University of Reading

Structural breaks in panel data: large number of panels and short length time series

Article

Accepted Version

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.

To link to this article DOI: http://dx.doi.org/10.1080/07474938.2018.1454378

Publisher: Taylor & Francis

www.reading.ac.uk/centaur

Structural breaks in panel data:
Large number of panels and short length time series

Jaromír Antoch, Jan Hanousek, Lajos Horváth

Marie Hušková, Shixuan Wang

Affiliation:

JA & MH, Charles University, Department of Probability and Mathematical Statistics, Sokolovská 83, CZ-18600 Praha, Czech Republic;
antoch@karlin.mff.cuni.cz, huskova@karlin.mff.cuni.cz

JH, CERGE-EI, a joint workplace of Charles University and the Economics Institute of the Czech Academy of Sciences, Prague, Czech Republic; and C.E.P.R., London; jan.hanousek@cerge-ei.cz

LH, University of Utah, Department of Mathematics, Salt Lake City, UT 84112–0090 USA; horvath@math.utah.edu

SW, Cardiff Business School, Cardiff University, Cardiff, CF10 3EU, UK; and Department of Economics, University of Birmingham, Birmingham, B15 2TT, UK;
shixuan_wang@hotmail.com

Abstract: The detection of (structural) breaks or the so called change point problem has drawn increasing attention from the theoretical, applied economic and financial fields. Much of the existing research concentrates on the detection of change points and asymptotic properties of their estimators in panels when N, the number of panels, as well as T, the number of observations in each panel are large. In this paper we pursue a different approach, i.e., we consider the asymptotic properties when $N \to \infty$ while keeping T fixed. This situation is typically related to large (firm-level) data containing financial information about an immense number of firms/stocks across a limited number of years/quarters/months. We propose a general approach for testing for break(s) in this setup. In particular, we obtain the asymptotic behavior of test statistics. We also propose a wild bootstrap procedure that could be used to generate the critical values of the test statistics. The theoretical approach is supplemented by numerous simulations and by an empirical illustration. We demonstrate that the testing procedure works well in the framework of the four factors CAPM model. In particular, we estimate the breaks in the monthly returns of US mutual funds during the period January 2006 to February 2010 which covers the subprime crises.

1. Introduction

Structural changes and model stability in panel data are of general concern in empirical economics and finance research. Model parameters are assumed to be stable over time if there is no reason to believe otherwise. It is well-known that various economic and political events can cause structural breaks in financial data. Such events include, for example, the change of parameters associated with the introduction of a single currency in Europe, price liberalization in emerging markets, and world integration of equity markets. Further examples and discussions can be found in Hansen (2001), Andreou and Ghysels (2009) among others. The exact time of change is usually unknown and therefore the construction of tests for stability is a challenging question. Andrews (1993) (cf. corridendum Andrews, 2003) provides the first econometric insights into break detection. Brodsky and Darkhovskii (2000) and Csörgő and Horváth (1997) cover the early development from statistical point of view.

In both the statistics and econometrics literature we can find very many of papers related to the detection of changes and structural breaks. Recent developments in the field consider a general class of time series processes and panel data models with errors that can exhibit temporal dependence, heteroskedasticity, trending variables, possible unit roots, cointegration and long memory processes. For a literature survey we refer the reader to Arellano (2004), Perron (2006) and Aue and Horváth (2012).

A model for panel data with possible breaks in the mean was introduced by Joseph and Wolfson (1992, 1993). Bai (2010) significantly generalized the earlier results, still assuming that the panels are independent of each other, thanks to allowing serial correlation within each panel. This approach was extended by Horváth and Hušková (2012), who discussed dependence between the panels via common factors. Results on breaks in the means and/or in the variances have been extended to various forms of linear regression by Kim (2011), Baltagi et al. (2016) and Kao et al. (2012, 2014), among others.

Existing research mostly concentrates on testing for structural breaks in the panel data when T, the number of observations in each panel, goes to infinity and N, the number of panels/firms/assets/individuals, also goes to infinity, but slower than N (Bai, 2010, Bai and Carrion-i-Silvestre, 2009, Kim, 2011, Baltagi at al., 2016).

In this paper we study a different asymptotic case when T is finite and N goes to infinity. This situation is typically related to large firm-level data panels containing financial information about a very substantial number of firms/stocks across a limited number of years/quarters. The issue of model stability in such a setup applies to broadly defined research in empirical asset pricing as well as in corporate finance. For example, asset pricing models hinge heavily on the stability of factor parameters to explain stock returns, while the financial industry eyes the generation of alpha returns (Fama and French, 1993 and 2014, Ghysels, 1998, Guercio and Reuter, 2014).

Similarly, in corporate finance the leverage literature relies on several parameters to estimate target leverage ratios, and on the stability of fixed effects to show persistence in capital structure choices (DeAngelo and Roll, 2015, Lemmon at al., 2008). Pešta and Peštová (2015) use this setup for an application to insurance mathematics, while Torgovitski (2015) focuses on multiple changes in spatial functional data. Cho and Fryzlewicz (2015) and Jirák (2015) study estimates and tests for multiple change points in high dimensional data.

There is evidence of structural breaks in the financial market that affect financial indicators, including returns and volatility, changes in exchange rate regimes, asset allocations, equity premium, value at risk (VaR), expected shortfall (ES), credit risk models,

etc. Recently, researchers have started to analyze structural breaks in the panel data setup, which reveal an increasing variety of potential applications with a strong impact in financial economics (cf., for example Antoch and Jarušková, 2017, DeAngelo and Roll, 2015, Graham et al, 2015, Whited and Wu, 2006). Hence the stability of parameter estimates in panel data, and the estimation of the time of change where T is fixed and N goes to infinity have interesting applications.

In this paper we consider the model

$$(1.1) \qquad y_{i,t} = \mathbf{x}_{i,t}^\top (\boldsymbol{\beta}_i + \boldsymbol{\delta}_i I\{t \geq t_0\}) + e_{i,t}, \quad 1 \leq i \leq N \text{ and } 1 \leq t \leq T,$$

where $\mathbf{x}_{i,t} = \left(x_{i,t}(1), \ldots, x_{i,t}(d)\right)^\top$ are vectors of explanatory variables (design points that can be random or non-random), $\boldsymbol{\beta}_i \in R^d$ are unknown regression (vector) parameters in the ith panel and $^\top$ denotes the transpose of vectors and matrices. Finally, we assume that $e_{i,t}$ are errors that can be cross-sectionally correlated. As previously mentioned, we consider the detection of the unknown time of change. The regression parameters in the ith panel change from $\boldsymbol{\beta}_i$ to $\boldsymbol{\beta}_i + \boldsymbol{\delta}_i$ at unknown time t_0, usually called either a *change point* or a *break point*.

The paper is structured as follows. Section 2 introduces our notation, settings and lists the main assumptions. Section 3 contains the main results (detailed proofs are provided in the technical appendix). Section 4 describes the (wild) bootstrap procedure for one of the proposed test statistics. We use Monte Carlo simulations to illustrate the finite sample properties of the test. Section 5 contains an application to real data. Finally, Section 6 concludes the paper. Proofs are presented in Section 7.

2. Notation, settings and general assumptions

First we introduce some notation that will be used through the paper. Since we are dealing with detection and estimation of an unknown change point, we will analyze various estimators, matrices, and parameters before and after a certain time t, $1 \leq t \leq T$.

In particular, $\widehat{\boldsymbol{\beta}}_{i,t}$ and $\widetilde{\boldsymbol{\beta}}_{i,t}$ denote the OLS estimators computed from the first t and the last $T - t$ observations of the ith panel, respectively. Hence,

$$(2.1) \qquad \widehat{\boldsymbol{\beta}}_{i,t} = \left(\mathbf{X}_{i,t}^\top \mathbf{X}_{i,t}\right)^{-1} \mathbf{X}_{i,t}^\top \mathbf{Y}_{i,t}, \quad t_1 \leq t \leq T,$$

$$(2.2) \qquad \widetilde{\boldsymbol{\beta}}_{i,t} = \left(\widetilde{\mathbf{X}}_{i,t}^\top \widetilde{\mathbf{X}}_{i,t}\right)^{-1} \widetilde{\mathbf{X}}_{i,t}^\top \widetilde{\mathbf{Y}}_{i,t}, \quad 1 \leq t \leq T - t_2,$$

where $\mathbf{Y}_{i,t} = (y_{i,1}, y_{i,2}, \ldots, y_{i,t})^\top$, $\widetilde{\mathbf{Y}}_{i,t} = (y_{i,t+1}, y_{i,t+2}, \ldots, y_{i,T})^\top$,

$$\mathbf{X}_{i,t} = \begin{pmatrix} x_{i,1}(1), x_{i,1}(2), \ldots, x_{i,1}(d) \\ x_{i,2}(1), x_{i,2}(2), \ldots, x_{i,2}(d) \\ \vdots \\ x_{i,t}(1), x_{i,t}(2), \ldots, x_{i,t}(d) \end{pmatrix} = \begin{pmatrix} \mathbf{x}_{i,1}^\top \\ \mathbf{x}_{i,2}^\top \\ \vdots \\ \mathbf{x}_{i,t}^\top \end{pmatrix}$$

and

$$\widetilde{\mathbf{X}}_{i,t} = \begin{pmatrix} x_{i,t+1}(1), x_{i,t+1}(2), \ldots, x_{i,t+1}(d) \\ x_{i,t+2}(1), x_{i,t+2}(2), \ldots, x_{i,t+2}(d) \\ \vdots \\ x_{i,T}(1), \ x_{i,T}(2), \ldots, \quad x_{i,T}(d) \end{pmatrix} = \begin{pmatrix} \mathbf{x}_{i,t+1}^\top \\ \mathbf{x}_{i,t+2}^\top \\ \vdots \\ \mathbf{x}_{i,T}^\top \end{pmatrix}.$$

We also denote by "tr" the trace of a matrix and put

$$\mathbf{Z}_{i,t} = \mathbf{X}_{i,t}^\top \mathbf{X}_{i,t}, \quad \widetilde{\mathbf{Z}}_{i,t} = \widetilde{\mathbf{X}}_{i,t}^\top \widetilde{\mathbf{X}}_{i,t}, \quad t = 1, \ldots, T.$$

2.1. **Test statistics.** We propose several test statistics for the change point problem in the panel data when T is fixed and $N \to \infty$. We are interested in testing if there is a change in the regression parameters. We introduce two processes and their functionals can be used as test statistics.

Under the null hypothesis there is no change in the coefficients, i.e., the null hypothesis is stated as

$$(2.3) \qquad\qquad H_0 : \sum_{i=1}^{N} \|\boldsymbol{\delta}_i\| = 0,$$

where $\| \cdot \|$ denotes the Euclidean norm vectors and matrices.

We consider two types of test procedures; the first is based on quadratic forms of least squares estimators and the second employs sums of squares of residuals. First, let us consider

$$(2.4) \quad U_N(t) = \sum_{i=1}^{N} (\widehat{\boldsymbol{\beta}}_{i,t} - \widehat{\boldsymbol{\beta}}_{i,T})^\top \mathbf{C}_{i,t} (\widehat{\boldsymbol{\beta}}_{i,t} - \widehat{\boldsymbol{\beta}}_{i,T}), \quad \max\{d, t_1\} \le t \le \min\{T - d, T - t_2\}.$$

$U_N(t)$ is a Wald-type test statistic (with weighting matrices $\boldsymbol{C}_{i,t}$'s) for each t to compare the least squares estimators computed from the first t observations with the estimator from the full sample. To keep the detection procedure broad we do not restrict $\boldsymbol{C}_{i,t}$ to a particular form. Instead we list a set of assumptions that $\boldsymbol{C}_{i,t}$ should satisfy and we will discuss pros and cons of particular choices of $\boldsymbol{C}_{i,t}$ later.

The second test procedure is based on the process

$$(2.5) \qquad V_N(t) = \sum_{i=1}^{N} \sum_{s=1}^{t} \widehat{e}_{i,s}^2, \quad \widehat{e}_{i,s} = y_{i,s} - \mathbf{x}_{i,s}^\top \widehat{\boldsymbol{\beta}}_{i,T}, \quad 1 \le i \le N, \ 1 \le t \le T,$$

representing CUSUM-type statistics based on partial sums of the squares of the residuals.

In general, large values of $U_N(t)$ and/or $V_N(t)$ for at least some t indicate that the null hypothesis is violated. Required critical values can be obtained either asymptotically or through a proper version of the bootstrap.

2.2. **Assumptions.** This subsection contains a set of assumptions needed to prove the main theorems describing asymptotic properties of the suggested test statistics. We start with assumptions about the error terms and assume that they can be decomposed as

$$(2.6) \qquad\qquad e_{i,t} = \boldsymbol{\gamma}_i^\top \boldsymbol{\Lambda}_t + \varepsilon_{i,t},$$

where $\boldsymbol{\Lambda}_t \in R^q, 1 \le t \le T$ is the common factor, $\varepsilon_{i,t}, \ 1 \le i \le N, \ 1 \le t \le T$ are the innovations and $\boldsymbol{\gamma}_i \in R^q, 1 \le i \le N$, is loading in the ith panel satisfying:

Assumption 2.1. $E \boldsymbol{\Lambda}_t = \mathbf{0} \quad$ and $\quad E \|\boldsymbol{\Lambda}_t\|^2 < \infty, \ 1 \le t \le T.$

Assumption 2.2. (i) *The innovations* $\boldsymbol{\varepsilon}_i = (\varepsilon_{i,1}, \varepsilon_{i,2}, \ldots, \varepsilon_{i,T})^\top, \ i = 1, \ldots, N$ *are independent;*
(ii) $E \varepsilon_{i,t} = 0, \ E \varepsilon_{i,t}\varepsilon_{i,s} = 0$ *and* $c_3 \le \sigma_i^2 = E \varepsilon_{i,t}^2 \le c_4$ *for all* $1 \le i \le N, 1 \le t \ne s \le T$ *with some constants* $0 < c_3 \le c_4 < \infty;$

(iii) *there is $\kappa > 4$ such that* $\limsup_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} E |\varepsilon_{i,t}|^{\kappa} < \infty, \ 1 \leq t \leq T.$

Note that Assumption 2.2 (ii) is usually satisfied in financial applications, since ARCH, GARCH and similar volatility models assume uncorrelated innovations (cf. Francq and Zakoian (2010)).

Assumption 2.3. $\{\boldsymbol{\Lambda}_t, 1 \leq t \leq T\}$ *and* $\{\boldsymbol{\varepsilon}_i, 1 \leq i \leq N\}$ *are independent.*

Let us note that serial correlation of the error disturbances in the original model (1.1) can be obtained via common factors defined in error term decomposition (2.6). The serial correlation observed in $e_{i,t}$ can then be viewed as impure correlation caused by omitting common factors in the main regression model (1.1).

Assumption 2.4.

$$\lim_{N \to \infty} \frac{1}{N^{1/2}} \sum_{i=1}^{N} \|\boldsymbol{\gamma}_i\|^2 = 0.$$

As usual we assume that the common factors are negligible. Inference with non-negligible common factors in change-point detection has been mentioned by several authors in different setups when both N and T go to infinity but N is much smaller than T (cf. Horváth and Hušková, 2012 and Baltagi et al., 2016). We discuss the behavior of $U_N(t)$ and $V_N(t)$ when Assumption 2.4 fails in Section 3.4.

Now we continue with assumptions on the design matrix.

Assumption 2.5. *There is c_0 such that* $\|\mathbf{x}_{i,t}\| \leq c_0$ *for all* $1 \leq i \leq N, \ 1 \leq t \leq T.$

Assumption 2.6. (i) *There are positive constants t_1 and c_1 such that* $\left(\sum_{j=1}^{t_1} \mathbf{x}_{i,j} \mathbf{x}_{i,j}^{\top} \right)^{-1}$ *exists and* $\|\left(\sum_{j=1}^{t_1} \mathbf{x}_{i,j} \mathbf{x}_{i,j}^{\top} \right)^{-1}\| \leq c_1$ *for all i,* $1 \leq i \leq N$;
(ii) *there are positive constants t_2 and c_2 such that* $\left(\sum_{j=T-t_2}^{T} \mathbf{x}_{i,j} \mathbf{x}_{i,j}^{\top} \right)^{-1}$ *exists and* $\|\left(\sum_{j=T-t_2}^{T} \mathbf{x}_{i,j} \mathbf{x}_{i,j}^{\top} \right)^{-1}\| \leq c_2$ *for all i,* $1 \leq i \leq N$.

Notice that Assumption 2.6 (i) implies $\|\left(\sum_{j=1}^{t} \mathbf{x}_{i,j} \mathbf{x}_{i,j}^{\top} \right)^{-1}\| \leq c_1$ for all $t \geq t_1$ and that similarly to Assumption 2.6 (ii) we find that $\|\left(\sum_{j=t+1}^{T} \mathbf{x}_{i,j} \mathbf{x}_{i,j}^{\top} \right)^{-1}\| \leq c_2$ for all $t \leq T - t_2$. Assumption 2.6 (i) guaranties that the OLS estimators in (2.1) and (2.2) are uniquely defined.

For the sake of simplicity we assume that the $\boldsymbol{x}_{i,t}$'s are non-random. We briefly discuss modifications of Assumptions 2.5 and 2.6 to allow random design vectors in Section 3.5.

Finally, we present our assumptions on weighting matrices.

Assumption 2.7. $\mathbf{C}_{i,t} = \mathbf{C}_{i,t}^{\top}$, *it is non–negative definite for all* $1 \leq i \leq N, d \leq t \leq T-d$, *and there exist constants c_5 and t_3, t_4 such that* $\|\mathbf{C}_{i,t}\| \leq c_5$ *for all* $1 \leq i \leq N, t_3 \leq t \leq T - t_4$.

Instead of a particular choice of weighting matrixes, we consider a general condition which would guarantee the validity of theorems and testing procedures specified later. Nevertheless, by a specific choice of $\mathbf{C}_{i,t}$ we can obtain variants/generalizations of standard tests. In particular, by choosing

$$(2.7) \qquad \qquad \mathbf{C}_{i,t} = \mathbf{Z}_{i,T}^{-1} \widetilde{\mathbf{Z}}_{i,t} \boldsymbol{D}_{i,t} \widetilde{\mathbf{Z}}_{i,t} \mathbf{Z}_{i,T}^{-1},$$

where $\boldsymbol{D}_{i,t}$ are positive semidefinite matrices and $\widehat{\boldsymbol{\beta}}_{i,t}$ and $\widetilde{\boldsymbol{\beta}}_{i,t}$ are defined by (2.1) and (2.2), we have that

$$(2.8) \qquad U_{1,N}(t) \equiv U_{1,N}(t; \boldsymbol{D}_{i,t}) = U_N(t; \mathbf{Z}_{i,T}^{-1} \widetilde{\mathbf{Z}}_{i,t} \boldsymbol{D}_{i,t} \widetilde{\mathbf{Z}}_{i,t} \mathbf{Z}_{i,T}^{-1})$$

$$= \sum_{i=1}^{N} (\widehat{\boldsymbol{\beta}}_{i,t} - \widetilde{\boldsymbol{\beta}}_{i,t})^{\top} \boldsymbol{D}_{i,t} (\widehat{\boldsymbol{\beta}}_{i,t} - \widetilde{\boldsymbol{\beta}}_{i,t}).$$

Another useful choice of the weighting matrices $\boldsymbol{C}_{i,t}$ is $\mathbf{C}_{i,t} = \mathbf{Z}_{i,t} \boldsymbol{G}_{i,t} \mathbf{Z}_{i,t}$ gives

$$(2.9) \qquad U_{2,N}(t) = U_N(t; \mathbf{Z}_{i,t} \boldsymbol{G}_{i,t} \mathbf{Z}_{i,t})$$

$$= \sum_{i=1}^{N} \left(\sum_{v=1}^{t} \mathbf{x}_{i,v} (y_{i,v} - \mathbf{x}_{i,v}^{\top} \widehat{\boldsymbol{\beta}}_{i,T}) \right)^{\top} \boldsymbol{G}_{i,t} \left(\sum_{v=1}^{t} \mathbf{x}_{i,v} (y_{i,v} - \mathbf{x}_{i,v}^{\top} \widehat{\boldsymbol{\beta}}_{i,T}) \right),$$

where $\boldsymbol{G}_{i,t}$ is a positive semidefinite weighting matrix.

We note that $U_{2,N}(t)$ requires only the existence of $\mathbf{Z}_{i,T}^{-1}$. Also, $U_{2,N}(t)$ can be interpreted as a weighted version of $V_N(t)$.

3. Main results

3.1. **Results under H_0.** First we present the main results on the limit behavior of $U_N(t)$ and $V_N(t)$ under the null hypothesis.

Define

$$(3.1) \quad a_{i,t,1}^2 = \sigma_i^2 \left(\sum_{s=1}^{t} \mathbf{x}_{i,s}^{\top} (\mathbf{Z}_{i,t}^{-1} - \mathbf{Z}_{i,T}^{-1}) \mathbf{C}_{i,t} (\mathbf{Z}_{i,t}^{-1} - \mathbf{Z}_{i,T}^{-1}) \mathbf{x}_{i,s} + \sum_{s=t+1}^{T} \mathbf{x}_{i,s}^{\top} \mathbf{Z}_{i,T}^{-1} \mathbf{C}_{i,t} \mathbf{Z}_{i,T}^{-1} \mathbf{x}_{i,s} \right)$$

$$= \sigma_i^2 \left\{ \mathrm{tr} \left((\mathbf{Z}_{i,t}^{-1} - \mathbf{Z}_{i,T}^{-1}) \mathbf{C}_{i,t} (\mathbf{Z}_{i,t}^{-1} - \mathbf{Z}_{i,T}^{-1}) \mathbf{Z}_{i,t} \right) + \mathrm{tr} \left(\mathbf{Z}_{i,T}^{-1} \mathbf{C}_{i,t} \mathbf{Z}_{i,T}^{-1} \widetilde{\mathbf{Z}}_{i,T-t} \right) \right\}$$

$$= \sigma_i^2 \, \mathrm{tr} \left(\mathbf{C}_{i,t} (\mathbf{Z}_{i,t}^{-1} - \mathbf{Z}_{i,T}^{-1}) \right),$$

$$(3.2) \qquad A_N^{(1)}(t) = \sum_{1 \le i \le N} a_{i,t,1}^2,$$

$$(3.3) \qquad a_{i,t,2}^2 = \sigma_i^2 \sum_{s=1}^{t} \left(1 - \mathbf{x}_{i,s}^{\top} \mathbf{Z}_{i,T}^{-1} \mathbf{x}_{i,s} \right) = \sigma_i^2 \left(t - \mathrm{tr} \left(\mathbf{Z}_{i,t} \mathbf{Z}_{i,T}^{-1} \right) \right),$$

$$(3.4) \qquad A_N^{(2)}(t) = \sum_{i=1}^{N} a_{i,t,2}^2,$$

$$(3.5) \qquad \mathbf{S}_{i,t} = \sum_{s=1}^{t} \mathbf{x}_{i,s} \varepsilon_{i,s}.$$

To state the asymptotic distributions of $U_N(t)$ and $V_N(t)$ we need the existence of their asymptotic covariance functions. This is specified in Assumptions 3.1 and 3.2.

Assumption 3.1. *The function*

$$\Gamma^{(1)}(t, t') = \lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} \left[E \left\{ (\mathbf{Z}_{i,t}^{-1} \mathbf{S}_{i,t} - \mathbf{Z}_{i,T}^{-1} \mathbf{S}_{i,T})^{\top} \mathbf{C}_{i,t} (\mathbf{Z}_{i,t}^{-1} \mathbf{S}_{i,t} - \mathbf{Z}_{i,T}^{-1} \mathbf{S}_{i,T}) \right. \right.$$

$$\left. \left. \times (\mathbf{Z}_{i,t'}^{-1} \mathbf{S}_{i,t'} - \mathbf{Z}_{i,T}^{-1} \mathbf{S}_{i,T})^{\top} \mathbf{C}_{i,t'} (\mathbf{Z}_{i,t'}^{-1} \mathbf{S}_{i,t'} - \mathbf{Z}_{i,T}^{-1} \mathbf{S}_{i,T}) \right\} - a_{i,t,1}^2 a_{i,t',1}^2 \right]$$

exists for all $\underline{t}_0 \leq t, t' \leq T - \bar{t}_0$, *where* $\underline{t}_0 = \max(d, t_1, t_3)$ *and* $\bar{t}_0 = max(d, t_2, t_4)$.

Assumption 3.2. *The function*

$$\Gamma^{(2)}(t,t') = \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \left[\sum_{s=1}^{t} \sum_{s'=1}^{t'} E\left\{ (\varepsilon_{i,s} - \mathbf{x}_{i,s}^\top \mathbf{Z}_{i,T}^{-1} \mathbf{S}_{i,T})(\varepsilon_{i,s'} - \mathbf{x}_{i,s'}^\top \mathbf{Z}_{i,T}^{-1} \mathbf{S}_{i,T}) \right\}^2 - a_{i,t,2}^2 a_{i,t',2}^2 \right]$$

exists for all $1 \leq t, t' \leq T$.

Using these assumptions, we can state two theorems on asymptotic distribution of the test statistics $U_N(t)$ and $V_N(t)$ under the null hypothesis.

Theorem 3.1. *If H_0, Assumptions 2.1–2.7 and 3.1 hold, then we have that, as $N \to \infty$,*

$$(3.6) \qquad \left\{ N^{-1/2} \left(U_N(t) - A_N^{(1)}(t) \right), \ \underline{t}_0 \leq t \leq T - \bar{t}_0 \right\} \overset{\mathcal{D}}{\to} \left\{ \xi_t^{(1)}, \underline{t}_0 \leq t \leq T - \bar{t}_0 \right\},$$

where $\xi_t^{(1)}$, $\underline{t}_0 \leq t \leq T - \bar{t}_0$, is jointly normal with zero mean, covariance $\Gamma^{(1)}(t,t')$ and $\underline{t}_0, \bar{t}_0$ are defined in Assumption 3.1.

Theorem 3.2. *If H_0, Assumptions 2.1 – 2.6 and 3.2 hold, then we have that, as $N \to \infty$,*

$$(3.7) \qquad \left\{ N^{-1/2} \left(V_N(t) - A_N^{(2)}(t) \right), \ 1 \leq t \leq T \right\} \overset{\mathcal{D}}{\to} \left\{ \xi_t^{(2)}, 1 \leq t \leq T \right\},$$

where $\xi_t^{(2)}, 1 \leq t \leq T$, is jointly normal with zero mean and covariance $\Gamma^{(2)}(t,t')$ defined in Assumption 3.2.

Proofs of both theorems are given in Section 7.

Theorem 3.1 and Theorem 3.2 can be used to find the asymptotic distributions of the test statistics based on $U_N(t)$ and $V_N(t)$ when H_0 holds.

3.2. **Results under alternatives.** We briefly discuss the behavior of the proposed test statistics when a change occurs. We focus on the statistics based on $U_N(t)$, $\underline{t}_0 \leq t \leq \bar{t}_0$, since the behavior of tests based on $V_N(t)$, $1 \leq t \leq T$, is similar.

We obtain asymptotically consistent tests of rejecting H_0 for large values of $N^{-1/2} \max_t |U_N(t) - A_N^{(1)}(t)|$, if some of the changes $\boldsymbol{\delta}_i$'s, the sizes of the changes, are sufficiently large. Particularly, if Assumptions 2.1 – 2.7 are satisfied and for some $\underline{t}_0 \leq t, t_1 \leq T - \bar{t}_0$

$$(3.8) \qquad \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \boldsymbol{\delta}_i^\top \left(\boldsymbol{Z}_{it_1}(\boldsymbol{Z}_{i,\max(t_1,t)}^{-1} - \boldsymbol{Z}_{iT}^{-1}) \boldsymbol{C}_{it}(\boldsymbol{Z}_{i,\max(t_1,t)} - \boldsymbol{Z}_{iT}^{-1}) \boldsymbol{Z}_{it_1} \right) \boldsymbol{\delta}_i \to \infty$$

then we have that

$$(3.9) \qquad N^{-1/2} \max_t |U_N(t) - A_N^{(1)}(t)| \to^P \infty,$$

which together with Theorem 3.1 ensures asymptotic consistency of the test. One can show that (3.9) holds under other type of alternatives including multiple changes, gradual changes, or when larger changes occur only in a fraction of panels.

To find an estimator for t_0 we choose $\boldsymbol{C}_{i,t}$ such that for all t and i

$$\boldsymbol{Z}_{i,t}^{-1} \boldsymbol{C}_{i,t} \boldsymbol{Z}_{i,t}^{-1} = \boldsymbol{G}_T$$

with some positive definite matrix $\boldsymbol{G}_T$. Then under the alternative hypothesis the behavior of $|U_N(t) - A_N^{(1)}(t)|$ is asymptotically equivalent with

$$\frac{1}{\sqrt{N}} \sum_{i=1}^{N} \boldsymbol{\delta}_i^\top \left(\boldsymbol{Z}_{i,t_0} \boldsymbol{Z}_{i,T}^{-1} \widetilde{\boldsymbol{Z}}_{i,t} \boldsymbol{G}_T \widetilde{\boldsymbol{Z}}_{i,t} \boldsymbol{Z}_{i,T}^{-1} \boldsymbol{Z}_{i,t_0} \right) \boldsymbol{\delta}_i, \quad t \geq t_0,$$

$$\frac{1}{\sqrt{N}} \sum_{i=1}^{N} \boldsymbol{\delta}_i^\top \left(\widetilde{\boldsymbol{Z}}_{i,t_0} \boldsymbol{Z}_{i,T}^{-1} \boldsymbol{Z}_{i,t} \boldsymbol{G}_T \boldsymbol{Z}_{i,t} \boldsymbol{Z}_{i,T}^{-1} \widetilde{\boldsymbol{Z}}_{i,t_0} \right) \boldsymbol{\delta}_i, \quad t \leq t_0.$$

Note that the elements of the sums defined above are non-decreasing for $t \leq t_0$, while they are non-increasing for $t \geq t_0$. Then the condition for the asymptotic consistency (3.8) can be replaced by

$$\frac{1}{\sqrt{N}} \sum_{i=1}^{N} \boldsymbol{\delta}_i^\top \left(\boldsymbol{Z}_{it_0} \boldsymbol{Z}_{i,T}^{-1} \widetilde{\boldsymbol{Z}}_{it_0} \boldsymbol{G}_T \widetilde{\boldsymbol{Z}}_{i,t_0} \boldsymbol{Z}_{i,T}^{-1} \boldsymbol{Z}_{it_0} \right) \boldsymbol{\delta}_i \to \infty,$$

and an estimator $\widehat{t}_N$ of t_0 can be defined as

$$\widehat{t}_N = \operatorname*{argmax}_{t=d,\dots,T-d} \left| U_N(t) - A_N^{(1)}(t) \right|.$$

The asymptotic properties of $\widehat{t}_N$, as $N \to \infty$, are the subject of future research.

3.3. **Remarks.** The test procedures studied in the previous section depend on some unknown quantities. We briefly discuss their possible estimators.

The normalizing sequences $A_N^{(1)}(t)$ and $A_N^{(2)}(t)$ in Theorem 3.1 and 3.2 depend on the unknown variances σ_i^2, $i = 1, \dots, N$. To estimate them we can employ the following options:

(1) σ_i^2 is estimated by the sample variance of the ith panel

$$(3.10) \qquad \widehat{\sigma}_i^2 = \frac{1}{T-d} \sum_{t=1}^{T} \left(y_{i,t} - \boldsymbol{x}_{i,t}^\top \widehat{\boldsymbol{\beta}}_{i,T} \right)^2;$$

(2) σ_i^2 is estimated by

$$(3.11) \qquad \widehat{\widehat{\sigma}}_i^2 = \min_{d<v<T-d} \frac{1}{T-2d} \left\{ \sum_{t=1}^{v} \left(y_{i,t} - \boldsymbol{x}_{i,t}^\top \widehat{\boldsymbol{\beta}}_{i,v} \right)^2 + \sum_{t=v+1}^{T} \left(y_{i,t} - \boldsymbol{x}_{i,t}^\top \widetilde{\boldsymbol{\beta}}_{i,v} \right)^2 \right\}.$$

The estimator $\widehat{\widehat{\sigma}}_i^2$ is asymptotically consistent under the one change point alternative. However, it is not stable since it includes $\widehat{\boldsymbol{\beta}}_{i,v}$ and $\widetilde{\boldsymbol{\beta}}_{i,v}$ when they are computed from a few observations. Hence in the simulations and the data example we use $\widehat{\sigma}_i^2$.

The covariances $\Gamma^{(1)}$ and $\Gamma^{(2)}$ in Assumptions 3.1 and 3.2 are complicated functions of the observations. Nevertheless, they can be simplified for certain choices of the weighting matrices $\mathbf{C}_{i,t}$ and moderate values of T. It follows from Assumptions 2.2 and 2.5 that

$$(3.12) \qquad \frac{1}{N} \sum_{i=1}^{N} E \left\| \sum_{v=s}^{t} \mathbf{x}_{i,v} \varepsilon_{i,v} \right\|^4 \leq c_{2,1} (t - s + 1)^2$$

and

$$\left\| \sum_{v=s}^{t} \mathbf{x}_{i,v} \mathbf{x}_{i,v}^{\top} \right\| \leq c_{2,2}(t - s + 1)$$

where $c_{2,1}$ and $c_{2,2}$ are some positive numbers. We note that if the $\mathbf{x}_{i,t}$'s realizations of stationary, ergodic sequences with finite mean, then

$$\frac{1}{T}\mathbf{Z}_{i,T} \approx \mathbf{Z}_i,$$

assuming that $\mathbf{Z}_i$ is nonsingular, we arrive at the natural conclusion that

$$(3.13) \qquad \qquad \|\mathbf{Z}_{i,T}^{-1}\| \leq c_{2,3}/T.$$

Let us assume that $\mathbf{C}_{i,t} = \mathbf{Z}_{i,t}^2$ in the definition of $U_N(t)$. In this case

$$\Gamma^{(1)}(t, t') = \lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} E \left\{ (\mathbf{S}_{i,t} - \mathbf{Z}_{i,t}\mathbf{Z}_{i,T}^{-1}\mathbf{S}_{i,T})^{\top}(\mathbf{S}_{i,t} - \mathbf{Z}_{i,t}\mathbf{Z}_{i,T}^{-1}\mathbf{S}_{i,T}) \right.$$

$$\left. \times (\mathbf{S}_{i,t'} - \mathbf{Z}_{i,t'}\mathbf{Z}_{i,T}^{-1}\mathbf{S}_{i,T})^{\top}(\mathbf{S}_{i,t'} - \mathbf{Z}_{i,t'}\mathbf{Z}_{i,T}^{-1}\mathbf{S}_{i,T}) - a_{i,t,1}^2 a_{i,t',1}^2 \right\},$$

and therefore by (3.12) and (3.13) we have for moderate and large T's that

$$\Gamma^{(1)}(t, t') \approx \lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} E \left\{ \left(\mathbf{S}_{i,t} - \frac{t}{T}\mathbf{S}_{i,T}\right)^{\top} \left(\mathbf{S}_{i,t} - \frac{t}{T}\mathbf{S}_{i,T}\right) \right.$$

$$\left. \times \left(\mathbf{S}_{i,t'} - \frac{t}{T}\mathbf{S}_{i,T}\right)^{\top} \left(\mathbf{S}_{i,t'} - \frac{t'}{T}\mathbf{S}_{i,T}\right) - a_{i,t,1}^2 a_{i,t',1}^2 \right\}.$$

The approximation for $\Gamma^{(1)}(t, t')$ also shows that $U_N(t)$ is related to CUSUM processes used extensively in change point analysis.

Under Assumption 2.4 (negligible influence of common factors) the residual $\widehat{e}_{i,t} = y_{i,t} - \mathbf{x}_{i,t}\widehat{\boldsymbol{\beta}}_{i,T}$ is a good approximation for the innovation $\varepsilon_{i,t}$. Hence, under H_0, we could use

$$\widehat{\Gamma}_N^{(1)}(t, t') = \frac{1}{N} \sum_{i=1}^{N} \left\{ \widehat{\mathbf{S}}_{i,t}^{\top}\widehat{\mathbf{S}}_{i,t}\widehat{\mathbf{S}}_{i,t'}^{\top}\widehat{\mathbf{S}}_{i,t'} - \widehat{a}_{i,t,1}^2\widehat{a}_{i,t',1}^2 \right\},$$

where

$$\widehat{\mathbf{S}}_{i,t} = \sum_{v=1}^{t} \mathbf{x}_{i,v}\widehat{e}_{i,v}, \quad \widehat{a}_{i,t,1}^2 = \widehat{\sigma}_{i,T}^2 \operatorname{tr}\left(\mathbf{Z}_{i,t}\mathbf{Z}_{i,T}^{-1}\widetilde{\mathbf{Z}}_{i,t}\right) \quad \text{and} \quad \widehat{\sigma}_{i,T}^2 = \frac{1}{T} \sum_{v=1}^{T} \widehat{e}_{i,v}^2,$$

to approximate the covariance matrix $\Gamma^{(1)}(t, t')$, since $\widehat{\mathbf{S}}_{i,T} = \mathbf{0}$.

In some models commonly used in econometric finance, e.g., CAPM type models, $\boldsymbol{x}_{i,t}$'s do not depend on i. The model and also the test statistics become much simpler, because the indices i in $\boldsymbol{x}_{it}$'s and $\boldsymbol{Z}_{it}$'s can be omitted and our weighting matrix also becomes independent on i, i.e., $\boldsymbol{C}_{it} = \boldsymbol{C}_t$, $i = 1, \ldots, N$. Hence we have

$$A_N^{(1)}(t) = \operatorname{tr}(\boldsymbol{C}_t(\boldsymbol{Z}_t^{-1} - \boldsymbol{Z}_T^{-1})) \sum_{i=1}^{N} \sigma_i^2 \quad \text{and} \quad A_N^{(2)}(t) = (t - \operatorname{tr}(\boldsymbol{Z}_t\boldsymbol{Z}_T^{-1})) \sum_{i=1}^{N} \sigma_i^2.$$

Then it suffices to estimate $\sum_{i=1}^{N} \sigma_i^2$ instead of a set of single σ_i^2's, $i = 1, \ldots, N$.

3.4. **Non-negligible common factors.** In this section we show that if Assumption 2.4 (negligible common factors) fails, then $N^{-1/2}|U_N(t) - A_N^{(1)}(t)| \to \infty$ and $N^{-1/2}|V_N(t) - A_N^{(2)}(t)| \to \infty$ in probability for some t. These are immediate consequences of Theorem 3.3 and 3.4 stated below.

Assumption 3.3.

$$\frac{1}{r_N} \sum_{i=1}^{N} \|\boldsymbol{\gamma}_i\|^2 = O(1) \text{ with some numerical sequence satisfying } r_N/N^{1/2} \to \infty.$$

Assumption 3.4. *The limit*

$$\boldsymbol{Q}(s, v, z) = \lim_{N \to \infty} \frac{1}{r_N} \sum_{i=1}^{N} \boldsymbol{\gamma}_i \mathbf{x}_{i,s}^{\top} \left(\mathbf{Z}_{i,z}^{-1} - \mathbf{Z}_{i,T}^{-1} \right) \mathbf{C}_{i,z} \left(\mathbf{Z}_{i,z}^{-1} - \mathbf{Z}_{i,T}^{-1} \right) \mathbf{x}_{i,v} \boldsymbol{\gamma}_i^{\top}, \ 1 \le s, v \le z \le T,$$

exists and there are s_0, v_0 and z_0 such that $1 \le s_0, v_0 \le z_0 \le T$, and at least one element of $\boldsymbol{Q}(s_0, v_0, z_0)$ is different from 0.

Assumption 3.5. *The limit*

$$\overline{\boldsymbol{Q}}(s, v, z) = \lim_{N \to \infty} \frac{1}{r_N} \sum_{i=1}^{N} \left(\frac{1}{T} \boldsymbol{\gamma}_i - \boldsymbol{\gamma}_i \mathbf{x}_{i,s}^{\top} \mathbf{Z}_{i,T}^{-1} \mathbf{x}_{i,z} \right) \left(\frac{1}{T} \boldsymbol{\gamma}_i - \boldsymbol{\gamma}_i \mathbf{x}_{i,v}^{\top} \mathbf{Z}_{i,T}^{-1} \mathbf{x}_{i,z} \right)^{\top}, \ 1 \le s, v \le z \le T,$$

exists, and there are s_0, v_0 and z_0 such that $1 \le s_0, v_0 \le z_0 \le T$, and at least one element of $\overline{\boldsymbol{Q}}(s_0, v_0, z_0)$ is different from 0.

It follows from Assumption 3.3 that $\|\boldsymbol{Q}(s, v, z)\| < \infty$. Similarly, Assumption 3.5 implies that $\|\overline{\boldsymbol{Q}}(s, v, z)\| < \infty$. The rate r_N is the exact rate in Assumptions 3.3 and 3.4, i.e. the matrix $\boldsymbol{Q}$ is different from the zero matrix. The term describing the dependence between the panels dominates the limit since $r_N/N^{1/2} \to \infty$.

Theorem 3.3. *If H_0 and Assumptions 2.1–2.3, 2.5 – 2.7, 3.1 and 3.4 hold, then we have*
$$(3.14)$$
$$\left\{ \frac{1}{r_N} \left(U_N(t) - A_N^{(1)}(t) \right), \ \underline{t}_0 \le t \le T - \bar{t}_0 \right\} \xrightarrow{\mathcal{D}} \left\{ \xi_t^{(3)}, \underline{t}_0 \le t \le T - \bar{t}_0 \right\}, \quad N \to \infty,$$

where

$$\xi_t^{(3)} = \sum_{s,v=1}^{t} \boldsymbol{\Lambda}_s^{\top} \boldsymbol{Q}(s, v, t) \boldsymbol{\Lambda}_v, \ \underline{t}_0 \le t \le T - \bar{t}_0,$$

non-zero matrix $\boldsymbol{Q}$ is defined in Assumption (3.4) and $\underline{t}_0, \bar{t}_0$ are defined in Assumption 3.1 and r_N in Assumption (3.3).

Now we formulate the analogue of Theorem 3.3 for $V_N(t)$.

Theorem 3.4. *If H_0 and Assumptions 2.1–2.3, 2.5 – 2.6 and 3.5 hold, then we have*

$$(3.15) \qquad \left\{ \frac{1}{r_N} \left(V_N(t) - A_N^{(2)}(t) \right), \ 1 \le t \le T \right\} \xrightarrow{\mathcal{D}} \left\{ \xi_t^{(4)}, 1 \le t \le T \right\}, \quad N \to \infty,$$

where

$$\xi_t^{(4)} = \sum_{s,v=1}^{t} \boldsymbol{\Lambda}_s^{\top} \overline{\boldsymbol{Q}}(s, v, t) \boldsymbol{\Lambda}_v, \ 1 \le t \le T,$$

and the non-zero matrix $\overline{\boldsymbol{Q}}$ is defined in Assumption 3.5.

Proofs of both theorems are given in Section 7.

Under the Assumptions of Theorems 3.3 and 3.4 the limit distributions are fully determined by the common factor $\boldsymbol{\Lambda}_s$, $s \in T$, and are not normally distributed.

3.5. **Random design vectors.** For the sake of simplicity, Theorem 3.1 and 3.2 with their proofs are provided for non-random $\boldsymbol{x}_{i,t}$'s. However, the results can be extended to random design points. We assume:

Assumption 3.6. $\{\varepsilon_{i,t}, \boldsymbol{\Lambda}_t, 1 \leq i \leq N, 1 \leq t \leq T\}$ *and* $\{\boldsymbol{x}_{i,t}, 1 \leq i \leq N, 1 \leq t \leq T\}$ *are independent.*

Assumptions 2.5 and 2.6 are modified so the uniform boundness is replaced with the uniform boundness of moments:

Assumption 3.7. *There is c_0 such that $E\|\boldsymbol{x}_{i,t}\|^8 \leq c_0$ for all $1 \leq i \leq N$, $1 \leq t \leq T$*

and

Assumption 3.8. *(i) There are positive constants c_1 and t_1 such that $E\|(\sum_{t=1}^{t_1} \boldsymbol{x}_{i,t}\boldsymbol{x}_{i,t}^T)^{-1}\|^8 \leq c_1$ for all $1 \leq i \leq N$.*
(ii) There are positive constants c_2 and t_2 such that $E\|(\sum_{t=T-t_2}^{T} \boldsymbol{x}_{i,t}\boldsymbol{x}_{i,t}^T)^{-1}\|^8 \leq c_2$ for all $1 \leq i \leq N$.

It follows from Assumption 3.8 that

$$P\left(\left(\sum_{t=1}^{t} \boldsymbol{x}_{i,t}\boldsymbol{x}_{i,t}^T \right)^{-1} \text{exists for all } t_1 \leq t \leq T \right) = 1$$

and

$$P\left(\left(\sum_{t=T-t}^{T} \boldsymbol{x}_{i,t}\boldsymbol{x}_{i,t}^T \right)^{-1} \text{exists for all } t_2 \geq t \right) = 1.$$

We also need to replace $a_{i,t,1}^2$ and $a_{i,t,2}^2$, $1 \leq i \leq N$, $\underline{t}_o \leq t \leq T - \overline{t}_0$ with their expected values.

4. Bootstrap tests and simulations

4.1. **Bootstrap tests.** In this subsection we describe the settings of the wild bootstrap procedure that could be used to test for the change point.

Due to Assumption 2.4 (negligible common factors) we could also apply the bootstrap to get approximations to the distribution functions of $U_N(t)$ and $V_N(t)$. We demonstrate the application of the wild bootstrap for $U_N(t)$, and $V_N(t)$ can be resampled in a similar way.

For $\underline{t}_0 \leq t \leq T - \overline{t}_0$ and $1 \leq i \leq N$ let

$$\phi_{i,t} = (\widehat{\boldsymbol{\beta}}_{i,t} - \widehat{\boldsymbol{\beta}}_{i,T})^\top \mathbf{C}_{i,t}(\widehat{\boldsymbol{\beta}}_{i,t} - \widehat{\boldsymbol{\beta}}_{i,T}) - \frac{1}{N}\sum_{j=1}^{N}(\widehat{\boldsymbol{\beta}}_{j,t} - \widehat{\boldsymbol{\beta}}_{j,T})^\top \mathbf{C}_{j,t}(\widehat{\boldsymbol{\beta}}_{j,t} - \widehat{\boldsymbol{\beta}}_{j,T}).$$

Assumption 4.1. *$\zeta_1, \ldots, \zeta_N$ are independent standard normal variables, independent of $\{\varepsilon_{i,t}, \boldsymbol{\Lambda}_t, 1 \leq i \leq N, 1 \leq t \leq T\}$*

The bootstrapped observations are defined by

$$(4.1) \qquad \phi_{i,t}^* = \zeta_i \, \phi_{i,t}, \quad \underline{t}_0 \le t \le T - \bar{t}_0, \ 1 \le i \le N.$$

and the respective bootstrapped process by

$$(4.2) \qquad U_N^*(t) = N^{-1/2} \sum_{i=1}^{N} \phi_{i,t}^*, \quad \underline{t}_0 \le t \le T - \bar{t}_0.$$

Theorem 4.1. *We assume that Assumptions 2.6 – 3.1 and 4.1 hold.*
(i) If H_0 holds, then

$$\{U_N^*(t), \ \underline{t}_0 \le t \le T - \bar{t}_0\} \ \overset{\mathcal{D}}{\to} \ \{\xi_t^{(1)}, \ \underline{t}_0 \le t \le T - \bar{t}_0\}$$

where $\xi_t^{(1)}, \underline{t}_0, \bar{t}_0$ are defined in Theorem 3.1.
(ii) If H_A holds, then

$$\sup_{\underline{t}_0 \le t \le T - \bar{t}_0} |U_N^*(t)| = O_P(1).$$

Proof: See Section 7.

4.2. **Test procedure and algorithm.** Based on the results in the previous sections we describe here the test procedures for testing H_0 specified in (2.3) of our model (1.1). Our test procedures are based on the maximum functionals of

$$(4.3) \qquad \overline{U}_N(t) = N^{-1/2}(U_N(t) - A_N^{(1)}(t))$$

where $U_N(t)$ and $A_N^{(1)}(t)$ is defined in (2.4) and (3.2), respectively, and

$$(4.4) \qquad \widehat{U}_N(t) = N^{-1/2}(U_N(t) - \widehat{A}_N^{(1)}(t))$$

with

$$\widehat{A}_N^{(1)}(t) = \sum_{i=1}^{N} \widehat{a}_{i,t,1}^2, \quad \widehat{a}_{i,t,1}^2 = \widehat{\sigma}_i^2 \mathrm{tr}\big(\boldsymbol{C}_{i,t}(\boldsymbol{Z}_{i,t}^{-1} - \boldsymbol{Z}_{i,T}^{-1}) \big)$$

and for $\widehat{\sigma}_i^2$ we recommend the estimator defined in (3.10). In our study we use the test statistics

$$\overline{u}_N = \max_t |\overline{U}_N(t)|, \quad \text{and} \quad \widehat{u}_N = \max_t |\widehat{U}_N(t)|.$$

Large values of either test statistics indicate that the null hypothesis is violated. The respective critical values are constructed via wild bootstrap. Notice that $\overline{u}_N$ depends on σ_i^2's when these are replaced by their estimators $\widehat{\sigma}_i^2$'s. Since σ_i^2's are usually unknown the test statistic $\widehat{u}_N$ is more practical. In application we use $\widehat{u}_N$, but in the simulations we also present selected results for $\overline{u}_N$ in order to see influence of unknown σ_i^2's.

Bellow we provide the algorithm for the test based on $\widehat{u}_N$ and level α

Algorithm

(1) Compute $\widehat{A}_N^{(1)}(t)$, $\widehat{U}_N(t)$ and $\widehat{u}_N$ as described above and $\widehat{\sigma}_i^2$ using formula (3.10).
(2) Generate bootstrap version $u_N^\star$:
    (a) Calculate $\phi_{i,t}$ using

$$\phi_{i,t} = (\widehat{\boldsymbol{\beta}}_{i,t} - \widehat{\boldsymbol{\beta}}_{i,T})^\top \boldsymbol{C}_{i,t}(\widehat{\boldsymbol{\beta}}_{i,t} - \widehat{\boldsymbol{\beta}}_{i,T}) - \frac{1}{N} \sum_{j=1}^{N} (\widehat{\boldsymbol{\beta}}_{j,t} - \widehat{\boldsymbol{\beta}}_{j,T})^\top \boldsymbol{C}_{j,t}(\widehat{\boldsymbol{\beta}}_{j,t} - \widehat{\boldsymbol{\beta}}_{j,T}).$$

(b) Generate ζ_i from standard normal distribution $N(0,1)$.

(c) Generate bootstrap observations $\phi_{i,t}^*$ using relation

$$\phi_{i,t}^* = \zeta_i\,\phi_{i,t}, \quad \underline{t}_0 \le t \le T - \bar{t}_0,\ 1 \le i \le N.$$

(d) Obtain bootstrap process

$$U_N^*(t) = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \phi_{i,t}^*.$$

(e) Get the bootstrap version $u_N^\star$ by taking maximum values of $U_N^*(t)$, i.e.

$$u_N^\star = \max_{d < t < T-d} \left| U_N^*(t) \right|.$$

(3) Find critical values from bootstrap.
- Repeat step (2) $B = 1000$ times, obtain $(1-\alpha)100\%$ quantile value of $u_N^\star$ and denote it $u_{N,\alpha}^\star$.
- Reject H_0 if $\widehat{u}_N > u_{N,\alpha}^\star$.

4.3. **Simulations.** For our simulations we used the model

$$y_{i,t} = \boldsymbol{x}_{i,t}^\top\big(\boldsymbol{\beta}_i + \boldsymbol{\delta}_i I\{t \ge t_0\}\big) + e_{i,t}, \quad 1 \le i \le N,\ 1 \le t \le T.$$

with $\boldsymbol{x}_{i,t}^\top = \big(1, x_{i,t}(2), \ldots, x_{i,t}(d)\big)^\top$, $d = 2$ and 5. The $x_{i,t}(j)$'s, $j = 2, \ldots, d$, has normal distribution $N(1,1)$, and are independent of $e_{i,t}, 1 \le t \le T, 1 \le i \le N$.

The errors $e_{i,t}, \ t = 1, \ldots, T$, are independent sequences. We consider three cases:

- *The "iid case"*: $e_{i,t}, i = 1, \ldots, N, \ t = 1, \ldots, T$, are independent and identically distributed and follow standard normal distribution.
- *The "unequal variances"*: $e_{i,t}, i = 1, \ldots, N, \ t = 1, \ldots, T$, are independent and normally distributed with zero mean and variance σ_i^2, that are generated from uniform distribution $U(0.5, 1.5)$.
- *The "GARCH(1,1) process"*: $e_{i,t} = \nu_{i,t}\varepsilon_{i,t}$, where $\nu_{i,t}^2 = 0.25 + 0.25e_{i,t-1}^2 + 0.5\nu_{i,t-1}^2$ and $\varepsilon_{i,t}$ are independent and follow the standard normal distribution $N(0,1)$.

In the sequel we report on two model specifications:

S1: Linear model with $d = 2$, $\boldsymbol{\beta} = (1,2)^\top$ and $\boldsymbol{C}_{i,t} = \boldsymbol{Z}_{it}\boldsymbol{Z}_{it}$.

S2: Linear model with $d = 5$, $\boldsymbol{\beta} = (1,1,2,3,4)^\top$ and $\boldsymbol{C}_{i,t} = \boldsymbol{Z}_{i,t}\boldsymbol{Z}_{i,T}^{-1}\boldsymbol{Z}_{i,t}$.

We use extensive simulations to explore finite sample properties of the underlying test procedure. In particular, we analyze the significance level and the empirical power of the test for different pairs of (N, T), different assumptions on t_0 (the change point) and various assumptions on the errors terms. Overall, we keep the number of simulations equal to $1\,000$ and the number of bootstrap steps also equals $1\,000$. In our simulation setup we also include cases when both N and T are small, to explore small sample properties of the test statistics $\overline{u}_N$ and $\widehat{u}_N$ and the behavior of the proposed wild bootstrap procedure. In simulations we first consider the case when non zero changes occur in all panels at the same time t_0. Later we analyze the sensitivity of the detection of change in the case when only a fraction of panels ($\vartheta = 0.25$) experience change. Selected results are presented bellow.

A. Empirical versus theoretical significance level

Our results show that, even for rather small N and T the nominal rejection rate was close to the corresponding significance level of the tests. Table 4.1 demonstrates how well the

nominal rejection of the null hypothesis (no change) in the case of model S2 corresponds to the chosen significance level, with four independent explanatory variables and intercept. A similar result is reported in Table 4.2 for model S1 for small N and T. Other model assumptions (different number of regressors, different data generating process, different distributions) lead to similar results; additional results are available upon request.

$T \setminus N$	σ_i^2 known ($\bar{u}_N$)				σ_i^2 estimated by $\widehat{\sigma}_i^2$ ($\widehat{u}_N$)			
	100	200	500	1000	100	200	500	1000
50	0.047	0.049	0.049	0.048	0.052	0.048	0.051	0.047
100	0.054	0.052	0.051	0.052	0.056	0.053	0.052	0.053
200	0.053	0.051	0.049	0.051	0.053	0.054	0.048	0.051

Table 4.1. Nominal rejection of the H_0; test based on $\bar{u}_N$ (left part of the table) and on $\widehat{u}_N$ (right part of the table) at significance level 5%; model S2 (four regressors and intercept).

	σ_i^2 known ($\bar{u}_N$)				σ_i^2 estimated ($\widehat{u}_N$)			
	number of panels				number of panels			
	100	200	500	1000	100	200	500	1000
iid case								
$T = 50$	0.045	0.049	0.051	0.055	0.037	0.040	0.052	0.048
$T = 100$	0.049	0.044	0.044	0.051	0.044	0.045	0.041	0.033
$T = 200$	0.056	0.045	0.047	0.043	0.039	0.049	0.049	0.043
unequal variances								
$T = 50$	0.042	0.041	0.039	0.047	0.032	0.036	0.032	0.043
$T = 100$	0.039	0.028	0.046	0.045	0.039	0.041	0.03	0.040
$T = 200$	0.040	0.041	0.039	0.033	0.037	0.035	0.051	0.039
GARCH(1,1)								
$T = 50$	0.045	0.037	0.056	0.048	0.020	0.036	0.027	0.029
$T = 100$	0.038	0.048	0.051	0.042	0.031	0.025	0.037	0.045
$T = 200$	0.040	0.044	0.040	0.037	0.054	0.047	0.049	0.040

Table 4.2. Nominal rejection of the H_0; test based on $\bar{u}_N$ (left part of the table) and on $\widehat{u}_N$ (right part of the table) at significance level 5%; model S1 (one regressor and intercept).

B. Empirical power of the test statistic $\widehat{u}_N$

In our simulation setup we set the time of change t_0 equal to $T/2$ and $T/4$ under H_A. Table 4.3 shows the empirical power of the tests based on $\overline{u}_N$ and $\widehat{u}_N$ at a 5% significance level in case of model S2, for several combinations of N and T. The empirical power is an increasing function of N and T. It is high even for small and moderate N and T, which can be explained by changes in all panels at the same time. We note that the estimation of the σ_i^2's reduces power but very mildly. By increasing N for each T we see the increasing empirical power of the test. Similar results of the simpler model S1 with different data generating processes are depicted in Table 4.4

	N = 100		N = 200		N = 500		N = 1000	
	$t_0 = \frac{T}{2}$	$t_0 = \frac{T}{4}$	$t_0 = \frac{T}{2}$	$t_0 = \frac{T}{4}$	$t_0 = \frac{T}{2}$	$t_0 = \frac{T}{4}$	$t_0 = \frac{T}{2}$	$t_0 = \frac{T}{4}$
$T = 50$	0.699	0.678	0.876	0.855	0.941	0.929	0.971	0.932
$T = 100$	0.847	0.832	0.929	0.917	0.956	0.941	0.992	0.983
$T = 200$	0.908	0.889	0.947	0.925	0.965	0.957	0.995	0.989

Table 4.3. Power of the test based on $\widehat{u}_N$ when the change occurs at $t_0 = T/2$ or $T/4$ in all regressors of model S2 ($\boldsymbol{\delta} = (0, 1, 0.3, 0.2, 0.1)^\top$); the significance level 5%.

	change in intercept				change in slope			
	number of panels				number of panels			
	100	200	500	1000	100	200	500	1000
iid case								
$T = 50$	0.042	0.071	0.172	0.350	0.139	0.350	0.765	0.981
$T = 100$	0.113	0.264	0.625	0.904	0.569	0.914	0.999	1
$T = 200$	0.408	0.791	0.998	1	0.986	1	1	1
unequal variances								
$T = 50$	0.036	0.046	0.149	0.284	0.130	0.307	0.765	0.976
$T = 100$	0.088	0.210	0.576	0.888	0.522	0.885	0.999	1
$T = 200$	0.376	0.754	0.997	1	0.987	1	1	1
GARCH(1,1)								
$T = 50$	0.027	0.043	0.124	0.269	0.104	0.277	0.666	0.960
$T = 100$	0.085	0.190	0.541	0.892	0.504	0.878	1	1
$T = 200$	0.380	0.785	0.990	1	0.988	1	1	1

Table 4.4. Power of the test based on $\widehat{u}_N$ when the change occurs at $t_0 = T/2$ in either intercept or slope of model S1 ($\delta = 0.1$); the significance level 5%.

C. Sensitivity analysis to fraction of panels experiencing a change

We also investigate a situation when only a fraction of the panels experience a change. The fraction of the panels with change is denoted by ϑ. Given the space limitations we present this sensitivity analysis only for the model S1 and for the case of $\widehat{u}_N$ (σ_i^2 estimated by $\widehat{\sigma}_i^2$). The other setups lead to similar results. Bellow we present the results for model S1 when all changes occurred at time $t_0 = T/2$. We applied $\widehat{u}_N$ to detect instability at a 5% significance level and the size of changes was $\boldsymbol{\delta} = (0.25, 0)$ (change in intercept) and $\boldsymbol{\delta} = (0, 0.25)$ (change in the slope). The results are presented in Table 4.5 when $\vartheta = 0.25$. For other ϑ and δ not presented here we observed the following pattern. The power of the test decreases with ϑ (fraction of panels with changes) and increases with δ (size of the change). The level of detection is very good, even if only a quarter of the panels contains changes in the regression parameters; one can see that the power of the test increases with T and with N. Additional results are available upon request.

	change in intercept				change in slope			
	number of panels				number of panels			
	100	200	500	1000	100	200	500	1000
iid case								
$T=50$	0.061	0.117	0.390	0.726	0.302	0.699	0.987	1
$T=100$	0.229	0.544	0.940	0.999	0.858	0.994	1	1
$T=200$	0.734	0.986	1	1	1	1	1	1
unequal variances								
$T=50$	0.042	0.103	0.331	0.683	0.269	0.653	0.987	1
$T=100$	0.187	0.489	0.911	0.999	0.839	0.998	1	1
$T=200$	0.703	0.976	1	1	0.999	1	1	1
GARCH(1,1)								
$T=50$	0.040	0.094	0.273	0.601	0.249	0.578	0.972	1
$T=100$	0.169	0.461	0.902	0.997	0.838	0.995	1	1
$T=200$	0.707	0.973	1	1	1	1	1	1

Table 4.5. Power of the test based on $\widehat{u}_N$ when the change occurs at $t_0 = T/2$ in either intercept or slope of model S1 ($\delta = 0.25$); fraction of panels with change $\vartheta = 0.25$; the significance level 5%.

5. Empirical Illustration

In this section we present an application of our theoretical results to the capital asset pricing model (CAPM). We apply our approach to the CAPM, using the conventional factor model, which includes the Fama-French three-factor model augmented with the Carhart (1997) momentum factor.

Specifically, we test for breaks in coefficients for the US mutual fund return data around the sub-prime crisis. Our time span covers the period from January 2006 to February 2010, which is consistent with other sub-prime crisis studies (e.g. Dick-Nielsen et al., 2012). The sub-prime crisis is generally defined as the period between the fourth quarter of 2007 and the end of 2008 (Santos, 2011).

The Fama-French (1993) three factor model augmented with the Carhart (1997) momentum factor is defined as

$$R_{i,t} - R_t^f = \alpha_{i,t} + \left(R_t^M - R_t^f\right)\beta_{i,t}^M + R_t^{HML}\beta_{i,t}^{HML} + R_t^{SMB}\beta_{i,t}^{SMB} + R_t^{MOM}\beta_{i,t}^{MOM} + e_{i,t},$$

$1 \le t \le T$, $1 \le i \le N$, where $R_{i,t} - R_t^f$ denotes the excess return on the mutual fund; $R_t^M - R_t^f$ is the market risk premium; R_t^{HML} is the value factor, calculated as the return difference between portfolios with the highest decile of stocks and the lowest decile of stocks in terms of the ratio of book equity-to-market equity (HML); R_t^{SMB} is the value factor, calculated as the return difference between portfolios with the smallest decile of stocks and the largest decile of stocks in terms of size (SMB); R_t^{MOM} is the momentum factor calculated as the return difference between portfolios with the highest decile of stocks and lowest decile of stocks in terms of recent return (i.e., momentum, or MOM); and $e_{i,t}$ the random errors.

In our investigation we use $\widehat{u}_N$ as our test statistics and we display the graphs of $|\widehat{U}_N(t)|$. The unknown σ_i^2 is estimated by $\widehat{\sigma}_i^2$ of (3.10). The critical values were obtained by the wild bootstrap of Section 4.1.

5.1. **Data Information.** The four factors can be downloaded from Ken French's data library[1]. The raw dataset on mutual funds contains monthly return data of 6 190 US mutual funds from January 1984 to November 2014[2]. However, there are many missing values in the mutual fund dataset because different mutual funds have different start dates and some of them have already been terminated. For the purpose of our illustration we select the mutual funds that have no missing returns for the period of the subprime crisis (January 2006 to February 2010).

Using the Yahoo finance classification, we consider nine categories of the US mutual funds. These are *Large Blend, Large Growth, Large Value, Middle Blend, Middle Growth, Middle Value, Small Blend, Small Growth* and *Small Value*. These categories are combinations of the mutual fund size and their investment strategies as illustrated in Table 5.1.

			Large
			Medium
			Small
Blend	Growth	Value	

Table 5.1. Categories of mutual funds.

5.2. **Results for the sub-prime crisis.** Table 5.2 shows the test results for the sub-prime crisis. It covers the period from January 2006 to February 2010, which is equal to 50 periods. $\widehat{u}_N$ is the test statistic, CV stands for critical values which are reported for 10%, 5% , and 1% significance levels obtained by the wild bootstrap. The *break point* column shows the month (time period) in which a break point was detected and N shows the number of observations per fund category. Between the brackets in the breakpoint column the estimated time of the change is reported.

For all but one fund categories the test statistic $\widehat{u}_N$ is above the 5% critical values during the sub-prime crisis period (middle of 2008 to early 2009). Interestingly, the test cannot detect change for the category of Small Value mutual funds at the 5% significance level. However the Small Value mutual fund category is significant at the 10% significance level.

The coefficients for the four factors changed first for the Large Blend mutual fund category with estimated break point in August 2008, and the Small Growth mutual fund category has estimated change point just a month later in September 2008. The coefficients for Large Growth, Middle Blend, Middle Growth, and Small Blend changed last, in March 2009.

It is interesting to note that only the Large Blend indicated a change in the environment out of all large-type mutual funds. Recall that the Large Blend is defined as a balanced mix of growth and value stocks. We hypothesize that the Large Blend more closely resembles the market and that is why it indicated a shift in the underlying structure.

---

[1]The factors data is available at
http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html
[2]The monthly return history of US mutual funds is available for the period January 1984 to November 2014 at the web site http://finance.yahoo.com

Category	$\widehat{u}_N$	10 % CV	5 % CV	1 % CV	break point $\widehat{t_0}$	N
Large Blend	47 125	18 361	21 964	28 961	Aug 2008 (32)	659
Large Growth	149 624	36 993	43 449	56 892	Mar 2009 (39)	748
Large Value	48 210	9 138	10 576	13 239	Feb 2009 (38)	528
Middle Blend	81 238	38 007	44 299	56 424	Mar 2009 (39)	145
Middle Growth	151 827	23 034	26 999	34 196	Mar 2009 (39)	329
Middle Value	36 235	14 910	17 269	21 878	Oct 2008 (34)	173
Small Blend	37 067	14 893	17 364	22 394	Mar 2009 (39)	279
Small Growth	67 004	14 742	16 944	21 575	Sep 2008 (33)	313
Small Value	40 583	35 504	41 433	52 510	-	135

Table 5.2. Detection of the break point during the sub-prime crisis

Figure 5.1 plots the estimated break points for the coefficients for the Mutual Fund categories with the levels of the S&P 500. The model employed is the Fama and French three factor model (1993) augmented with the Carhart (1997) momentum factor.
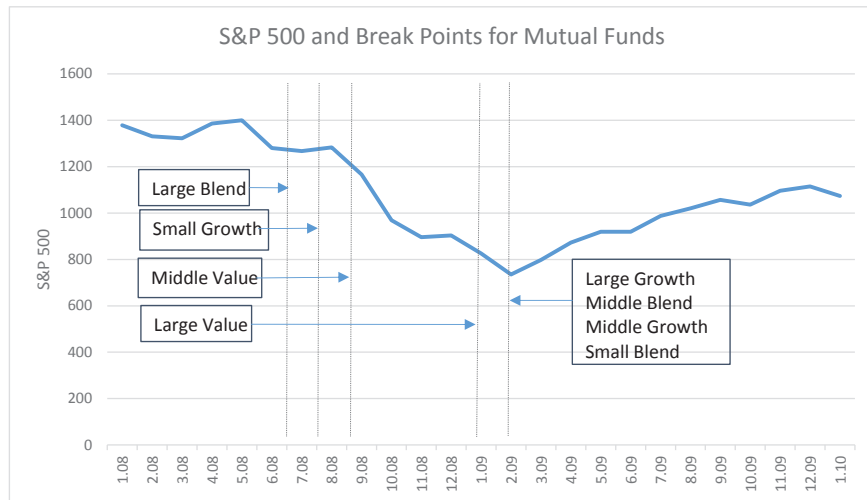
Figure 5.1. Estimated break points of the mutual fund returns and the level of the S&P 500.

The Large Blend category indicated a structural change when the S&P 500 was well above the 1200 value. This could have been used as a potential trading signal. The lowest point of the S&P 500 was February 2009, below the 800 mark. Here the change point was detected for the Large Growth, Middle Blend, Middle Growth, and Small Blend. Again, as for the earlier estimated break point for Large Blend, this time could also be used as a trading signal, obviously indicating an upturn of the market.

Figures 5.2–5.10 graphically depict the processes $\left|\widehat{U}_N(t)\right|$, estimated break point $\widehat{t}_N$, and the estimated simulated critical values at the 10%, 5% , and 1% significance levels for all categories of mutual funds.
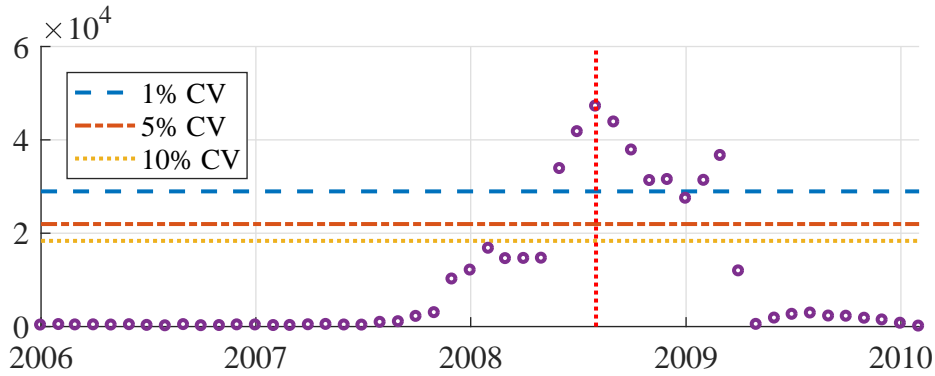
Figure 5.2. Graphs of $\left|\widehat{U}_N(t)\right|$ and simulated critical values for Large Blend mutual funds.

Figure 5.2 shows $\left|\widehat{U}_N(t)\right|$ and simulated critical values for the category of Large Blend mutual funds. $\left|\widehat{U}_N(t)\right|$ stays below the critical level up until late 2007 and then in 2008 it starts slowly breaking off. The estimated break point (corresponding to the highest value of $\left|\widehat{U}_N(t)\right|$) is in August of 2008, but it is clear that in the preceding months an increase is visible. After the break in the coefficients, $\left|\widehat{U}_N(t)\right|$ slowly decreased to regular levels in late 2009 and early 2010. Another peak in the first quarter of 2009, could suggest an influence of the European sovereign debt crisis, i.e., potentially two break points.
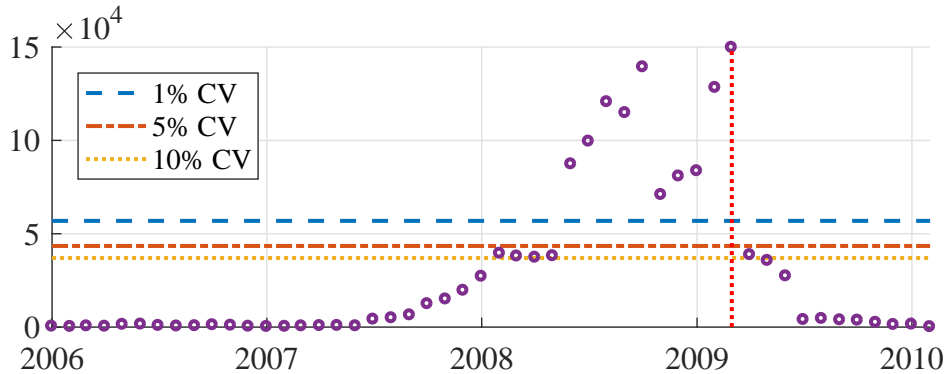
Figure 5.3. Graphs of $\left|\widehat{U}_N(t)\right|$ and simulated critical values for Large Growth mutual funds.

Figure 5.3 shows $\left|\widehat{U}_N(t)\right|$ and simulated critical values for the category of Large Growth mutual funds. The estimated break point is in March of 2009, which is at the deepest point of the sub-prime crisis. However, the graph tells a slightly different story. $\left|\widehat{U}_N(t)\right|$ is at zero in 2006 and 2007, and slowly increases in 2008. The process $\left|\widehat{U}_N(t)\right|$ is well above the simulated critical values in late 2008 and all 2009, but the estimator of the time of a change is quite late due the magnitude of the March 2009 value. Immediately after the detection, $\left|\widehat{U}_N(t)\right|$ quickly decreases below critical levels.
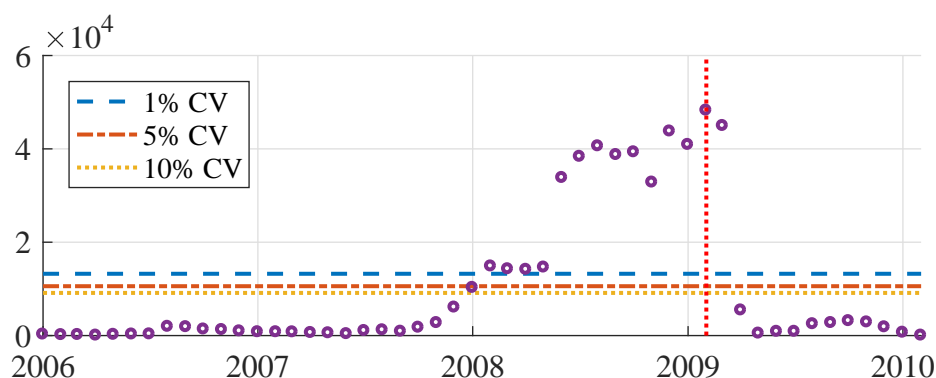
Figure 5.4. Graphs of $\left|\widehat{U}_N(t)\right|$ and simulated critical values for Large Value mutual funds.

Figure 5.4 shows $\left|\widehat{U}_N(t)\right|$ and simulated critical values for the category of Large Value mutual funds. The estimated break point is in February 2009. Obviously large mutual funds could be strongly influenced by the European sovereign debt crisis that started in early 2009. Again, a similar story emerges to the Large Value mutual fund. Values of $\left|\widehat{U}_N(t)\right|$ are around zero and well below the simulated critical values in 2006, 2007, and early 2008. For most of 2008 the process $\left|\widehat{U}_N(t)\right|$ is above the simulated critical values but the break point is again indicated late. After the estimated break in February $\left|\widehat{U}_N(t)\right|$ drops below critical value levels.
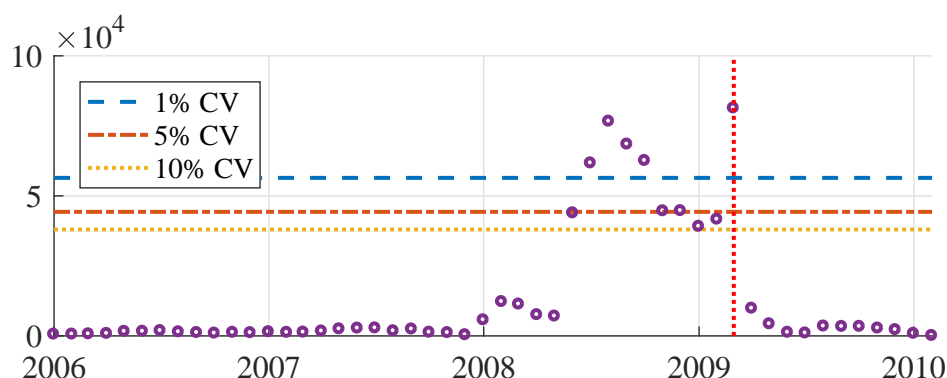

Figure 5.5. Graphs of $\left|\widehat{U}_N(t)\right|$ and simulated critical values for Middle Blend mutual funds.
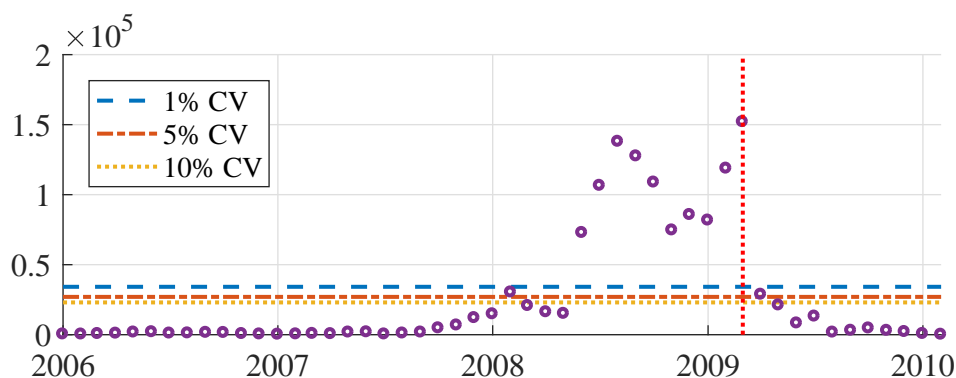

Figure 5.6. Graphs of $\left|\widehat{U}_N(t)\right|$ and simulated critical values for Middle Growth mutual funds.
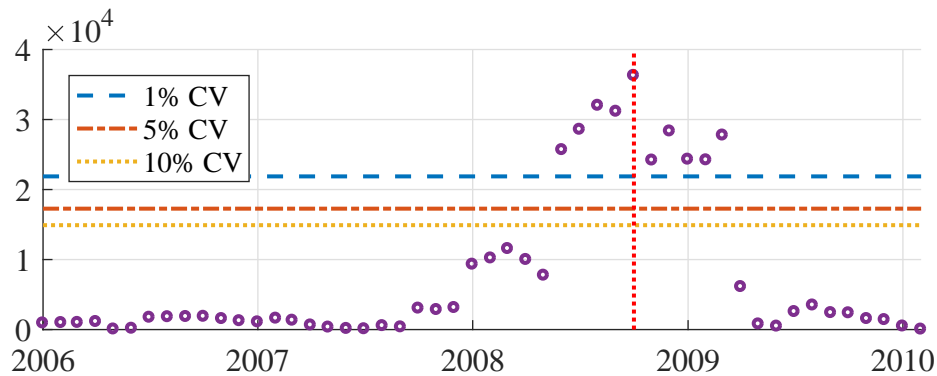
Figure 5.7. Graphs of $\left|\widehat{U}_N(t)\right|$ and simulated critical values for Middle Value mutual funds.

Figures 5.5 – 5.7 show $\left|\widehat{U}_N(t)\right|$ and simulated critical values for the category of Middle Blend, Middle Growth, and Middle Value mutual funds. The Middle Blend and Middle Growth mutual fund categories have an estimated break point in March 2009. For both mutual fund categories, $\left|\widehat{U}_N(t)\right|$ remains fairly low up until early 2008. Even though the test statistic was above the simulated critical values in late 2008 and early 2009 the break point is indicated fairly late, in March 2009 for both funds. $\left|\widehat{U}_N(t)\right|$ decreases to below simulated critical values shortly after the break detection. Middle Blend and Middle Growth mutual funds show a very similar pattern, suggesting two breakpoints. The first relates to to the US crisis and the second to the European sovereign debt crisis. The European crisis is detected as the main break point for Middle Blend and Middle Growth mutual funds. Similar characteristics of $\left|\widehat{U}_N(t)\right|$ are observed. The values remain very low and well below the simulated critical values up until late 2008. After the estimated break the values of $\left|\widehat{U}_N(t)\right|$ remain quite high for a few months before dropping in late 2009.
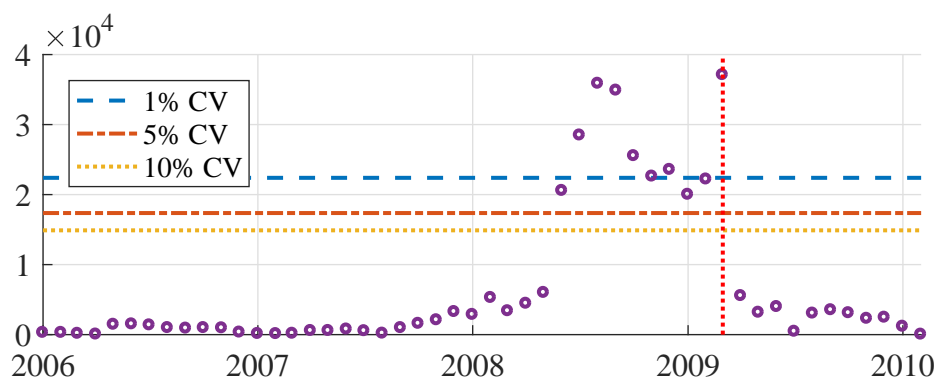

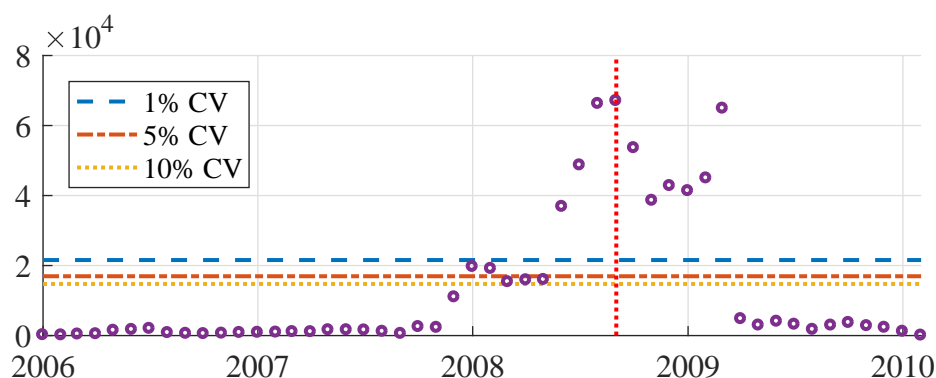Figure 5.8. Graphs of $\left|\widehat{U}_N(t)\right|$ and simulated critical values for Small Blend mutual funds.

Figure 5.9. Graphs of $\left|\widehat{U}_N(t)\right|$ and simulated critical values for Small Growth mutual funds.

Figures 5.8 – 5.10 show $\left|\widehat{U}_N(t)\right|$ and simulated critical values for the category of Small Blend, Small Growth, and Small Value mutual funds. The estimated break points for the Small Blend and Small Growth mutual funds are detected in March 2009 and September 2008, respectively. As is clearly visible, both graphs suggest possibility of two break points, as before. One is linked to the US mortgage crisis and the other to the European sovereign debt crisis. The test statistics for the maximum at these peaks are very similar. The simulated critical values were low until late 2008 and well below the significance levels.
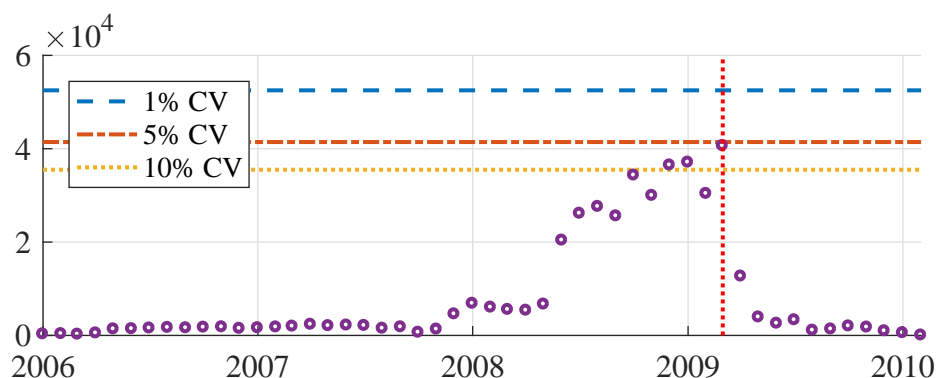

Figure 5.10. Graphs of $\left|\widehat{U}_N(t)\right|$ and simulated critical values for Small Value mutual funds.

The Small Value mutual fund category does not indicate a change point because the test statistic never exceeded the 5% critical value. However, a change is detected at the 10% level.

Overall, the result shows the sensitivity of large and medium mutual funds to the 2008 economic crisis, as well as the subsequent European sovereign debt crisis. For a detailed financial analysis of the associated change point it would be useful to classify the mutual funds according to their international portfolio holdings. The portfolio structure and its diversification could likely be attributed to a possible change. We leave these points for further research.

6. Conclusion and suggestions for further research

The detection of (structural) breaks has been at the center of both theoretical and applied economic and financial research for many years. Part of the existing research concentrates on the detection of breaks in panels when both N, the number of panels

and T, the time dimension are large. In contrast, in this paper we concentrate on a qualitatively different situation. We elaborate a general approach for testing for the break(s) when N is large and T is fixed and rather small. This situation is typically related to large (firm-level) data containing financial information about an immense number of firms/stocks across a limited period of time. More precisely, we study the asymptotic properties of suggested test procedures when $N \to \infty$ while keeping T fixed. Our approach also allows the estimation of the time of the change(s).

The goal of this paper is to present a background for constructing a set of quite general test statistics which are suitable in the given context, and to prove theorems about their asymptotic behavior. We also aim to show that several special choices of the weighting matrices used in our test statistics resulted in a generalization of some well-known test statistics.

In the rich simulation section we study the finite sample properties of two processes. The maximum norm of these processes can be used as test statistics and critical values obtained by a wild bootstrap. We present a selected set of tables to document the accuracy of the significance levels and the power of the test under several simulation schemes and alternatives, including cases when only a fraction of the panels experienced a break.

For a practical application we apply the testing procedure in the framework of the four factors CAPM model. In particular, we estimate breaks in monthly returns of US mutual funds during the period January 2006 to February 2010 which covers the subprime crises.

The theoretical considerations concerning multiple breaks are beyond the scope of this paper and we plan to address them in future research.

References

[1] Andreou, E., and Ghysels, E. *Structural breaks in financial time series*. In Handbook of Financial Time Series, Andersen, T.G. eds., 839–870, Springer, 2009.

[2] Andrews, D.W.K. Tests for parameters instability and structural changewith unknown change point. *Econometrica* **61**, 821–856, 1993.

[3] Andrews, D.W.K. Tests for parameter instability and structural change with unknown change point: A corrigendum. *Econometrica* **71**, 395–397, 2003.

[4] Antoch, J., and Jarušková, D. Detection of breaks in a capital structure: A case study. *Statistika: Statistics and Economy Journal* **97**, 32–43, 2017.

[5] Arellano, M. *Panel Data Econometrics*. Oxford University Press, Oxford, 2004.

[6] Aue, A., and Horváth, L. Structural breaks in time series. *Journal of Time Series Analysis* **23**, 1–16, 2012.

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

[8] Bai, J., and Carrion-i-Silvestre, J.L. Structural changes, common stochastic trends, and unit roots in panel data. *Review of Economic Studies* **76**, 471–501, 2009.

[9] Baltagi, B.H., Kao, C., and Liu, L. Estimation and identification of change points in panel models with nonstationary or stationary regressors and error terms. *Journal of Econometrics* **191**, 176–195, 2016.

[10] Banerjee, A., and Urga, G. Modelling structural breaks, long memory and stock market volatility: An overview. *Journal of Econometrics* **129**, 1–34, 2005.

[11] Billingsley, P. *Convergence of Probability Measures*. J. Wiley, New York, 1968.

[12] Brodsky, B.E., and Darkhovskii, B. *Non-Parametric Statistical Diagnosis*. Kluwer, Dordrecht, 2000.

[13] Carhart, M.M. On persistence in mutual fund performance. *Journal of Finance* **52**, 57–82, 1997.

[14] Cho H., and Fryzlewicz P. Multiple-change-point detection for high dimensional time series via sparsified binary segmentation. *JRSS B* **77**, 475–507, 2015.

[15] Csörgő, M., and Horváth, L. *Limit Theorems in Change–Point Analysis*. J. Wiley, Chichester, 1997.

[16] DeAngelo, H., and Roll, R. How stable are corporate capital structures? *Journal of Finance* **70**, 373–418, 2015.

[17] Dick-Nielsen, J., Feldhütter, P., Lando, D. Corporate bond liquidity before and after the onset of the subprime crisis. *Journal of Financial Economics* **103**, 2012, 471–492,

[18] Fama, E.F., and French, K.R.F., A five-factor asset pricing model. Fama-Miller Working Paper, September 2014. Available at http://dx.doi.org/10.2139/ssrn.2287202

[19] Fama, E.F., and French, K.R.F. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* **33**, 3–56, 1993.

[20] Francq, C., and Zakoian, J.M. *GARCH Models.* J. Wiley, Chichester, 2010.

[21] Ghysels, E. On stable factor structures in the pricing of risk: Do time-varying betas help or hurt? *Journal of Finance* **53**, 549–573, 1998.

[22] Graham, J.R., Leary, M.T., and Roberts, M.R. A century of capital structure: The leveraging of corporate America. *Journal of Financial Economics* **118**, 658–683, 2015.

[23] Guercio, D. D., and Reuter, J. Mutual fund performance and the incentive to generate alpha. *The Journal of Finance* **69**, 1673–1704, 2014.

[24] Hansen, B.E. The new econometrics of structural change: Dating breaks in US labor productivity. *The Journal of Economic Perspectives* **15**, 117–128, 2001.

[25] Horváth, L., and Hušková, M. Change–point detection in panel data. *Journal of Time Series Analysis* **33**, 631–648, 2012.

[26] Horváth, L., Hušková, M., Rice, G. and Wang, J. CUSUM estimator for the time of change in linear panel data models. *Econometric Theory* **33**, 366-412, 2017.

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

[28] Hsiao, C. *Analysis of Panel Data, 2nd ed.* Cambridge University Press, Cambridge, 2003.

[29] Hsiao, C. Panel data analysis – advantages and challenges. *Test* **16**, 1–22, 2007.

[30] Jirák, M. Uniform change point tests in high dimension. *Annals of Statistics* **43**, 2451–2483, 2015.

[31] Joseph, L., and Wolfson, D.B. Estimation in multi–path change–point problems. *Communications in Statistics – Theory and Methods* **21**, 897–913, 1992.

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

[33] Kao, C., Trapani, L., and Urga, G. Asymptotics for panel models with common shocks. *Econometric Reviews* **31**, 390–439, 2012.

[34] Kao, C., Trapani, L., and Urga, G. Testing for breaks in cointegrated panels. *Econometric Reviews* 2014 (to appear).

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

[36] Lemmon, M.L., Roberts, M.R., and Zender, J.F. Back to the beginning: Persistence and the cross-section of corporate capital structure. *Journal of Finance* **63**, 1575–1608, 2008.

[37] Perron, P. Dealing with structural breaks. In: Palgrave Handbook of Econometrics, Vol. 1: Econometric Theory, Patterson, K. and Mills, T.C. (eds.), Palgrave Macmillan, 278-352, 2006.

[38] Pešta, M, and Peštová, B. Testing structural changes in panel data with small fixed panel size and bootstrap. *Metrika* **78**, 665–689, 2015.

[39] Petrov, V.V. *Limit Theorems of Probability Theory.* Clarendon Press, Oxford, 1995.

[40] Santos Joo, A.C. Bank Corporate Loan Pricing Following the Subprime Crisis. *Review Financial Studies* **24 (6)**, 1916–1943, 2011.

[41] Torgovitski, L. Panel data segmentation under finite time horizon. *Journal of Statistical Planning and Inference* **167**, 69-89, 2015.

[42] Whited, T.M., and Wu, G. Financial constraints risk. *Review of Financial Studies* **19**, 531–559, 2006.

[43] Wooldridge, J.M. *Econometric Analysis of Cross Section and Panel Data, 2nd ed.* MIT Press, Boston, 2010.

7. Proofs

We assume that H_0 holds. Using (1.1) we get

$$\widehat{\boldsymbol{\beta}}_{i,t} = \boldsymbol{\beta}_i + \mathbf{Z}_{i,t}^{-1}\big(\mathbf{S}_{i,t} + \mathbf{R}_{i,t}\big),$$

where $\mathbf{S}_{i,t}, \mathbf{R}_{i,t} \in R^d$, $\mathbf{S}_{i,t}$ is defined in (3.5) and $\mathbf{R}_{i,t} = \sum_{s=1}^t \mathbf{x}_{i,s} \boldsymbol{\gamma}_i^\top \boldsymbol{\Lambda}_s$, $1 \le i \le N, 1 \le t \le T$. It also holds that

$$
\begin{aligned}
(7.1) \quad (\widehat{\boldsymbol{\beta}}_{i,t} &- \widehat{\boldsymbol{\beta}}_{i,T})^\top \mathbf{C}_{i,t}(\widehat{\boldsymbol{\beta}}_{i,t} - \widehat{\boldsymbol{\beta}}_{i,T}) \\
&= \mathbf{S}_{i,t}^\top \mathbf{C}_{i,t,1} \mathbf{S}_{i,t} + \mathbf{S}_{i,T}^\top \mathbf{C}_{i,t,2} \mathbf{S}_{i,T} - 2\mathbf{S}_{i,t}^\top \mathbf{C}_{i,t,3} \mathbf{S}_{i,T} + 2\mathbf{S}_{i,t}^\top \mathbf{C}_{i,t,1} \mathbf{R}_{i,t} \\
&\quad + 2\mathbf{S}_{i,T}^\top \mathbf{C}_{i,t,2} \mathbf{R}_{i,T} - 2\mathbf{S}_{i,t}^\top \mathbf{C}_{i,t,3} \mathbf{R}_{i,T} - 2\mathbf{R}_{i,t}^\top \mathbf{C}_{i,t,3} \mathbf{S}_{i,T} + \mathbf{R}_{i,t}^\top \mathbf{C}_{i,t,1} \mathbf{R}_{i,t} \\
&\quad + \mathbf{R}_{i,T}^\top \mathbf{C}_{i,t,2} \mathbf{R}_{i,T} - 2\mathbf{R}_{i,t}^\top \mathbf{C}_{i,t,3} \mathbf{R}_{i,T},
\end{aligned}
$$

where

$$
\mathbf{C}_{i,t,1} = \mathbf{Z}_{i,t}^{-1} \mathbf{C}_{i,t} \mathbf{Z}_{i,t}^{-1}, \quad \mathbf{C}_{i,t,2} = \mathbf{Z}_{i,T}^{-1} \mathbf{C}_{i,t} \mathbf{Z}_{i,T}^{-1} \quad \text{and} \quad \mathbf{C}_{i,t,3} = \mathbf{Z}_{i,t}^{-1} \mathbf{C}_{i,t} \mathbf{Z}_{i,T}^{-1}.
$$

Lemma 7.1. *If H_0 and Assumptions 2.1 – 2.3 and 2.5 – 2.7 hold, then we have that*

$$
E\, \mathbf{S}_{i,t}^\top \mathbf{C}_{i,t,1} \mathbf{S}_{i,t} = \sigma_i^2 \sum_{s=1}^t \mathbf{x}_{i,s}^\top \mathbf{C}_{i,t,1} \mathbf{x}_{i,s} = \sigma_i^2 \operatorname{tr}\left(\mathbf{C}_{i,t,1} \mathbf{Z}_{i,t}\right),
$$

$$
E\, \mathbf{S}_{i,T}^\top \mathbf{C}_{i,t,2} \mathbf{S}_{i,T} = \sigma_i^2 \sum_{s=1}^T \mathbf{x}_{i,s}^\top \mathbf{C}_{i,t,2} \mathbf{x}_{i,s} = \sigma_i^2 \operatorname{tr}\left(\mathbf{C}_{i,t,2} \mathbf{Z}_{i,T}\right)
$$

and

$$
E\, \mathbf{S}_{i,t}^\top \mathbf{C}_{i,t,3} \mathbf{S}_{i,T} = \sigma_i^2 \sum_{s=1}^t \mathbf{x}_{i,s}^\top \mathbf{C}_{i,t,3} \mathbf{x}_{i,s} = \sigma_i^2 \operatorname{tr}\left(\mathbf{C}_{i,t,3} \mathbf{Z}_{i,t}\right).
$$

Proof. With the notation $\mathbf{C}_{i,t,1} = \{c_{i,t,1}(\ell, \ell'), 1 \le \ell, \ell' \le d\}$ and $\mathbf{S}_{i,t} = (S_{i,t}(1), \ldots, S_{i,t}^d)^\top$ we have

$$
\begin{aligned}
(7.2) \quad \mathbf{S}_{i,t}^\top \mathbf{C}_{i,t,1} \mathbf{S}_{i,t} &= \sum_{s=1}^t \mathbf{x}_{i,s}^\top \, \varepsilon_{i,s} \mathbf{Z}_{i,t}^{-1} \mathbf{C}_{it} \mathbf{Z}_{i,t}^{-1} \sum_{v=1}^t \mathbf{x}_{i,v} \varepsilon_{i,v} \\
&= \sum_{\ell=1}^d \sum_{\ell'=1}^d S_{i,t}(\ell) c_{i,t,1}(\ell, \ell') S_{i,t}(\ell') \\
&= \sum_{\ell=1}^d \sum_{\ell'=1}^d \sum_{s=1}^t \sum_{s'=1}^t x_{i,s}(\ell) x_{i,s'}(\ell') \varepsilon_{i,s} \varepsilon_{i,s'} c_{i,t,1}(\ell, \ell'),
\end{aligned}
$$

which implies the first statement of the lemma since $E\, \varepsilon_{i,s} \varepsilon_{i,s'} = \sigma_i^2$ if $s = s'$, and zero otherwise. The same argument can be used to prove the rest of Lemma 7.1. $\qquad \square$

Let

$$
H_{i,t} = \mathbf{S}_{i,t}^\top \mathbf{C}_{i,t,1} \mathbf{S}_{i,t} + \mathbf{S}_{i,T}^\top \mathbf{C}_{i,t,2} \mathbf{S}_{i,T} - 2\mathbf{S}_{i,t}^\top \mathbf{C}_{i,t,3} \mathbf{S}_{i,T}.
$$

Lemma 7.2. *If H_0 and Assumptions 2.1 – 2.7 and 3.1 hold, then we have*

$$
(7.3) \quad \left\{N^{-1/2} \sum_{i=1}^N (H_{i,t} - a_{i,t,1}^2), \ d \le t \le T\right\} \xrightarrow{\mathcal{D}} \left\{\xi_t^{(1)}, \ \underline{t}_0 \le t \le T - \bar{t}_0\right\},
$$

where $\xi_t^{(1)}$, $\underline{t}_0 \le t \le T - \bar{t}_0$, is jointly normal with zero mean and covariance $\Gamma^{(1)}(t, t')$ and $\underline{t}_0, \bar{t}_0$ are defined in Assumption 3.1.

Proof. It follows from Assumptions 2.2, 2.6 and 2.7 that the random vectors $H_{i,t}, \underline{t}_0 \leq t \leq T - \bar{t}_0$ are independent and by Lemma 7.1 $E\,H_{i,t} = a_{i,t,1}^2$. Also, that $E\,|H_{i,t} - a_{i,t,1}^2|^{\kappa/2} \leq cE\,|\varepsilon_{i,0}|^\kappa$ and therefore we have

$$(7.4) \qquad \sum_{i=1}^N E\,|H_{i,t} - a_{i,t,1}^2|^{\kappa/2} = O(N).$$

It is easy to see that

$$\mathrm{cov}\left(H_{i,t}, H_{i,t'}\right) = E\left[\left(\mathbf{Z}_{i,t}^{-1}\mathbf{S}_{i,t} - \mathbf{Z}_{i,T}^{-1}\mathbf{S}_{i,T}\right)^\top \mathbf{C}_{i,t}\left(\mathbf{Z}_{i,t}^{-1}\mathbf{S}_{i,t} - \mathbf{Z}_{i,T}^{-1}\mathbf{S}_{i,T}\right)\right.$$
$$\left.\times \left(\mathbf{Z}_{i,t'}^{-1}\mathbf{S}_{i,t'} - \mathbf{Z}_{i,T}^{-1}\mathbf{S}_{i,T}\right)^\top \mathbf{C}_{i,t'}(\mathbf{Z}_{i,t'}^{-1}\mathbf{S}_{i,t'} - \mathbf{Z}_{i,T}^{-1}\mathbf{S}_{i,T})\right] - a_{i,t,1}^2 a_{i,t',1}^2.$$

We can assume that $\{\Gamma^{(1)}(t,t'),\ \underline{t}_0 \leq t,t' \leq T - \bar{t}_0\}$ is a non–singular matrix. If $\Gamma^{(1)}(t,t')$ is singular, one needs to prove the central limit theorem for a subvector of $\left(H_{i,\underline{t}_0}, \ldots, H_{i,T-\bar{t}_0}\right)^\top$ such that the related submatrix of $\Gamma^{(1)}(t,t')$ is regular and has the same rank as $\Gamma^{(1)}(t,t')$. Hence for any constants $\left(c_{\underline{t}_0}, c_{\underline{t}_0+1}, \ldots, c_{T-\bar{t}_0}\right) \neq \mathbf{0}$ we have

$$\lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^N \mathrm{var}\left(\sum_{v=\underline{t}_0}^{T-\bar{t}_0} c_v H_{i,v}\right) > 0.$$

Equation (7.4) implies that

$$\sum_{i=1}^N E\left|\sum_{v=\underline{t}_0}^{T-\bar{t}_0} c_v(H_{i,v} - a_{i,v,1}^2)\right|^{\kappa/2} = O(N).$$

Hence Lemma 7.2 follows from Lyapunov's theorem $\left(\text{cf. Petrov (1995, p. 122)}\right)$. □

Lemma 7.3. *If H_0 and Assumptions 2.1–2.7 hold, then we have for all $\underline{t}_0 \leq t \leq T - \bar{t}_0$ that*

$$\sum_{i=1}^N \mathbf{S}_{i,t}^\top \mathbf{C}_{i,t,1}\mathbf{R}_{i,t} = o_P(N^{1/2}), \quad \sum_{i=1}^N \mathbf{S}_{i,t}^\top \mathbf{C}_{i,t,3}\mathbf{R}_{i,T} = o_P(N^{1/2})$$

and

$$\sum_{i=1}^N \mathbf{S}_{i,T}^\top \mathbf{C}_{i,t,2}\mathbf{R}_{i,T} = o_P(N^{1/2}), \quad \sum_{i=1}^N \mathbf{R}_{i,t}^\top \mathbf{C}_{i,t,3}\mathbf{S}_{i,T} = o_P(N^{1/2}),$$

where $\underline{t}_0$ and $\bar{t}_0$ are defined in Assumption 3.1.

Proof. Using Assumptions 2.1–2.3 we conclude that $E\,\mathbf{S}_{i,t}^\top \mathbf{C}_{i,t,1}\mathbf{R}_{i,t} = 0$ and

$$(7.5) \qquad \mathrm{var}\left(\sum_{i=1}^N \mathbf{S}_{i,t}^\top \mathbf{C}_{i,t,1}\mathbf{R}_{i,t}\right) = \sum_{i=1}^N E\,(\mathbf{S}_{i,t}^\top \mathbf{C}_{i,t,1}\mathbf{R}_{i,t})^2.$$

Next we note that $(\mathbf{S}_{i,t}^\top \mathbf{C}_{i,t,1}\mathbf{R}_{i,t})^2 \leq \|\mathbf{S}_{i,t}\|^2\|\mathbf{C}_{i,t,1}\|^2\|\mathbf{R}_{i,t}\|^2 \leq c\|\mathbf{S}_{i,t}\|^2\|\mathbf{R}_{i,t}\|^2$ with some constant on account of Assumption 2.6–2.7. Using again Assumption 2.6 we get that $E\,\|\mathbf{S}_{i,t}\|^2 \leq c\sigma_i^2$ and $E\,\|\mathbf{R}_{i,t}\|^2 \leq c\|\boldsymbol{\gamma}_i\|^2$ with some constant c. On account of Assumption 2.2(ii) we conclude

$$\text{var}\left(\sum_{i=1}^{N}\mathbf{S}_{i,t}^{\top}\mathbf{C}_{i,t,1}\mathbf{R}_{i,t}\right) = O(1)\sum_{i=1}^{N}\sigma_i^2\|\boldsymbol{\gamma}_i\|^2 = O(1)\sum_{i=1}^{N}\|\boldsymbol{\gamma}_i\|^2,$$

and therefore Chebyshev's inequality and Assumption 2.4 imply the first statement in Lemma 7.3. The same argument can be used to prove the other upper bounds.

$\square$

Lemma 7.4. *If H_0 and Assumptions 2.1 and 2.3 – 2.7 hold, then we have for all $\underline{t}_0 \le t \le T - \bar{t}_0$ that*

$$\sum_{i=1}^{N}\mathbf{R}_{i,t}^{\top}\mathbf{C}_{i,t,1}\mathbf{R}_{i,t} = o_P(N^{1/2}), \quad \sum_{i=1}^{N}\mathbf{R}_{i,T}^{\top}\mathbf{C}_{i,t,2}\mathbf{R}_{i,T} = o_P(N^{1/2})$$

and

$$\sum_{i=1}^{N}\mathbf{R}_{i,t}^{\top}\mathbf{C}_{i,t,3}\mathbf{R}_{i,T} = o_P(N^{1/2}),$$

where $\underline{t}_0$ and $\bar{t}_0$ are defined in Assumption 3.1.

Proof. Using Assumptions 2.6 and 2.7 we get that

$$\left|\sum_{i=1}^{N}\mathbf{R}_{i,t}^{\top}\mathbf{C}_{i,t,\ell}\mathbf{R}_{i,t}\right| \le \sum_{i=1}^{N}\sum_{s=1}^{t}\sum_{s'=1}^{t}\left|\mathbf{x}_{i,s}^{\top}\mathbf{C}_{i,t,1}\mathbf{x}_{i,s'}\boldsymbol{\gamma}_i^{\top}\boldsymbol{\Lambda}_s\boldsymbol{\gamma}_i^{\top}\boldsymbol{\Lambda}_{s'}\right|$$

$$= O_P(1)\sum_{i=1}^{N}\|\boldsymbol{\gamma}_i\|^2,$$

and therefore the first part follows from Assumption 2.4. The remaining three cases can be proven in a similar way and therefore the details are omitted.

$\square$

Proof of Theorem 3.1. The result follows immediately from Lemmas 7.1–7.4. $\square$

To prove Theorem 3.2 we note that

$$\widehat{e}_{i,t} = w_{i,t} + r_{i,t}, \text{ where } w_{i,t} = \varepsilon_{i,t} - \mathbf{x}_{i,t}^{\top}\mathbf{Z}_{i,T}^{-1}\mathbf{S}_{i,T} \text{ and } r_{i,t} = \boldsymbol{\gamma}_i^{\top}\boldsymbol{\Lambda}_t - \mathbf{x}_{i,t}^{\top}\mathbf{Z}_{i,T}^{-1}\mathbf{J}_{i,T}$$

with

$$\mathbf{J}_{i,t} = \sum_{v=1}^{t}\mathbf{x}_{i,v}\boldsymbol{\gamma}_i^{\top}\boldsymbol{\Lambda}_v.$$

Hence

$$\widehat{e}_{i,t}^2 = w_{i,t}^2 + 2w_{i,t}r_{i,t} + r_{i,t}^2.$$

Lemma 7.5. *If H_0 and Assumptions 2.1 – 2.6 and 3.2 hold, then we have*

$$\left\{N^{-1/2}\sum_{i=1}^{N}\left(\sum_{s=1}^{t}w_{i,s}^2 - a_{i,t,2}^2\right), 1 \le t \le T\right\} \xrightarrow{\mathcal{D}} \left\{\xi_t^{(2)}, 1 \le t \le T\right\},$$

where $\xi_t^{(2)}, 1 \le t \le T$ is jointly normal with zero mean and covariance $\Gamma^{(2)}(t,t')$.

Proof. It is easy to see that $E\,w_{i,s} = \sigma_i^2(1 - \mathbf{x}_{i,s}^\top \mathbf{Z}_{i,T}^{-1}\mathbf{x}_{i,s})$. It follows from Assumptions 2.2, 2.5 and 2.6 that $E\,|w_{i,s}^2 - E\,w_{i,s}^2|^{\kappa/2} \le cE\,\varepsilon_{i,s}^\kappa$ with some c and therefore

$$(7.6) \qquad \sum_{i=1}^N E\,\left|\sum_{s=1}^t w_{i,s}^2 - a_{i,t,2}^2\right|^{\kappa/2} = O(N).$$

As in the proof of Lemma 7.2, we can assume without loss of generality that $\Gamma^{(2)}(t,r')$ is non–singular. Hence for all $(c_1, c_2, \ldots, c_T)^\top \neq \mathbf{0}$ we have that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^N \mathrm{var}\left(\sum_{t=1}^T c_t\left(\sum_{s=1}^t w_{i,s}^2 - a_{i,t,2}^2\right)\right) > 0.$$

We get from (7.6) that

$$\sum_{i=1}^N E\,\left|\sum_{t=1}^T c_t\left(\sum_{s=1}^t w_{i,s}^2 - a_{i,t,2}^2\right)\right|^{\kappa/2} = O(N).$$

Now the lemma follows from Lyapunov's theorem (cf. Petrov (1995, p. 122)) and the Cramér–Wold lemma (cf. Billingsley (1968)). $\qquad\square$

Lemma 7.6. *If H_0 and Assumptions 2.1–2.6 and 3.2 hold, then we have for all $1 \le t \le T$ that*

$$\sum_{i=1}^N r_{i,t}^2 = o_P(N^{1/2})$$

and

$$\sum_{i=1}^N r_{i,t}w_{i,t} = o_P(N^{1/2}).$$

Proof. We note that

$$r_{i,t}^2 \le 2\|\boldsymbol{\gamma}_i\|^2\|\boldsymbol{\Lambda}_t\|^2 + 2\left(\mathbf{x}_{i,t}^\top \mathbf{Z}_{i,T}^{-1}\mathbf{J}_{i,T}\right)^2$$

and by Assumption 2.4

$$\sum_{i=1}^N \|\boldsymbol{\gamma}_i\|^2\|\boldsymbol{\Lambda}_t\|^2 = o_P(N^{1/2}).$$

Also, by Assumptions 2.5 and 2.6 we have

$$\left(\mathbf{x}_{i,t}^\top \mathbf{Z}_{i,T}^{-1}\mathbf{J}_{i,T}\right)^2 \le c\|\mathbf{J}_{i,T}\|^2 \le c$$

and therefore

$$\sum_{i=1}^N \left(\mathbf{x}_{i,t}^\top \mathbf{Z}_{i,T}^{-1}\mathbf{J}_{i,T}\right)^2 = O(1)\max_{1\le s\le T}\|\boldsymbol{\Lambda}_s\|^2 \sum_{i=1}^N \|\boldsymbol{\gamma}_i\|^2 = o_P(N^{1/2}),$$

completing the proof of the first part of Lemma 7.6.

We write that

$$w_{i,t}r_{i,t} = \varepsilon_{i,t}\boldsymbol{\gamma}_i^\top \boldsymbol{\Lambda}_t - \varepsilon_{i,t}\mathbf{x}_{i,t}^\top\,\mathbf{Z}_{i,T}^{-1}\mathbf{J}_{i,T} - \mathbf{x}_{i,t}^\top\mathbf{Z}_{i,T}^{-1}\mathbf{S}_{i,T}\boldsymbol{\gamma}_i^\top\boldsymbol{\Lambda}_t + \mathbf{x}_{i,t}^\top\mathbf{Z}_{i,T}^{-1}\mathbf{S}_{i,T}\mathbf{x}_{i,t}^\top\mathbf{Z}_{i,T}^{-1}\mathbf{J}_{i,T}.$$

Using Assumptions 2.1–2.3 we get that $E\,\varepsilon_{i,t}\boldsymbol{\gamma}_{i,t}^\top\boldsymbol{\Lambda}_t = 0$ and

$$\mathrm{var}\left(\sum_{i=1}^N \varepsilon_{i,t}\boldsymbol{\gamma}_i^\top\boldsymbol{\Lambda}_t\right) = O(1)\sum_{i=1}^N \|\boldsymbol{\gamma}_i\|^2 = O(N^{1/2})$$

on account of Assumption 2.4. Similarly, $E\,\varepsilon_{i,t}\mathbf{x}_{i,t}^\top\mathbf{Z}_{i,T}^{-1}\mathbf{J}_{i,T}=0$ and

$$\mathrm{var}\left(\sum_{i=1}^{N}\varepsilon_{i,t}\mathbf{x}_{i,t}^\top\mathbf{Z}_{i,T}^{-1}\mathbf{J}_{i,T}\right)=O(N^{1/2}).$$

Repeating the arguments above, one can verify that

$$\left|\sum_{i=1}^{N}\mathbf{x}_{i,t}\mathbf{Z}_{i,T}^{-1}\mathbf{S}_{i,T}\boldsymbol{\gamma}_i^\top\boldsymbol{\Lambda}_t\right|=O_P(N^{1/4})$$

and

$$\left|\sum_{i=1}^{N}\mathbf{x}_{i,t}\mathbf{Z}_{i,T}^{-1}\mathbf{S}_{i,T}\mathbf{x}_{i,t}\mathbf{Z}_{i,T}^{-1}\mathbf{J}_{i,T}\right|=O_P(N^{1/4}),$$

completing the proof of the second part of Lemma 7.6.

$\square$

Proof of Theorem 3.2. It is an immediate consequence of Lemmas 7.5 and 7.6. $\square$

Proof of Theorem 3.3. It follows from the proof of Lemma 7.2 that

$$\sum_{i=1}^{N}H_{i,t}-A_N^{(1)}(t)=O_P(N^{1/2}),$$

without assuming the existence of the limit in Assumption 3.1, since under the conditions of Theorem 3.3

$$\limsup_{N\to\infty}\frac{1}{N}\left|\sum_{i=1}^{N}\left[E\left\{(\mathbf{Z}_{i,t}^{-1}\mathbf{S}_{i,t}-\mathbf{Z}_{i,T}^{-1}\mathbf{S}_{i,T})^\top\,bC_{i,t}(\mathbf{Z}_{i,t}^{-1}\mathbf{S}_{i,t}-\mathbf{Z}_{i,T}^{-1}\mathbf{S}_{i,T})\right.\right.\right.$$

$$\left.\left.\left.\times(\mathbf{Z}_{i,t'}^{-1}\mathbf{S}_{i,t'}-\mathbf{Z}_{i,T}^{-1}\mathbf{S}_{i,T})^\top\mathbf{C}_{i,t'}(\mathbf{Z}_{i,t'}^{-1}\mathbf{S}_{i,t'}-\mathbf{Z}_{i,T}^{-1}\mathbf{S}_{i,T})\right\}-a_{i,t,1}^2a_{i,t',1}^2\right]\right|<\infty.$$

Elementary arguments give

$$\overline{H}_{i,t}=\sum_{s,v=1}^{t}\boldsymbol{\Lambda}_s^\top\boldsymbol{\gamma}_i\mathbf{x}_{i,s}^\top(\mathbf{Z}_{i,t}^{-1}-\mathbf{Z}_{i,T}^{-1})\mathbf{C}_{i,t}(\mathbf{Z}_{i,t}^{-1}-\mathbf{Z}_{i,T}^{-1})\boldsymbol{\gamma}_i^\top\boldsymbol{\Lambda}_v,$$

where

$$\overline{H}_{i,t}=\mathbf{R}_{i,t}^\top\mathbf{C}_{i,t,1}\mathbf{R}_{i,t}+\mathbf{R}_{i,T}^\top\mathbf{C}_{i,t,2}\mathbf{R}_{i,T}-2\mathbf{R}_{i,t}^\top\mathbf{C}_{i,t,3}\mathbf{R}_{i,T}$$

$$=\boldsymbol{\gamma}_i^T\left(\sum_{j=1}^{t}\boldsymbol{\Lambda}_j\boldsymbol{x}_{i,j}^T\mathbf{Z}_{i,t}^{-1}-\sum_{j=1}^{T}\boldsymbol{\Lambda}_j\boldsymbol{x}_{i,j}^T\mathbf{Z}_{i,T}^{-1}\right)\mathbf{C}_{i,t}\left(\mathbf{Z}_{i,t}^{-1}\sum_{v=1}^{t}\boldsymbol{x}_{iv}\boldsymbol{\Lambda}_v^T-\mathbf{Z}_{i,T}^{-1}\sum_{v=1}^{T}\boldsymbol{x}_{iv}\boldsymbol{\Lambda}_v^T\right)\boldsymbol{\gamma}_i$$

$$=\sum_{j=1}^{t}\sum_{v=1}^{t}\boldsymbol{\Lambda}_j^T\boldsymbol{\gamma}_i\boldsymbol{x}_{i,j}^T\left(\mathbf{Z}_{i,t}^{-1}I\{1\le j\le t\}-\mathbf{Z}_{i,T}^{-1}\right)\mathbf{C}_{i,t}\left(\mathbf{Z}_{i,t}^{-1}I\{1\le v\le t\}-\mathbf{Z}_{i,T}^{-1}\right)\boldsymbol{x}_{iv}\boldsymbol{\gamma}_i^T\boldsymbol{\Lambda}_v$$

Hence the definition of the function $\boldsymbol{Q}$ yields

$$\frac{1}{r_N}\sum_{i=1}^{N}\overline{H}_{i,t}\;\to\;\sum_{s,v=1}^{t}\boldsymbol{\Lambda}_s^\top\boldsymbol{Q}(s,v,t)\boldsymbol{\Lambda}_v\quad\text{a.s.}.$$

We showed in the proof of Lemma 7.3 that

$$\left|\sum_{i=1}^{N}\mathbf{S}_{i,t}^{\top}\mathbf{C}_{i,t,1}\mathbf{R}_{i,t}\right| = O_P(1)\sum_{i=1}^{N}\|\boldsymbol{\gamma}_i\|^2, \qquad \left|\sum_{i=1}^{N}\mathbf{S}_{i,T}^{\top}\mathbf{C}_{i,t,2}\mathbf{R}_{i,T}\right| = O_P(1)\sum_{i=1}^{N}\|\boldsymbol{\gamma}_i\|^2$$

and

$$\left|\sum_{i=1}^{N}\mathbf{S}_{i,t}^{\top}\mathbf{C}_{i,t,3}\mathbf{R}_{i,T}\right| = O_P(1)\sum_{i=1}^{N}\|\boldsymbol{\gamma}_i\|^2, \qquad \left|\sum_{i=1}^{N}\mathbf{R}_{i,t}^{\top}\mathbf{C}_{i,t,3}\mathbf{S}_{i,T}\right| = O_P(1)\sum_{i=1}^{N}\|\boldsymbol{\gamma}_i\|^2.$$

The proof of Theorem 3.3 is now complete. $\qquad\square$

Proof of Theorem 3.4. Elementary arguments give

$$r_{i,t}^2 = \boldsymbol{\Lambda}_t^{\top}\boldsymbol{\gamma}_i\boldsymbol{\gamma}_i^{\top}\boldsymbol{\Lambda}_t - 2\sum_{v=1}^{T}\boldsymbol{\Lambda}_v^{\top}\boldsymbol{\gamma}_i\mathbf{x}_{i,v}^{\top}\mathbf{Z}_{i,T}^{-1}\mathbf{x}_{i,t}\boldsymbol{\gamma}_i^{\top}\boldsymbol{\Lambda}_t + \sum_{s,v=1}^{T}\boldsymbol{\Lambda}_s^{\top}\boldsymbol{\gamma}_i\mathbf{x}_{i,s}^{\top}\mathbf{Z}_{i,t}^{-1}\mathbf{x}_{i,t}\mathbf{x}_{i,t}^{\top}\mathbf{Z}_{i,t}^{-1}\mathbf{x}_{i,v}\boldsymbol{\gamma}_i^{\top}\boldsymbol{\Lambda}_v.$$

Using Assumption 3.5 we obtain that

$$\left\{\frac{1}{r_N}\sum_{i=1}^{N}r_{i,t}^2, \ 1 \le t \le T\right\} \to \left\{\xi_t^{(4)}, \ 1 \le t \le T\right\} \quad \text{a.s..}$$

It follows from the proof of Theorem 3.2 that

$$\sum_{i=1}^{N}w_{i,t}^2 - A_N^{(2)}(t) = O_P(N^{1/2}).$$

Minor modification of the proof of Theorem 3.3 gives

$$\sum_{i=1}^{N}w_{i,t}r_{i,t} = o_P(r_N),$$

and therefore (3.15) is proven. $\qquad\square$

Proof of Theorem 4.1 Let $\overline{\boldsymbol{\beta}}_{i,t}$ and $\overline{\boldsymbol{\beta}}_{i,T}$ denote the OLS estimators when $\boldsymbol{\gamma}_i = \mathbf{0}$, $1 \le i \le N$ and define

$$\overline{\Phi}_{i,t} = (\overline{\boldsymbol{\beta}}_{i,t} - \overline{\boldsymbol{\beta}}_{i,T})^{\top}\boldsymbol{C}_{it}(\overline{\boldsymbol{\beta}}_{i,t} - \overline{\boldsymbol{\beta}}_{i,T}) - \frac{1}{N}\sum_{j=1}^{N}(\overline{\boldsymbol{\beta}}_{j,t} - \overline{\boldsymbol{\beta}}_{j,T})^{\top}\boldsymbol{C}_{jt}(\overline{\boldsymbol{\beta}}_{j,t} - \overline{\boldsymbol{\beta}}_{j,T}).$$

Using Assumption 2.4 one can show that

$$(7.7) \qquad \left|U_N^*(t) - N^{-1/2}\sum_{i=1}^{N}\zeta_i\overline{\Phi}_{i,t}\right| = o_P(1), \quad \underline{t}_0 \le t \le \overline{t}_0.$$

Note that $\{\zeta_i\overline{\Phi}_{i,t}, \underline{t}_0 \le t \le \overline{t}_0\}$ are independent vectors with zero mean. Let $\lambda_{\underline{t}_0}, \ldots, \lambda_{T-\overline{t}_0}$ be arbitrary constant and define linear combinations

$$z_i = \zeta_i \sum_{t=\underline{t}_0}^{T-\overline{t}_0}\lambda_t\overline{\Phi}_{i,t}.$$

It follows from Assumption 3.1 along the lines of the proof of Theorem 3.1 that

$$E\left(\frac{1}{\sqrt{N}}\sum_{i=1}^{N}z_i\right)^2 \to \sum_{t=\underline{t}_0}^{T-\overline{t}_0}\sum_{s=\underline{t}_0}^{T-\overline{t}_0}\lambda_t\Gamma^{(1)}(t,s)\lambda_s$$

Also, on account of Assumption 2.2 we have that $E|z_i|^{1+\kappa/2}$ is uniformly bounded. Hence by the Cramér-Wold device (cf. Billingsley, 1968)

$$\left\{ N^{-1/2} \sum_{i=1}^{N} \zeta_i \overline{\Phi}_{i,t}, \ \underline{t}_0 \le t \le \overline{t}_0 \right\} \xrightarrow{\mathcal{D}} \left\{ \xi_t^{(1)}, \ \underline{t}_0 \le t \le \overline{t}_0 \right\}.$$

To prove the second part of the theorem we note that (7.7) holds under alternatives. It is easy to see that

$$E\left(N^{-1/2} \sum_{i=1}^{N} \zeta_i \overline{\Phi}_{i,t} \right)^2 = O(1).$$

The proof is complete. $\qquad\qquad\square$