n \to 0. \tag{4.10}$$

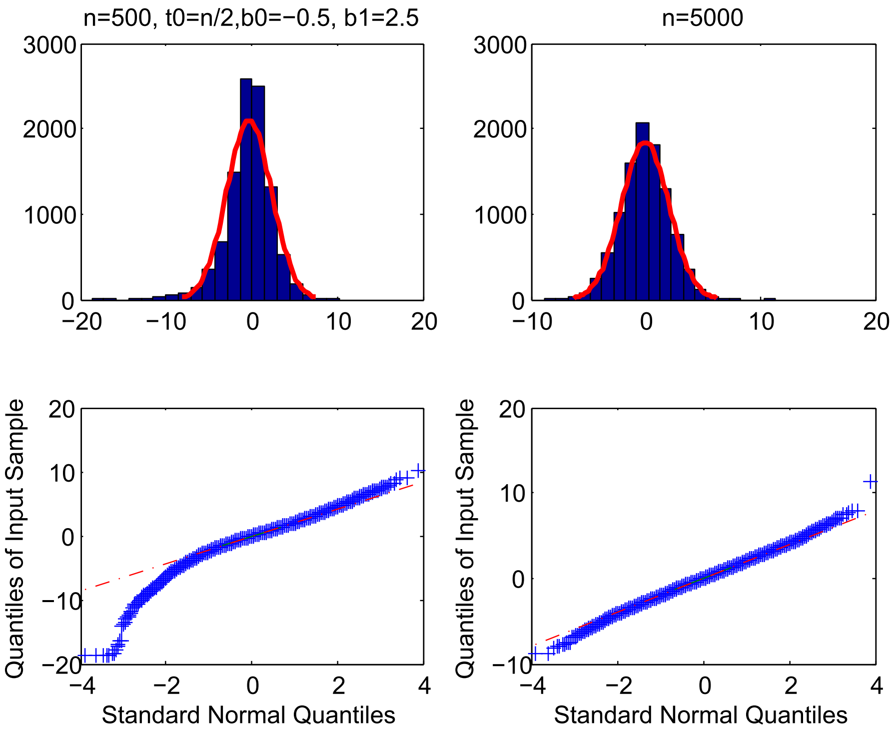

Fig. 6 Histograms and Q-Q plots of $\sqrt{n}(\widehat{\tau}_0 - \tau_0)$, left: $n = 500$, right: $n = 5\,000$, $t_0 = n/2$, $\beta_0 = -0.5$, $\beta_1 = 2.5$

Proof (a) Let $\mathcal{F}_{j-1}$ denote the σ-field generated by $e_0, \ldots, e_{j-1}$. Then, in view of our assumptions on $\{e_t\}$ and (4.2), as $n \to \infty$,

$$E\Big(\sum_{j=1}^{n} e_j X_{j-1}\Big)^2 = \sum_{j=1}^{n} E\big(e_j^2 X_{j-1}^2\big) + 2\sum_{i<j} E\big(e_i X_{i-1} e_j X_{j-1}\big)$$
$$= \sum_{j=1}^{n} E\big(E\big[e_j^2 X_{j-1}^2 | \mathcal{F}_{j-1}\big]\big) + 2\sum_{i<j} E\big(E\big[e_i X_{i-1} e_j X_{j-1} | \mathcal{F}_{j-1}\big]\big) = \sigma^2 \sum_{j=1}^{n} E X_{j-1}^2 = O(n),$$

which suffices to prove (4.8).

(b) Via similar arguments, along the lines of proof of (4.3), as $n \to \infty$,

$$E\Big(\sum_{t=1}^{n} e_t X_{t-1} g(t, t_*)\Big)^2 = \sigma^2 \sum_{t=1}^{n} g^2(t, t_*) E X_{t-1}^2 = O(n),$$

which proves (4.9).

(c) Consider, with $t_* = \lfloor n\tau_* \rfloor$, the sequence of stochastic processes $\{X_n(\cdot)\}_{n=1,2,\dots}$ in $D[0,1]$, where

$$X_n(\tau_*) = \frac{\varepsilon_n}{\sqrt{n}} \sum_{t=1}^{n} e_t X_{t-1} g(t, t_*), \quad 0 \le \tau_* \le 1.$$

In view of (4.9), the finite-dimensional distributions of X_n tend to 0. Moreover, for $0 \le \tau_1 \le \tau_* \le \tau_2 \le 1$,

$$E|X_n(\tau_*) - X_n(\tau_1)||X_n(\tau_2) - X_n(\tau_*)| \le \big(E|X_n(\tau_*) - X_n(\tau_1)|^2\big)^{1/2}\big(E|X_n(\tau_2) - X_n(\tau_*)|^2\big)^{1/2},$$

and, according to our assumptions on g, with $t_1 = \lfloor n\tau_1 \rfloor$,

$$\begin{aligned} E|X_n(\tau_*) - X_n(\tau_1)|^2 &= \sigma^2 \frac{\varepsilon_n^2}{n} \sum_{t=1}^{n} |g(t, t_*) - g(t, t_1)|^2 E X_{t-1}^2 \\ &\le C\, \varepsilon_n^2 \left|\frac{t_* - t_1}{n}\right|^2 \le C\, \varepsilon_n^2\, |\tau_2 - \tau_1|^2. \end{aligned}$$

Analogously, with $t_2 = \lfloor n\tau_2 \rfloor$,

$$E|X_n(\tau_2) - X_n(\tau_*)|^2 \le C\, \varepsilon_n^2 \left|\frac{t_2 - t_*}{n}\right|^2 \le C\, \varepsilon_n^2\, |\tau_2 - \tau_1|^2,$$

so that

$$E|X_n(\tau_*) - X_n(\tau_1)||X_n(\tau_2) - X_n(\tau_*)| \le C\, \varepsilon_n^2\, |\tau_2 - \tau_1|^2.$$

In view of Billingsley (1968), Theorem 15.6, this proves that

$$X_n \xrightarrow{\mathcal{D}[0,1]} 0 \quad \text{as} \quad n \to \infty,$$

which suffices for the proof of (4.10).

□

Lemma 4 *Under the assumptions of Theorem 1, as $n \to \infty$, with $t_* = \lfloor n\tau_* \rfloor$,*

$$\left|\sum_{t=1}^{n} \big(X_{t-1}^2 - E X_{t-1}^2\big)\right| = O_P\big(\sqrt{n}\big), \tag{4.11}$$

$$\left|\sum_{t=1}^{n} g(t, t_0) g(t, t_*)\big(X_{t-1}^2 - E X_{t-1}^2\big)\right| = O_P\big(\sqrt{n}\big), \quad \textit{for every fixed } \tau_* \in [0, 1], \tag{4.12}$$

$$\left|\sum_{t=1}^{n} g(t, t_*)\big(X_{t-1}^2 - E X_{t-1}^2\big)\right| = O_P\big(\sqrt{n}\big), \quad \textit{for every fixed } \tau_* \in [0, 1], \tag{4.13}$$

$$\Big|\sum_{t=1}^{n} g^2(t,t_*)\big(X_{t-1}^2 - EX_{t-1}^2\big)\Big| = O_P\big(\sqrt{n}\big), \quad \textit{for every fixed} \quad \tau_* \in [0,1], \tag{4.14}$$

and, for every sequence $\{\varepsilon_n\}_{n=1,2,\ldots}$ *with* $\varepsilon_n \to 0$,

$$\frac{\varepsilon_n}{\sqrt{n}} \max_{t_*} \Big|\sum_{t=1}^{n} g(t,t_0) g(t,t_*)\big(X_{t-1}^2 - EX_{t-1}^2\big)\Big| = o_P(1), \tag{4.15}$$

$$\frac{\varepsilon_n}{\sqrt{n}} \max_{t_*} \Big|\sum_{t=1}^{n} g(t,t_*)\big(X_{t-1}^2 - EX_{t-1}^2\big)\Big| = o_P(1), \tag{4.16}$$

$$\frac{\varepsilon_n}{\sqrt{n}} \max_{t_*} \Big|\sum_{t=1}^{n} g^2(t,t_*)\big(X_{t-1}^2 - EX_{t-1}^2\big)\Big| = o_P(1). \tag{4.17}$$

Proof We only give a proof of (4.13) and (4.16) here. The other assertions can be shown in a similar manner.

(a) In view of (1.5),

$$\begin{aligned}
\sum_{t=1}^{n} g(t,t_*)X_{t-1}^2 &= \sum_{t=1}^{n} g(t,t_*)\Big(e_{t-1} + \sum_{j=1}^{t-1} e_{t-1-j} \prod_{i=1}^{j-1} \big(\beta_0 + \beta_1 g(t-1-i,t_0)\big)\Big)^2 \\
&= \sum_{t=1}^{n} g(t,t_*)e_{t-1}^2 + \sum_{t=1}^{n} g(t,t_*) \sum_{j=1}^{t-1} e_{t-1-j}^2 \prod_{i=1}^{j-1} \big(\beta_0 + \beta_1 g(t-1-i,t_0)\big)^2 \\
&\quad + 2\sum_{t=1}^{n} g(t,t_*)e_{t-1} \sum_{j=1}^{t-1} e_{t-1-j} \prod_{i=1}^{j-1} \big(\beta_0 + \beta_1 g(t-1-i,t_0)\big) \\
&\quad + 2\sum_{t=1}^{n} g(t,t_*) \sum_{j_1<j_2} e_{t-1-j_1} e_{t-1-j_2} \prod_{i_1=1}^{j_1-1} \big(\beta_0 + \beta_1 g(t-1-i_1,t_0)\big) \\
&\quad \times \prod_{i_2=1}^{j_2-1} \big(\beta_0 + \beta_1 g(t-1-i_2,t_0)\big) \\
&=: S_1 + S_2 + 2S_3 + 2S_4.
\end{aligned} \tag{4.18}$$

Since g is bounded and $\{e_t\}$ is an i.i.d. sequence,

$$E(S_1 - ES_1)^2 = \sum_{t=1}^{n} g^2(t,t_*) E\big(e_t^2 - Ee_t^2\big)^2 \le C\,n. \tag{4.19}$$

Next, with $v^2 := Var(e_0^2)$ and b as in (1.4), due to the independence of $e^2_{t_1-1-j}$ and $e^2_{t_2-1-k}$, if $t_1-1-j \neq t_2-1-k$, i.e., if $k \neq j+t_2-t_1$, for sufficiently large n,

$$
\begin{aligned}
E(S_2-ES_2)^2 &= E\Big(\sum_{t=1}^{n} g(t,t_*)\sum_{j=1}^{t-1}\big(e^2_{t-1-j}-Ee^2_{t-1-j}\big)\prod_{i=1}^{j-1}\big(\beta_0+\beta_1 g(t-1-i,t_0)\big)^2\Big)^2\\
&= v^2\sum_{t=1}^{n} g^2(t,t_*)\sum_{j=1}^{t-1}\prod_{i=1}^{j-1}\big(\beta_0+\beta_1 g(t-1-i,t_0)\big)^4\\
&\quad + 2v^2\sum_{t_1<t_2} g(t_1,t_*)g(t_2,t_*)\sum_{j=1}^{t_1-1}\prod_{i_1=1}^{j-1}\big(\beta_0+\beta_1 g(t_1-1-i_1,t_0)\big)^2\prod_{i_2=1}^{j-1+t_2-t_1}\big(\beta_0+\beta_1 g(t_2-1-i_2,t_0)\big)^2\\
&\leq C\Big(\sum_{t=1}^{n} g^2(t,t_*)\sum_{j=1}^{t-1} b^{4(j-1)}+\sum_{t_2=2}^{n} g(t_2,t_*)\sum_{t_1=1}^{t_2-1} g(t_1,t_*)\sum_{j=1}^{t_1-1} b^{4(j-1)+2(t_2-t_1)}\Big)\leq Cn. \qquad (4.20)
\end{aligned}
$$

Similarly, since $ES_3 = 0$,

$$
\begin{aligned}
E(S_3-ES_3)^2 &= E\Big(\sum_{t=1}^{n} g(t,t_*)e_{t-1}\sum_{j=1}^{t-1} e_{t-1-j}\prod_{i=1}^{j-1}\big(\beta_0+\beta_1 g(t-1-i,t_0)\big)^2\Big)^2\\
&= \sigma^2\sum_{t=1}^{n} g^2(t,t_*)E\Big(\sum_{j=1}^{t-1} e_{t-1-j}\prod_{i=1}^{j-1}\big(\beta_0+\beta_1 g(t-1-i,t_0)\big)^2\Big)^2\\
&= \sigma^4\sum_{t=1}^{n} g^2(t,t_*)\sum_{j=1}^{t-1}\prod_{i=1}^{j-1}\big(\beta_0+\beta_1 g(t-1-i,t_0)\big)^4 \leq C\sum_{t=1}^{n} g^2(t,t_*)\sum_{j=1}^{t-1} b^{4(j-1)}\leq Cn. \qquad (4.21)
\end{aligned}
$$

Finally, via a corresponding estimation,

$$
\begin{aligned}
&E(S_4-ES_4)^2\\
&= \sum_{t=1}^{n} g^2(t,t_*)E\Big(\sum_{j_1<j_2} e_{t-1-j_1}e_{t-1-j_2}\prod_{i_1=1}^{j_1-1}\big(\beta_0+\beta_1 g(t-1-i_1,t_0)\big)\prod_{i_2=1}^{j_2-1}\big(\beta_0+\beta_1 g(t-1-i_2,t_0)\big)\Big)^2\\
&\leq C\sum_{t=1}^{n} g^2(t,t_*)\leq Cn. \qquad (4.22)
\end{aligned}
$$

On combining (4.18)–(4.22), we see that

$$
E\Big(\sum_{t=1}^{n} g(t,t_*)\big(X^2_{t-1}-EX^2_{t-1}\big)\Big)^2 \leq Cn,
$$

which suffices to prove (4.13).

b) The proof of (4.16) can be given analogously to that of (4.10) in Lemma 3. Consider, with $t_* = \lfloor n\tau_*\rfloor$, the sequence of stochastic processes $\{\widetilde{X}_n(\cdot)\}_{n=1,2,\ldots}$ in $D[0,1]$, where

$$
\widetilde{X}_n(\tau_*) = \frac{\varepsilon_n}{\sqrt{n}}\sum_{t=1}^{n} g(t,t_*)\big(X^2_{t-1}-EX^2_{t-1}\big), \quad 0\leq\tau_*\leq 1.
$$

In view of (4.13), the finite-dimensional distributions of $\widetilde{X}_n$ tend to 0. Moreover, for $0 \le \tau_1 \le \tau_* \le \tau_1 \le 1$, by just replacing $g(t, t_*)$ in the estimations (4.18)-(4.22) by $g(t, t_*) - g(t, t_1)$ and taking the Lipschitz continuity of g_0 into account,

$$E\big(\widetilde{X}_n(\tau_*) - \widetilde{X}_n(\tau_1)\big)^2 \le C\,\frac{\varepsilon_n^2}{n}\sum_{t=1}^{n}|g(t,t_*) - g(t,t_1)|^2 \le C\,\varepsilon_n^2\,|\tau_2 - \tau_1|^2.$$

Analogously,

$$E\big(\widetilde{X}_n(\tau_2) - \widetilde{X}_n(\tau_*)\big)^2 \le C\,\varepsilon_n^2\,|\tau_2 - \tau_1|^2,$$

and, via the Cauchy-Schwarz inequality,

$$E\big|\widetilde{X}_n(\tau_*) - \widetilde{X}_n(\tau_1)\big|\big|\widetilde{X}_n(\tau_2) - \widetilde{X}_n(\tau_*)\big| \le C\,\varepsilon_n^2\,|\tau_2 - \tau_1|^2.$$

In view of Billingsley (1968), Theorem 15.6, this proves that

$$\widetilde{X}_n \overset{\mathcal{D}[0,1]}{\longrightarrow} 0 \qquad \text{as} \qquad n \to \infty,$$

which suffices for the proof of (4.16).

□

The following two lemmas from real analysis will also be used in the proof of Theorem 1.

Lemma 5 *Let $\{b_n\}_{n=1,2,\ldots}$ be a real sequence such that $b_n \to \infty$ as $n \to \infty$. Then there is a sequence $\{\varepsilon_n\}_{n=1,2,\ldots}$ of positive reals, with $\varepsilon_n \to 0$, such that still $\varepsilon_n b_n \to \infty$ as $n \to \infty$.*

Proof For n sufficiently large, there exists an integer k_n such that $2^{k_n} \le b_n < 2^{k_n+1}$. Obviously, $k_n \to \infty$ as $n \to \infty$, so that a choice of $\varepsilon_n = 1/k_n$ completes the proof.

□

Lemma 6 *Let f be a continuous real function on a compact set K and x_0 be a unique maximizer of f, i.e., $x_0 = \arg\max_x f(x)$. Furthermore assume that $\lim_{n\to\infty} \max_{x\in K} |f_n(x) - f(x)| = 0$ and let $\widehat{x}_n = \arg\max_x f_n(x)$ be a maximizer of f_n (not necessarily unique). Then,*

$$\widehat{x}_n \to x_0 \qquad as \qquad n \to \infty.$$

Proof Suppose $\widehat{x}_n \nrightarrow x_0$ as $n \to \infty$. Since K is compact, there exists a subsequence $\{\widehat{x}_{k_n}\}$ and an $x_1 \neq x_0$ such that $\widehat{x}_{k_n} \to x_1$, hence

$$\big|f_{k_n}(\widehat{x}_{k_n}) - f(x_1)\big| \le \max_x \big|f_{k_n}(x) - f(x)\big| + \big|f(\widehat{x}_{k_n}) - f(x_1)\big| \to 0 \quad \text{as} \quad n \to \infty,$$

because of the continuity of f. Since x_0 is the unique maximizer of f, it holds $f(x_1) < f(x_0)$ implying that

$$\limsup_{n\to\infty} \big(f_{k_n}(\widehat{x}_{k_n}) - f(x_0)\big) < 0.$$

This, however, contradicts our assumptions, since

$$\big|f_n(\widehat{x}_n) - f(x_0)\big| = \big|\max_x f_n(x) - \max_x f(x)\big| \le \max_x \big|f_n(x) - f(x)\big| \to 0 \quad \text{as} \quad n \to \infty,$$

so that the proof is complete. □

Next, we need some extensions of Lemmas 1–4. Particularly, we study properties of the following quantities for $|t_0 - t_*| \le b_n$, with $b_n \to \infty,\ b_n/n \to 0$:

$$L_0(t_0, t_*) = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} e_t X_{t-1} \frac{g(t, t_*) - g(t, t_0)}{(t_0 - t_*)/n}$$

$$L_1(t_0, t_*) = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} e_t X_{t-1} \widetilde{h}_n(t, t_*, t_0), \quad \text{with } \widetilde{h}_n(t, t_*, t_0) \text{ as in (4.33) below,}$$

$$L_2(t_0, t_*) = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} e_t X_{t-1} \frac{\widetilde{g}_n(t, t_*) - \widetilde{g}_n(t, t_0)}{(t_* - t_0)/n},$$

$$L_3(t_0, t_*) = \frac{1}{n} \sum_{t=1}^{n} X_{t-1}^2\, \widetilde{g}_n(t, t_0) \frac{\widetilde{g}_n(t, t_0) - \widetilde{g}_n(t, t_*)}{(t_* - t_0)/n},$$

$$L_4(t_0, t_*) = \frac{1}{n} \sum_{t=1}^{n} X_{t-1}^2 \frac{(\widetilde{g}_n(t, t_0) - \widetilde{g}_n(t, t_*))^2}{((t_* - t_0)/n)^2}.$$

We start with an extension of Lemma 3.

Lemma 7 *Let the assumptions of Theorem 3 be satisfied and let t_* be such that $|t_* - t_0| \le r_n \sqrt{n}\, |\beta_1|^{-1}$, with*

$$r_n \to \infty, \quad \frac{|\beta_1|\sqrt{n}}{r_n} \to \infty. \tag{4.23}$$

Then, for $j = 0, 1, 2$, $L_j(t_0, t_*)$ has an asymptotically normal limit distribution, with zero mean and variance σ_j^2, where

$$\sigma_0^2 = \frac{\sigma^4}{1 - \beta_0^2} \int_0^1 g_0'^2(x - \tau_0)dx,$$

$$\sigma_1^2 = \frac{\sigma^4}{1 - \beta_0^2} \Big(\int_0^1 \widetilde{g}_0'^2(x - \tau_0)dx - \frac{\big(\int_0^1 \widetilde{g}_0(x - \tau_0)\widetilde{g}_0'(x - \tau_0)dx\big)^2}{\int_0^1 \widetilde{g}_0^2(x - \tau_0)dx} \Big),$$

$$\sigma_2^2 = \frac{\sigma^2}{1-\beta_0^2}\int_0^1 \widetilde{g}_0'^2(x-\tau_0)dx.$$

Moreover, as $n \to \infty$,

$$\max_{|t_0-t_{**}|\le b_n} |L_j(t_0,t_*) - L_j(t_0,t_{**})| = o_P(c_n), \quad j=0,1,2, \tag{4.24}$$

for some $c_n \to 0$ *and any* t_* *such that* $|t_0 - t_*| \le b_n$, $b_n \to \infty$, $b_n/n \to 0$.

Proof We focus on $L_0(t_0,t_*)$. The desired results for $L_1(t_0,t_*)$ and $L_2(t_0,t_*)$ can be derived in the same way, the obtained expressions are just somewhat more complex, but can be omitted.

Note that, for fixed t_* and t_0, $L_j(t_0,t_*)$, $j=0,1,2$, are sums of martingale difference arrays. Hence to obtain their limit properties, we can apply Theorem 24.3 in Davidson (1994), p. 383, which means for $L_0(t_0,t_*)$ to verify validity of the conditions

$$\frac{\frac{1}{n}\sum_{t=1}^n \left[e_t X_{t-1}\frac{g(t,t_*)-g(t,t_0)}{(t_0-t_*)/n}\right]^2}{\frac{1}{n}\sum_{j=1}^n E\left[e_j X_{j-1}\frac{g(j,t_*)-g(j,t_0)}{(t_0-t_*)/n}\right]^2} \xrightarrow{P} 1, \tag{4.25}$$

$$\max_t \frac{\left[e_t X_{t-1}\frac{g(t,t_*)-g(t,t_0)}{(t_0-t_*)/n}\right]^2}{\sum_{j=1}^n E\left[e_j X_{j-1}\frac{g(j,t_*)-g(j,t_0)}{(t_0-t_*)/n}\right]^2} \xrightarrow{P} 0. \tag{4.26}$$

We make repeated use of assumption (A.4), particularly

$$\left|\frac{g(j,t_*)-g(j,t_0)}{(t_0-t_*)/n} - g_0'((j-t_0)/n)\right| \le D|((t_*-t_0)/n)|^{\Delta}.$$

Consider first the denominator in (4.25). Direct calculations give

$$\begin{aligned}&\frac{1}{n}\sum_{j=1}^n E\left[e_j X_{j-1}\frac{g(j,t_*)-g(j,t_0)}{(t_0-t_*)/n}\right]^2\\ &= \frac{\sigma^4}{1-\beta_0^2}\frac{1}{n}\sum_{j=1}^n\left[\frac{g(j,t_*)-g(j,t_0)}{(t_0-t_*)/n}\right]^2 + \frac{\sigma^2}{n}\sum_{j=1}^n\left[EX_{j-1}^2 - \frac{\sigma^2}{1-\beta_0^2}\right]\left[\frac{g(j,t_*)-g(j,t_0)}{(t_0-t_*)/n}\right]^2.\end{aligned}$$

By Assumption (A.4) and using the same arguments as in the proofs of Lemmas 2 and 4 , we get

$$\begin{aligned}&\frac{1}{n}\sum_{j=1}^n\left[\frac{g(j,t_*)-g(j,t_0)}{(t_0-t_*)/n} - g_0'((j-t_0)/n))\right]^2 \le C\,|(t_*-t_0)/n|^{2\Delta} \le C\left(\frac{r_n}{|\beta_1|\sqrt{n}}\right)^{2\Delta},\\ &\left|\frac{1}{n}\sum_{j=1}^n\left[EX_{t-1}^2 - \frac{\sigma^2}{1-\beta_0^2}\right]\left[\frac{g(j,t_*)-g(j,t_0)}{(t_0-t_*)/n} - g_0'((j-t_0)/n)\right]^2\right| \le C|(t_*-t_0)/n|^{2\Delta}\\ &\le C\left(\frac{r_n}{|\beta_1|\sqrt{n}}\right)^{2\Delta},\end{aligned}$$

$$\frac{1}{n}\sum_{t=1}^{n} g_0'^2((t-t_0)/n) = \int_0^1 g_0'^2(x-\tau_0)dx + o(1),$$
$$\frac{1}{n}\sum_{t=1}^{n} \Big(EX_{t-1}^2 - \frac{\sigma^2}{1-\beta_0^2}\Big) g_0'^2((t-t_0)/n)\big)^2 = O(|\beta_1|),$$

as $n \to \infty$, and combining all these results we have

$$\frac{1}{n}\sum_{t=1}^{n} E\Big[e_t X_{t-1}\frac{g(t,t_*)-g(t,t_0)}{(t_0-t_*)/n}\Big]^2 = \frac{\sigma^4}{1-\beta_0^2}\int_0^1 g_0'^2(x-\tau_0)dx + o(1). \tag{4.27}$$

For the numerator in (4.25) we have

$$\frac{1}{n}\sum_{t=1}^{n} \Big[e_t X_{t-1}\frac{g(t,t_*)-g(t,t_0)}{(t_0-t_*)/n}\Big]^2$$
$$= \frac{1}{n}\sum_{t=1}^{n} [e_t^2-\sigma^2] X_{t-1}^2\Big[\frac{g(t,t_*)-g(t,t_0)}{(t_0-t_*)/n}\Big]^2 + \frac{\sigma^2}{n}\sum_{t=1}^{n} X_{t-1}^2\Big[\frac{g(t,t_*)-g(t,t_0)}{(t_0-t_*)/n}\Big]^2.$$

The first term on the r.h.s. is the sum of martingale difference arrays, with zero mean and variance

$$E\Big[\frac{1}{n}\sum_{t=1}^{n}\Big((e_t^2-\sigma^2)X_{t-1}^2\frac{g(t,t_*)-g(t,t_0)}{(t_0-t_*)/n}\Big)^2\Big]^2 = O\big(n^{-1}\big),$$

which follows from the finiteness of Ee_t^4, Assumption (A.3) and the uniform boundedness of EX_{t-1}^4 (cf. (4.1)). Thus,

$$\frac{1}{n}\sum_{t=1}^{n} [e_t^2-\sigma^2] X_{t-1}^2\Big[\frac{g(t,t_*)-g(t,t_0)}{(t_0-t_*)/n}\Big]^2 = O_P\big(n^{-1/2}\big). \tag{4.28}$$

Next, proceeding in the same way as in the proofs of Lemmas 1 and 4, we get

$$\frac{\sigma^2}{n}\sum_{t=1}^{n} X_{t-1}^2\Big[\frac{g(t,t_*)-g(t,t_0)}{(t_0-t_*)/n}\Big]^2 = \frac{\sigma^4}{1-\beta_0^2}\int_0^1 g_0'^2(x-\tau_0)dx + o_P(1), \tag{4.29}$$

and combining (4.28), (4.29) and (4.27) we obtain (4.25).

To verify (4.26) note that, due to (4.27), it suffices to prove

$$\max_t \frac{1}{n}\Big[e_t X_{t-1}\frac{g(t,t_*)-g(t,t_0)}{(t_0-t_*)/n}\Big]^2 \xrightarrow{P} 0.$$

For this we have, using similar arguments as above,

$$P\Big(\max_t \Big[e_t X_{t-1} \frac{g(t,t_*)-g(t,t_0)}{(t_0-t_*)/n}\Big]^2 \geq n\epsilon\Big) \leq \sum_{t=1}^n P\Big(\Big[e_t X_{t-1} \frac{g(t,t_*)-g(t,t_0)}{(t_0-t_*)/n}\Big]^2 \geq n\epsilon\Big)$$
$$\leq \frac{1}{n^2\epsilon^2} \sum_{t=1}^n E e_t^4 E X_{t-1}^4 \Big[\frac{g(t,t_*)-g(t,t_0)}{(t_0-t_*)/n}\Big]^4 = O(n^{-1}),$$

and we can conclude that the asymptotic normality of $L_0(t_0, t_*)$ holds for fixed t_*.

Next we show (4.24) for $j = 0$. We proceed as in the proof of Lemma 3, part c). Toward this we study for $t_0 < t_* < t_{**}$ and $|t_0 - t_*| + |t_* - t_{**}| \leq b_n$, satisfying $b_n \to \infty$, $b_n/n \to 0$, the quantity

$$E\big[L_0(t_0,t_*) - L_0(t_0,t_{**})\big]^2 = \frac{1}{n}\sum_{t=1}^n \sigma^2 E(X_{t-1}^2)\Big(\frac{g(t,t_*)-g(t,t_0)}{(t_0-t_*)/n} - \frac{g(t,t_{**})-g(t,t_0)}{(t_0-t_{**})/n}\Big)^2$$
$$\leq C\frac{1}{n}\sum_{t=1}^n \Big(\frac{g(t,t_*)-g(t,t_0)}{(t_0-t_*)/n} - \frac{g(t,t_{**})-g(t,t_0)}{(t_0-t_{**})/n}\Big)^2,$$

for some $C > 0$. By Assumption (A.4)

$$\big|g(t,t_{**}) - g(t,t_*) - g_0'((t-t_*)/n)((t_*-t_{**})/n)\big| \leq C\big|(t_{**}-t_*)/n\big|^{1+\Delta}$$

which gives

$$\frac{1}{n}\sum_{t=1}^n \Big(\frac{g(t,t_*)-g(t,t_0)}{(t_0-t_*)/n} - \frac{g(t,t_{**})-g(t,t_0)}{(t_0-t_{**})/n}\Big)^2$$
$$= \frac{1}{n}\sum_{t=1}^n \Big[\frac{g(t,t_*)-g(t,t_0)}{(t_0-t_*)/n}\Big(1 - \frac{(t_0-t_*)/n}{(t_0-t_{**})/n}\Big) - \frac{g(t,t_{**})-g(t,t_*)}{(t_0-t_{**})/n}\Big]^2$$
$$= \frac{1}{n}\sum_{t=1}^n \Big[\Big(\frac{g(t,t_*)-g(t,t_0)}{(t_0-t_*)/n} - g_0'((t-t_*)/n)\Big)\frac{(t_*-t_{**})/n}{(t_0-t_{**})/n} + O\Big(\frac{|(t_*-t_{**})/n|^{1+\Delta}}{|(t_0-t_{**})/n|}\Big)\Big]^2$$
$$\leq C\Big[\big|(t_0-t_*)/n\big|^{\Delta}\frac{(t_*-t_{**})/n}{(t_0-t_{**})/n} + \frac{|(t_*-t_{**})/n|^{1+\Delta}}{|(t_0-t_{**})/n|}\Big]^2$$
$$\leq C\Big[\frac{|(t_*-t_{**})/n|}{|(t_0-t_{**})/n|^{1-\Delta}} + \frac{|(t_*-t_{**})/n|^{1+\Delta}}{|(t_0-t_{**})/n|}\Big]^2 \leq C\big|(t_*-t_{**})/n\big|^{2\Delta},$$

since $|t_0 - t_*| \vee |t_* - t_{**}| \leq |t_0 - t_{**}|$ and $\Delta \leq 1$.

In view of $2\Delta > 1$ by Assumption (A.4), assertion (4.24) can now be finished by again making use of Billingsley (1968), Theorem 15.6, as in the proof of Lemma 3, part c). □

The next lemma is an extension of Lemma 4.

Lemma 8 *Under the assumptions of Theorem 3 we have*

$$L_4(t_0,t_*) = \frac{\sigma^2}{1-\beta_0^2}\Big[\frac{1}{n}\sum_{t=1}^n\Big(g_0'((t-t_0)/n) - \frac{1}{n}\sum_{j=1}^n g_0'((j-t_0)/n)\Big)^2\Big] + O_P\Big(\big|(t_0-t_*)/n\big|^{\Delta}\Big) + o_P(1)$$

$$= \frac{\sigma^2}{1-\beta_0^2}\int_0^1 \widetilde{g}_0'^2(x-\tau_0)dx + O_P\Big(\big|(t_0-t_*)/n)\big|^{\Delta}\Big) + o_P(1), \quad |t_*-t_0| \le b_n, \tag{4.30}$$

$$L_3(t_0,t_*) = \frac{\sigma^2}{1-\beta_0^2}\int_0^1 \widetilde{g}_0(x-\tau_0)\,\widetilde{g}_0'(x-\tau_0)dx + o_P(1), \quad |t_*-t_0| \le b_n. \tag{4.31}$$

Moreover, as $n \to \infty$,

$$\max_{|t_0-t_{**}|\le b_n} \big|L_j(t_0,t_*) - L_j(t_0,t_{**})\big| = o_P(c_n), \quad j = 3, 4, \tag{4.32}$$

for some $c_n \to 0$ *and any* $|t_*-t_0| \le b_n$, *with* $b_n \to \infty,\ b_n/n \to 0$.

Proof Note that

$$\begin{aligned}
L_4(t_0,t_*) &= \frac{1}{n}\sum_{t=1}^n X_{t-1}^2 \frac{\big(\widetilde{g}_n(t,t_0)-\widetilde{g}_n(t,t_*)\big)^2}{((t_*-t_0)/n)^2}\\
&= \frac{1}{n}\sum_{t=1}^n X_{t-1}^2 \frac{\big(g(t,t_0)-g(t,t_*)\big)^2}{\big((t_*-t_0)/n\big)^2} - \frac{\Big(\frac{1}{n}\sum_{j=1}^n X_{j-1}^2\big(g(j,t_0)-g(j,t_*)\big)\Big)^2}{\frac{1}{n}\sum_{j=1}^n X_{j-1}^2\big((t_*-t_0)/n\big)^2}\\
&=: L_{41}(t_0,t_*) - L_{42}(t_0,t_*).
\end{aligned}$$

We have

$$\begin{aligned}
L_{41}(t_0,t_*) = \frac{1}{n}&\sum_{j\ge\max(t_0,t_*)} X_{j-1}^2 \frac{\big(g(j,t_0)-g(j,t_*)\big)^2}{\big((t_*-t_0)/n\big)^2}\\
&+ \frac{1}{n}\sum_{\min(t_0,t_*)\le j\le\max(t_0,t_*)} X_{j-1}^2 \frac{\big(g(j,t_0)-g(j,t_*)\big)^2}{\big((t_*-t_0)/n\big)^2}.
\end{aligned}$$

The latter sum on the r.h.s has only $|t_0-t_*| \le 2b_n$ summands, with $b_n \to \infty,\ b_n/n \to 0$, and the terms $\big(g(j,t_0)-g(j,t_*)\big)\big/\big((t_*-t_0)/n\big)$ are bounded in j, thus the respective terms are not influential. By (A.4),

$$\begin{aligned}
&\frac{1}{n}\sum_{j\ge\max(t_0,t_*)} X_{j-1}^2\Big(\frac{g(j,t_*)-g(j,t_0)}{(t_0-t_*)/n)} - g_0'((j-t_0)/n)\Big)^2\\
&\le C\,\frac{1}{n}\sum_{j=1}^n X_{j-1}^2\Big|\frac{t_*-t_0}{n}\Big|^{2\Delta} = O_P\Big(\Big|\frac{t_*-t_0}{n}\Big|^{2\Delta}\Big),
\end{aligned}$$

where the last relation follows from (4.1) and (4.11). Proceeding as in the proofs of Lemmas 1 and 2, we get

$$\frac{1}{n}\sum_{j=1}^n X_{j-1}^2 g_0'^2((j-t_0)/n) = \frac{\sigma^2}{1-\beta_0^2}\int_0^1 g_0'^2(x-\tau_0)dx + o_P(1)$$

and hence

$$L_{41}(t_*, t_0) = \frac{\sigma^2}{1-\beta_0^2}\int_0^1 g_0'^2(x-\tau_0)dx + o_P(1).$$

The result for $L_{42}(t_*, t_0)$ is obtained in the same way and thus the assertion on $L_4(t_*, t_0)$ follows.

The proof for $L_3(t_*, t_0)$ follows the same lines and can therefore be omitted.

To prove (4.32) we proceed as in the proof of (4.24). To avoid too many technicalities we study only

$$\big(L_5(t_*, t_0) - L_5(t_{**}, t_0)\big)^2$$

where

$$L_5(t_*, t_0) = \frac{1}{n}\sum_{t=1}^n X_{t-1}^2 \frac{(g(t,t_0)-g(t,t_*))^2}{((t_*-t_0)/n)^2}.$$

Now

$$\begin{aligned} L_5(t_*, t_0) - L_5(t_{**}, t_0) &= \frac{1}{n}\sum_{t=1}^n (X_{t-1}^2 - EX_{t-1}^2)\Big[\frac{(g(t,t_0)-g(t,t_*))^2}{((t_*-t_0)/n)^2} - \frac{(g(t,t_0)-g(t,t_{**}))^2}{((t_{**}-t_0)/n)^2}\Big] \\ &\quad + \frac{1}{n}\sum_{t=1}^n EX_{t-1}^2\Big[\frac{(g(t,t_0)-g(t,t_*))^2}{((t_*-t_0)/n)^2} - \frac{(g(t,t_0)-g(t,t_{**}))^2}{((t_{**}-t_0)/n)^2}\Big]. \end{aligned}$$

As in the proof of Lemma 4 and utilizing Assumption (A.4),

$$\begin{aligned} &E\Big[\frac{1}{n}\sum_{t=1}^n \big(X_{t-1}^2 - EX_{t-1}^2\big)\Big(\frac{(g(t,t_0)-g(t,t_*))^2}{((t_*-t_0)/n)^2} - \frac{(g(t,t_0)-g(t,t_{**}))^2}{((t_{**}-t_0)/n)^2}\Big)\Big]^2 \\ &\le C\frac{1}{n^2}\sum_{t=1}^n \Big[\frac{(g(t,t_0)-g(t,t_*))^2}{((t_*-t_0)/n)^2} - \frac{(g(t,t_0)-g(t,t_{**}))^2}{((t_{**}-t_0)/n)^2}\Big]^2 \\ &\le C\frac{1}{n^2}\sum_{t=1}^n \Big|\frac{g(t,t_0)-g(t,t_*)}{(t_*-t_0)/n} - \frac{g(t,t_0)-g(t,t_{**})}{(t_{**}-t_0)/n}\Big|^2 \le C\,\frac{1}{n}\,\big|(t_*-t_{**})/n\big|^{2\Delta}, \end{aligned}$$

where we also used arguments from the proof of Lemma 7. Finally,

$$\begin{aligned} &\frac{1}{n}\Big|\sum_{t=1}^n EX_{t-1}^2\Big[\frac{(g(t,t_0)-g(t,t_*))^2}{((t_*-t_0)/n)^2} - \frac{(g(t,t_0)-g(t,t_{**}))^2}{((t_{**}-t_0)/n)^2}\Big]\Big| \\ &\le C\frac{1}{n}\sum_{t=1}^n \Big|\frac{g(t,t_0)-g(t,t_*)}{(t_*-t_0)/n} - \frac{g(t,t_0)-g(t,t_{**})}{(t_{**}-t_0)/n}\Big| \le C\,\big|(t_*-t_{**})/n\big|^{\Delta}. \end{aligned}$$

On combining all these estimates, we can conclude that (4.32) holds true. □

By Theorem 1 and since $Q(t_0)$ (defined below) does not depend on t_*, the estimator $\widehat{t_0}$ has the same limit distribution as

$$\widehat{t_0}(b_n) = \underset{|t_* - t_0)| \le b_n}{\arg\max} \big(Q(t_*) - Q(t_0)\big),$$

for some $b_n \to \infty, \; b_n/n \to 0$, and

$$\begin{aligned}
Q(t_*) &= \frac{1}{\frac{1}{n}\sum_{t=1}^n \widetilde{g}_n^2(t, t_*)X_{t-1}^2}\left(\beta_1 \frac{1}{n}\sum_{t=1}^n \widetilde{g}_n(t, t_0)\widetilde{g}_n(t, t_*)X_{t-1}^2 + \frac{1}{n}\sum_{t=1}^n e_t X_{t-1}\widetilde{g}_n(t, t_*)\right)^2, \\
&= \beta_1^2 Q_1(t_*) + 2\beta_1 Q_2(t_*) + Q_3(t_*), \\
Q_1(t_*) &= \frac{\left(\frac{1}{n}\sum_{t=1}^n \widetilde{g}_n(t, t_0)\widetilde{g}_n(t, t_*)X_{t-1}^2\right)^2}{\frac{1}{n}\sum_{t=1}^n \widetilde{g}_n^2(t, t_*)X_{t-1}^2}, \\
Q_2(t_*) &= \frac{\left(\frac{1}{n}\sum_{t=1}^n \widetilde{g}_n(t, t_0)\widetilde{g}_n(t, t_*)X_{t-1}^2\right)\left(\frac{1}{n}\sum_{t=1}^n e_j X_{t-1}\widetilde{g}_n(t, t_*)\right)}{\frac{1}{n}\sum_{t=1}^n \widetilde{g}_n^2(t, t_*)X_{t-1}^2}, \\
Q_3(t_*) &= \frac{\left(\frac{1}{n}\sum_{t=1}^n e_j X_{t-1}\widetilde{g}_n(t, t_*)\right)^2}{\frac{1}{n}\sum_{t=1}^n \widetilde{g}_n^2(t, t_*)X_{t-1}^2}.
\end{aligned}$$

We need to study the properties of $Q_j(t_*) - Q_j(t_0), \;\; j = 1, 2, 3$, separately. This is formulated in the next three Propositions.

Proposition 1 *Under the Assumptions (A.1)– (A.4) we get*

$$\max_{1 \le |t_* - t_0| \le b_n} \left| \frac{Q_1(t_*) - Q_1(t_0)}{(\frac{t_* - t_0}{n})^2} + \frac{\sigma^2}{1 - \beta_0^2}\widetilde{H} \right| = o_p(1).$$

Proof Direct but long calculations give

$$\begin{aligned}
&\Big(Q_1(t_*) - Q_1(t_0)\Big)\left(\frac{1}{n}\sum_{t=1}^n X_{t-1}^2 \widetilde{g}_n^2(t, t_*)\right) = \left(\frac{1}{n}\sum_{t=1}^n X_{t-1}^2 \widetilde{g}_n(t, t_0)\big(\widetilde{g}_n(t, t_0) - \widetilde{g}_n(t, t_*)\big)\right)^2 \\
&- \left(\frac{1}{n}\sum_{t=1}^n X_{t-1}^2\big(\widetilde{g}_n(t, t_0) - \widetilde{g}_n(t, t_*)\big)^2\right)\left(\frac{1}{n}\sum_{t=1}^n X_{t-1}^2 \widetilde{g}_n^2(t, t_0)\right).
\end{aligned}$$

Using the assertions in Lemma 1, 2, 4 and the calculations in Lemma 8, we get after some steps, uniformly in $|t_* - t_0)| \le b_n$,

$$\begin{aligned}
&Q_1(t_*) - Q_1(t_0) \\
&= \big((t_* - t_0)/n\big)^2 \left[\frac{\big(\frac{1}{n}\sum_{t=1}^n X_{t-1}^2 \widetilde{g}_0((t - t_0)/n)\widetilde{g}_0'((t - t_0)/n)\big)^2}{\frac{1}{n}\sum_{t=1}^n X_{t-1}^2 \widetilde{g}_0^2((t - t_0)/n)} - \frac{1}{n}\sum_{t=1}^n X_{t-1}^2 \widetilde{g}_0'^2((t - t_0)/n)\right] \\
&\times \big(1 + O_P(|\beta_1| + (|t_* - t_0|/n)^{\Delta}\big) \\
&= \frac{\sigma^2}{1 - \beta_0^2}\Big(\frac{t_0 - t_*}{n}\Big)^2 \left[\frac{\big(\int_0^1 \widetilde{g}_0(x - \tau_0)\widetilde{g}_0'(x - \tau_0)dx\big)^2}{\int_0^1 \widetilde{g}_0^2(x - \tau_0)dx} - \int_0^1 \widetilde{g}_0'^2(x - \tau_0)dx\right]
\end{aligned}$$

$$\times \big(1 + O_P(|\beta_1| + (|t_* - t_0|/n)^{\Delta}\big),$$

which suffices for the proof. □

Proposition 2 *Under Assumptions (A.1)–(A.4) we get*

$$\max_{r_n\beta_1^{-1}\sqrt{n}<|t_*-t_0|\le b_n} \Big|\frac{Q_2(t_*)-Q_2(t_0)}{\beta_1(\frac{t_*-t_0}{n})^2}\Big| = o_p\big(n^{-1/2}+r_n^{-1}\big) \quad \textit{and}$$

$$\max_{1\le|t_*-t_0|\le r_n\beta_1^{-1}\sqrt{n}} \Big|\frac{Q_2(t_*)-Q_2(t_0)}{\frac{t_*-t_0}{n}} - \frac{1}{\sqrt{n}}\frac{1}{\sqrt{n}}\sum_{t=1}^{n} e_t X_{t-1}\widetilde{h}_n(t,t_*,t_0)\Big|$$
$$= O_P\Big(\max_{1\le|t_*-t_0|\le r_n\beta_1^{-1}\sqrt{n}} \frac{1}{\sqrt{n}}\frac{|t_*-t_0|}{n}\Big) = O_P\Big(\frac{1}{\sqrt{n}}\frac{r_n}{|\beta_1|\sqrt{n}}\Big),$$

where r_n satisfies (4.23) and

$$\widetilde{h}_n(t,t_*,t_0) = \frac{1}{\big(\frac{t_*-t_0}{n}\big)}\bigg(\big(\widetilde{g}_n(t,t_*)-\widetilde{g}_n(t,t_0)\big) - \widetilde{g}_n(t,t_0)\frac{\frac{1}{n}\sum_{t=1}^{n}\widetilde{g}_n(t,t_*)\big(\widetilde{g}(t,t_*)-\widetilde{g}_n(t,t_0)\big)X_{t-1}^2}{\frac{1}{n}\sum_{t=1}^{n}\widetilde{g}_n^2(t,t_*)X_{t-1}^2}\bigg). \tag{4.33}$$

Moreover, the limit distribution of $\frac{1}{\sqrt{n}}\sum_{t=1}^{n} e_t X_{t-1}\widetilde{h}_n(t,t_*,t_0)$ is normal $\mathcal{N}\big(0, \frac{\sigma^4}{1-\beta_0^2}\widetilde{H}\big)$ for each $|t_*-t_0| \le r_n|\beta_1|^{-1}\sqrt{n}$, which follows from Lemma 7.

Proof Direct calculations give

$$Q_2(t_*) - Q_2(t_0) = \frac{\frac{1}{n}\sum_{t=1}^{n}\widetilde{g}_n(t,t_0)\widetilde{g}_n(t,t_*)X_{t-1}^2}{\frac{1}{n}\sum_{t=1}^{n}\widetilde{g}_n^2(t,t_*)X_{t-1}^2}\,\frac{1}{n}\sum_{t=1}^{n} e_t X_{t-1}\widetilde{g}_n(t,t_*) - \frac{1}{n}\sum_{t=1}^{n} e_j X_{t-1}\widetilde{g}_n(t,t_0)$$
$$= \frac{1}{n}\sum_{t=1}^{n} e_t X_{t-1}\Big[\big(\widetilde{g}_n(t,t_*)-\widetilde{g}_n(t,t_0)\big) - \widetilde{g}_n(t,t_0)\frac{\frac{1}{n}\sum_{j=1}^{n}\widetilde{g}_n(j,t_*)\big(\widetilde{g}_n(j,t_*)-\widetilde{g}_n(j,t_0)\big)X_{j-1}^2}{\frac{1}{n}\sum_{j=1}^{n}\widetilde{g}_n^2(j,t_*)X_{j-1}^2}\Big]$$
$$+ \frac{1}{n}\sum_{t=1}^{n} e_t X_{t-1}\big(\widetilde{g}_n(t,t_*)-\widetilde{g}_n(t,t_0)\big)\frac{\frac{1}{n}\sum_{j=1}^{n}\widetilde{g}_n(j,t_*)\big(-\widetilde{g}_n(j,t_*)+\widetilde{g}_n(j,t_0)\big)X_{j-1}^2}{\frac{1}{n}\sum_{j=1}^{n}\widetilde{g}_n^2(j,t_*)X_{j-1}^2}.$$

On applying Lemmas 1–3 and 7 to the above sums, we get that the first term on the r.h.s. is influential, while the latter one is negligible. So, both assertions follow from here. □

Proposition 3 *Under Assumptions (A.1)–(A.4) we get*

$$\max_{\sqrt{n}|\beta_1|^{-1}r_n<|t_*-t_0|\le b_n} \frac{|Q_3(t_*)-Q_3(t_0)|}{\beta_1^2\big(\frac{t_*-t_0}{n}\big)^2} = O_P\Big(\frac{1}{\sqrt{n}\,|\beta_1|r_n}\Big) \quad \textit{and}$$

$$\max_{1\le|t_*-t_0|\le\sqrt{n}|\beta_1|^{-1}r_n} \frac{|Q_3(t_*)-Q_3(t_0)|}{|\beta_1||\frac{t_*-t_0}{n}|} = O_P\Big(\frac{1}{n\,|\beta_1|}\Big),$$

where r_n satisfies (4.23).

Proof Direct calculations give

$$
\begin{aligned}
Q_3(t_*) - Q_3(t_0) &= \frac{\big(\frac{1}{n}\sum_{t=1}^n e_t X_{t-1}\widetilde{g}_n(t,t_*)\big)^2}{\frac{1}{n}\sum_{t=1}^n \widetilde{g}_n^2(t,t_*)X_{t-1}^2} - \frac{\big(\frac{1}{n}\sum_{t=1}^n e_t X_{t-1}\widetilde{g}_n(t,t_0)\big)^2}{\frac{1}{n}\sum_{t=1}^n X_{t-1}^2\widetilde{g}_n^2(t,t_0)} \\
&= \frac{\big[\frac{1}{n}\sum_{t=1}^n e_t X_{t-1}(\widetilde{g}_n(t,t_*) - \widetilde{g}_n(t,t_0))\big]\big[\frac{1}{n}\sum_{t=1}^n e_t X_{t-1}(\widetilde{g}_n(t,t_*) + \widetilde{g}_n(t,t_0))\big]}{\frac{1}{n}\sum_{t=1}^n \widetilde{g}_n^2(t,t_*)X_{t-1}^2} \\
&\quad + \Big(\frac{1}{n}\sum_{t=1}^n e_t X_{t-1}\widetilde{g}_n(t,t_0)\Big)^2 \frac{\frac{1}{n}\sum_{t=1}^n X_{t-1}^2\big(\widetilde{g}_n^2(t,t_0) - \widetilde{g}_n^2(t,t_*)\big)}{\big(\frac{1}{n}\sum_{t=1}^n X_{t-1}^2\widetilde{g}_n^2(t,t_0)\big)\big(\frac{1}{n}\sum_{t=1}^n X_{t-1}^2\widetilde{g}_n^2(t,t_*)\big)} \\
&= O_P\Big(\frac{1}{n}\frac{|t_* - t_0|}{n}\Big),
\end{aligned}
$$

uniformly for $|t_* - t_0| \le b_n$, where the last relation is implied by a combination of Lemmas 7 and 8. Then both assertions follow immediately. □

5 Proofs of the main theorems

Now we are ready to turn to the proofs of Theorems 1–3. We first prove the consistency of our least squares estimator $\widehat{t_0}$.

Proof of Theorem 1 In view of (1.13), we consider, for $t_* = 0, 1, \ldots, \lfloor n(1-\delta)\rfloor$,

$$
h_n(t_*) = \frac{\Big[\beta_1\Big(\frac{1}{n}\sum_{t=1}^n g(t,t_0)g(t,t_*)X_{t-1}^2 - \frac{\frac{1}{n}\sum_{j=1}^n g(j,t_0)X_{j-1}^2 \frac{1}{n}\sum_{j=1}^n g(j,t_*)X_{j-1}^2}{\frac{1}{n}\sum_{j=1}^n X_{j-1}^2}\Big) + R_n(t_*)\Big]^2}{\frac{1}{n}\sum_{t=1}^n g^2(t,t_*)X_{t-1}^2 - \frac{\big(\frac{1}{n}\sum_{j=1}^n g(j,t_*)X_{j-1}^2\big)^2}{\frac{1}{n}\sum_{j=1}^n X_{j-1}^2}}, \tag{5.1}
$$

where

$$
R_n(t_*) = \frac{1}{n}\sum_{t=1}^n e_t X_{t-1} g(t,t_*) - \frac{\frac{1}{n}\sum_{j=1}^n e_j X_{j-1}\frac{1}{n}\sum_{j=1}^n g(j,t_*)X_{j-1}^2}{\frac{1}{n}\sum_{j=1}^n X_{j-1}^2}. \tag{5.2}
$$

In view of Lemma 5 and the assumptions on β_1, there exists a sequence $\{\varepsilon_n\}$ of positive reals such that $\varepsilon_n \to 0$, but $|\beta_1|\,\varepsilon_n\sqrt{n} \to \infty$ as $n \to \infty$. Then, on combining Lemmas 1–4, as $n \to \infty$,

$$
\max_{t_*}\big|R_n(t_*)\big| = o_P\Big(\frac{1}{\varepsilon_n\sqrt{n}}\Big) + O_P\Big(\frac{1}{\sqrt{n}}\Big) = o_P\big(|\beta_1|\big), \tag{5.3}
$$

and, with $t_* = \lfloor n\tau_*\rfloor$, for every fixed $\tau_* \in [0, 1-\delta]$,

$$
f_n(\tau_*) := \frac{1}{\beta_1^2}h_n(t_*) \xrightarrow{P} \frac{\sigma^2}{1-\beta_0^2}\frac{\big[\int_0^1 g_0(x-\tau_0)g_0(x-\tau_*)dx - \int_0^1 g_0(x-\tau_0)dx\int_0^1 g_0(x-\tau_*)dx\big]^2}{\int_0^1 g_0^2(x-\tau_*)dx - \big(\int_0^1 g_0(x-\tau_*)dx\big)^2}
$$

$$=: f(\tau_*), \tag{5.4}$$

as a consequence of (5.1)–(5.3) in combination with the approximations obtained in Lemmas 1–4.

Note that the denominator in f is bounded away from 0 on $[0, 1-\delta]$, with $0 < \delta < 1$, since, via Jensen's inequality, first for $\tau_* \in [\tilde{\delta}, 1-\delta]$, with $0 < \tilde{\delta} < 1-\delta$,

$$\begin{aligned}
&\int_{\tau_*}^1 g_0^2(x-\tau_*)dx - (1-\tau_*)^2\Big(\int_{\tau_*}^1 g_0(x-\tau_*)\frac{dx}{1-\tau_*}\Big)^2 \\
&\quad \geq \int_{\tau_*}^1 g_0^2(x-\tau_*)dx - (1-\tau_*)^2\int_{\tau_*}^1 g_0^2(x-\tau_*)\frac{dx}{1-\tau_*} = \tau_*\int_{\tau_*}^1 g_0^2(x-\tau_*)dx \\
&\quad \geq \tilde{\delta}\int_{\tau_*}^1 g_0^2(x-\tau_*)dx = \tilde{\delta}\int_0^{1-\tau_*} g_0^2(x)dx \geq \tilde{\delta}\int_0^{\delta} g_0^2(x)dx > 0.
\end{aligned} \tag{5.5}$$

Secondly, since $\int_0^1 g_0^2(x)dx - \big(\int_0^1 g_0(x)dx\big)^2 > 0$ and the denominator in f is continuous in τ_*, also

$$\int_0^1 g_0^2(x-\tau_*)dx - \Big(\int_0^1 g_0(x-\tau_*)dx\Big)^2 > 0, \tag{5.6}$$

for $\tau_* \in [0, \tilde{\delta})$, with some $\tilde{\delta} > 0$.

So, (5.5) and (5.6) prove the positivity on $[0, 1-\delta]$ and, in view of Lemmas 1–4, the latter positivity also implies that the convergence in (5.4) is uniform on $[0, 1-\delta]$, i.e.

$$\max_{\tau_* \in [0,1-\delta]} \big|f_n(\tau_*) - f(\tau_*)\big| \xrightarrow{P} 0 \quad (n \to \infty). \tag{5.7}$$

We finally show that the limit function f has a *unique* maximum at $\tau_* = \tau_0$. First, via the Cauchy-Schwarz inequality,

$$\begin{aligned}
&\Big[\int_0^1 g_0(x-\tau_0)g_0(x-\tau_*)dx - \int_0^1 g_0(x-\tau_0)dx\int_0^1 g_0(x-\tau_*)dx\Big]^2 \\
&\leq \Big[\int_0^1 g_0^2(x-\tau_0)dx - \Big(\int_0^1 g_0(x-\tau_0)dx\Big)^2\Big]\Big[\int_0^1 g_0^2(x-\tau_*)dx - \Big(\int_0^1 g_0(x-\tau_*)dx\Big)^2\Big].
\end{aligned} \tag{5.8}$$

Hence

$$f(\tau_*) \leq \frac{\sigma^2}{1-\beta_0^2}\Big[\int_0^1 g_0^2(x-\tau_0)dx - \Big(\int_0^1 g_0^2(x-\tau_0)dx\Big)^2\Big],$$

and the bound is attained for $\tau_* = \tau_0$.

It remains to prove that τ_0 is the *unique* maximizer of f. To do so, we show that *strict* inequality holds in (5.8), if $\tau_* \neq \tau_0$. Assume equality. Then, there is a $\lambda \neq 0$ such that, for almost every $x \in [0, 1]$,

$$\widetilde{g}_0(x-\tau_*) := g_0(x-\tau_*) - \int_0^1 g_0(x-\tau_*)dx = \lambda\Big(g_0(x-\tau_0) - \int_0^1 g_0(x-\tau_0)dx\Big)$$
$$=: \lambda\widetilde{g}_0(x-\tau_0).$$

This, however, is impossible, since, e.g., for $\tau_* > \tau_0$,

$$\widetilde{g}_0(\cdot - \tau_*) \text{ is constant on } [\tau_0, \tau_*], \text{ but } \lambda\widetilde{g}_0(\cdot - \tau_0) \text{ is not.}$$

Similarly, for $\tau_* < \tau_0$,

$$\lambda\widetilde{g}_0(\cdot - \tau_0) \text{ is constant on } [\tau_*, \tau_0], \text{ but } \widetilde{g}_0(\cdot - \tau_*) \text{ is not,}$$

which proves that τ_0 is the unique maximizer of f.

Via the subsequence principle for convergence in probability, the proof of Theorem 1 can now be completed from (5.7) by applying Lemma 6. □

Proof of Remark 5 In case of an unknown change function g, it is obvious from the proof of Theorem 1 that, under (2.2), $g(t, t_*) = g_0((t - t_*)/n)$ in (1.11) resp. (1.13) can be replaced by $\widehat{g}_n(t, t_*) = \widehat{g}_0((t-t_*)/n)$. The reason is that, in view of (5.1)–(5.3) and the rate $o_P(\beta_1)$, the convergence in (5.4) still holds with estimated $g(t, t_*)$'s, so that the proof can be completed as before.

If $g_0(x) = x_+^\kappa$, with some $\kappa \geq 1$, and $|\widehat{\kappa}_n - \kappa| = o_P(\beta_1)$ as $n \to \infty$ for an estimator $\widehat{\kappa}_n$, note that on $\{\widehat{\kappa}_n > \kappa_1\}$, with $0 < \kappa_1 < \kappa$, by the mean-value theorem, for some $\bar{\kappa}_n$ between $\widehat{\kappa}_n$ and κ,

$$\max_{x\in(0,1]} |x^{\widehat{\kappa}_n} - x^\kappa| = \max_{x\in(0,1]} |\log x||x^{\bar{\kappa}_n}||\widehat{\kappa}_n - \kappa| \leq \max_{x\in(0,1]} |\log x||x^{\kappa_1}||\widehat{\kappa}_n - \kappa|.$$

In view of $|\log x||x^{\kappa_1}| \to 0$ as $x \downarrow 0$, $|\log x||x^{\kappa_1}|$ is bounded on $(0, 1]$, which suffices to prove that

$$\max_{x\in(0,1]} |x^{\widehat{\kappa}_n} - x^\kappa|/\beta_1 = \max_{x\in[0,1]} |x^{\widehat{\kappa}_n} - x^\kappa|/\beta_1 \xrightarrow{P} 0,$$

since the max is attained for $x > 0$ and $P(\widehat{\kappa}_n > \kappa_1) \to 1$ as $n \to \infty$. □

Proof of Theorem 2 In view of the consistency obtained in Theorem 1, it suffices to concentrate on a small neighbourhood $[\tau_1, \tau_2]$ of τ_0. With the notations in (5.4), we have

$$f_n(\widehat{\tau}_0) - f(\tau_0) = \max_{\tau_*\in[\tau_1,\tau_2]} f_n(\tau_*) - \max_{\tau_*\in[\tau_1,\tau_2]} f(\tau_*) = \big(f_n(\widehat{\tau}_0) - f(\widehat{\tau}_0)\big) + \big(f(\widehat{\tau}_0) - f(\tau_0)\big)$$
$$= \big(f_n(\widehat{\tau}_0) - f(\widehat{\tau}_0)\big) + \Big(f'(\tau_0)(\widehat{\tau}_0 - \tau_0) + f''(\tau_n)\frac{(\widehat{\tau}_0 - \tau_0)^2}{2}\Big),$$

where τ_n is between $\widehat{\tau}_0$ and τ_0.

Since $f'(\tau_0) = 0$ and $|f''(\tau_n)| \ge C$, for some $C > 0$, this results in the estimate

$$\begin{aligned}|\widehat{\tau}_0 - \tau_0|^2 &\le \frac{2}{C}\Big(\big|\max_{\tau_*\in[\tau_1,\tau_2]} f_n(\tau_*) - \max_{\tau_*\in[\tau_1,\tau_2]} f(\tau_*)\big| + \max_{\tau_*\in[\tau_1,\tau_2]} \big|f_n(\tau_*) - f(\tau_*)\big|\Big)\\ &\le \frac{4}{C}\max_{\tau_*\in[\tau_1,\tau_2]} \big|f_n(\tau_*) - f(\tau_*)\big|.\end{aligned}$$

Now, on checking the steps in the proof of Theorem 1 more carefully, one can show that

$$\max_{\tau_*\in[\tau_1,\tau_2]} \big|f_n(\tau_*) - f(\tau_*)\big| = O_P\big(|\beta_1|\big) + o_P\Big(\frac{1}{|\beta_1|\,\varepsilon_n\sqrt{n}}\Big). \tag{5.9}$$

Note that, in view of Lemmas 2–4, the uniform rate of approximation for the denominator of f_n in (5.4) by that of f is $O_P(|\beta_1|) + o_P(1/(\varepsilon_n\sqrt{n}))$, whereas the one for the numerator is $O_P(|\beta_1|) + o_P(1/(|\beta_1|\,\varepsilon_n\sqrt{n}))$. This, together with the positivity of the denominator of f on $[\tau_1, \tau_2]$, results in (5.9) after some elementary calculation, which completes the proof. □

Proof of Remark 7 Under (2.4), the approximation rate $O_P\big(|\beta_1|\big)+o_P\big(1/(|\beta_1|\,\varepsilon_n\sqrt{n})\big)$ in (5.9) still holds, since only an additional rate $O_P\big(1/\sqrt{n}\big)$ would have to be added in the denominator of f_n and $O_P\big(1/(|\beta_1|\sqrt{n})\big)$ in the numerator, which are negligible compared to the other terms. □

Proof of Theorem 3 The proof follows the usual lines of proofs of the limit behavior of estimators under a gradual change (see, e.g., Hušková (1998b) or Jarušková (1998a)). Therefore we will focus on the main steps only. □

Gathering the assertions in Propositions 1 –3 and Theorem 1, we can conclude that $\widehat{t}_0$ has the same asymptotic distribution as

$$\begin{aligned}\widehat{t}_0(r_n\sqrt{n}|\beta_1|^{-1}) = &\operatorname*{arg\,max}_{|t_*-t_0|\le\sqrt{n}|\beta_1|^{-1}r_n}\\ &\Big\{-\beta_1^2\Big(\frac{|t_0-t_*|}{n}\Big)^2\frac{\sigma^2}{1-\beta_0^2}\,\widetilde{H} + \beta_1\Big(\frac{t_0-t_*}{n}\Big)\frac{2}{\sqrt{n}}\frac{1}{\sqrt{n}}\sum_{t=1}^{n} e_t X_{t-1}\widetilde{h}_n(t, t_*, t_0)\Big\}.\end{aligned}$$

By Lemma 7, particularly by (4.24), its limit behavior does not change if $\widetilde{h}_n(t, t_*, t_0)$ is replaced by $\widetilde{h}_n(t, t_{**}, t_0)$, with any fixed t_{**} such that $|t_{**} - t_0| \le r_n\sqrt{n}\,|\beta_1|^{-1}$. Also, by Lemma 7,

$$Z_n = \frac{1}{\sqrt{n}}\sum_{t=1}^{n} e_t X_{t-1}\widetilde{h}_n(t, t_{**}, t_0)$$

has an asymptotic normal distribution with zero mean and variance

$$E\Big(\frac{1}{\sqrt{n}}\sum_{t=1}^{n} e_t X_{t-1}\widetilde{h}_n(t, t_{**}, t_0)\Big)^2 = \frac{\sigma^4}{1-\beta_0^2}\widetilde{H}\big(1+o(1)\big).$$

Thus it suffices to study the estimator

$$\widetilde{t}_0\left(r_n\sqrt{n}\,|\beta_1|^{-1}\right) = \underset{|t_*-t_0|\le\sqrt{n}|\beta_1|^{-1}r_n}{\arg\max}\Big\{-\beta_1^2\Big(\frac{|t_0-t_*|}{n}\Big)^2\frac{\sigma^2}{1-\beta_0^2}\,\widetilde{H}+\beta_1\sqrt{\widetilde{H}}\,\frac{t_0-t_*}{n}\,\frac{2}{\sqrt{n}}\,\frac{Z_n}{\sqrt{\widetilde{H}}}\Big\}.$$

The boundary can only be a attained with probability tending to 0, since $Z_n = O_P(1)$, and, on plugging the upper bound into the above expression, one gets

$$-\beta_1^2\Big(\frac{\sqrt{n}\,|\beta_1|^{-1}r_n}{n}\Big)^2\frac{\sigma^2}{1-\beta_0^2}\widetilde{H}+\beta_1\,\frac{\sqrt{n}\,|\beta_1|^{-1}r_n}{n}\,\frac{2}{\sqrt{n}}\,Z_n = -\frac{r_n^2}{n}\,\frac{\sigma^2}{1-\beta_0^2}\widetilde{H}+\frac{r_n}{n}\beta_1|\beta_1|^{-1}\,2Z_n.$$

Using the assumption on r_n (see (4.23)) together with $Z_n = O_P(1)$, it is now straightforward to check that this estimator satisfies

$$\sqrt{n}\,\beta_1\,\frac{\sigma^2}{1-\beta_0^2}\,\widetilde{H}\,\frac{\widetilde{t}_0\left(r_n\sqrt{n}\,|\beta_1|^{-1}\right)-t_0}{n} = Z_n\left(1+o_P(1)\right).$$

From here and the asymptotic normality of Z_n it can be concluded that the assertion of Theorem 3 holds true. □

Acknowledgements The authors thank two anonymous referees for their constructive suggestions, which led to improve the presentation of their results. They also gratefully acknowledge support by the grants GAČR 18-08888S (M. Hušková) and GX19-28231X (Z. Prášková) of the Czech Science Foundation and by the partnership program of Charles University in Prague with the University of Cologne.

Funding Open Access funding enabled and organized by Projekt DEAL.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

Albin JMP, Jarušková D (2003) On a test statistic for linear trend. Extremes 6:247–258
Aue A, Horváth L (2013) Structural breaks in time series. J Time Ser Anal 34:1–16

Aue A, Steinebach J (2002) A note on estimating the change-point of a gradually changing stochastic process. Stat Probab Lett 56:177–191
Bai J (1994) Least squares estimation of a shift in linear processes. J Time Ser Anal 15:453–472
Bai J (2000) Vector autoregressive models with structural changes in regression coefficients and in variance-covariance matrices. Ann Econom Finance 1:303–339
Billingsley P (1968) Convergence of probability measures. Wiley, New York
Davidson D (1994) Stochastic limit theory. Oxford University Press, Oxford
Davis RA, Huang D, Yao YC (1995) Testing for a change in the parameter value and order of an autoregressive model. Ann Stat 23:282–304
Döring M (2015a) Asymmetric cusp estimation in regression models. Statistics 49:1279–1297
Döring M (2015b) Rate of convergence of a change point estimator in a misspecified regression model. In: A. Steland, E. Rafajłowicz, K. Szajowski (eds.) Stochastic Models, Statistics and Their Applications (Wrocław, 2015), *Springer Proc. Math. Statist.*, vol. 122, pp. 49–55. Springer, Cham
Döring M, Jensen U (2015) Smooth change point estimation in regression models with random design. Ann Inst Stat Math 67:595–619
Dümbgen L (1991) The asymptotic behavior of some nonparametric change-point estimators. Ann Stat 19:1471–1495
Dvořák M (2015) Stability in Autoregressive Time Series Models. Ph.D. thesis, Charles University, Prague
Gombay E (2008) Change detection in autoregressive time series. J Multiv Anal 99:451–464
He C, Teräsvirta T, González A (2008) Testing parameter constancy in stationary vector autoregressive models against continuous change. Econom Rev 28:225–245
Hlávka Z, Hušková M (2017) Two-sample gradual change analysis. Revstat 15:355–372
Hoffmann M, Dette H (2019) On detecting changes in the jumps of arbitrary size of a time-continuous stochastic process. Electr J Stat 13:3654–3709
Hoffmann M, Vetter M, Dette H (2018) Nonparametric inference of gradual changes in the jump behaviour of time-continuous processes. Stoch Process Appl 128:3679–3723
Horváth L (1993) Change in autoregressive processes. Stoch Process Appl 44:221–242
Hušková M (1998a) Estimators in the location model with gradual changes. Comment Math Univ Carolin 39:147–157
Hušková M (1998b) Remarks on test procedures for gradual changes. In: Szyszkowicz B (ed) Asymptotic methods in probability and statistics (Ottawa, 1997). North-Holland, Amsterdam, pp. 577–583
Hušková M (1999) Gradual changes versus abrupt changes. J Statist Plann Infer 76:109–125
Hušková M (2001) A note on estimators of gradual changes. In: M. de Gunst, C. Klaassen, A. van der Vaart (eds.) State of the Art in Probability and Statistics (Leiden, 1999), *IMS Lecture Notes Monogr. Ser.*, vol. 36, pp. 345–358. Institute of Mathematical Statistics, Beachwood, OH
Hušková M, Kirch C, Prášková Z, Steinebach J (2008) On the detection of changes in autoregressive time series. II. Resampling procedures. J Stat Plann Infer 138:1697–1721
Hušková M, Prášková Z, Steinebach J (2007) On the detection of changes in autoregressive time series. I. Asymptotics. J Stat Plann Infer 137:1243–1259
Hušková M, Steinebach J (2000) Limit theorems for a class of tests of gradual changes. J Stat Plann Infer 89:57–77
Hušková M, Steinebach J (2002) Asymptotic tests for gradual changes. Stat Decis 20:137–151
Jarušková D (1998a) Change-point estimator in gradually changing sequences. Comment Math Univ Carolin 39:551–561
Jarušková D (1998b) Testing appearance of linear trend. J Stat Plann Infer 70:263–276
Jarušková D (1999) Testing appearance of polynomial trend. Extremes 2:25–37
Jarušková D (2001) Change-point estimator in continuous quadratic regression. Comment Math Univ Carolin 42:741–752
Jarušková D (2002) Change in polynomial regression and related processes. Theory Stoch Process 8:162–168
Jarušková D (2003) Asymptotic distribution of a statistic testing a change in simple linear regression with equidistant design. Stat Probab Lett 64:89–95
Kirch C, Muhsal B, Ombao H (2015) Detection of changes in multivariate time series with application to EEG data. J Am Stat Assoc 110:1197–1216
Kirch C, Steinebach J (2006) Permutation principles for the change analysis of stochastic processes under strong invariance. J Comput Appl Math 186:64–88
Picard D (1985) Testing and estimating change-points in time series. Adv Appl Probab 17:841–867

Quessy JF (2019) Consistent nonparametric tests for detecting gradual changes in the marginals and the copula of multivariate time series. Stat Papers 60:717–746
Račkauskas A, Tamulis A (2013) Modeling of gradual epidemic changes. Liet Mat Rink, Proc Lith Math Soc, Ser A 54:55–60
Salazar D (1982) Structural changes in time series models. J Econom 19:147–163
Slabý A (2001) Limit theorems for rank statistics detecting gradual changes. Comment Math Univ Carolin 42:591–600
Slama A, Saggou H (2017) A Bayesian analysis of a change in the parameters of autoregressive time series. Comm Stat Simul Comput 46:7008–7021
Steinebach J (2000) Some remarks on the testing of smooth changes in the linear drift of a stochastic process. Theory Probab Math Stat 61:173–185
Steinebach J, Timmermann H (2011) Sequential testing of gradual changes in the drift of a stochastic process. J Stat Plann Infer 141:2682–2699
Timmermann H (2014) Monitoring Procedures for Detecting Gradual Changes. Ph.D. thesis, University of Cologne
Timmermann H (2015) Sequential detection of gradual changes in the location of a general stochastic process. Stat Probab Lett 99:85–93
Venkatesan D, Arumugam P (2007) Bayesian analysis of structural changes in autoregressive models. Am J Math Manag Sci 27:153–162
Vogt M, Dette H (2015) Detecting gradual changes in locally stationary processes. Ann Stat 43:713–740
Wang L (2007) Gradual changes in long memory processes with applications. Statistics 41:221–240