Econometric Theory, **33**, 2017, 366–412.
doi:10.1017/S0266466615000468

ASYMPTOTIC PROPERTIES OF THE CUSUM ESTIMATOR FOR THE TIME OF CHANGE IN LINEAR PANEL DATA MODELS

LAJOS HORVÁTH
University of Utah

MARIE HUŠKOVÁ
Charles University

GREGORY RICE
University of Waterloo

JIA WANG
University of Utah

We consider the problem of estimating the common time of a change in the mean parameters of panel data when dependence is allowed between the cross-sectional units in the form of a common factor. A CUSUM type estimator is proposed, and we establish first and second order asymptotics that can be used to derive consistent confidence intervals for the time of change. Our results improve upon existing theory in two primary directions. Firstly, the conditions we impose on the model errors only pertain to the order of their long run moments, and hence our results hold for nearly all stationary time series models of interest, including nonlinear time series like the ARCH and GARCH processes. Secondly, we study how the asymptotic distribution and norming sequences of the estimator depend on the magnitude of the changes in each cross-section and the common factor loadings. The performance of our results in finite samples is demonstrated with a Monte Carlo simulation study, and we consider applications to two real data sets: the exchange rates of 23 currencies with respect to the US dollar, and the GDP per capita in 113 countries.

1. INTRODUCTION

In this paper, we consider the problem of estimating the time of a change in the mean present in panel data in which there are N cross-sectional units comprised of time series data of length T. A common structural break in panel data is a quite natural occurrence. For example, if the panel data under consideration is

We would like to thank four reviewers, the editor, Peter C.B. Phillips as well as the co–editor, Liangjun Su, for their helpful comments that led to substantial improvements to this paper. This project was supported by NSF grant DMS 1305858 and grants GAČR 15-096635, GAČR 201/12/1277. Address correspondence to Lajos Horváth, Department of Mathematics, University of Utah, Salt Lake City, UT 84112–0090 USA; e-mail: horvath@math.utah.edu.

comprised of exchange rates of various currencies with respect to US dollars, then a crisis in the US would be expected to simultaneously affect each cross-sectional unit. Similar phenomena may be produced by governmental policy changes, the introduction of a new technology, etc., and in these cases it is of interest to estimate the time at which such occurrences are manifested in sample data. The theory of change point analysis has been extensively developed to study problems of this nature; see Csörgő and Horváth (1997), Brodsky and Darkhovskii (2002), Aue and Horváth (2012), and Horváth and Rice (2014) for reviews of the field.

Classical methods in change point analysis consider univariate and multivariate data of a fixed dimension. In many panel data examples, however, the number of cross-sections N is comparable in size to the length of the series T. In these cases, asymptotics as T remains fixed and N tends to infinity, or as N and T jointly tend to infinity, are more appropriate. Although in principle one could detect the common change present in each cross-sectional unit by examining a single series, an analysis that utilizes all available series should provide improved detection and estimation.

The literature on structural breaks in panel data has grown considerably in the last two decades. We refer to Arellano (2003), Hsiao (2003, 2007), and Baltagi (2013) for surveys of several panel data models and their applications to econometrics and finance. The early foundations for estimating structural changes in panel data were developed in Joseph and Wolfson (1992, 1993), and many aspects of the problem have now seen at least some consideration; Li, Qian, and Su (2015) and Qian and Su (2015) consider multiple structural breaks in panel data, and Kao Trapani, and Urga (2015) considers break testing under cointegration.

Bai (2010), Kim (2011, 2014), Baltagi, Kao, and Liu (2012), Baltagi, Feng, and Kao (2015), and Horváth and Hušková (2012) are most closely related to the present paper. Bai (2010) considers the problem of estimating a common break in the means of panel data that do not exhibit cross sectional dependence. A least squares estimator is proposed that is shown to be consistent when N tends to infinity, and its asymptotic properties are derived as N and T jointly tend to infinity. Kim (2011, 2014) and Baltagi et al. (2012, 2015) extend the methodology of the least squares estimator of Bai (2010) to panels exhibiting cross-sectional dependence modeled by common factors, and to detect a change in the time trend or slope parameters of a panel regression. Horváth and Hušková (2012) study testing for the presence of a common change in the mean using a CUSUM estimator under cross-sectional dependence. In each of these papers, asymptotics are derived assuming the model errors are linear processes, and that the rates of divergence relative to N and T of the size of the changes and the magnitudes of the factor loadings are fixed.

In this paper, we extend the existing theory in two primary directions. We derive second order asymptotics for the CUSUM change point estimator assuming only an order condition on the long run moments. This extends the asymptotic theory of change point estimation to a wide variety of error processes, notably many nonlinear time series examples like the ARCH and GARCH processes. We also show explicitly how the asymptotic distribution and norming sequences of the estimator depend on the magnitude of the changes in each panel and the common factor loadings. This allows for the computation of the limit distribution under several conceivable rates of divergence for the magnitudes of changes and factor loadings.

The remainder of the paper is organized as follows. In Section 2, we present our assumptions and the main results of the paper. Section 3 contains examples of error processes that satisfy the assumptions of Section 2. Estimators for the norming sequences that appear in the results of Section 2 are developed and studied in Section 4. The implementation of the results of the paper as well as a Monte Carlo simulation study and data applications are detailed in Section 5. The proofs of all results are contained in Appendices A and B.

2. ASSUMPTIONS, AND MAIN RESULTS

We consider the panel data model

$$X_{i,t} = \mu_i + \delta_i I\{t > t_0\} + \gamma_i \eta_t + e_{i,t}, \quad 1 \le i \le N, 1 \le t \le T, \tag{2.1}$$

where the idiosyncratic errors $e'_{i,t}s$ have mean zero, η_t denotes the common factor with loadings γ_i, $1 \le i \le N$, and δ_i denotes the change in the mean of panel i that occurs at the common, and unknown, change point t_0.

Assumption 2.1.

(i) The sequences $\{e_{i,t}, -\infty < t < \infty\}, 1 \le i \le N$ are independent, and

(ii) $\{\eta_t, -\infty < t < \infty\}$ and $\{e_{i,t}, -\infty < t < \infty\}$, $1 \le i \le N$ are stationary.

According to Assumption 2.1(i), the only source of dependence between the panels is the common factor η_t. The idiosyncratic errors form a stationary time series, similar to the assumption in Bai (2010) and Kao Trapani, and Urga (2012, 2015). Throughout this paper δ_i and γ_i, $1 \le i \le N$, are allowed to depend on N and T. For the sake of simplicity, we consider the case when $\gamma_i \in R$, but our results could be extended to the more general case of a vector valued factor loading and common factor.

Assumption 2.2. The time of change in the mean t_0 satisfies

$$t_0 = \lfloor T\theta \rfloor \quad \text{with some} \quad 0 < \theta < 1.$$

Assumption 2.2 is standard in change point analysis, and corresponds with the assumptions of Bai (2010), Kim (2011, 2014), and Horváth and Hušková (2012). It is of interest in some econometric applications to allow for θ to depend on N and T and tend to the end points 0 or 1 at a certain rate; see Andrews (2003) and, in the panel data setting, Qian and Su (2015). The consideration of this problem in generality for our estimator is not a goal of the present paper, and requires a thorough study.

The estimator that we use for t_0 is defined as the location of the maximum of the sum of the CUSUM processes over cross–sections:

$$\hat{t}_{N,T} = \text{argmax}_{1\le t<T} \sum_{i=1}^{N} \left(S_i(t) - \frac{t}{T} S_i(T) \right)^2,$$

where

$$S_i(t) = \sum_{s=1}^{t} X_{i,s}.$$

The estimator of Bai (2010) is

$$t^*_{N,T} = \text{argmax}_{1\le t<T} \sum_{i=1}^{N} \left(S_i(t) - \frac{t}{T} S_i(T) \right)^2 \frac{1}{(t(T-t))}, \qquad \textbf{(2.2)}$$

which is the maximum likelihood estimator for t_0 assuming that the panels are independent and normally distributed with the same variance. The estimator $\hat{t}_{N,T}$ maximizes the weighted log likelihood; see e.g. Hawkins (1986), Csörgő and Horváth (1997, Sects. 2.1 and 2.8.1) for the case when $N = 1$, and Chan et al. (2013). Due to the weight $(t(T-t))^{-1}$ in the definition of $t^*_{N,T}$, the limit distribution of $t^*_{N,T}$ does not depend on the break fraction θ; see Bai (2010).

We impose only conditions on the long run moments of the error processes for our asymptotic results. The long run moments of the errors in panel i are defined by

$$U_{i,\nu}(t) = E\left| \sum_{s=1}^{t} e_{i,s} \right|^{\nu}, \quad 1 \le i \le N,$$

and we assume that they satisfy the following conditions:

Assumption 2.3.

(i) There exists $\sigma_i, 1 \le i \le N$ such that

$$\max_{1\le i\le N} \sup_{1\le t\le T} \left| \frac{1}{t} U_{i,2}(t) - \sigma_i^2 \right| = o(1),$$

where $C_1 \le \sigma_i \le C_2$ for all $1 \le i \le N$ with some $0 < C_1 \le C_2 < \infty$, and

(ii)

$$\frac{1}{N} \sup_{1\le t\le T} \sum_{i=1}^{N} \left(\frac{1}{t^{\kappa/2}} U_{i,\kappa}(t) \right)^2 = O(1) \text{ with some } \kappa > 4.$$

Additionally, we must assume an analogous condition on the common factors:

Assumption 2.4. With some $\bar{\kappa} > 2$.

$$E\eta_t = 0,\ E\left(\sum_{s=1}^{t}\eta_s\right)^2 = t + o(t),\ \text{ and } E\left|\sum_{s=1}^{t}\eta_s\right|^{\bar{\kappa}} = O(t^{\bar{\kappa}/2}),\ \text{as } t \to \infty.$$

Assumptions 2.3 and 2.4 do not assume any specific structure on the error terms, in contrast to the structural break literature with panel data to date. We provide several examples in Section 3, including linear and nonlinear time series, martingales, and mixing sequences, where Assumptions 2.3 and 2.4 are satisfied.

The size of the changes and the correlation between the panels will play a crucial role in the asymptotic distribution of the estimator, and these quantities will be measured by

$$\Delta_{N,T} = \sum_{i=1}^{N}\delta_i^2,\quad \Gamma_{N,T} = \sum_{i=1}^{N}\gamma_i^2 \text{ and } \Sigma_{N,T} = \sum_{i=1}^{N}\delta_i\gamma_i.$$

The limit results below are proven when $\min(N,T) \to \infty$.

Assumption 2.5. As $\min(N,T) \to \infty$,

(i) $$\frac{T\Delta_{N,T}}{N} \to \infty,$$

and

(ii) $$\frac{\Gamma_{N,T}}{(T\Delta_{N,T})^{1/2}} \to 0.$$

Assumption 2.5 means that the sizes of all changes cannot be too small and that the factor loadings cannot be much larger than the sample size and the size of the changes. Bai (2010) assumes that $\Delta_{N,T}/N$ converges to a positive limit while under the assumptions of Kim (2011), the common factor dominates. A primary goal of our paper is to show how the relationship between the loadings and the sizes of the changes affect the limit distribution of the time of change estimator.

Our first result pertains to the asymptotic distribution of $\hat{t}_{N,T}$ when $\Delta_{N,T}$ is large.

THEOREM 2.1. *If Assumptions 2.1–2.5 hold,*

$$\Delta_{N,T} \to \infty \tag{2.3}$$

https://doi.org/10.1017/S0266466615000468 Published online by Cambridge University Press

and

$$\frac{\Sigma_{N,T}}{\Delta_{N,T}} = o(1), \tag{2.4}$$

as $N, T \to \infty$*, then we have that*

$$P\left\{\hat{t}_{N,T} = t_0\right\} \to 1. \tag{2.5}$$

The assumption in (2.4) may seem somewhat restrictive since it rules out the example of simultaneous fixed break sizes and factor loadings. We note that due to the result on page 635 of Horváth and Hušková (2012), when the factor loadings are fixed, the CUSUM test for the presence of a change point will reject with probability tending to one regardless of if a change exists or not, and so something along the lines of (2.4) must be assumed for the CUSUM estimator of the time of change to be consistent.

Remark 2.1. Assume that T is fixed. If Assumptions 2.1–2.4 are satisfied and $\Gamma_{N,T}/\Delta_{N,T} \to 0$, and $\Delta_{N,T}/N \to \infty$, then (2.5) holds.

Remark 2.2. In order to establish the consistency of Bai's (2010) estimator in (2.2) for fixed T and under our assumptions, we must assume in addition that for each i, $e_{i,t}$ and η_t are uncorrelated random variables, and that $\{e_{i,t}, 0 \le t < \infty, 1 \le i \le N\}$ and $\{\eta_t, 0 \le t < \infty\}$ are independent. If in addition to Assumptions 2.1–2.4, $E\eta_0^4 < \infty$, $\Delta_{N,T}/N^{1/2} \to \infty$ and $\Gamma_{N,T}/\Delta_{N,T} \to 0$ hold, then we have that

$$\lim_{N\to\infty} P\left\{t^*_{N,T} = t_0\right\} = 1. \tag{2.6}$$

We provide a proof of Remark 2.2 in Appendix A.

The main difference between Remarks 2.1 and 2.2 is in the assumption that $\Delta_{N,T}/N \to \infty$ and $\Delta_{N,T}/N^{1/2} \to \infty$. Remark 2.2 allows smaller changes to establish consistency, but much stronger assumptions on the sequences $e_{i,t}, 1 \le i \le N$ and η_t. If we cannot assume that $e_{i,t}, 1 \le i \le N$ and η_t are sequences of uncorrelated random variables, and the independence of $\{\eta_t, t \ge 0\}$ and $\{e_{i,t}, t \ge 0, 1 \le i \le N\}$, then (2.6) can be proven under conditions of Remark 2.1. In this case (2.5) and (2.6) can hold only if $\Delta_{N,T}/N \to \infty$ when T is fixed.

We now turn to the asymptotic distribution of $\Delta_{N,T}(\hat{t}_{N,T} - t_0)$ when (2.3) does not hold, i.e. the sizes of the changes are small, or occur in only a few panels:

Assumption 2.6.

(i) $\Delta_{N,T} = O(1)$,
(ii) $T\Delta_{N,T}(1+\log(T/\Delta_{N,T}))^{-2/\bar{\kappa}} \to \infty$, where $\bar{\kappa}$ is defined in Assumption 2.4, and
(iii) $T^{1-2/\kappa}\Delta_{N,T}/N^{1/2} \to \infty$, where κ is defined in Assumption 2.3(ii).

By Assumption 2.5, we have that $T\Delta_{N,T} \to \infty$, so Assumption 2.6(ii) holds if, $\Delta_{N,T}T(\log T)^{-2/\bar{\kappa}} \to \infty$. If $N/T^{4/\kappa} \to \infty$, i.e., the number of the cross-sections is large, then Assumption 2.6(iii) follows from Assumption 2.5. However, Assumption 2.5(ii) also holds if the number of cross-sectional units is relatively small with respect to the length of the panels and the sizes of the changes.
Next we introduce an assumption that is a companion to Assumption 2.3:

Assumption 2.7.

(i) $\frac{1}{N}\sum_{i=1}^{N}|Ee_{i,0}e_{i,t}| = O(t^{-\tau})$ with some $\tau > 2$, and,

(ii) $\max_{1\le i\le N} U_{i,\bar{\tau}}(t) = O\left(t^{\bar{\tau}/2}\right)$ with some $\bar{\tau} > 2$.

Assumption 2.7 requires an upper bound for the average autocorrelation of the idiosyncratic errors, and a uniformity condition that augments Assumption 2.3(ii).

Our first result in this direction covers the case when the sizes of the changes are small and the effect of the correlation between the panels is negligible or moderate. We measure the dependence between the panels with respect to the sizes of the changes by

$$\mathfrak{s} = \lim_{\min(N,T)\to\infty} \frac{\Sigma_{N,T}}{\Delta_{N,T}^{1/2}}.$$

To describe the limit distribution of $\hat{t}_{N,T}$ we need to introduce a drift function

$$g_\theta(u) = \begin{cases} (1-\theta)|u|, & \text{if } u < 0 \\ \theta u, & \text{if } u \ge 0 \end{cases}$$

and an asymptotic variance term

$$\sigma^2 = \lim_{N,T\to\infty} \frac{1}{\Delta_{N,T}} \sum_{i=1}^{N} \delta_i^2 \sigma_i^2.$$

We note that by Assumption 2.3(i) we get that $C_1 \le \sigma \le C_2$. Let

$$\mathcal{U}_i(t) = \begin{cases} \sum_{s=1}^{t} e_{i,s}, & \text{if } t = 1,2,3,\ldots \\ 0, & \text{if } t = 0 \\ -\sum_{s=t}^{-1} e_{i,s}, & \text{if } t = -1,-2,-3,\ldots \end{cases}$$

https://doi.org/10.1017/S0266466615000468 Published online by Cambridge University Press

We define the function $\mathfrak{u}(s,t)$ as the asymptotic covariance of $\sum_{i=1}^{N}\delta_i\mathcal{U}_i(t)$, i.e. for all integers s and t

$$\mathfrak{u}(s,t)=\lim_{N\to\infty}E\left(\sum_{j=1}^{N}\delta_j\mathcal{U}_j(s)\right)\left(\sum_{i=1}^{N}\delta_i\mathcal{U}_i(t)\right). \tag{2.7}$$

We note that $\mathfrak{u}(s,t)=\lim_{N\to\infty}\sum_{i=1}^{N}\delta_i^2 E\mathcal{U}_i(s)\mathcal{U}_i(t)$. It follows from Assumption 2.3(i) that the function $\mathfrak{u}(s,t)$ is finite for all integers s and t. This function only appears in Theorem 2.2 when $\Delta_{N,T}$ is above some positive bound for all N and T. In this case $\mathfrak{u}(t,t)>0$, if $t\neq 0$.
The next result considers the case when the common factors are negligible.

THEOREM 2.2. *We assume that Assumptions 2.1–2.7 hold,*

$$\Delta_{N,T}^{-\bar{\tau}/2}\sum_{i=1}^{N}|\delta_i|^{\bar{\tau}}\to 0, \tag{2.8}$$

where $\bar{\tau}$ is defined in Assumption 2.7(ii), and

$$\mathfrak{s}=0. \tag{2.9}$$

(a) If

$$\Delta_{N,T}\to 0, \tag{2.10}$$

then we have

$$\frac{\Delta_{N,T}\left(\hat{t}_{N,T}-t_0\right)}{\sigma^2}\ \stackrel{\mathcal{D}}{\to}\ \operatorname{argmax}_u\{W(u)-g_\theta(u)\}, \tag{2.11}$$

where $W(u),-\infty<u<\infty$ is a two–sided Wiener process.
(b) If

$$\Delta_{N,T}\to\mathfrak{d}\in(0,\infty), \tag{2.12}$$

then we have

$$\hat{t}_{N,T}-t_0\ \stackrel{\mathcal{D}}{\to}\ \operatorname{argmax}_t\{\mathfrak{G}(t)-\mathfrak{d}g_\theta(t)\}, \tag{2.13}$$

where $\mathfrak{G}(t),t=0,\pm1,\pm2,\ldots$ is Gaussian with $E\mathfrak{G}(t)=0$ and $E\mathfrak{G}(s)\mathfrak{G}(t)=\mathfrak{u}(s,t)$.

Remark 2.3. Since the proofs of Theorem 2.2 and the results to follow depend on normal approximations for the sums $\sum_{i=1}^{N}\sum_{s=1}^{t}e_{i,s}$ and $\sum_{i=1}^{N}\left(\sum_{s=1}^{t}e_{i,s}\right)^2$, the independence of $\{e_{i,t},t\geq 0\}$ on i could be relaxed, as pointed out by Bai (2010) and Kim (2011). This would be an important consideration, for example, if the cross-sectional units of the panels are indexed by location, i.e. $i=\mathbf{i}$, a vector describing the location of each cross-section. In this case a spatial structure could be assumed on the errors. If Assumption 2.1(i) is replaced by a weak dependence or spatial assumption, the norming constants would change in our limit theorems.

Remark 2.4. If $\Delta_{N,T}^{-\bar{\tau}/2}\sum_{i=1}^{N}|\delta_i|^{\bar{\tau}} \to 0$ with some $\bar{\tau} > 2$ does not hold, then the limit in (2.13) might not be the argmax of a normal process with a drift. For example, if $\delta_1 = 1$ and $\delta_i = 0$ for all $i \geq 2$, then the limit in (2.13) is determined by the error terms $e_{1,t}, -\infty < t < \infty$.

Remark 2.5. The distribution of $\text{argmax}_u\{W(u) - g_\theta(u)\}$ is known explicitly. Its density was derived by Ferger (1994) from Corollary 4 of Bhattacharya and Brockwell (1976) (cf. Csörgő and Horváth (1997, p. 177)).

Remark 2.6. If (2.12) holds, $\gamma_i = 0$ for all $1 \leq i \leq N$ and $Ee_{i,0}e_{i,t} = 0$ for all $1 \leq i \leq N$ and $t \neq 0$, then we get the analogue of Theorem 4.2 of Bai (2010). In this case $\mathfrak{G}(s)$ is a Wiener process on integers, so the main difference between the limits in (2.11) and (2.13) is that the argmax is computed on the real line or on integers.

So far, in this paper, the common factor was treated as a part of the error term with a negligible contribution to the limiting behavior of the estimator. However, it has been observed in testing for changes in panel data that the effect of strong correlation between the cross-sectional units of the panels might make standard statistical procedures invalid (cf. Horváth and Hušková (2012)). Our next result covers the case when the order of the common factors and the sizes of the changes are essentially the same. Since the contribution of the η_t's to the limit will not be negligible, we need to specify the relation between the errors and the common factors:

Assumption 2.8. The sequences $\{\eta_t, -\infty < t < \infty\}$ and $\{e_{i,t}, -\infty < t < \infty, 1 \leq i \leq N\}$ are independent.

Similarly to $\mathcal{U}_i(t)$, we introduce

$$\mathcal{V}(t) = \begin{cases} \sum_{s=1}^{t} \eta_s, & \text{if } t = 1, 2, 3, \ldots \\ 0, & \text{if } t = 0 \\ -\sum_{s=t}^{-1} \eta_s, & \text{if } t = -1, -2, -3, \ldots, \end{cases}$$

which will be part of the limit distribution when (2.12) holds or $\Sigma_{N,T}$ is proportional to $\Delta_{N,T}$. In all the other cases we assume the asymptotic normality of $\mathcal{V}(t)$:

Assumption 2.9.

$$T^{-1/2}\sum_{t=1}^{\lfloor Tu \rfloor} \eta_t \xrightarrow{\mathcal{D}[0,1]} W(u), \text{ where } W \text{ is a Wiener process.}$$

https://doi.org/10.1017/S0266466615000468 Published online by Cambridge University Press

THEOREM 2.3. *We assume that Assumptions 2.1–2.9, and (2.8) hold, and*

$$0 < |\mathfrak{s}| < \infty. \tag{2.14}$$

(a) If (2.10) holds, then we have

$$\frac{\Delta_{N,T}(\hat{t}_{N,T} - t_0)}{\sigma^2 + \mathfrak{s}^2} \stackrel{\mathcal{D}}{\rightarrow} \operatorname{argmax}_u \{W(u) - g_\theta(u)\}, \tag{2.15}$$

where $W(u), -\infty < u < \infty$ is a two–sided Wiener process.
(b) If (2.12) holds, then we have

$$\hat{t}_{N,T} - t_0 \stackrel{\mathcal{D}}{\rightarrow} \operatorname{argmax}_t \{\mathfrak{G}(t) + \mathfrak{s}\mathfrak{d}^{1/2}\mathcal{V}(t) - \mathfrak{d} g_\theta(t)\}, \tag{2.16}$$

where $\mathfrak{G}(t), t = 0, \pm 1, \pm 2, \ldots$ is the Gaussian process defined in Theorem 2.2, independent of $\mathcal{V}(t), t = 0, \pm 1, \pm 2, \ldots$

The effect of correlation between the cross-sectional units of the panels is demonstrated in Theorem 2.3. The limit distribution in (2.15) remains the same as in (2.11) but the variance of the estimator increases by $\mathfrak{s}^2$. The effect of the common factor is more transparent in (2.16) since an additional term appears in the limit which depends on the distribution of the common factors.

THEOREM 2.4. *If Assumptions 2.1–2.9, and (2.8) hold, and*

$$|\mathfrak{s}| = \infty, \tag{2.17}$$

then we have

$$\frac{\Delta_{N,T}^2}{\Sigma_{N,T}^2}\left(\hat{t}_{N,T} - t_0\right) \stackrel{\mathcal{D}}{\rightarrow} \operatorname{argmax}_u \{W(u) - g_\theta(u)\}, \tag{2.18}$$

where $W(u), -\infty < u < \infty$ is a two–sided Wiener process.

Theorem 2.4 covers the case when the limit distribution of the estimator for the time of change is completely determined by the common factors. The limit distribution in (2.18) is the same as in (2.11) and (2.15) but the rate of convergence is much slower. The effect of having several panel cross-sections with changes is overshadowed by the strong influence of the common factor. For further results when the common factor is dominant in the limit distribution, we refer to Kim (2011).

3. EXAMPLES

In this section, we study some examples of error processes that satisfy the assumptions of Section 2. We restrict our attention to establishing Assumption 2.3 for examples of possible model error sequences $e_{i,t}$'s, but the same sequences could be used for the common factors as well.

Example 3.1
Let $\{e_{i,t}, -\infty < t < \infty\}$ be independent, identically distributed random variables with $Ee_{i,0} = 0$, $Ee_{i,0}^2 = \sigma_i^2$ and $E|e_{i,0}|^\kappa < \infty$. Due to independence we have that

$$U_{i,2}(t) = t\sigma_i^2 \quad \text{for all } t = 0, 1, 2 \ldots \tag{3.1}$$

By the Rosenthal inequality (cf. Petrov (1995, p. 59)) we obtain for all $t \geq 1$ that

$$U_{i,\kappa}(t) \leq C\left\{tE|e_{i,0}|^\kappa + t^{\kappa/2}\sigma_i^\kappa\right\} \leq Ct^{\kappa/2}\{E|e_{i,0}|^\kappa + \sigma_i^\kappa\},$$

where C is an absolute constant, depending on $\kappa > 2$ only. If the error terms in each panel are independent and identically distributed, then, assuming $C_1 \leq \sigma_i \leq C_2$ for all $1 \leq i \leq N$, Assumption 2.3(i) holds; Assumption 2.3(ii) is satisfied if

$$\frac{1}{N}\sum_{i=1}^N (E|e_{i,0}|^\kappa)^2 = O(1) \quad \text{with some } \kappa > 4. \tag{3.2}$$

If $\max_{1\leq i\leq N} E|e_{i,0}|^{\bar{\tau}} \leq C_0$ with some $\bar{\tau} > 2$ and $C_0 > 0$, then Assumption 2.7(ii) is also fulfilled.

ARMA processes are very often used in classical time series analysis and our next example shows that stationary ARMA processes satisfy the basic assumptions of the first section. We consider the more general case of linear processes, which are investigated by Bai (2010), Kim (2011), and Horváth and Hušková (2012).

Example 3.2
We assume that $\{\varepsilon_{i,t}, -\infty < t < \infty\}$ are independent and identically distributed random variables with $E\varepsilon_{i,0} = 0$ and $E|\varepsilon_{i,0}|^\kappa < \infty$ with some $\kappa > 4$. The error terms $e_{i,t}$ form a linear process given by

$$e_{i,t} = \sum_{\ell=0}^\infty c_{i,\ell}\varepsilon_{i,t-\ell},$$

where $\sup_\ell \ell^{-2-\alpha_i}|c_{i,\ell}| \leq C_i$ with some $C_i > 0$ and $\alpha_i > 0$. By the Phillips and Solo (1992) representation we get

$$\sum_{s=1}^t e_{i,s} - \bar{C}_i\sum_{s=1}^t \varepsilon_{i,s} = \sum_{j=-\infty}^\infty\left(\sum_{k=1}^t \bar{c}_{k-j}\right)\varepsilon_{i,j},$$

where $\bar{C}_i = \sum_{\ell=0}^\infty c_{i,\ell} \neq 0, \bar{c}_{i,0} = c_{i,0} - \bar{C}_i, \bar{c}_{i,\ell} = c_{i,\ell}$, if $\ell \geq 1$ and $\bar{c}_{i,\ell} = 0$, if $\ell \leq -1$. Minkowski's inequality and the discussion in Example 3.1 indicate that we need to choose $\sigma_i^2 = \bar{C}_i^2 E\varepsilon_{i,0}^2$ in Assumption 2.3(i) and we also have $U_{i,\kappa}(t) = O\left(t^{\kappa/2}\right)$ and $|Ee_{i,0}e_{i,t}| = O\left(t^{-2-\alpha_i}\right)$, as $t \to \infty$.

Example 3.3
Let us assume that $\{e_{i,t}, -\infty < t < \infty\}$ is a stationary orthogonal martingale difference sequence with respect to some filtration with $Ee_{i,0} = 0$, $Ee_{i,0}^2 = \sigma_i^2$ and $E|e_{i,0}|^\kappa < \infty$. Then (3.1) as well as Assumption 2.7 hold. By Li (2003) we also have

$$U_{i,\kappa}(t) \le Ct^{\kappa/2} E|e_{i,0}|^\kappa \quad \text{with some constant } C \text{ depending only on } \kappa,$$

completing the proof of (3.2). Under assumption $\max_{1\le i \le N} E|e_{i,0}|^{\bar{\tau}} \le C_0$ with some $\bar{\tau} > 2$ and $C_0 > 0$, we obtain Assumption 2.7.

Since the early 1980's, ARCH, GARCH processes and their various extensions have become extremely popular models in the analysis of macroeconomic and financial data. For a survey and detailed study of volatility models we refer to Francq and Zakoïan (2010). The next example shows that a large class of volatility processes satisfies the assumptions in Section 2.

Example 3.4
We assume that $\{\varepsilon_{i,t}, -\infty < t < \infty\}$ are independent and identically distributed random variables with $E\varepsilon_{i,0} = 0$ and $E\varepsilon_{i,0}^2 = 1$. The error terms are defined by

$$e_{i,t} = h_{i,t}\varepsilon_{i,t}, \tag{3.3}$$

where the volatility process $h_{i,t} > 0$ is measurable with respect to the σ–algebra generated by $\varepsilon_{i,s}, s \le t-1$. Usually, $h_{i,t}$ is given by a recursion involving $e_{i,s}, h_{i,s}, s \le t-1$. Francq and Zakoïan (2010) provide conditions for the existence of a stationary solution of (3.3) in several models and establish their basic properties. Assuming that $E|e_{i,0}|^\kappa < \infty$ with some $\kappa > 4$, $\{e_{i,t}, -\infty < t < \infty\}$ is a stationary orthogonal martingale satisfying the conditions in Example 3.3. In case of the most popular GARCH(p,q) model $h_{i,t} = \omega + \sum_{\ell=1}^{p} \alpha_\ell e_{i,t-\ell}^2 + \sum_{\ell=1}^{q} \beta_j h_{i,t-j}$, $\omega > 0, \alpha_\ell \ge 0, \beta_j \ge 0, 1 \le \ell \le p, 1 \le j \le q$. The necessary and sufficient condition for the existence of the higher moments in case of GARCH(1,1) is given in Nelson (1990). He and Teräsvirta (1999), Ling and McAleer (2002) and Berkes, Horváth, and Kokoszka (2003) partially extend his results to the more general case. The existence of moments of augmented GARCH sequences is discussed in Carrasco and Chen (2002) and Hörmann (2008).

Linear processes and the volatility models of Example 3.4 are in the class of m–decomposable processes.

Example 3.5
We say the $e_{i,t}$ is a Bernoulli shift if it can be written as

$$e_{i,t} = f_i(\varepsilon_{i,t}, \varepsilon_{i,t-1}, \varepsilon_{i,t-2}, \ldots)$$

with some functional f_i, where $\{\varepsilon_{i,t}, -\infty < t < \infty\}$ are independent and identically distributed random variables. The conditions of Section 2 are satisfied if the

https://doi.org/10.1017/S0266466615000468 Published online by Cambridge University Press

Bernoulli shift is m–decomposable, i.e. if

$$\sum_{m=1}^{\infty}\left(E\left|e_{i,t}-e_{i,t}^{(m)}\right|^{\kappa}\right)^{1/\kappa}<\infty \quad \text{with some } \kappa>4,$$

where $e_{i,t}^{(m)}=f_i(\varepsilon_{i,t},\varepsilon_{i,t-1},\ldots,\varepsilon_{i,t-m+1},\varepsilon_{i,t-m}^{*},\varepsilon_{i,t-m-1}^{*}\ldots)$, and the $\varepsilon_{i,t}^{*}$'s are independent copies of $\varepsilon_{i,0}$, independent of $\varepsilon_{i,t}, 1\le i\le N, -\infty<t<\infty$. Berkes, Hörmann, and Schauer (2011) prove that there is constant C_i such that $U_{i,\kappa}(t)\le C_i t^{\kappa/2}$, $|Ee_{i,0}e_{i,t}|\le C_i t^{-\kappa/2}$ and $U_{i,2}(t)=t\sigma_i^2+(t)$, as $t\to\infty$, with some σ_i^2. They also provide several examples for m–decomposable Bernoulli shifts.

Example 3.6
There is a well developed theory of partial sums of mixing random variables where the long run moments $U_{i,\kappa}(t)$ play a crucial role. It has been established under various conditions that $U_{i,2}(t)=t\sigma_i^2+o(t)$ and $U_{i,\kappa}(t)=O(t^{\kappa/2})$, as $t\to\infty$. For surveys on mixing processes we refer to Bradley (2007) and Dedecker et al. (2007).

Example 3.7
We assumed in Examples 3.2 and 3.4 that the innovations $\varepsilon_{i,t}, -\infty<t<\infty$ are independent and identically distributed. However, this assumption can be replaced with the less restrictive requirement that $\{\varepsilon_{i,t}, -\infty<t<\infty\}$ is a stationary sequence. Rosenthal–type inequalities for sums of functionals of stationary processes are developed in Wu (2002) and Merlevéde, Peligrad, and Utev (2006). These results can be used to establish Assumptions 2.3, and 2.7.

4. ESTIMATION OF NORMING SEQUENCES

Theorems 2.2–2.4 contain the limit distribution of $\hat{t}_{N,T}$ with different normalizations that reflect the effects of the sizes of changes and the loading factors. However, in case of finite N and T it is impossible to check which specific condition on the growths of $\Delta_{N,T}$ and $\Sigma_{N,T}$ holds. Therefore it is useful to produce norming sequences that would work in all possible cases. Let

$$\Xi_{N,T}=\sum_{i=1}^{N}\sigma_i^2\delta_i^2+\Sigma_{N,T}^2.$$

Under the conditions of Theorems 2.2–2.4 we have that

$$\frac{\Delta_{N,T}^2}{\Xi_{N,T}}\left(\hat{t}_{N,T}-t_0\right) \quad \text{converges in distribution.} \tag{4.1}$$

The limit distribution in (4.1) is $\operatorname{argmax}_u\{W(u)-g_\theta(u)\}$ except in the special cases of (2.13) and (2.16). In these cases, the limit distribution is the argmax of a process defined on integers. The limits in (2.13) and (2.16) depend on the distributions of $\{X_{i,t}, 1 \le i \le N, 1 \le t \le T\}$. If $\mathfrak{d}$ is close to zero in (2.13) or (2.16), the limiting distributions can be well approximated with $\operatorname{argmax}_u (W(u) - g_\theta(u))$. As numerically investigated by Bai (2010) in the context of the least squares estimator, $\operatorname{argmax}_u \{W(u) - g_\theta(u)\}$ gives a reasonable approximation for the limit in (2.13) when $\mathfrak{u}(s,t) = \min(t,s)$. Hence we recommend that $\operatorname{argmax}_u (W(u) - g_\theta(u))$ can be used as the limit in (4.1) in practice.

The limit result in (4.1) can only be used for hypothesis testing or confidence intervals if the norming factor can be consistently estimated from the sample. We estimate $\Delta_{N,T}$ with

$$\hat{\Delta}_{N,T} = \sum_{i=1}^{N} \left(\frac{1}{\hat{t}_{N,T}} \sum_{1 \le t \le \hat{t}_{N,T}} X_{i,t} - \frac{1}{T - \hat{t}_{N,T}} \sum_{\hat{t}_{N,T} < t \le T} X_{i,t} \right)^2.$$

It is more difficult to estimate $\Xi_{N,T}$. Let

$$U_N(t) = \sum_{i=1}^{N} \left(S_i(t) - \frac{t}{T} S_i(T) \right)^2, \tag{4.2}$$

$$\hat{r}_{N,T}(t) = -t \frac{T - \hat{t}_{N,T}}{T} + \left(t - \hat{t}_{N,T}\right) I\left\{t > \hat{t}_{N,T}\right\},$$

and

$$\hat{r}_{N,T} = \hat{r}_{N,T}\left(\hat{t}_{N,T}\right) = -\hat{t}_{N,T} \frac{T - \hat{t}_{N,T}}{T}. \tag{4.3}$$

The estimator for $\Xi_{N,T}$ is defined as

$$\hat{\Xi}_{N,T} = \frac{1}{2(M_2 - M_1)} \sum_{M_1 < |v| \le M_2} \frac{1}{4|v| \hat{r}_{N,T}^2} \Big(U_N\left(\hat{t}_{N,T} + v\right) - U_N\left(\hat{t}_{N,T}\right) - \hat{\Delta}_{N,T} \left(\hat{r}_{N,T}^2\left(\hat{t}_{N,T} + v\right) - \hat{r}_{N,T}^2 \right) \Big)^2.$$

THEOREM 4.1. *(i) We assume that the conditions of Theorem 2.2 or 2.3 are satisfied. If $M_1 < M_2$, $M_1 \to \infty$ and $M_2/T \to 0$, then*

$$\frac{\hat{\Delta}_{N,T}}{\Delta_{N,T}} \xrightarrow{P} 1 \tag{4.4}$$

and

$$\frac{\hat{\Xi}_{N,T}}{\Xi_{N,T}} \xrightarrow{P} 1. \tag{4.5}$$

https://doi.org/10.1017/S0266466615000468 Published online by Cambridge University Press

(ii) We assume that the conditions of Theorem 2.4 are satisfied. If $M_1 < M_2$, $M_1 \to \infty$ and

$$M_2/\min\left(T, \Sigma_{N,T}^2/\Delta_{N,T}^2\right) \to 0,$$

then (4.4) and (4.5) hold.

Since (4.4) and (4.5) hold, the limit result in (4.1) remains true when the norming is replaced with the corresponding estimators, i.e.

$$\frac{\hat{\Delta}_{N,T}^2}{\hat{\Xi}_{N,T}}\left(\hat{t}_{N,T} - t_0\right) \quad \text{converges in distribution.} \tag{4.6}$$

If the interaction between the cross-sectional units of the panels is small, i.e. $\Sigma_{N,T}^2/\sum_{i=1}^N \sigma_i^2\delta_i^2 \to 0$, as $N, T \to \infty$, then we need to estimate only $\Delta_{N,T}$ and $\sum_{i=1}^N \sigma_i^2\delta_i^2$. Using $\hat{\sigma}_i^2, 1 \le i \le N$, the long run variance estimators for σ_i^2, a possible estimator for $\sum_{i=1}^N \sigma_i^2\delta_i^2$ is $\sum_{i=1}^N \hat{\sigma}_i^2\big[\sum_{1\le s\le \hat{t}_{N,T}} X_{i,s}/\hat{t}_{N,T} - \sum_{\hat{t}_{N,T}<s\le T} X_{i,s}/(T-\hat{t}_{N,T})\big]^2$.

5. SIMULATIONS, AND DATA EXAMPLES

5.1. Simulations

Using the estimators defined in Section 4, we computed the empirical percentages when the variable defined in (4.6) was below the asymptotic quantiles for several N and T. We considered the case when there was no interaction between the panel's cross-sections, i.e. $\gamma_i = 0$, and also examples when cross-sectional correlation was present. We tried various values of the break point $t_0 = \lfloor T\theta \rfloor$, and we observed that the applicability of the limit results presented in Section 2 did not depend on θ. Therefore, in Tables 1 and 2, we present the results when $\theta = 1/2$. We used 1,000 independent repetitions to produce each table entry. The 90^{th}, 95^{th}, and 99^{th} percentiles of the distribution of $\text{argmax}_u(W(u) - g_{1/2}(u))$ are 4.70, 7.69, and 15.89, respectively. Table 1 illustrates that $\Delta_{N,T}$ must be small in order to use Theorem 2.2(a), and the approximation improves when T increases. In case of larger $\Delta_{N,T}$, when the conditions of Theorem 2.1 are more applicable, we observed that the distribution of $\hat{t}_{N,T}$ was more concentrated than what we would expect from the asymptotics in Theorem 2.2(a). We observed that the coverage at the 99% level was inflated in the examples that we considered, indicating that the 99^{th} percentile of $\text{argmax}_u(W(u) - g_{1/2}(u))$ may be conservative in practice to approximate the 99^{th} percentile of the normalized changepoint estimator.

Our comments also hold when interaction between the cross-sections of the panel is allowed as illustrated by Table 2, but, due to the dependence, larger T is needed to use the limit results.

TABLE 1. Empirical percentages of $\hat{\Delta}^2_{N,T}(\hat{t}_{N,T}-t_0)/\hat{\Xi}_{N,T}$ below the asymptotic quantiles for various values of N, T, δ_i when $\gamma_i = 0$ for all $1 \le i \le N$.

N/T	δ_i	90%	95%	99%
25/100	0.150	88.7%	94.7%	100%
25/250	0.100	87.6%	94.4%	99.8%
25/500	0.060	89.9%	96.4%	100%
50/100	0.100	89.1%	97.1%	100%
50/250	0.070	88.1%	95.4%	100%
50/500	0.050	86.5%	94.9%	100%
100/100	0.085	86.1%	94.5%	100%
100/250	0.055	86.9%	94.3%	99.9%
100/500	0.035	88.5%	96.5%	100%

TABLE 2. Empirical percentages of $\hat{\Delta}^2_{N,T}(\hat{t}_{N,T}-t_0)/\hat{\Xi}_{N,T}$ below the asymptotic quantiles for various values of N, T, δ_i when $\gamma_i = 0.03$ for all $1 \le i \le N$.

N/T	δ_i	90%	95%	99%
25/100	0.150	86.9%	94.9%	100%
25/250	0.100	90.7%	95.6%	100%
25/500	0.060	89.2%	97.2%	100%
50/100	0.100	87.6%	95.6%	100%
50/250	0.070	91.0%	96.4%	100%
50/500	0.050	88.3%	96.3%	100%
100/100	0.085	89.4%	95.9%	100%
100/250	0.055	85.1%	94.0%	99.9%
100/500	0.035	90.1%	96.5%	100%

Figure 1 shows that the density function of the limit follows the shape of the histogram of $\hat{t}_{N,T}$ closely.

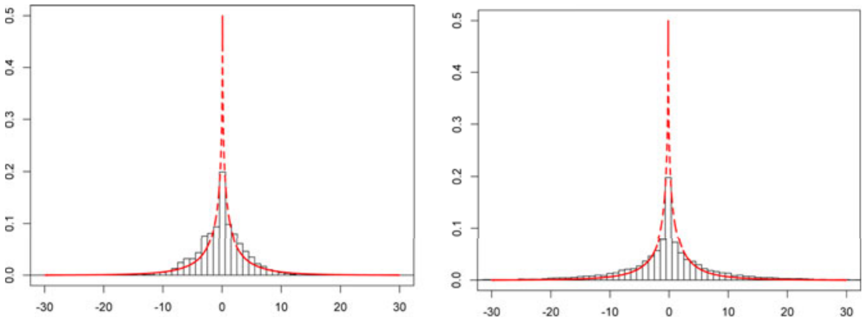

FIGURE 1. The histogram of $\hat{t}_{100,500}$ with $\delta_i = 0.07, \gamma_i = 0, 1 \le i \le 100$ (left panel) and the histogram of $\hat{t}_{50,100}$ with $\delta_i = 0.1, \gamma_i = 0.03, 1 \le i \le 50$ (right panel) and the density of the limiting random variable

https://doi.org/10.1017/S0266466615000468 Published online by Cambridge University Press

5.2. Applications

In the first example we consider the exchange rates between the US dollar and 23 other currencies. The data can be found at the website `www.federalreserve.gov/releases/h10/hist/`. Figure 2 contains the graphs of the exchange rates between the United Kingdom (UK), Canada (CA), Singapore (SI), Switzerland (SW), Denmark (DN), Norway (NO) and Sweden (SD). In our study we used the time period 03/13/2001–03/11/2003 so we have $N = 23$ panels and each panel has $T = 500$ observations. Using the testing method in Horváth and Hušková (2013) the no change in the mean of the panels null hypothesis is rejected. The estimated time of change is $\hat{t}_{23,500} = 297$ so the change is indicated on 05/16/2002. We also constructed confidence intervals using (4.6). Since $\hat{\Delta}_{23,500}$ is very large, the 90%, 95%, and 99% confidence intervals contain only a single element, $\hat{t}_{23,500} = 297$, i.e. the conditions of Theorem 2.1 hold in this case. It is clear from Figure 2 that the exchange rates are between 1.3 and 11, so if the same proportional change occurs in a panel with high values, this change will give a very large δ^2 compared to the other panels. Hence a single panel can disproportionately contribute to $\hat{\Delta}_{N,T}$. To overcome this problem we rescaled the observations in each panel with the first observation, i.e. with the exchange rate on 03/13/2001. Figure 3 contains the graphs of the relative changes in exchange rates with respect to the US dollar for the same countries as in Figure 2. We repeated our analysis for the relative changes (rescaled) in the exchange rates with respect to the US dollar, resulting in $\hat{t}_{23,500} = 303$ which corresponds to 05/24/2002. In the definition of $\hat{\Xi}_{23,500}$ we used $M_2 - M_1 \approx \lfloor T^{1/2}/\hat{\Delta}_{23,500} \rfloor$. The 90%, 95%, and 99% confidence intervals are $[292, 330]$, $[287, 341]$, and $[274, 371]$. Note that all these confidence intervals contain 297 which was obtained for the non–scaled exchange rates. Some of the graphs show a linear trend after the time of change instead of changing to another constant mean. The limit distribution of the time of change is local, i.e. it is determined by the observations in the neighborhood of t_0. Replacing the linear trend with an average value close to t_0 can justify the asymptotic validity of the confidence intervals. Also, after the change point was found and the means of the corresponding segments were removed, the stationarity of the residuals could not be rejected.

The exchange rates data and the scaled exchange rates contain further level shifts. Using the binary segmentation method one can divide the data into homogenous

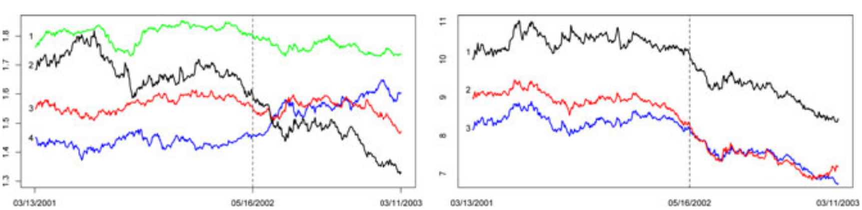

FIGURE 2. The graphs exchange rates, 1=UK, 2=SI, 3=CA, 4=SW (left panel); 1=DN, 2=NO, 3=SD (right panel) with respect to the US dollar.

https://doi.org/10.1017/S0266466615000468 Published online by Cambridge University Press

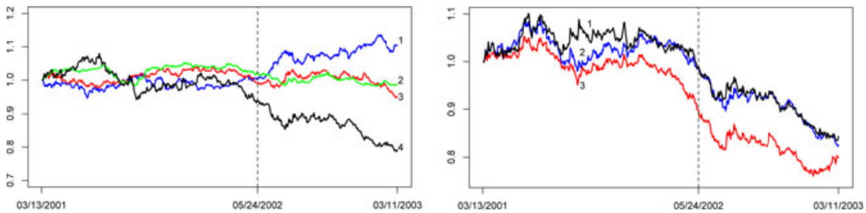

FIGURE 3. The graphs of relative exchange rates, 1=UK, 2=SI, 3=CA, 4=SW (left panel); 1=DN, 2=NO, 3=SD (right panel) with respect to the US dollar.

segments and provide confidence intervals for the time of changes. Sato (2013) investigates the number and the location of the changes in daily log returns in 30 currency pairs between 04/01/2001 and 30/12/2011 and points out that the study of the individual pairs might not find all the changes in the exchange rates mechanism.

In the second example we compare the growth rates of the GDP/capita for $N = 113$ countries. The data can be found at the website `www://data.worldbank.org/indicator/NY.GDP.MKTP.CD`. The data are recorded in current US dollars. We removed some countries from the data set due to a large number of missing values, so we used $N = 113$ panels with $T = 51$ covering the time period 1961–2012. To achieve stationarity of the errors we transformed the data by taking log differences. The graphs of the log transformed GDP's are exhibited in Figure 4. We used the CUSUM test of Horváth and Hušková (2013) to test the stability of the means of the panels which was rejected at a very high significance level. The estimator for the time of change is $\hat{t}_{113,51} = 19$

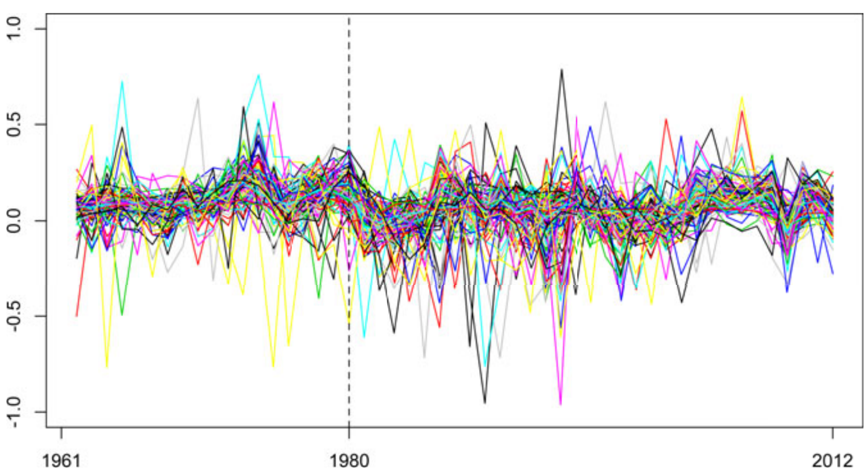

FIGURE 4. The graphs of the log differences of the GDP/capita for 113 countries between 1961 and 2012.

https://doi.org/10.1017/S0266466615000468 Published online by Cambridge University Press

which corresponds to 1979/1980. Applying the limit result in (4.6), [16, 21] is the asymptotic 90% and 95% confidence interval, while the 99% confidence interval is [12, 22].

6. CONCLUSION

We established the first and second order asymptotic properties of a CUSUM estimator of the time of change in the mean of panel data. Our results are derived under long run moment conditions, which serve to extend the asymptotic theory to a broader family of error processes than had been previously considered in the literature, and we provided an in depth study on how the rates of divergence of the sizes of changes and the common factor loadings are manifested in the asymptotic behavior of the test statistic. Our results were demonstrated with a Monte Carlo simulation study, and we considered application to two real data sets.

REFERENCES

Andrews, D.W.K. (2003) End-of-sample instability tests. *Econometrica* 71, 1661–1694.

Arellano, M. (2003) *Panel Data Econometrics.* Oxford University Press.

Aue, A. & L. Horváth (2013) Structural breaks in time series. *Journal of Time Series Analysis* 23, 1–16.

Bai, J. (2010) Common breaks in means and variances for panel data. *Journal of Econometrics* 157, 78–92.

Baltagi, B.H. (2013) *Econometric Analysis of Panel Data*, 5th ed. Wiley.

Baltagi, B.H., Q. Feng, & C. Kao (2015) Estimation of heterogeneous panels with structural breaks. *Journal of Econometrics*, forthcoming.

Baltagi, B.H., C. Kao, & L. Liu (2012) Estimation and identification of change points in panel models with nonstationary or stationary regressors and error terms, preprint.

Berkes, I., S. Hörmann, & J. Schauer (2011) Split invariance principles for stationary processes. *The Annals of Probability* 39, 2441–2473.

Berkes, I., L. Horváth, & P. Kokoszka (2003) GARCH processes: Structure and estimation. *Bernoulli* 9, 201–227.

Bhattacharya, P.K. & P.J. Brockwell (1976) The minimum of an additive process with applications to signal estimation and storage theory. *Zeitschrift für Wahrscheinlichkeitstheorie verwandte Gebiete* 37, 51–75.

Billingsley, P. (1968) *Convergence of Probability Measures.* Wiley.

Bradley, R.C. (2007) *Introduction to Strong Mixing Conditions*, vol. 1–3. Kendrick Press.

Brodsky, B.E. & B. Darkhovskii (2000) *Non-Parametric Statistical Diagnosis.* Kluwer.

Carrasco, M. & X. Chen (2002) Mixing and moment properties of various GARCH and stochastic volatility models. *Econometric Theory* 18, 17–39.

Chan, J., L. Horváth, & M. Hušková (2013) Darling–Erdős limit results for change-point detetction in panel data. *Journal of Statistical Planning and Inference* 143, 955–970.

Csörgő, M. & L. Horváth (1997) *Limit Theorems in Change-Point Analysis.* Wiley.

Dedecker, I., P. Doukhan, G. Lang, J.R.R. León, S. Louhichi, & C. Prieur (2007) *Weak Dependence with Examples and Applications.* Lecture Notes in Statistics. Springer.

Ferger, D. (1994) Change-point estimators in case of small disorders. *Journal of Statistical Planning and Inference* 40, 33–49.

Francq, C. & J.-M. Zakoian (2010) *GARCH Models*. Wiley.

Hawkins, D.L. (1986) A simple least squares method for estimating a change in mean. *Communications in Statistics* 15, 655–679.

He, C. & T. Teräsvirta (1999) Properties of moments of a family of GARCH processes. *Journal of Econometrics* 92, 173–192.

Hörmann, S. (2008) Augmented GARCH sequences: Dependence structure and asymptotics. *Bernoulli* 14, 543–561.

Horváth, L. & M. Hušková (2012) Change-point detection in panel data. *Journal of Time Series Analysis* 33, 631–648.

Horváth, L. & G. Rice (2014) Extensions of some classical methods in change point analysis (with discussions). *Test* 23, 219–290.

Hsiao, C. (2003) *Analysis of Panel Data*, 2nd ed. Cambridge University Press.

Hsiao, C. (2007) Panel data analysis—advantages and challenges. *Test* 16, 1–22.

Joseph, L. & D.B. Wolfson (1992) Estimation in multi-path change-point problems. *Communications in Statistics-Theory and Methods* 21, 897–913.

Joseph, L. & D.B. Wolfson (1993) Maximum likelihood estimation in the multi-path changepoint problem. *Annals of the Institute of Statistical Mathematics* 45, 511–530.

Kao, C., L. Trapani, & G. Urga (2012) Asymptotics for panel models with common shocks. *Econometric Reviews* 31, 390–439.

Kao, C., L. Trapani, & G. Urga (2015) Testing for breaks in cointegrated panels. *Econometric Reviews*, to appear.

Kim, D. (2011) Estimating a common deterministic time trend break in large panels with cross-sectional dependence. *Journal of Econometrics* 164, 310–330.

Kim, D. (2014) Common breaks in time trends for large panel data with a factor structure. *Econometrics Journal* 17, 301–337.

Li, D., J. Qian, & L. Su (2015) Panel data models with interactive fixed effects and multiple structrual breaks. *Journal of the American Statistical Association*, to appear.

Li, Y. (2003) A martingale inequality and large deviations. *Statistics & Probability Letters* 62, 317–321.

Ling, S. & M. McAleer (2002) Necessary and sufficient moment conditions for GARCH(r,s) and asymmetric GARCH(r,s) model. *Econometric Theory* 18, 722–729.

Merlevéde, F., M. Peligrad, & S. Utev (2006) Recent advances in invariance principles for stationary sequences. *Probability Surveys* 3, 1–36.

Móricz, F., R. Serfling, & W. Stout (1982) Moment and probability bounds with quasi-superadditive structure for the maximal partial sum. *Annals of Probability* 10, 1032–1040.

Nelson, D.B. (1990) Stationarity and persistance in the GARCH(1,1) model. *Econometric Theory* 6, 318–334.

Petrov, V.V. (1995) *Limit Theorems of Probability Theory*. Clarendon Press.

Phillips, P.C.P. & V. Solo (1992) Asymptotics for linear processes. *Annals of Statistics* 20, 971–1001.

Qian, J. & L. Su (2015) Shrinkage estimation of common breaks in panel data models via adaptive group fused lasso. *Journal of Econometrics* 191, 86–109.

Sato, A.-H. (2013) Recursive segmentation procedure based on the Akaike information criterion test. *2013 IEEE 37th Annual Computer Software and Applications Conference*, 226–233.

Wu, W.B. (2002) Central limit theorems for functionals of linear processes and their applications. *Statistica Sinica* 12, 635–649.

APPENDIX A. Proofs of Theorems 2.1–2.4 and Remarks 2.1 and 2.2

Throughout the proofs c denotes unimportant constants whose values might change from line to line. Using (2.1) we have

$$S_i(t) - \frac{t}{T} S_i(T) = Q_i(t) + \gamma_i V(t) + \delta_i r(t),$$

where

$$Q_i(t) = \sum_{s=1}^{t} e_{i,s} - \frac{t}{T}\sum_{s=1}^{T} e_{i,s}, \quad V(t) = \sum_{s=1}^{t} \eta_s - \frac{t}{T}\sum_{s=1}^{T} \eta_s \tag{A.1}$$

and

$$r(t) = -t\frac{T-t_0}{T} + (t-t_0)I\{t > t_0\}. \tag{A.2}$$

Hence we have

$$\left(S_i(t) - \frac{t}{T}S_i(T)\right)^2 = Q_i^2(t) + \gamma_i^2 V^2(t) + \delta_i^2 r^2(t) + 2V(t)\gamma_i Q_i(t) + 2r(t)\delta_i Q_i(t) + 2V(t)r(t)\gamma_i\delta_i. \tag{A.3}$$

Let $0 < \alpha < \theta < 1-\alpha$ and define

$$\tilde{t}_{N,T}(\alpha) = \text{argmax}_{\lfloor T\alpha\rfloor \le t \le T - \lfloor T\alpha\rfloor} \sum_{i=1}^{N}\left(S_i(t) - \frac{t}{T}S_i(T)\right)^2.$$

LEMMA A.1. *If Assumptions 2.1–2.5 are satisfied, then we have*

$$\lim_{\min(N,T)\to\infty} P\left\{\hat{t}_{N,T} = \tilde{t}_{N,T}(\alpha)\right\} = 1 \quad \textit{for all} \quad 0 < \alpha < \theta < 1-\alpha \tag{A.4}$$

and

$$\frac{\hat{t}_{N,T}}{T} \xrightarrow{P} \theta \quad \textit{as} \quad \min(T,N) \to \infty. \tag{A.5}$$

Proof. It is easy to see that for every $0 < \alpha < \theta < 1-\alpha$

$$\frac{1}{T^2}\max_{\lfloor T\alpha\rfloor \le t \le T-\lfloor T\alpha\rfloor} r^2(t) \;\to\; \theta^2(1-\theta)^2,$$

$$\frac{1}{T^2}\max_{1 \le t \le \lfloor T\alpha\rfloor} r^2(t) \;\to\; \alpha^2(1-\theta)^2,$$

$$\frac{1}{T^2}\max_{T-\lfloor T\alpha\rfloor \le t \le T} r^2(t) \;\to\; \alpha^2\theta^2.$$

We prove that

$$\sup_{1\le t\le T}\sum_{i=1}^{N} Q_i^2(t) = O_P(NT). \tag{A.6}$$

Elementary arguments give

$$EQ_i^2(t) \le U_{i,2}(t) + 2\left\{U_{i,2}(t)U_{i,2}(T)\right\}^{1/2} + U_{i,2}(T)$$

https://doi.org/10.1017/S0266466615000468 Published online by Cambridge University Press

and therefore by Assumption 2.3(i) we have

$$\max_{1\le t\le T}\sum_{i=1}^{N} EQ_i^2(t) = O(NT).$$

Let $q_i(u) = (Q_i^2(uT) - EQ_i^2(uT))/T$. Using Assumption 2.1(i), for every $0 \le u < v \le 1$ by the Rosenthal inequality (cf. Petrov (1995, p. 59)) we have with $\nu = \kappa/2$

$$E\left|\sum_{i=1}^{N}(q_i(v)-q_i(u))\right|^{\nu} \le c\left\{\sum_{i=1}^{N}E|q_i(v)-q_i(u)|^{\nu} + \left(\sum_{i=1}^{N}E(q_i(v)-q_i(u))^2\right)^{\nu/2}\right\}.$$

By the Cauchy–Schwarz inequality we conclude for all $1 \le s \le t \le T$ that

$$\begin{aligned} E(Q_i^2(t)-Q_i^2(s))^2 &\le E\left\{(Q_i(t)-Q_i(s))^2\,(|Q_i(t)|+|Q_i(s)|)^2\right\} \\ &\le 4\left(E(Q_i(t)-Q_i(s))^4\right)^{1/2}\left(EQ_i^4(t)+EQ_i^4(s)\right)^{1/2}. \end{aligned}$$

Using the definition of $Q_i(t)$ and Assumption 2.1(ii) we get that

$$E(Q_i(t)-Q_i(s))^4 \le 2^3\left\{U_{i,4}(t-s) + \left(\frac{t-s}{T}\right)^4 U_{i,4}(T)\right\}.$$

and similarly

$$EQ_i^4(t) \le 2^3\left(U_{i,4}(t)+U_{i,4}(T)\right).$$

Also,

$$\begin{aligned} \left(EQ_i^2(t)-EQ_i^2(s)\right)^2 &\le 2E(Q_i(t)-Q_i(s))^2\left(EQ_i^2(t)+EQ_i^2(s)\right) \\ &\le 8\left(U_{i,2}(t-s)+\left(\frac{t-s}{T}\right)^2 U_{i,2}(T)\right)\left(U_{i,2}(t)+U_{i,2}(s)+2U_{i,2}(T)\right). \end{aligned}$$

Thus applying Assumption 2.3(ii) we get that with some $0 < c < \infty$

$$\left(\frac{1}{N}\sum_{i=1}^{N}E(q_i(v)-q_i(u))^2\right)^{\nu/2} \le c|u-v|^{\nu/2} \text{ for all } 0 \le u < v \le 1.$$

Repeating the arguments used above we obtain that

$$\begin{aligned} E\left|Q_i^2(t)-Q_i^2(s)\right|^{\nu} &\le E\left\{|Q_i(t)-Q_i(s)|^{\nu}\,(|Q_i(t)|+|Q_i(s)|)^{\nu}\right\} \\ &\le 2^{\nu}\left(E|Q_i(t)-Q_i(s)|^{2\nu}\right)^{1/2}\left(E|Q_i(t)|^{2\nu}+E|Q_i(s)|^{2\nu}\right)^{1/2}, \end{aligned}$$

$$E|Q_i(t)-Q_i(s)|^{2\nu} \le 2^{2\nu}\left\{U_{i,2\nu}(t-s)+\left(\frac{t-s}{T}\right)^{2\nu}U_{i,2\nu}(T)\right\},$$

https://doi.org/10.1017/S0266466615000468 Published online by Cambridge University Press

$$E|Q_i(t)|^{2\nu} \le 2^{2\nu}\left(U_{i,2\nu}(t)+U_{i,2\nu}(T)\right)$$

and

$$\begin{aligned}\left|EQ_i^2(t)-EQ_i^2(s)\right|^{2\nu} &\le E\left|Q_i(t)-Q_i(s)\right|^{\nu}\left(E|Q_i(t)|^{\nu}+E|Q_i(s)|^{\nu}\right)\\ &\le 2^{2\nu}\left(U_{i,\nu}(t-s)+\left(\frac{t-s}{T}\right)^{\nu}U_{i,\nu}(T)\right)\left(U_{i,\nu}(t)+U_{i,\nu}(s)+2U_{i,\nu}(T)\right)\end{aligned}$$

resulting in

$$\frac{1}{N}\sum_{i=1}^{N}E|q_i(u)-q_i(v)|^{\nu} \le c|u-v|^{\nu/2} \quad \text{for all } 0\le u,v\le 1$$

with some c. Using Billingsley (1968, pp. 95 and 127) we conclude that the process $\sum_{i=1}^{N} q_i(u)/N$ is tight in $\mathcal{D}[0,1]$ and therefore (A.6) holds.

The moment assumption in Assumption 2.4 with the maximal inequality of Móricz, Serfling, and Stout (1982) yields that $E(\max_{1\le t\le T}|V(t)|)^{\bar{\kappa}} = O(T^{\bar{\kappa}})$ and therefore by Markov's inequality we conclude

$$\max_{1\le t\le T}|V(t)| = O_P\left(T^{1/2}\right). \tag{A.7}$$

By (A.7) we get immediately that

$$\sup_{1\le t\le T}\sum_{i=1}^{N}\gamma_i^2V^2(t) = O_P(T\Gamma_{N,T}). \tag{A.8}$$

Following the proof of (A.6) we get

$$\sup_{1\le t\le T}\left|\sum_{i=1}^{N}\gamma_iQ_i(t)\right| = O_P\left(T^{1/2}\Gamma_{N,T}^{1/2}\right)$$

and therefore by (A.7)

$$\sup_{1\le t\le T}\left|\sum_{i=1}^{N}V(t)\gamma_iQ_i(t)\right| = O_P\left(T\Gamma_{N,T}^{1/2}\right). \tag{A.9}$$

Similarly to (A.9) we have that

$$\sup_{1\le t\le T}\left|\sum_{i=1}^{N}r(t)\delta_iQ_i(t)\right| = O_P\left(T^{3/2}\Delta_{N,T}^{1/2}\right). \tag{A.10}$$

Using again Assumption 2.4 and the definition of $r(t)$, one can easily verify that

$$\max_{1\le t\le T}\left|\sum_{i=1}^{N}V(t)r(t)\gamma_i\delta_i\right| = O_P\left(T^{3/2}\left|\Sigma_{N,T}\right|\right). \tag{A.11}$$

https://doi.org/10.1017/S0266466615000468 Published online by Cambridge University Press

It follows from (A.6)–(A.11) that

$$\frac{1}{T^2\Delta_{N,T}}\max_{1\le t\le T}\sum_{i=1}^{N}\left(S_i(t)-\frac{t}{T}S_i(T)\right)^2 \xrightarrow{P} \theta^2(1-\theta)^2,$$

$$\frac{1}{T^2\Delta_{N,T}}\max_{1\le t\le \lfloor \alpha T\rfloor}\sum_{i=1}^{N}\left(S_i(t)-\frac{t}{T}S_i(T)\right)^2 \xrightarrow{P} \alpha^2(1-\theta)^2$$

and

$$\frac{1}{T^2\Delta_{N,T}}\max_{T-\lfloor \alpha T\rfloor\le t\le T}\sum_{i=1}^{N}\left(S_i(t)-\frac{t}{T}S_i(T)\right)^2 \xrightarrow{P} \alpha^2\theta^2,$$

which immediately implies Lemma A.1. ∎

According to (A.4), it is enough to consider the asymptotic behavior of

$$\hat{t}_{N,T}=\operatorname{argmax}_{\lfloor \alpha T\rfloor\le t\le\lfloor(1-\alpha)T\rfloor}\{U_N(t)-U_N(t_0)\},$$

with any $0<\alpha<\theta<1-\alpha<1$, where $U_N(t)$ is defined in (4.2). It is easy to see that

$$\begin{aligned}U_N(t)-U_N(t_0)&\\ =\sum_{i=1}^{N}\Big\{&\delta_i^2\left(r^2(t)-r^2(t_0)\right)+Q_i^2(t)-Q_i^2(t_0)+\gamma_i^2\left(V^2(t)-V^2(t_0)\right)\\ &+2\gamma_i(Q_i(t)V(t)-Q_i(t_0)V(t_0))+2\delta_i(r(t)Q_i(t)-r(t_0)Q_i(t_0))\\ &+2\gamma_i\delta_i(V(t)r(t)-V(t_0)r(t_0))\Big\}. \qquad \textbf{(A.12)}\end{aligned}$$

LEMMA A.2. *If Assumption 2.2 holds, then for all* $0<\alpha<\theta<1-\alpha$ *there are* $0<c_1,c_2<\infty$ *such that*

$$-c_1|t_0-t|T\le r^2(t)-r^2(t_0)\le -c_2|t_0-t|T$$

for all $1\le t\le T$.

Proof. The result follows from Assumption 2.2 and the definition of $r(t)$. ∎

Throughout Lemmas A.3–A.7 we assume that $1\le M\le T$.

LEMMA A.3. *If Assumptions 2.1–2.3 hold, then*

$$\max_{|t-t_0|\ge M}\frac{1}{|t-t_0|}\left|\sum_{i=1}^{N}\left[Q_i^2(t)-Q_i^2(t_0)\right]\right|=O_P\left(N+T^{2/\kappa}N^{1/2}+(NT/M)^{1/2}\right). \qquad \textbf{(A.13)}$$

Proof. We write

$$Q_i^2(t)-Q_i^2(t_0)=2Q_i(t_0)(Q_i(t)-Q_i(t_0))+(Q_i(t)-Q_i(t_0))^2.$$

Using Assumption 2.3(i) we get that

$$\sup_{1\le t\le T}\sum_{i=1}^{N}\frac{\left|EQ_i^2(t)-EQ_i^2(t_0)\right|}{|t-t_0|}=O(N). \tag{A.14}$$

Let $\chi_i(t)=(Q_i(t)-Q_i(t_0))^2-E(Q_i(t)-Q_i(t_0))^2$. Elementary arguments give

$$\begin{aligned}
&P\left\{\max_{1\le t\le T}\frac{1}{|t-t_0|}\left|\sum_{i=1}^{N}\chi_i(t)\right|\ge x\right\}\\
&\quad=P\left\{\left|\sum_{i=1}^{N}\chi_i(t)\right|\ge x|t-t_0| \text{ for at least one } 1\le t\le T\right\}\\
&\quad\le\sum_{1\le t\le T}P\left\{\left|\sum_{i=1}^{N}\chi_i(t)\right|\ge x|t-t_0|\right\}\\
&\quad\le\sum_{1\le t\le T}(x|t-t_0|)^{-\nu}E\left|\sum_{i=1}^{N}\chi_i(t)\right|^{\nu},
\end{aligned} \tag{A.15}$$

where in the last step we used Markov's inequality. Let $\nu=\kappa/2$, where κ is given in Assumption 2.3(ii). The processes $\chi_i(t)$ are independent in i, so using the Rosenthal inequality (cf. Petrov (1995, p. 59)) we conclude with some $c>0$, not depending on t,

$$E\left|\sum_{i=1}^{N}\chi_i(t)\right|^{\nu}\le c\left\{\sum_{i=1}^{N}E|\chi_i(t)|^{\nu}+\left(\sum_{i=1}^{N}E\chi_i^2(t)\right)^{\nu/2}\right\}.$$

Assumptions 2.1(i) and 2.3(ii) yield

$$\sum_{i=1}^{N}E|\chi_i(t)|^{\nu}\le cN|t-t_0|^{\nu} \tag{A.16}$$

and

$$\left(\sum_{i=1}^{N}E\chi_i^2(t)\right)^{\nu/2}\le cN^{\nu/2}|t-t_0|^{\nu}. \tag{A.17}$$

Thus we conclude via (A.15)–(A.17)

$$P\left\{\max_{1\le t\le T}\frac{1}{|t-t_0|}\left|\sum_{i=1}^{N}\chi_i(t)\right|\ge x\right\}\le c\sum_{1\le t\le T}\frac{N^{\nu/2}|t-t_0|^{\nu}}{(x|t-t_0|)^{\nu}}\le\frac{c}{x^{\nu}}TN^{\nu/2}.$$

Choosing $x=c_*N^{1/2}T^{1/\nu}$ with a large enough c_* we get that

$$\max_{1\le t\le T}\frac{1}{|t-t_0|}\left|\sum_{i=1}^{N}\chi_i(t)\right|=O_P\left(T^{1/\nu}N^{1/2}\right).$$

https://doi.org/10.1017/S0266466615000468 Published online by Cambridge University Press

Let

$$\mathcal{S}_i(t) = \sum_{s=1}^{t} e_{i,s}. \tag{A.18}$$

It follows from the definition of $Q_i(t)$ that

$$\begin{aligned}
&\left|\sum_{i=1}^{N}(Q_i(t_0)(Q_i(t)-Q_i(t_0))-E[Q_i(t_0)(Q_i(t)-Q_i(t_0))])\right| \\
&\quad \le \left|\sum_{i=1}^{N}\left(Q_i(t_0)(\mathcal{S}_i(t_0)-\mathcal{S}_i(t))-E\left[Q_i(t_0)(\mathcal{S}_i(t_0)-\mathcal{S}_i(t))\right]\right)\right| \\
&\qquad + \frac{|t-t_0|}{T}\left|\sum_{i=1}^{N}\left(Q_i(t_0)\mathcal{S}_i(T)-E\left[Q_i(t_0)\mathcal{S}_i(T)\right]\right)\right|.
\end{aligned}$$

Using Assumptions 2.1(i), and 2.3(ii) with the Cauchy–Schwarz inequality we get that

$$\operatorname{var}\left(\sum_{i=1}^{N}Q_i(t_0)\mathcal{S}_i(T)\right) = \sum_{i=1}^{N}\operatorname{var}(Q_i(t_0)\mathcal{S}_i(T)) \le \sum_{i=1}^{N}\left(EQ_i^4(t_0)E\mathcal{S}_i^4(T)\right)^{1/2} = O\left(NT^2\right)$$

and therefore

$$\left|\sum_{i=1}^{N}\left(Q_i(t_0)\mathcal{S}_i(T)-E\left[Q_i(t_0)\mathcal{S}_i(T)\right]\right)\right| = O_P\left(N^{1/2}T\right).$$

With $\zeta_i(t) = Q_i(t_0)e_{i,t} - E[Q_i(t_0)e_{i,t}]$ we can write for $1 \le t \le t_0$ that

$$\sum_{i=1}^{N}\left(Q_i(t_0)(\mathcal{S}_i(t_0)-\mathcal{S}_i(t))-E\left[Q_i(t_0)(\mathcal{S}_i(t_0)-\mathcal{S}_i(t))\right]\right) = \sum_{s=t+1}^{t_0}\sum_{i=1}^{N}\zeta_i(s).$$

By the Markov inequality we have

$$\begin{aligned}
&P\left\{\max_{1\le t\le t_0-M}\frac{1}{t_0-t}\left|\sum_{s=t+1}^{t_0}\sum_{i=1}^{N}\zeta_i(s)\right| \ge x(NT/M)^{1/2}\right\} \\
&\quad \le P\left\{\max_{\log M\le k\le \log t_0}\max_{e^k\le \ell\le e^{k+1}}\frac{1}{\ell}\left|\sum_{s=t_0-\ell}^{t_0}\sum_{i=1}^{N}\zeta_i(s)\right| \ge x(NT/M)^{1/2}\right\} \\
&\quad \le P\left\{\max_{e^k\le \ell\le e^{k+1}}\frac{1}{\ell}\left|\sum_{s=t_0-\ell}^{t_0}\sum_{i=1}^{N}\zeta_i(s)\right| \ge x(NT/M)^{1/2}\right. \\
&\qquad\qquad \left.\text{for at least one } \log M \le k \le \log t_0\right\}
\end{aligned}$$

https://doi.org/10.1017/S0266466615000468 Published online by Cambridge University Press

$$\le \sum_{k=\log M}^{\log t_0} P\left\{ \max_{e^k \le \ell \le e^{k+1}} \left| \sum_{s=t_0-\ell}^{t_0} \sum_{i=1}^{N} \zeta_i(s) \right| \ge x(NT/M)^{1/2} e^k \right\}$$

$$\le \left(x(NT/M)^{1/2}\right)^{-\nu} \sum_{k=\log M}^{\log t_0} e^{-k\nu} E \max_{e^k \le \ell \le e^{k+1}} \left| \sum_{s=t_0-\ell}^{t_0} \sum_{i=1}^{N} \zeta_i(s) \right|^{\nu}. \tag{A.19}$$

Next we need a maximal inequality for the double sum in the last term above. With $\bar{\zeta}_i(s) = \zeta_i(t_0 - s + 1)$ we get that

$$\sum_{s=t_0-\ell+1}^{t_0} \sum_{i=1}^{N} \zeta_i(s) = \sum_{s=1}^{\ell} \sum_{i=1}^{N} \bar{\zeta}_i(s).$$

By the independence of the processes $\bar{\zeta}_i(s)$ in i, Rosenthal's inequality (cf. Petrov (1995, p.59)) implies that

$$E\left| \sum_{i=1}^{N} \sum_{s=u}^{v} \bar{\zeta}_i(s) \right|^{\nu} \le c \left\{ \sum_{i=1}^{N} E\left| \sum_{s=u}^{v} \bar{\zeta}_i(s) \right|^{\nu} + \left(\sum_{i=1}^{N} E\left(\sum_{s=u}^{v} \bar{\zeta}_i(s) \right)^2 \right)^{\nu/2} \right\}. \tag{A.20}$$

We have via the Cauchy–Schwarz inequality

$$E\left(\sum_{s=u}^{v} \bar{\zeta}_i(s) \right)^2 \le E\left(Q_i(t_0) \sum_{s=u}^{v} e_{i,s} \right)^2 \le \left(E Q_i^4(t_0) \right)^{1/2} \left(E\left(\sum_{s=u}^{v} e_{i,s} \right)^4 \right)^{1/2}$$

and therefore

$$\sum_{i=1}^{N} E\left(\sum_{s=u}^{v} \bar{\zeta}_i(s) \right)^2 \le \sum_{i=1}^{N} \left(E Q_i^4(t_0) \right)^{1/2} \left(E\left(\sum_{s=u}^{v} e_{i,s} \right)^4 \right)^{1/2}$$

$$\le \left\{ \sum_{i=1}^{N} E Q_i^4(t_0) \sum_{i=1}^{N} E\left(\sum_{s=u}^{v} e_{i,s} \right)^4 \right\}^{1/2}.$$

Using that the $e_{i,t}$'s have mean zero and Assumption 2.3, we conclude

$$\sum_{i=1}^{N} E\left(\sum_{s=u}^{v} \bar{\zeta}_i(s) \right)^2 \le cNT|u - v|.$$

Similarly, by the definition of $\bar{\zeta}_i$ we get

$$E\left| \sum_{s=u}^{v} \bar{\zeta}_i(s) \right|^{\nu} \le 2^{\nu} \left\{ E\left| Q_i(t_0) \sum_{s=u}^{v} e_{i,s} \right|^{\nu} + \left| E\left[Q_i(t_0) \sum_{s=u}^{v} e_{i,s} \right] \right|^{\nu} \right\},$$

and by applications of the Cauchy–Schwarz inequality we have

$$\left| E\left[Q_i(t_0) \sum_{s=u}^{v} e_{i,s} \right] \right| \le \left\{ E Q_i^2(t_0) E\left[\sum_{s=u}^{v} e_{i,s} \right]^2 \right\}^{1/2},$$

https://doi.org/10.1017/S0266466615000468 Published online by Cambridge University Press

$$E\left|Q_i(t_0)\sum_{s=u}^{v} e_{i,s}\right|^{\nu} \leq \left\{E|Q_i(t_0)|^{2\nu} E\left|\sum_{s=u}^{v} e_{i,s}\right|^{2\nu}\right\}^{1/2}$$

resulting in

$$\sum_{i=1}^{N} E\left|\sum_{s=u}^{v} \bar{\zeta}_i(s)\right|^{\nu} \leq cNT^{\nu/2}|u-v|^{\nu/2}.$$

Using the inequalities above, we get the upper bound for the moment in (A.20):

$$E\left|\sum_{s=u}^{v}\sum_{i=1}^{N} \bar{\zeta}_i(s)\right|^{\nu} \leq cN^{\nu/2}T^{\nu/2}|u-v|^{\nu/2}. \tag{A.21}$$

Applying the maximal inequality in Móritz et al. (1982) to (A.21) we conclude

$$E \max_{e^k \leq \ell \leq e^{k+1}} \left|\sum_{s=1}^{\ell}\sum_{i=1}^{N} \bar{\zeta}_i(s)\right|^{\nu} \leq cN^{\nu/2}T^{\nu/2}e^{k\nu/2}.$$

Hence (A.19) implies that

$$\begin{aligned} P&\left\{\max_{1\leq t\leq t_0-M} \frac{1}{t_0-t}\left|\sum_{s=t+1}^{t_0}\sum_{i=1}^{N} \zeta_i(s)\right| \geq x(NT/M)^{1/2}\right\} \\ &\leq \frac{c}{x^{\nu}(NT/M)^{\nu/2}} \sum_{k=\log M}^{\infty} e^{-k\nu/2}N^{\nu/2}T^{\nu/2} \\ &\leq \frac{c}{x^{\nu}}, \end{aligned}$$

resulting in

$$\max_{1\leq t\leq t_0-M} \frac{1}{t_0-t}\left|\sum_{s=t+1}^{t_0}\sum_{i=1}^{N} \zeta_i(s)\right| = O_P((NT/M)^{1/2}).$$

Similar arguments yield

$$\max_{t_0+M\leq t\leq T} \frac{1}{t_0-t}\left|\sum_{s=t+1}^{t_0}\sum_{i=1}^{N} \zeta_i(s)\right| = O_P((NT/M)^{1/2}),$$

which completes the proof of the lemma. ∎

LEMMA A.4. *If Assumptions 2.1–2.3 hold, then*

$$\max_{|t-t_0|\geq M} \frac{1}{|t-t_0|}\left|\sum_{i=1}^{N} \delta_i(r(t)Q_i(t)-r(t_0)Q_i(t_0))\right| = O_P\left(T^{1/2}\Delta_{N,T}^{1/2} + T(\Delta_{N,T}/M)^{1/2}\right).$$

https://doi.org/10.1017/S0266466615000468 Published online by Cambridge University Press

Proof. First we write

$$\delta_i(r(t)Q_i(t) - r(t_0)Q_i(t_0)) = r(t)\delta_i(Q_i(t) - Q_i(t_0)) + \delta_i Q_i(t_0)(r(t) - r(t_0)).$$

Applying the definition of $r(t)$ with Assumptions 2.1(i) and 2.3(i), we get

$$\max_{|t-t_0|\le T} \frac{|r(t)-r(t_0)|}{|t-t_0|}\left|\sum_{i=1}^{N}\delta_i Q_i(t_0)\right| = O_P\left(T^{1/2}\Delta_{N,T}^{1/2}\right).$$

It follows from the definition of $Q_i(t)$ that for all i

$$Q_i(t_0) - Q_i(t) = Z_i(t,t_0) - \frac{t_0 - t}{T}\sum_{s=1}^{T} e_{i,s}, \ \text{ if } 1 \le t \le T,$$

where

$$Z_i(t,t_0) = \begin{cases} \displaystyle\sum_{s=t+1}^{t_0} e_{i,s}, & \text{if } 1 \le t < t_0 \\ 0, & \text{if } t = t_0 \\ \displaystyle -\sum_{s=t_0+1}^{t} e_{i,s}, & \text{if } t_0 < t \le T. \end{cases}$$

Clearly, $\left(\sum_{i=1}^N \delta_i \sum_{s=1}^T e_{i,s}\right)^2 = O\left(T\Delta_{N,T}^{1/2}\right)$ on account of Assumptions 2.1(i) and 2.3(i) and therefore

$$\max_{|t-t_0|\le T}\frac{|r(t)|}{|t-t_0|}\left|\frac{t-t_0}{T}\sum_{i=1}^{N}\delta_i\sum_{s=1}^{T}e_{i,s}\right| \le \left|\sum_{i=1}^{N}\delta_i\sum_{s=1}^{T}e_{i,s}\right| = O_P\left(T^{1/2}\Delta_{N,T}^{1/2}\right).$$

Repeating the arguments used in (A.19), by Markov's inequality we have

$$\begin{aligned} &P\left\{\max_{1\le t\le t_0-M}\frac{1}{t_0-t}\left|\sum_{i=1}^{N}\delta_i Z_i(t,t_0)\right| \ge x(\Delta_{N,T}/M)^{1/2}\right\} \\ &\le P\left\{\max_{\log M\le j\le \log T}\max_{e^j\le \ell\le e^{j+1}}\frac{1}{\ell}\left|\sum_{i=1}^{N}\delta_i Z_i(t_0-\ell,t_0)\right| \ge x(\Delta_{N,T}/M)^{1/2}\right\} \\ &\le \sum_{j=\log M}^{\infty} P\left\{\max_{e^j\le \ell\le e^{j+1}}\left|\sum_{i=1}^{N}\delta_i Z_i(t_0-\ell,t_0)\right| \ge x e^j(\Delta_{N,T}/M)^{1/2}\right\} \\ &\le \frac{(M/\Delta_{N,T})^{\nu/2}}{x^\nu}\sum_{j=\log M}^{\infty} e^{-j\nu} E\max_{e^j\le \ell\le e^{j+1}}\left|\sum_{i=1}^{N}\delta_i Z_i(t_0-\ell,t_0)\right|^{\nu}. \end{aligned} \tag{A.22}$$

With $\bar{e}_{i,s} = e_{i,t_0-s+1}$ we get

$$\max_{e^j\le \ell\le e^{j+1}}\left|\sum_{i=1}^{N}\delta_i Z_i(t_0-\ell,t_0)\right| = \max_{e^j\le \ell\le e^{j+1}}\left|\sum_{s=1}^{\ell+1}\sum_{i=1}^{N}\delta_i \bar{e}_{i,s}\right|.$$

https://doi.org/10.1017/S0266466615000468 Published online by Cambridge University Press

Using Assumptions 2.1(i), 2.3(ii) and 2.5(ii) with Rosenthal's inequality we conclude for all $\nu > 2$ that

$$E\left|\sum_{s=u}^{v}\sum_{i=1}^{N}\delta_i\bar{e}_{i,s}\right|^{\nu} \le c(v-u)^{\nu/2}\left\{\sum_{i=1}^{N}|\delta_i|^{\nu}+\Delta_{N,T}^{\nu/2}\right\} \le 2c(v-u)^{\nu/2}\Delta_{N,T}^{\nu/2}, \qquad \textbf{(A.23)}$$

since by the multinomial theorem

$$\sum_{i=1}^{N}|\delta_i|^{\nu} \le \Delta_{N,T}^{\nu/2}.$$

The maximal inequality of Móricz et al. (1982) and (A.23) imply that

$$E\max_{e^j\le\ell\le e^{j+1}}\left|\sum_{i=1}^{N}\delta_i Z_i(t_0-\ell,t_0)\right|^{\nu} \le c\Delta_{N,T}^{\nu/2}e^{j\nu/2},$$

and therefore by (A.23) we have

$$\begin{aligned}P&\left\{\max_{1\le t\le t_0-M}\frac{1}{t_0-t}\left|\sum_{i=1}^{N}\delta_i Z_i(t,t_0)\right| \ge x(\Delta_{N,T}/M)^{1/2}\right\}\\ &\le c\frac{(M/\Delta_{N,T})^{\nu/2}}{x^{\nu}}\sum_{j=\log M}^{\infty}e^{-j\nu/2}\Delta_{N,T}^{\nu/2}\\ &\le \frac{c}{x^{\nu}}. \qquad \textbf{(A.24)}\end{aligned}$$

Thus we conclude that

$$\max_{1\le t\le t_0-M}\frac{1}{t_0-t}\left|\sum_{i=1}^{N}\delta_i Z_i(t,t_0)\right| = O_P\left((\Delta_{N,T}/M)^{1/2}\right)$$

and by similar arguments we have

$$\max_{t_0+M\le t\le T}\frac{1}{t-t_0}\left|\sum_{i=1}^{N}\delta_i Z_i(t,t_0)\right| = O_P\left((\Delta_{N,T}/M)^{1/2}\right),$$

which also completes the proof of the lemma. ■

LEMMA A.5. *If Assumptions 2.1–2.4 hold, then*

$$\max_{|t-t_0|\ge M}\frac{1}{|t-t_0|}\left|\sum_{i=1}^{N}\gamma_i^2(V^2(t)-V^2(t_0))\right| = O_P\left(\Gamma_{N,T}\left(1+(\log(T/M))^{2/\bar{\kappa}}+M^{-1/2}T^{1/2}\right)\right).$$

https://doi.org/10.1017/S0266466615000468 Published online by Cambridge University Press

Proof. We write $|V^2(t)-V^2(t_0)| \le (V(t)-V(t_0))^2+2|V(t_0)||V(t)-V(t_0)|$. If $1 \le t \le t_0$, then

$$|V(t)-V(t_0)| \le \left|\sum_{s=t+1}^{t_0}\eta_s\right| + \frac{|t-t_0|}{T}\left|\sum_{s=1}^{T}\eta_s\right|$$

and therefore

$$(V(t)-V(t_0))^2 \le 4\left(\sum_{s=t+1}^{t_0}\eta_s\right)^2 + 4\frac{(t-t_0)^2}{T^2}\left(\sum_{s=1}^{T}\eta_s\right)^2.$$

Thus we get from Assumption 2.4 that

$$\max_{|t-t_0|\ge M}\frac{(V(t)-V(t_0))^2}{t_0-t} = O_P(1)+2\left(\max_{|t-t_0|\ge M}\frac{1}{(t_0-t)^{1/2}}\left|\sum_{s=t+1}^{t_0}\eta_s\right|\right)^2.$$

Repeating the arguments used in (A.19) we get that

$$\begin{aligned}
&P\left\{\max_{t_0-t\ge M}\frac{1}{(t_0-t)^{1/2}}\left|\sum_{s=t+1}^{t_0}\eta_s\right|\ge x\right\}\\
&\quad\le P\left\{\max_{\log M\le k\le\log T}\max_{e^k\le u\le e^{k+1}}\frac{1}{u^{1/2}}\left|\sum_{s=t_0-u+1}^{t_0}\eta_s\right|\ge x\right\}\\
&\quad\le \sum_{k=\log M}^{\log T}P\left\{\max_{e^k\le u\le e^{k+1}}\left|\sum_{s=t_0-u+1}^{t_0}\eta_s\right|\ge xe^{k/2}\right\}\\
&\quad\le \frac{1}{x^{\bar\kappa}}\sum_{k=\log M}^{\log T}e^{-k\bar\kappa/2}E\max_{e^k\le u\le e^{k+1}}\left|\sum_{s=t_0-u+1}^{t_0}\eta_s\right|^{\bar\kappa}.
\end{aligned}$$

Following the arguments used in the proofs of Lemmas A.3 and A.4 one can verify that

$$E\max_{e^k\le u\le e^{k+1}}\left|\sum_{s=t_0-u+1}^{t_0}\eta_s\right|^{\bar\kappa}\le ce^{k\bar\kappa/2}$$

which implies that

$$P\left\{\max_{t_0-t\ge M}\frac{1}{(t_0-t)^{1/2}}\left|\sum_{s=t+1}^{t_0}\eta_s\right|\ge x\right\}\le c\frac{\log(T/M)}{x^{\bar\kappa}}.$$

Similar computations can be performed for $t-t_0\ge M$ and thus we conclude

$$\max_{|t_0-t|\ge M}\frac{1}{(t_0-t)^{1/2}}\left|\sum_{s=t+1}^{t_0}\eta_s\right| = O_P\left((\log(T/M))^{1/\bar\kappa}\right).$$

As in the proof of Lemma A.3 we have that

$$\sup_{t_0 - t \geq M} \frac{1}{t_0 - t} \left| \sum_{s=t+1}^{t_0} \eta_s \right| = O_P\left(M^{-1/2}\right) \quad \text{and}$$
$$\sup_{t - t_0 \geq M} \frac{1}{t - t_0} \left| \sum_{s=t_0+1}^{t} \eta_s \right| = O_P\left(M^{-1/2}\right). \tag{A.25}$$

The proof of the lemma is now complete. ■

LEMMA A.6. *If Assumptions 2.1–2.4 hold, then*

$$\max_{|t-t_0| \geq M} \frac{1}{|t - t_0|} \left| \sum_{i=1}^{N} \gamma_i (Q_i(t)V(t) - Q_i(t_0)V(t_0)) \right| = O_P\left(\Gamma_{N,T}^{1/2} T^{1/2} M^{-1/2}\right).$$

Proof. We write

$$Q_i(t)V(t) - Q_i(t_0)V(t_0) = V(t)(Q_i(t) - Q_i(t_0)) + (V(t) - V(t_0))Q_i(t_0).$$

Assumption 2.4 implies that

$$\max_{1 \leq t \leq T} |V(t)| = O_P\left(T^{1/2}\right)$$

and by the arguments used in the proof of Lemma A.4 one can show that

$$\max_{|t-t_0| \geq M} \frac{1}{|t - t_0|} \left| \sum_{i=1}^{N} \gamma_i (Q_i(t) - Q_i(t_0)) \right| = O_P\left(M^{-1/2} \Gamma_{N,T}^{1/2}\right).$$

Similar arguments yield

$$\max_{|t-t_0| \geq M} \frac{|V(t) - V(t_0)|}{|t - t_0|} \left| \sum_{i=1}^{N} \gamma_i Q_i(t_0) \right| = O_P\left(T^{1/2} M^{-1/2} \Gamma_{N,T}^{1/2}\right).$$

■

LEMMA A.7. *If Assumptions 2.1–2.4 hold, then*

$$\max_{|t-t_0| \geq M} \frac{1}{|t - t_0|} \left| \sum_{i=1}^{N} \gamma_i \delta_i (V(t)r(t) - V(t_0)r(t_0)) \right| = O_P\left(\left(T^{1/2} + T M^{-1/2}\right) |\Sigma_{N,T}|\right).$$

Proof. Since $V(t)r(t) - V(t_0)r(t_0) = V(t_0)(r(t) - r(t_0)) + r(t)(V(t) - V(t_0))$, Lemma A.7 follows from

$$\max_{|t-t_0| \geq M} \left| V(t_0) \frac{r(t) - r(t_0)}{t - t_0} \right| = O_P\left(T^{1/2}\right) \tag{A.26}$$

https://doi.org/10.1017/S0266466615000468 Published online by Cambridge University Press

and

$$\max_{|t-t_0|\geq M}\left|r(t)\frac{V(t)-V(t_0)}{t-t_0}\right| = O_P\left(TM^{-1/2}\right). \qquad \textbf{(A.27)}$$

The claim in (A.26) is an immediate consequence of the definition of $r(t)$ and Assumption 2.4 while (A.27) is proven in (A.25). ■

Proof of Theorem 2.1. Under assumptions (2.3) and (2.4) we use Lemmas A.2–A.7 with $M = 1$. ■

Proof of Remark 2.1. The proof of this remark follows Bai (2010) closely. We use (A.3). Since T is fixed,

$$\max_{1\leq t\leq T}\sum_{i=1}^{N}\gamma_i^2 V^2(t) = O_P(\Gamma)$$

and by Assumption 2.3(i) and Markov's inequality we have

$$\max_{1\leq t\leq T}\sum_{i=1}^{N} Q_i^2(t) = O_P(N).$$

By the Cauchy–Schwarz inequality and Assumption 2.3(i) we conclude

$$E\left|\sum_{i=1}^{N} V(t)\gamma_i Q_i(t)\right| = O\left(\Gamma^{1/2}\right)$$

and therefore

$$\max_{1\leq t\leq T}\left|\sum_{i=1}^{N} V(t)\gamma_i Q_i(t)\right| = O_P\left(\Gamma^{1/2}\right).$$

Similar arguments give

$$\max_{1\leq t\leq T}\left|r(t)\sum_{i=1}^{N}\delta_i Q_i(t)\right| = O_P\left(\Delta^{1/2}\right)$$

and

$$\max_{1\leq t\leq T}\left|r(t)V(t)\sum_{i=1}^{N}\delta_i\gamma_i\right| = O_P(|\Sigma|).$$

The final term coming from (A.3) to consider is $\Sigma_{i=1}^{N} r^2(t)\delta_i^2 = \Delta r^2(t)$. Under the conditions of the remark, this is the asymptotically dominating term which has a unique maximum at t_0. Hence Remark 2.1 is proven. ■

https://doi.org/10.1017/S0266466615000468 Published online by Cambridge University Press

Proof of Remark 2.2. Let

$$f_i(t) = \frac{1}{(t(T-t))^{1/2}}\left(Q_i(t)+\gamma_i V(t)\right),$$

where $Q_i(t)$ and $V(t)$ are defined in (A.1). We note that due to the assumption that the $e_{i,s}$ and η_t are sequences of uncorrelated random variables we get that

$$Ef_i^2(t) = \sigma_i^2 + \gamma_i^2 \tag{A.28}$$

and

$$\operatorname{var}(f_i^2(t)) \le C_1\left(Ee_{i,0}^4 + \gamma_i^4\right) \tag{A.29}$$

with some constant C_1. We write

$$\sum_{i=1}^N \left(S_i(t) - \frac{t}{T}S_i(t)\right)^2 \frac{1}{t(T-T)} = \mathcal{H}_{1,N}(t) + \mathcal{H}_{2,N}(t) + \mathcal{H}_{3,N}(t),$$

with

$$\mathcal{H}_{1,N}(t) = \Delta\frac{r^2(t)}{t(T-T)}, \quad \mathcal{H}_{2,N}(t) = \sum_{i=1}^N f_i^2(t) \quad \text{and} \quad \mathcal{H}_{3,n}(t) = 2\sum_{i=1}^N f_i(t)\frac{\delta_i r(t)}{(t(T-t))^{1/2}},$$

where $r(t), t = 1, 2, \ldots, T$ is defined in (A.2). We show that for all $t \neq t_0$

$$\lim_{N\to\infty} P\{\mathcal{H}_{1,N}(t_0) - \mathcal{H}_{1,N}(t) \le \mathcal{H}_{2,N}(t) - \mathcal{H}_{2,N}(t_0) + \mathcal{H}_{2,N}(t) - \mathcal{H}_{2,N}(t_0)\} = 0, \tag{A.30}$$

which immediately implies Remark 2.2. We note that with some $C_2 > 0$ we have that $\mathcal{H}_{1,N}(t_0) - \mathcal{H}_{1,N}(t) \ge C_2\Delta$ for all $t \neq t_0$. By the independence of the processes $Q_i(t)$, $1 \le i \le N$ and V(t) we conclude

$$\begin{aligned} E\left(\mathcal{H}_{2,N}(t) - \mathcal{H}_{2,N}(t_0)\right)^2 &= \sum_{i=1}^N E\left(\frac{1}{t(T-t)}Q_i^2(t) - \frac{1}{t_0(T-t_0)}Q_i^2(t_0)\right)^2 \\ &\quad + 4\sum_{i=1}^N E\left(\frac{1}{t(T-t)}Q_i(t)\gamma_i V(t) - \frac{1}{t_0(T-t_0)}Q_i(t_0)\gamma_i V(t)\right)^2 \\ &\quad + E\left(\frac{V^2(t)}{t(T-t)} - \frac{V^2(t_0)}{t_0(T-t_0)}\right)^2\Gamma^2 \\ &= O\left(N + \Gamma + \Gamma^2\right) \end{aligned}$$

and therefore

$$\max_{1\le t<T}\left|\mathcal{H}_{2,N}(t) - \mathcal{H}_{2,N}(t_0))\right| = O_P\left(N^{1/2} + \Gamma^{1/2} + \Gamma\right).$$

Similarly,

$$\max_{1\le t<T}\left|\mathcal{H}_{3,N}(t) - \mathcal{H}_{3,N}(t_0))\right| = O_P\left(\Delta^{1/2} + |\Sigma|\right) = O_P\left(\Delta^{1/2} + \Delta^{1/2}\Gamma^{1/2}\right),$$

since $|\Sigma| \le \Delta^{1/2}\Gamma^{1/2}$, completing the proof of (A.30). ∎

https://doi.org/10.1017/S0266466615000468 Published online by Cambridge University Press

LEMMA A.8. *We assume that Assumptions 2.1–2.6 hold, and* $|\mathfrak{s}| < \infty$. *Then, as* $N, T \to \infty$ *we have that*

$$\Delta_{N,T}|\hat{t}_{N,T} - t_0| = O_P(1). \quad \textbf{(A.31)}$$

Proof. By Lemma A.1 it is enough to prove that for all $0 < \alpha < \theta$

$$\Delta|\tilde{t}_{N,T}(\alpha) - t_0| = O_P(1). \quad \textbf{(A.32)}$$

Under Assumption 2.6(i) we choose $M = C/\Delta_{N,T}$, where $C > 0$ is a constant. Using Lemmas A.2, A.3, A.5–A.7 and Assumption 2.5(ii) we obtain that

$$\Delta_{N,T}(r^2(t) - r^2(t_0)) + \sum_{i=1}^{N}\left(Q_i^2(t) - Q_i^2(t_0)\right) + \Gamma_{N,T}\left(V^2(t) - V^2(t_0)\right) \quad \textbf{(A.33)}$$
$$+2\sum_{i=1}^{N}\gamma_i(Q_i(t)V(t) - Q_i(t_0)V(t_0)) + 2\sum_{i=1}^{N}\gamma_i\delta_i(V(t)r(t) - V(t_0)r(t_0))$$
$$= \Delta_{N,T}\left(r^2(t) - r^2(t_0)\right)(1 + o_P(1)) \ \text{uniformly on} \ |t - t_0| \geq M$$

for all M. Also, by Lemmas A.2 and A.4 we obtain that

$$\lim_{C\to\infty}\liminf_{N,T\to\infty} P\left\{\sup_{|t-t_0|\geq C/\Delta_{N,T}}\sum_{i=1}^{N}\left[\delta_i^2\left(r^2(t) - r^2(t_0)\right) + 2\delta_i(r(t)Q_i(t) - r(t_0)Q_i(t_0))\right] < 0\right\} = 1.$$

Hence Lemma A.8 is established under Assumption 2.6(i) and $|\mathfrak{s}| < \infty$. ∎

LEMMA A.9. *We assume that Assumptions 2.1–2.6 hold, and* $|\mathfrak{s}| < \infty$. *Then, as* $N, T \to \infty$ *we have that*

$$\sup_{|t-t_0|\leq C/\Delta}\left|\frac{1}{T}\sum_{i=1}^{N}\delta_i^2\left(r^2(t_0) - r^2(t)\right) - 2\theta(1-\theta)\Delta g_\theta(t - t_0)\right| = o(1), \quad \textbf{(A.34)}$$

$$\sup_{|t-t_0|\leq C/\Delta}\left|\frac{1}{T}\sum_{i=1}^{N}\delta_i(Q_i(t)r(t) - Q_i(t_0)r(t_0)) + \theta(1-\theta)\sum_{i=1}^{N}\delta_i(\mathcal{S}_i(t) - \mathcal{S}_i(t_0))\right|$$
$$= o_P(1), \quad \textbf{(A.35)}$$

$$\sup_{|t-t_0|\leq C/\Delta}\left|\frac{1}{T}\sum_{i=1}^{N}\gamma_i\delta_i(V(t)r(t) - V(t_0)r(t_0)) + \theta(1-\theta)\Sigma_{N,T}(V(t) - V(t_0))\right|$$
$$= o_P(1), \quad \textbf{(A.36)}$$

$$\sup_{|t-t_0|\le C/\Delta}\left|\frac{1}{T}\sum_{i=1}^{N}\left(Q_i^2(t)-Q_i^2(t_0)\right)\right|=o_P(1), \tag{A.37}$$

$$\sup_{|t-t_0|\le C/\Delta}\left|\frac{1}{T}\sum_{i=1}^{N}\gamma_i^2\left(V^2(t)-V^2(t_0)\right)\right|=o_P(1) \tag{A.38}$$

and

$$\sup_{|t-t_0|\le C/\Delta}\left|\frac{1}{T}\sum_{i=1}^{N}\gamma_i(Q_i(t)V(t)-Q_i(t_0)V(t_0))\right|=o_P(1), \tag{A.39}$$

for all $C>0$*, where* $\Delta=\Delta_{N,T}$ *and* $\mathcal{S}_i(\cdot)$ *is defined in (A.18).*

Proof. First we note

$$\frac{1}{T}\sum_{i=1}^{N}\delta_i^2\left(r^2(t)-r^2(t_0)\right)=\frac{2r(t_0)}{T}\sum_{i=1}^{N}\delta_i^2(r(t)-r(t_0))+\frac{1}{T}\sum_{i=1}^{N}\delta_i^2(r(t)-r(t_0))^2.$$

Using the definition of $r(t)$ and Assumption 2.5(i) we conclude

$$\sup_{|t-t_0|\le C/\Delta}\left|\frac{1}{T}\sum_{i=1}^{N}\delta_i^2(r(t)-r(t_0))^2\right|=O(1/(T\Delta))=o(1)$$

and

$$\sup_{|t-t_0|\le C/\Delta}\left|\frac{2r(t_0)}{T}\sum_{i=1}^{N}\delta_i^2(r(t)-r(t_0))-2\theta(1-\theta)\Delta g_\theta(t-t_0)\right|=o(1),$$

completing the proof of (A.34).
Similarly,

$$\sum_{i=1}^{N}\delta_i(Q_i(t)r(t)-Q_i(t_0)r(t_0))=r(t_0)\sum_{i=1}^{N}\delta_i(Q_i(t)-Q_i(t_0))+\sum_{i=1}^{N}\delta_iQ_i(t)(r(t)-r(t_0))$$

and

$$\sum_{i=1}^{N}\delta_i(Q_i(t)-Q_i(t_0))=\sum_{i=1}^{N}\delta_i(\mathcal{S}_i(t)-\mathcal{S}_i(t_0))+\frac{t-t_0}{T}\sum_{i=1}^{N}\delta_i\mathcal{S}_i(T).$$

Computing the variance of $\sum_{i=1}^{N}\delta_i\mathcal{S}_i(T)$ we get

$$\sup_{|t-t_0|\le C/\Delta}\left|\frac{r(t_0)}{T}\frac{t-t_0}{T}\sum_{i=1}^{N}\delta_i\mathcal{S}_i(T)\right|=O_P\left(1/(T\Delta)^{1/2}\right)=o_P(1)$$

by Assumption 2.5(i), so (A.35) is proven.
Clearly,

$$V(t)r(t) - V(t_0)r(t_0) = (V(t) - V(t_0))r(t) + V(t_0)(r(t) - r(t_0)).$$

By Assumption 2.4 we get that $V(t_0) = O_P(T^{1/2})$ and therefore

$$\sup_{|t-t_0|\le C/\Delta} |V(t_0)(r(t) - r(t_0))| = O_P\left(T^{1/2}/\Delta\right).$$

We note that for all $t_0 \le t \le t_0 + C/\Delta$

$$|V(t) - V(t_0)| \le \left|\sum_{s=t_0+1}^{t} \eta_s\right| + \frac{|t-t_0|}{T}\left|\sum_{s=1}^{T}\eta_s\right|$$

and by Assumption 2.4 we have that $|\sum_{s=t_0+1}^{t}\eta_s| = O_P\left(1/\Delta^{1/2}\right)$ and the process $\Delta^{1/2}\sum_{s=1}^{\lfloor u/\Delta\rfloor}\eta_s$, $0 \le u \le 1$ is tight in $\mathcal{D}[0, C]$. Thus by stationarity we get

$$\sup_{t_0\le t\le C/\Delta}\left|\sum_{s=t_0+1}^{t}\eta_s\right| = O_P\left(1/\Delta^{1/2}\right)$$

and similar arguments can be used on $t_0 - C/\Delta \le t \le t_0$. We conclude that

$$\sup_{|t-t_0|\le C/\Delta}|V(t) - V(t_0)| = O_P\left(1/\Delta^{1/2} + T^{-1/2}/\Delta\right) = O_P\left(1/\Delta^{1/2}\right) \tag{A.40}$$

completing the proof of (A.36) on account $|\mathfrak{s}| < \infty$.
With $\psi_i(t) = Q_i^2(t) - Q_i^2(t_0) - E(Q_i^2(t) - Q_i^2(t_0))$ we can write

$$\sup_{|t-t_0|\le C/\Delta}\frac{1}{T}\left|\sum_{i=1}^{N}\left(Q_i^2(t) - Q_i^2(t_0)\right)\right|$$
$$\le \sup_{|t-t_0|\le C/\Delta}\frac{1}{T}\left|\sum_{i=1}^{N}\left(EQ_i^2(t) - EQ_i^2(t_0)\right)\right| + \sup_{|t-t_0|\le C/\Delta}\frac{1}{T}\left|\sum_{i=1}^{N}\psi_i(t)\right|.$$

We obtain from the proof of Lemma A.1 that

$$\sup_{|t-t_0|\le C/\Delta}\frac{1}{T}\left|\sum_{i=1}^{N}\left(EQ_i^2(t) - EQ_i^2(t_0)\right)\right| = O\left(\frac{N}{T\Delta}\right) = o(1).$$

For every $t \in [t_0 - C/\Delta, t_0 + C/\Delta]$ we have that

$$E\left(\frac{1}{T}\sum_{i=1}^{N}\psi_i(t)\right)^2 \le C\frac{1}{T^2}\sum_{i=1}^{N}E\psi_i^2(t)$$

https://doi.org/10.1017/S0266466615000468 Published online by Cambridge University Press

and

$$\sup_{|t-t_0|\le C/\Delta} E\psi_i^2(t) \le 9 \sup_{|t-t_0|\le C/\Delta}\left\{Q_i^2(t)\left(Q_i(t)-Q_i(t_0)\right)^2+Q_i^2(t_0)\left(Q_i(t)-Q_i(t_0)\right)^2\right\}$$
$$= O(1)\left\{\sup_{|t-t_0|\le C/\Delta}\left(EQ_i^4(t)\right)^{1/2}\sup_{|t-t_0|\le C/\Delta}\left(E(Q_i(t)-Q_i(t_0))^4\right)^{1/2}\right\}$$
$$= O(T/\Delta).$$

Next we show that $\sqrt{\Delta/(NT)}\sum_{i=1}^N \psi_i(u/\Delta)$ is tight in $\mathcal{D}[-C,C]$. Using Rosenthal's inequality (cf. Petrov (1995, p. 59)) we obtain that

$$E\left|\sum_{i=1}^N(\psi_i(t)-\psi_i(s))\right|^{\kappa/2} \le c\left\{\sum_{i=1}^N E|\psi_i(t)-\psi_i(s)|^{\kappa/2}+\left(\sum_{i=1}^N E(\psi_i(t)-\psi_i(s))^2\right)^{\kappa/4}\right\}$$

with some constant c. It is easy to see that

$$|\psi_i(t)-\psi_i(s)| \le \left\{|Q_i(t)(Q_i(t)-Q_i(s))|+|Q_i(s)(Q_i(t)-Q_i(s))|+|EQ_i^2(t)-EQ_i^2(s)|\right\}$$

and

$$|EQ_i^2(t)-EQ_i^2(s)| \le c|t-s|.$$

By the Cauchy–Schwarz inequality and Assumption 2.3(ii) we have for all $t,s\in[t_0-C/\Delta, t_0+C/\Delta]$

$$E(Q_i(t)(Q_i(t)-Q_i(s)))^2 \le \left(EQ_i^4(t)E(Q_i(t)-Q_i(s))^4\right)^{1/2} \le cT|t-s|$$

and therefore

$$E(\psi_i(t)-\psi_i(s))^2 \le cT|t-s|$$

where c is a constant. Also, for κ of Assumption 2.3(ii) we have

$$E|Q_i(t)(Q_i(t)-Q_i(s))|^{\kappa/2} \le \left\{E|Q_i(t)|^\kappa E|Q_i(t)-Q_i(s)|^\kappa\right\}^{1/2} \le c\left\{U_{i,\kappa}(t)U_{i,\kappa}(|t-s|)\right\}^{1/2}$$

with some constant c. Thus we get via Assumption 2.3(ii) that

$$E\left\{\left|\left(\frac{\Delta}{NT}\right)^{1/2}\sum_{i=1}^N(\psi_i(u/\Delta)-\psi_i(v/\Delta))\right|^{\kappa/2}\right\} \le c|u-v|^{\kappa/4},$$

establishing tightness by Billingsley (1968, pp. 95 and 127). This also completes the proof of (A.37).

Following the arguments in the proof of (A.36) one can show that

$$\sup_{|t-t_0|\le C/\Delta}\left|V^2(t)-V^2(t_0)\right| \le 2\sup_{|t-t_0|\le C/\Delta}|V(t)-V(t_0)|\sup_{|t-t_0|\le C/\Delta}|V(t)|$$
$$= O_P\left(T^{1/2}/\Delta^{1/2}\right),$$

and therefore (A.38) follows from Assumption 2.5(ii). To prove (A.39) we first write

$$Q_i(t)V(t) - Q_i(t_0)V(t_0) = V(t)(Q_i(t) - Q_i(t_0)) + Q_i(t_0)(V(t) - V(t_0)).$$

Repeating the arguments used in the proof of (A.37) we obtain that

$$\frac{1}{T} \sup_{|t-t_0|\le C/\Delta} \left| V(t) \sum_{i=1}^{N} \gamma_i (Q_i(t) - Q_i(t_0)) \right| = O_P\left((\Gamma_{N,T}/(T\Delta))^{1/2} \right) = o_P(1)$$

via applying Assumption 2.5(ii) and $T\Delta \to \infty$. ■

Let

$$R_{N,T}(u) = \sum_{i=1}^{N} \delta_i \mathcal{S}_i(u/\Delta_{N,T}), \quad u \ge 0,$$

where $\mathcal{S}_i(\cdot)$ is defined in (A.18).

LEMMA A.10. *If Assumptions 2.1, 2.3, 2.7, (2.8) and (2.10) hold, then we have*

$$R_{N,T}(u) \overset{\mathcal{D}[0,C]}{\longrightarrow} \sigma W(u),$$

for all $C > 0$*, where* $W(u)$ *stands for a Wiener process.*

Proof. For the sake of notational simplicity we write $\Delta = \Delta_{N,T}$. Let $0 = u_0 < u_1 < u_2 < \cdots < u_k \le C$ and $\alpha_1, \alpha_2, \ldots, \alpha_k$. Under (2.10) we write

$$\sum_{\ell=1}^{k} \alpha_\ell (R_{N,T}(u_\ell) - R_{N,T}(u_{\ell-1})) = \sum_{i=1}^{N} \sum_{\ell=1}^{k} \alpha_\ell \delta_i (\mathcal{S}_i(u_\ell/\Delta) - \mathcal{S}_i(u_{\ell-1}/\Delta)).$$

Using Assumptions 2.1, 2.3, 2.7(i), and (2.10), we get that

$$\sum_{i=1}^{N} E\left(\sum_{\ell=1}^{k} \alpha_\ell \delta_i (\mathcal{S}_i(u_\ell/\Delta) - \mathcal{S}_i(u_{\ell-1}/\Delta)) \right)^2 = \sigma^2 \sum_{\ell=1}^{k} \alpha_\ell^2 (u_\ell - u_{\ell-1})(1 + o(1)).$$

Also, Assumptions 2.1(ii) and 2.7(ii) imply

$$\begin{aligned} &\sum_{i=1}^{N} E\left| \delta_i \sum_{\ell=1}^{k} \alpha_\ell (\mathcal{S}_i(u_\ell/\Delta) - \mathcal{S}_i(u_{\ell-1}/\Delta)) \right|^{\bar{\tau}} \\ &\quad \le c \sum_{i=1}^{N} |\delta_i|^{\bar{\tau}} \max_{1\le\ell\le k} U_{i,\bar{\tau}}(|u_\ell - u_{\ell-1}|/\Delta) \\ &\quad \le c \left\{ \Delta^{-\bar{\tau}/2} \sum_{i=1}^{N} |\delta_i|^{\bar{\tau}} \right\} \end{aligned} \qquad \textbf{(A.41)}$$

https://doi.org/10.1017/S0266466615000468 Published online by Cambridge University Press

So using Lyapunov's theorem (cf. Petrov (1995, p. 154)) we conclude via (2.8) that

$$\sum_{\ell=1}^{k}\alpha_\ell(R_{N,T}(u_\ell)-R_{N,T}(u_{\ell-1}))\ \stackrel{\mathcal{D}}{\to}\ \sigma\sum_{\ell=1}^{k}\alpha_\ell(W(u_\ell)-W(u_{\ell-1})),$$

where W stands for a Wiener process. Applying the Cramér–Wold theorem (cf. Billingsley (1968, p. 49)) we obtain that the finite dimensional distributions of $R_{N,T}(u)$ converge to that of $\sigma W(u)$. Next we show that $R_{N,T}(u)$ is tight in $\mathcal{D}[0,C]$. Following the arguments in (A.41), Rosenthal's inequality yields for all $0\le u,v\le C$

$$\begin{aligned}
&E|R_{N,T}(u)-R_{N,T}(v)|^{\bar\tau}\\
&\quad\le c\Delta^{-\bar\tau/2}\left\{\sum_{i=1}^{N}|\delta_i|^{\bar\tau}E|\mathcal{S}_i(u/\Delta)-\mathcal{S}_i(v/\Delta)|^{\bar\tau}+\left(\sum_{i=1}^{N}\delta_i^2E(\mathcal{S}_i(u/\Delta)-\mathcal{S}_i(v/\Delta))^2\right)^{\bar\tau/2}\right\}\\
&\quad\le c\left\{\sum_{i=1}^{N}|\delta_i|^{\bar\tau}U_{i,\bar\tau}(|u-v|/\Delta)+\left(\sum_{i=1}^{N}\delta_i^2U_{i,2}(|u-v|/\Delta)\right)^{\bar\tau/2}\right\}\\
&\quad\le c\left\{\Delta^{-\bar\tau/2}\sum_{i=1}^{N}|\delta_i|^{\bar\tau}+1\right\}|u-v|^{\bar\tau/2}\\
&\quad\le c|u-v|^{\bar\tau/2}
\end{aligned}$$

on account of Assumption 2.7(ii) and (2.8). The tightness now follows from Billingsley (1968, p. 127). ■

LEMMA A.11. *If Assumptions 2.1, 2.3, 2.7 and (2.8) hold, then for all integers* $0<t_1<t_2<\ldots<t_K$ *we have that*

$$\left(\sum_{i=1}^{N}\delta_i\mathcal{S}_i(t_\ell),1\le\ell\le K\right)\ \stackrel{\mathcal{D}}{\longrightarrow}\ (\mathfrak{G}(t_\ell),1\le\ell\le K),$$

where the Gaussian process $\mathfrak{G}(t),t=0,\pm1,\pm2\ldots$ *is defined in Theorem 2.2.*

Proof. We repeat the first half of the proof of Lemma A.10. The result of Lemma A.11 follows from (A.41) and Lyapunov's central limit theorem due to assumption (2.8). ■

Proof of Theorem 2.2. Let $\Delta=\Delta_{N,T}$. It follows from Assumption 2.1(ii) and Lemma A.10 that for all $C>0$

$$\sum_{i=1}^{N}\delta_i(\mathcal{S}(t_0+u/\Delta)-\mathcal{S}(t_0))\ \stackrel{\mathcal{D}[-C,C]}{\longrightarrow}\ \sigma W(u), \tag{A.42}$$

where $W(u),-\infty<u<\infty$ is a two sided Wiener process. Also, $\Delta g_\theta(u/\Delta)=g_\theta(u)$ and since $\mathfrak{s}=0$ by (A.40) we have that

$$\Sigma_{N,T}\sup_{|t-t_0|\le\Delta}|V(t)-V(t_0)|=o_P(1). \tag{A.43}$$

https://doi.org/10.1017/S0266466615000468 Published online by Cambridge University Press

By Lemma A.9 we conclude that for all $C > 0$

$$\frac{1}{T}(U_N(t_0+u/\Delta)-U_N(t_0)) \overset{\mathcal{D}[-C,C]}{\longrightarrow} 2\theta(1-\theta)(\sigma W(u)-g_\theta(u)). \quad \textbf{(A.44)}$$

By the continuous mapping theorem we conclude from (A.44) that for all C

$$\text{argmax}_{|t-t_0|\le C/\Delta}(U_N(t_0+u/\Delta)-U_N(t_0)) \overset{\mathcal{D}}{\longrightarrow} \text{argmax}_{|u|\le C}(\sigma W(u)-g_\theta(u)). \quad \textbf{(A.45)}$$

According to the law of iterated logarithm, we have that

$$\lim_{C\to\infty} \text{argmax}_{|u|\le C}(\sigma W(u)-g_\theta(u)) \ \to \ \text{argmax}_u(\sigma W(u)-g_\theta(u)) \ \text{ a.s.} \quad \textbf{(A.46)}$$

Now (2.11) follows from Lemma A.8, (A.45) and (A.46).
It follows from Assumption 2.1(ii), (A.43) and Lemmas A.9, A.11 that for every integer $C > 0$

$$\left\{\frac{1}{T}(U_N(t_0+t)-U_N(t_0)),\ t=0,\pm1,\pm2,\ldots,\pm C\right\} \overset{\mathcal{D}}{\longrightarrow} \{2\theta(1-\theta)(\mathfrak{G}(t)-\mathfrak{d}g_\theta(t)),\ t=0,\pm1,\pm2,\ldots,\pm C\}. \quad \textbf{(A.47)}$$

Observing that $\mathfrak{u}(t,t)=O(t)$, the normality of $\mathfrak{G}(t)$ with the Borel–Cantelli lemma yields that $\lim_{|t|\to\infty}\mathfrak{G}(t)/t=0$, and therefore

$$\lim_{C\to\infty} \text{argmax}_{|t|\le C}(\mathfrak{G}(t)-\mathfrak{d}g_\theta(t)) = \text{argmax}_t(\mathfrak{G}(t)-\mathfrak{d}g_\theta(t)) \quad \text{a.s.}$$

The proof of (2.13) is now completed via Lemma A.8. ■

Proof of Theorem 2.3. It follows from Assumption 2.8 that the Wiener processes in Assumption 2.9 and (A.42) are independent. Hence Lemma A.9 yields for all $C > 0$ that

$$\frac{1}{T}(U_N(t_0+u/\Delta)-U_N(t_0)) \overset{\mathcal{D}[-C,C]}{\longrightarrow} 2\theta(1-\theta)\left(\sigma^2+\mathfrak{s}^2\right)^{1/2} W(u)-g_\theta(u)),$$

where $W(u), -\infty < u < \infty$ is a two–sided Wiener process. Arguments used in (A.45) and (A.46) could be repeated to finish the proof of (2.15).
Referring again to Assumption 2.8 it is immediate that the Gaussian process $\mathfrak{G}(t)$ and $\mathcal{V}(t)$ are independent. So applying Lemma A.9 we replace (A.47) with

$$\left\{\frac{1}{T}(U_N(t_0+t)-U_N(t_0)),\ t=0,\pm1,\pm2,\ldots,\pm C\right\} \overset{\mathcal{D}}{\longrightarrow} \left\{2\theta(1-\theta)(\mathfrak{G}(t)+\mathfrak{s}\mathfrak{d}^{1/2}\mathcal{V}(t)-\mathfrak{d}g_\theta(t)),\ t=0,\pm1,\pm2,\ldots,\pm C\right\}$$

for any $C > 0$. Observing that $\mathcal{V}(t)/t \to 0$ a.s. we need only minor modifications of the proof of (2.13) to complete the proof of (2.16). ■

https://doi.org/10.1017/S0266466615000468 Published online by Cambridge University Press

LEMMA A.12. *We assume that Assumptions 2.1–2.6 hold, and* $|\mathfrak{s}| = \infty$. *Then, as* $N, T \to \infty$ *we have that*

$$|\hat{t}_{N,T} - t_0| = O_P(M_{N,T}), \tag{A.48}$$

where $M_{N,T} = (\Sigma_{N,T}/\Delta_{N,T})^2$.

Proof. By Lemma A.1 it is enough to prove that for all $0 < \alpha < \theta$

$$|\tilde{t}_{N,T}(\alpha) - t_0| = O_P(M_{N,T}). \tag{A.49}$$

The result follows from Lemmas A.2–A.7 with $M = (\Sigma_{N,T}/\Delta_{N,T})^2$. ∎

Proof of Theorem 2.4. Let $M = (\Sigma_{N,T}/\Delta_{N,T})^2$. Since $\Delta_{N,T}$ is bounded, by (2.17) we have

$$\frac{\Sigma_{N,T}}{\Delta_{N,T}} \to \infty. \tag{A.50}$$

Following the proof of Lemma A.9 one can show that for all $C > 0$

$$\sup_{|t-t_0|\le CM} \left| \frac{1}{T} \sum_{i=1}^{N} \delta_i (Q_i(t)r(t) - Q_i(t_0)r(t_0)) \right| = o_P(1), \tag{A.51}$$

$$\sup_{|t-t_0|\le CM} \left| \frac{1}{T} \sum_{i=1}^{N} (Q_i^2(t) - Q_i^2(t_0)) \right| = o_P(1), \tag{A.52}$$

$$\sup_{|t-t_0|\le CM} \left| \frac{1}{T} \sum_{i=1}^{N} \gamma_i^2 (V^2(t) - V^2(t_0)) \right| = o_P(1) \tag{A.53}$$

and

$$\sup_{|t-t_0|\le CM} \left| \frac{1}{T} \sum_{i=1}^{N} \gamma_i (Q_i(t)V(t) - Q_i(t_0)V(t_0)) \right| = o_P(1). \tag{A.54}$$

It follows from Assumptions 2.4 and 2.9 that for all $C > 0$

$$\frac{1}{T}\left(\Delta_{N,T}\left(r^2(t_0+uM) - r^2(t_0)\right) + 2\Sigma_{N,T}(V(t_0+uM)r(t_0+uM) - V(t_0)r(t_0))\right) \xrightarrow{\mathcal{D}[-C,C]} 2\theta(1-\theta)(W(u) - g_\theta(u)),$$

where $W(u), -\infty < u < \infty$ denotes a two–sided Wiener process. Using now (A.51)–(A.54)

$$\frac{1}{T}(U_N(t_0+uM) - U_N(t_0)) \xrightarrow{\mathcal{D}[-C,C]} 2\theta(1-\theta)(W(u) - g_\theta(u)).$$

Arguments used in (A.45) and (A.46) could be repeated to complete the proof of Theorem 2.4. ∎

https://doi.org/10.1017/S0266466615000468 Published online by Cambridge University Press

APPENDIX B. Proof of Theorem 4.1

For the sake of brevity we use $\hat{t}$, Δ and Σ for $\hat{t}_{N,T}$, $\Delta_{N,T}$ and $\Sigma_{N,T}$, respectively . We start with the proof of (4.4). Let $M = M(N,T)$ be a sequence satisfying

$$M \to \infty, \quad M/T \to 0, \tag{B.1}$$

if the conditions of Theorem 2.2 or 2.3 hold and

$$M \to \infty, \quad M/\min\left(T, \Sigma^2/\Delta^2\right) \to 0 \tag{B.2}$$

under the assumptions of Theorem 2.4. First we show that for all M satisfying $M/T \to 0$ we have that

$$\sup_{|u|\le M}\left|\frac{1}{\Delta}\sum_{i=1}^{N}\left(\frac{1}{t_0+u}\sum_{1\le t\le t_0+u}X_{i,t} - \frac{1}{T-(t_0+u)}\sum_{t_0+u<t\le T}X_{i,t}\right)^2 - 1\right| = o_P(1). \tag{B.3}$$

Using (2.1) we have for all $0 < u \le M$

$$\frac{1}{t_0+u}\sum_{1\le t\le t_0+u}X_{i,t} - \frac{1}{T-(t_0+u)}\sum_{t_0+u<t\le T}X_{i,t} = \left(\frac{u}{t_0+u}-1\right)\delta_i + \frac{\gamma_i}{t_0+u}\sum_{1\le t\le t_0+u}\eta_t$$
$$+\frac{1}{t_0+u}\sum_{1\le t\le t_0+u}e_{i,t} - \frac{\gamma_i}{T-(t_0+u)}\sum_{T-(t_0+u)<t\le T}\eta_t - \frac{1}{T-(t_0+u)}\sum_{T-(t_0+u)<t\le T}e_{i,t}.$$

Since $M/T \to 0$, we have

$$\sup_{0<u\le M}\left|\left(\frac{t_0}{t_0+u}\right)^2 - 1\right| \to 0. \tag{B.4}$$

Applying Assumption 2.4 with Markov's inequality and the maximal inequality of Móritz et al. (1982) we obtained for all $z > 0$

$$P\left\{\frac{\Gamma}{T^2\Delta}\sup_{0<u\le M}\left(\sum_{1\le t\le t_0+u}\eta_t\right)^2 \ge z\right\} = P\left\{\sup_{0<u\le M}\left(\sum_{1\le t\le t_0+u}\eta_t\right)^{\bar{\kappa}} \ge (zT^2\Delta/\Gamma)^{\bar{\kappa}/2}\right\}$$
$$\le \left(\frac{\Gamma}{T^2\Delta}\right)^{\bar{\kappa}/2} E\sup_{0<u\le M}\left(\sum_{1\le t\le t_0+u}\eta_t\right)^{\bar{\kappa}}$$
$$= O\left((\Gamma/(T\Delta))^{\bar{\kappa}/2}\right) \to 0 \tag{B.5}$$

on account of Assumption 2.5. Following the proof of (B.5) but now using Assumptions 2.1(i) and 2.3, we conclude

$$P\left\{\frac{1}{T^2\Delta}\sum_{i=1}^{N}\left(\sum_{1\le t\le t_0}e_{i,t}\right)^2 \ge z\right\} \le \frac{1}{zT^2\Delta}\sum_{i=1}^{N}E\left(\sum_{1\le t\le t_0}e_{i,t}\right)^2$$
$$= \frac{1}{z}O\left(N/(T\Delta)\right) \to 0 \tag{B.6}$$

by Assumption 2.5(i). The stationarity in Assumption 2.1(ii) yields

$$
\begin{aligned}
&P\left\{\frac{1}{T^2\Delta}\sup_{0<u\le M}\sum_{i=1}^{N}\left(\sum_{t_0+1\le t\le t_0+u} e_{i,t}\right)^2 \ge z\right\}\\
&\quad = P\left\{\frac{1}{T^2\Delta}\sup_{0<u\le M}\sum_{i=1}^{N}\left(\sum_{1\le t\le u} e_{i,t}\right)^2 \ge z\right\}\\
&\quad \le \sum_{u=1}^{M} P\left\{\frac{1}{T^2\Delta}\sum_{i=1}^{N}\left(\sum_{1\le t\le u} e_{i,t}\right)^2 \ge z\right\}\\
&\quad = \frac{1}{z}O\left(\frac{NM}{T^2\Delta}\right) \to 0 \qquad \textbf{(B.7)}
\end{aligned}
$$

by Assumption 2.5(i) and the assumption that $M/T \to 0$. Putting together (B.6) and (B.7) we obtain that

$$
\sup_{0<u\le M}\frac{1}{\Delta}\sum_{i=1}^{N}\left(\frac{1}{t_0+u}\sum_{1\le t\le t_0+u} e_{i,t}\right)^2 = o_P(1). \qquad \textbf{(B.8)}
$$

Following the proofs of (B.5) and (B.8) one can prove that

$$
\sup_{0<u\le M}\frac{1}{\Delta}\sum_{i=1}^{N}\left(\frac{\gamma_i}{T-(t_0+u)}\sum_{T-(t_0+u)<t\le T} \eta_t\right)^2 = o_P(1) \qquad \textbf{(B.9)}
$$

and

$$
\sup_{0<u\le M}\frac{1}{\Delta}\sum_{i=1}^{N}\left(\frac{1}{T-(t_0+u)}\sum_{T-(t_0+u)<t\le T} e_{i,t}\right)^2 = o_P(1). \qquad \textbf{(B.10)}
$$

The result in (B.3) now follows from (B.4), (B.5), and (B.8)–(B.10). Since under our conditions

$$
|\hat{t}-t_0|/M = o_P(1), \qquad \textbf{(B.11)}
$$

the proof of (4.4) is complete.

By (A.3) we have

$$
U_N\left(\hat{t}+v\right) - U_N\left(\hat{t}\right) = U_N^{(1)}(v) + \ldots + U_N^{(6)}(v),
$$

https://doi.org/10.1017/S0266466615000468 Published online by Cambridge University Press

where

$$U_N^{(1)}(v) = \sum_{i=1}^{N} \delta_i^2 \left(r^2(\hat{t}+v) - r^2(\hat{t}) \right), \quad U_N^{(2)}(v) = \sum_{i=1}^{N} \left(Q_i^2(\hat{t}+v) - Q_i^2(\hat{t}) \right),$$

$$U_N^{(3)}(v) = \sum_{i=1}^{N} \gamma_i^2 \left(V^2(\hat{t}+v) - V^2(\hat{t}) \right),$$

$$U_N^{(4)}(v) = 2\sum_{i=1}^{N} \gamma_i \left(Q_i(\hat{t}+v)V(\hat{t}+v) - Q_i(\hat{t})V(\hat{t}) \right),$$

$$U_N^{(5)}(v) = 2\sum_{i=1}^{N} \delta_i \left(r(\hat{t}+v)Q_i(\hat{t}+v) - r(\hat{t})Q_i(\hat{t}) \right),$$

and

$$U_N^{(6)}(v) = 2\sum_{i=1}^{N} \delta_i \gamma_i \left(r(\hat{t}+v)V(\hat{t}+v) - r(\hat{t})V(\hat{t}) \right).$$

It follows from the proofs of Lemmas A.3, A.5, A.6, and A.9 that for all M satisfying (B.1) we have

$$\sup_{|v|\le M} \frac{1}{|v|\hat{r}_{N,T}^2} \left\{ \left| U_N^{(2)}(v) \right| + \left| U_N^{(3)}(v) \right| + \left| U_N^{(4)}(v) \right| \right\}^2 = o_P\left(\Xi_{N,T} \right),$$

where $\hat{r}_{N,T}$ is defined in (4.3). Applying now Lemmas A.4, A.7, and A.9 we conclude

$$\sup_{|v|\le M} \frac{1}{|v|\hat{r}_{N,T}^2} \left| U_N^{(5)}(v) - U_N^{(7)}(v) \right|^2 = o_P(\Xi_{N,T})$$

and

$$\sup_{|v|\le M} \frac{1}{|v|\hat{r}_{N,T}} \left| U_N^{(6)}(v) - U_N^{(8)}(v) \right|^2 = o_P(\Xi_{N,T}),$$

where

$$U_N^{(7)}(v) = 2r(t_0)\sum_{i=1}^{N} \delta_i \mathcal{A}_i\left(v;\hat{t}\right) \quad \text{and} \quad U_N^{(8)}(v) = 2r(t_0)\sum_{i=1}^{N} \delta_i \gamma_i \mathcal{B}\left(v;\hat{t}\right),$$

$$\mathcal{A}_i(v;t) = \begin{cases} \sum_{s=t+1}^{t+v} e_{i,s}, & v > 0 \\ 0, & v = 0 \\ \sum_{s=t+v}^{t-1} e_{i,s}, & v < 0, \end{cases}$$

and

$$\mathcal{B}(v;t) = \begin{cases} \sum_{s=t+1}^{t+v} \eta_s, & v > 0 \\ 0, & v = 0 \\ \sum_{s=t+v}^{t-1} \eta_s, & v < 0. \end{cases}$$

Using (B.1) and (B.11) one can verify that

$$\sup_{|v|\le M} \frac{1}{|v|\hat{r}_{N,T}} \left| U_N^{(7)}(v) - U_N^{(9)}(v) \right|^2 = o_P(\Xi_{N,T})$$

and

$$\sup_{|v|\le M} \frac{1}{|v|\hat{r}_{N,T}} \left| U_N^{(8)}(v) - U_N^{(10)}(v) \right|^2 = o_P(\Xi_{N,T}),$$

where

$$U_N^{(9)}(v) = 2r(t_0) \sum_{i=1}^{N} \delta_i \mathcal{A}_i(v;t_0) \quad \text{and} \quad U_N^{(10)}(v) = 2r(t_0) \sum_{i=1}^{N} \delta_i \gamma_i \mathcal{B}(v;t_0).$$

Let $m = m(N,T) \le M$ and $m \to \infty$, as $\min(N,T) \to \infty$. Using Assumptions 2.3 and 2.4 we get

$$\sup_{m\le|v|\le M} \left| \frac{1}{4|v|r^2(t_0)} E\left(U_N^{(9)}(v) + U_N^{(10)}(v)\right)^2 - \Xi_{N,T} \right| = o(\Xi_{N,T}).$$

and

$$\sup_{|v|\le m} \frac{1}{4|v|r^2(t_0)\Xi_{N,T}} \left(U_N^{(9)}(v) + U_N^{(10)}(v)\right)^2 = O_P(1). \qquad \textbf{(B.12)}$$

We claim that

$$\sup_{m\le|v|\le M} \left| \frac{1}{4|v|r^2(t_0)\Xi_{N,T}} \left(U_N^{(9)}(v) + U_N^{(10)}(v)\right)^2 - 1 \right| = o_P(1). \qquad \textbf{(B.13)}$$

The statement in (B.13) is a uniform weak law of large numbers, so we can repeat the proof of (B.4) to prove it. Namely, due to stationarity, it follows from Assumptions 2.3 and 2.4 that for every $v \in [-M,\ldots,-m,m,\ldots,M]$ that

$$\frac{1}{4|v|r^2(t_0)\Xi_{N,T}} \left(U_N^{(9)}(v) + U_N^{(10)}(v)\right)^2 \overset{P}{\to} 1. \qquad \textbf{(B.14)}$$

Now (B.13) follows from (B.14) if $\left(U_N^{(9)}(v) + U_N^{(10)}(v)\right)^2 / (|v|r^2(t_0)\Xi_{N,T}), m \le |v| \le M$ is tight. The tightness can be proven along the lines of the proofs of Lemmas A.4 and A.7.

https://doi.org/10.1017/S0266466615000468 Published online by Cambridge University Press

Our arguments show that

$$\sup_{|v|\le M}\left|\frac{1}{4|v|\hat{r}_{N,T}^2}\left(U_N(\hat{t}+v)-U_N(\hat{t})-\Delta_{N,T}\left(r^2(\hat{t}+v)-r^2(\hat{t})\right)\right)^2 - \frac{1}{|v|r(t_0)}\left(U_N^{(9)}(v)+U_N^{(10)}(v)\right)^2\right| = o_P(\Xi_{N,T}). \quad \textbf{(B.15)}$$

Putting together (B.12), (B.13), and (B.15), the result in (4.5) follows.

https://doi.org/10.1017/S0266466615000468 Published online by Cambridge University Press