Kybernetika

Matúš Maciak; Barbora Peštová; Michal Pešta

Structural breaks in dependent, heteroscedastic, and extremal panel data

Kybernetika, Vol. 54 (2018), No. 6, 1106–1121

Persistent URL: `http://dml.cz/dmlcz/147599`

KYBERNETIKA — VOLUME *54* (2018), NUMBER 6, PAGES 1106–1121

STRUCTURAL BREAKS IN DEPENDENT, HETEROSCEDASTIC, AND EXTREMAL PANEL DATA

Matúš Maciak, Barbora Peštová, and Michal Pešta

New statistical procedures for a change in means problem within a very general panel data structure are proposed. Unlike classical inference tools used for the changepoint problem in the panel data framework, we allow for mutually dependent panels, unequal variances across the panels, and possibly an extremely short follow up period. Two competitive ratio type test statistics are introduced and their asymptotic properties are derived for a large number of available panels. The proposed tests are proved to be consistent and their empirical properties are investigated in an extensive simulation study. The suggested testing approaches are also applied to a real data problem.

Keywords: panel data, dependence within panels, dependence between panels, changepoint, short panels, heteroscedasticity, ratio type statistics, consistency

Classification: 62H15, 62H10, 62E20, 62P05, 62F40

1. INTRODUCTION

Panel data typically occurs in situations where some covariate of interest is repeatedly measured over time simultaneously on multiple subjects—*panels* (for instance, a financial development of a set of companies, economic growth of some specific countries, or some qualitative performance of various industrial businesses). For such data generating mechanisms, it is also common that sudden changes can occur in the panels and especially the common breaks in means are wide spread phenomena. These changes are caused by some known or unknown causes and the statistical models used for the panel data estimation should have the ability to detect and estimate these structural breaks. Another crucial task is to decide whether the changepoints are indeed present in the underlying panels, or not.

From the statistical point of view, the panel data with changepoints are represented as some multivariate data points across different subjects and they are usually assessed using an ordinary least squares approach. Hypothetical changepoints are firstly detected, they are tested for their significance, and then the overall model structure is estimated using the knowledge about the existing changepoints. The available literature falls into two main categories: in the first one, the authors consider the changepoint detection problem within homogeneous panel data (see, for instance, [18, 19]) and, in the second

DOI: 10.14736/kyb-2018-6-1106

category, they deal with the changepoint detection and estimation in heterogeneous panel data [2, 9, 14]. In all these situations, however, the authors firstly need to detect existing changepoints and, later, they can adopt some tests to decide whether these structural breaks in the panels are statistically significant or not. Moreover, the panels are considered to be independent in the aforementioned literature.

On the other hand, the changepoint detection problem is mostly considered for situations where the number of panels $N \in \mathbb{N}$ and the number of observations in each panel $T \in \mathbb{N}$ are both sufficiently large—they are both supposed to tend to infinity (see [8, 4]). For practical applications, however, it may not be possible to have a long follow up period. Therefore, the changepoint estimation is also studied for the panel data, where the number of observations in the panel is fixed and does not depend on N (for instance, [1]) or it is even extremely small [16].

In this paper, we propose a statistical test where no changepoint estimation is needed apriori and the panel data are assumed to form a very general structure: the *panels are allowed to be dependent* with some common dependence factor; the panels are *heteroscedastic*; the follow-up period is *extremely short*; and *different jump magnitudes* are possible across the panels accounting also for a situation that only some panels contain the jump and the remaining ones do not. Finally, the observations within each panel may preserve some form of dependence (for instance, an autoregresive process or even GARCH sequence). We are specifically interested in testing the null hypothesis that there is no common change in the means of such general panels: the no changepoint situation can be expressed as $\tau = T$ and the corresponding alternative hypothesis is that there is at least one panel $i \in \{1, \dots, N\}$ with the jump in its mean, located at $\tau < T$, with a nonzero magnitude $\delta_i \neq 0$.

The remainder of the paper is organized as follows: in the next section, the panel data model is introduced and the main technical assumptions are provided. Statistical tests based on two competitive ratio type test statistics, which may be used to decide whether there is some common changepoint in the panel data structure or not, are introduced in Section 3 and the asymptotic properties of the tests are derived. Finite sample performance is investigated via a simulation study in Section 4 and a real data application on insurance business data is conducted in Section 5. Concluding remarks with some discussion are given in Section 6. Proofs are postponed to the Appendix.

2. CHANGEPOINT MODEL FOR PANEL DATA

The *motivation* for the model presented in this paper can be taken, for instance, from a *non-life insurance business*, where multiple insurance companies in various countries collect claim amounts paid by every insurance company each year. The data are represented by cumulative claim payments, which can be seen in terms of the panel data structure, where the given insurance company $i \in \{1, \dots, N\}$ provides the overall claim amount $Y_{i,t}$ paid at the given time $t \in \{1, \dots, T\}$ (i. e., annual payments). The follow-up period may be relatively very short (only 10–15 years) and it is not reasonable to assume that T tends to infinity as it can be assumed for the number of available companies N.

The model which we assume for the scenario described above can be expressed as

$$Y_{i,t} = \mu_i + \delta_i \mathbb{I}\{t > \tau\} + \zeta_i \xi_t + \sigma_i \varepsilon_{i,t}, \quad i = 1, \dots, N, \, t = 1, \dots, T; \qquad (1)$$

where $\mu_i \in \mathbb{R}$ are the panel specific mean parameters, $\tau \in \{1, \dots, T\}$ is some common *changepoint time* (same for all considered panels) with the corresponding jump magnitudes $\delta_i \in \mathbb{R}$. Thus, if there is some common changepoint in model (1) present at time $\tau < T$, then the corresponding panel specific means change from μ_i before the change to $\mu_i + \delta_i$ after the change. This formulation also allows for a specific case where $\delta_i = 0$ meaning no jump is present for some given panel i. The panel specific variance scaling parameters $\sigma_i > 0$ mimic heteroscedasticity of the panels. The random factors ξ_t's are used to introduce a mutual dependence between individual panels where the level of dependence is modeled by the magnitude of unknown loadings $\zeta_i \in \mathbb{R}$.

Assumption $\mathcal{A}1$. The vectors $[\varepsilon_{i,1}, \dots, \varepsilon_{i,T}]^\top$ and $[\xi_1, \dots, \xi_T]^\top$ exist on a probability space $(\Omega, \mathcal{F}, \mathsf{P})$ and are independent for $i = 1, \dots, N$. Moreover, $[\varepsilon_{i,1}, \dots, \varepsilon_{i,T}]^\top$ are iid for $i = 1, \dots, N$ with $\mathsf{E}\varepsilon_{i,t} = 0$ and $\mathsf{Var}\,\varepsilon_{i,t} = 1$, having the autocorrelation function

$$\rho_t = \mathsf{Corr}\,(\varepsilon_{i,s}, \varepsilon_{i,s+t}) = \mathsf{Cov}\,(\varepsilon_{i,s}, \varepsilon_{i,s+t}), \quad \forall s \in \{1, \dots, T-t\},$$

which is independent of the time s, the cumulative autocorrelation function

$$r(t) = \mathsf{Var} \sum_{s=1}^{t} \varepsilon_{i,s} = \sum_{|s|<t} (t - |s|)\rho_s,$$

and the shifted cumulative correlation function

$$R(t,v) = \mathsf{Cov}\left(\sum_{s=1}^{t} \varepsilon_{i,s}, \sum_{u=t+1}^{v} \varepsilon_{i,u} \right) = \sum_{s=1}^{t} \sum_{u=t+1}^{v} \rho_{u-s}, \quad t < v;$$

for all $i = 1, \dots, N$ and $t, v = 1, \dots, T$.

The sequence $\{\varepsilon_{i,t}\}_{t=1}^{T}$ can be viewed as a part of a *weakly stationary* process. Note that the dependent errors within each panel do not necessarily need to be linear processes. For example, GARCH processes as error sequences are allowed as well. The heteroscedastic random noise is modeled via the nuisance variance parameters σ_i's. For instance, they reflect the situation in actuarial practice, where bigger insurance companies are expected to have higher variability in the total claim amounts paid. The common factors ξ_t's introduce dependence among the panels. They can be though of outer drivers influencing the stochastic panel behavior in the common way. E.g., the common factors can represent impact of the economic/political/social situation on the market. On one hand, there are no moment conditions on ξ_t's whatsoever. On the other hand, if the common factors have finite variance, then the correlation between panel observations at the same time t, for $i \neq j$, becomes

$$\mathsf{Corr}\,(Y_{i,t}, Y_{j,t}) = \frac{\mathsf{Cov}\,(\zeta_i \xi_t, \zeta_j \xi_t)}{\sqrt{(\sigma_i^2 + \zeta_i^2 \mathsf{Var}\,\xi_t)(\sigma_j^2 + \zeta_j^2 \mathsf{Var}\,\xi_t)}} = \frac{\zeta_i \zeta_j}{\sqrt{(\sigma_i^2/\mathsf{Var}\,\xi_t + \zeta_i^2)(\sigma_j^2/\mathsf{Var}\,\xi_t + \zeta_j^2)}}.$$

Hence, the sign and the magnitude of the panel factor loadings ζ_i and ζ_j affect the correlation between panels $i \neq j$. If there is $\zeta_i = 0$ for some panel i, then the panel is independent of the remaining ones due to Assumption $\mathcal{A}1$.

3. TEST STATISTICS

Let us consider the model described in (1). For the practical utilization of the model, we would like to construct a statistical test to decide whether there is some common changepoint (with the corresponding jumps in the means located at the changepoint time $\tau < T$) across the given panels $i = 1, \ldots, N$, or not. The null hypothesis can be formulated as

$$H_0 : \tau = T \tag{2}$$

against a general alternative

$$H_A : \tau < T \quad \text{and} \quad \exists i \in \{1, \ldots, N\} \text{ such that } \delta_i \neq 0. \tag{3}$$

There are various types of test statistics which can be employed to perform the test given by the set of hypothesis in (2) and (3) (cumulative sum statistics, maximum type statistics, Cramér-von Mises statistics, etc.). For some practical reasons, we propose a *ratio type statistic* to test H_0 against H_A, because this type of statistic *does not require* estimation of the nuisance parameters for the common variance (only mutual ratios of σ_i's are sufficient to be known or estimated). We aim to construct a valid and completely data driven testing procedure without interfering any estimation and plugging-in estimates instead of the nuisance parameters. For a more detailed surveys on the ratio type test statistics, we refer to [5, 6, 7, 11], and [12]. Our particular *panel changepoint test statistics* are defined as

$$\mathcal{R}_N(T) = \max_{t=2,\ldots,T-2} \frac{\max_{s=1,\ldots,t} \left| \sum_{i=1}^N \left[\sum_{r=1}^s \left(Y_{i,r} - \bar{Y}_{i,t} \right) \right] \right|}{\max_{s=t,\ldots,T-1} \left| \sum_{i=1}^N \left[\sum_{r=s+1}^T \left(Y_{i,r} - \widetilde{Y}_{i,t} \right) \right] \right|}$$

and

$$\mathcal{S}_N(T) = \max_{t=2,\ldots,T-2} \frac{\sum_{s=1}^t \left\{ \sum_{i=1}^N \left[\sum_{r=1}^s \left(Y_{i,r} - \bar{Y}_{i,t} \right) \right] \right\}^2}{\sum_{s=t}^{T-1} \left\{ \sum_{i=1}^N \left[\sum_{r=s+1}^T \left(Y_{i,r} - \widetilde{Y}_{i,t} \right) \right] \right\}^2},$$

where $\bar{Y}_{i,t}$ is the average of the first t observations in panel i and $\widetilde{Y}_{i,t}$ is the average of the last $T - t$ observations in panel i, i. e.,

$$\bar{Y}_{i,t} = \frac{1}{t} \sum_{s=1}^t Y_{i,s} \quad \text{and} \quad \widetilde{Y}_{i,t} = \frac{1}{T-t} \sum_{s=t+1}^T Y_{i,s}.$$

An alternative way for testing the change in panel means could be a usage of CUSUM type statistics. For example, a maximum or minimum of a sum (not a ratio) of properly standardized or modified sums from our test statistics $\mathcal{R}_N(T)$ or $\mathcal{S}_N(T)$. The theory, which follows, can be appropriately rewritten for such cases.

3.1. Asymptotic results

Prior to deriving asymptotic properties of the test statistics, we provide assumptions on the relationship between the heterogeneous volatility and the mutual dependence of the panels.

Assumption $\mathcal{A}2$. For some $\chi > 0$,

$$\lim_{N\to\infty} \frac{\left(\sum_{i=1}^N \sigma_i^{2+\chi}\right)^2}{\left(\sum_{i=1}^N \sigma_i^2\right)^{2+\chi}} = 0$$

and $\mathsf{E}|\varepsilon_{1,t}|^{2+\chi} < \infty$, for $t \in \{1,\ldots,T\}$.

Assumption $\mathcal{A}3$.

$$\lim_{N\to\infty} \frac{\left(\sum_{i=1}^N \zeta_i\right)^2}{\sum_{i=1}^N \sigma_i^2} = 0.$$

If there exist constants $\underline{\sigma}, \overline{\sigma} > 0$, not depending on N, such that

$$\underline{\sigma} \leqslant \sigma_i \leqslant \overline{\sigma}, \quad i = 1\ldots N;$$

then the first part of Assumption $\mathcal{A}2$ is satisfied. Additionally, suppose that, e. g., $|\zeta_i| \leqslant CN^{-1/2-\kappa}$ for all i's and some $C, \kappa > 0$, then Assumption $\mathcal{A}3$ holds as well.

Now, we derive the behavior of the test statistics under the null hypothesis.

Theorem 3.1. (Under null) Under Assumptions $\mathcal{A}1-\mathcal{A}3$ and hypothesis H_0,

$$\mathcal{R}_N(T) \xrightarrow[N\to\infty]{\mathscr{D}} \max_{t=2,\ldots,T-2} \frac{\max_{s=1,\ldots,t}\left|X_s - \frac{s}{t}X_t\right|}{\max_{s=t,\ldots,T-1}\left|Z_s - \frac{T-s}{T-t}Z_t\right|}$$

and

$$\mathcal{S}_N(T) \xrightarrow[N\to\infty]{\mathscr{D}} \max_{t=2,\ldots,T-2} \frac{\sum_{s=1}^t \left(X_s - \frac{s}{t}X_t\right)^2}{\sum_{s=t}^{T-1}\left(Z_s - \frac{T-s}{T-t}Z_t\right)^2},$$

where $Z_t := X_T - X_t$ and $[X_1,\ldots,X_T]^\top$ is a multivariate normal random vector with zero mean and covariance matrix $\boldsymbol{\Lambda} = \{\lambda_{t,v}\}_{t,v=1}^{T,T}$ such that

$$\lambda_{t,t} = r(t) \quad \text{and} \quad \lambda_{t,v} = r(t) + R(t,v),\ t < v.$$

The limiting distribution depends on the unknown correlation structure of the panel changepoint model, which has to be estimated for testing purposes. The way of its estimation is shown in Subsection 3.2. Theorem 3.1 could be extended for the bootstrap version of the test, where the correlation structure need not to be known neither estimated. Thus, Theorem 3.1 can also be viewed as a *theoretical mid-step* for justification of the bootstrap add-on. Note, that in case of independent observations within the panel, the correlation structure and, hence, the covariance matrix $\boldsymbol{\Lambda}$, are both simplified such that $r(t) = t$ and $R(t,v) = 0$.

We proceed to the assumption that is needed for deriving the asymptotic behavior of the proposed test statistics under the alternative.

Assumption $\mathcal{A}4$.

$$\lim_{N\to\infty} \frac{\left(\sum_{i=1}^N \delta_i\right)^2}{\sum_{i=1}^N \sigma_i^2} = \infty.$$

Next, we show how the test statistics behave under the alternative.

Theorem 3.2. (Under alternative) If $\tau \leqslant T-3$, then under Assumptions $\mathcal{A}1$–$\mathcal{A}4$, and alternative H_A,

$$\mathcal{R}_N(T) \xrightarrow[N\to\infty]{\mathsf{P}} \infty \xleftarrow[N\to\infty]{\mathsf{P}} \mathcal{S}_N(T). \tag{4}$$

Assumption $\mathcal{A}4$ controls the trade-off between the size of breaks and the variability of errors. It may be considered as a *detectability assumption*, because it specifies the value of signal-to-noise ratio. Assumption $\mathcal{A}4$ is satisfied, for instance, if $0 < \delta \leqslant \delta_i$, $\forall i$ (a common lower changepoint threshold) and $\sigma_i \leqslant \underline{\sigma}$, $\forall i$ (a common upper variance threshold). Another suitable example of δ_i's, for the condition in Assumption $\mathcal{A}4$, can be $0 < \delta_i = KN^{-1/2+\eta}$ for some $K > 0$ and $\eta > 0$ together with $\lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^N \sigma_i^2 < \infty$. Or, a sequence $\{\sum_{i=1}^N \sigma_i^2/N\}_N$ equibounded away from infinity with $\delta_i = Ci^{\alpha-1}\sqrt{N}$ may be used as well, where $\alpha \geqslant 0$ and $C > 0$. The assumption $\tau \leqslant T-3$ means that there are at least three observations in the panel after the changepoint. It is also possible to redefine the test statistic by interchanging the numerator and the denominator of $\mathcal{R}_N(T)$ or $\mathcal{S}_N(T)$. Afterwards, Theorem 3.2 for the modified test statistic would require three observations before the changepoint, i. e., $\tau \geqslant 3$.

Theorem 3.2 says that in presence of a structural change in the panel means, the test statistics explode above all bounds. Hence, the procedures are *consistent* and the asymptotic distributions from Theorem 3.1 can be used to construct the tests.

3.2. Estimation of the correlation structure

Despite the fact that the aim of the paper is to establish the testing procedures for the detection of a panel mean change, it is necessary to construct a consistent estimate for a possible changepoint. The reason is that estimation of the covariance matrix from Theorem 3.1 requires panels as vectors with elements having common mean (i. e., without a jump). A consistent estimate of the changepoint in the panel data is proposed in [1], but under circumstances that the change occurred for sure. In our situation, we do not know whether a change occurs or not. Therefore, the estimate proposed by [1] has to be modified. If the panel means change somewhere inside $\{2,\ldots,T-1\}$, let the estimate consistently select this change. If there is no change in panel means, the estimate points out at the very last time point T with probability going to one. In other words, the value of the changepoint estimate can be equal to T, which means there is no change. This is in contrast with [1], where T is not reachable. To overcome such deficiency, we utilize the changepoint estimator

$$\widehat{\tau}_N := \arg\min_{t=1,\ldots,T} \sum_{i=1}^N \left\{ \frac{1}{t^2}\sum_{s=1}^t (Y_{i,s}-\bar{Y}_{i,t})^2 + \frac{1}{(T-t)^2}\sum_{s=t+1}^T (Y_{i,s}-\widetilde{Y}_{i,t})^2 \right\} \tag{5}$$

of the time of change τ, which was proposed by [17]. Since the number of panels may be sufficiently large, one can estimate the correlation structure of the errors $[\varepsilon_{1,1},\ldots,\varepsilon_{1,T}]^\top$ empirically. We base the errors' estimates on *residuals*

$$\widehat{e}_{i,t} := \begin{cases} Y_{i,t} - \bar{Y}_{i,\widehat{\tau}_N}, & t \leqslant \widehat{\tau}_N, \\ Y_{i,t} - \widetilde{Y}_{i,\widehat{\tau}_N}, & t > \widehat{\tau}_N. \end{cases} \tag{6}$$

One may notice that the estimators which cannot result in the last time point are less suitable in the calculation of residuals.

Then, the empirical version of the autocorrelation function is

$$\widehat{\rho}_t := \frac{1}{N(T-t)} \sum_{i=1}^{N} \frac{1}{\widehat{\sigma}_i^2} \sum_{s=1}^{T-t} \widehat{e}_{i,s}\widehat{e}_{i,s+t},$$

where $\widehat{\sigma}_i^2 := \frac{1}{T}\sum_{s=1}^{T} \widehat{e}_{i,s}^2$ is the estimate of the variance parameter σ_i^2. Consequently, the cumulative autocorrelation function and shifted cumulative correlation function are estimated by

$$\widehat{r}(t) = \sum_{|s|<t} (t-|s|)\widehat{\rho}_s \quad \text{and} \quad \widehat{R}(t,v) = \sum_{s=1}^{t} \sum_{u=t+1}^{v} \widehat{\rho}_{u-s}, \quad t<v.$$

4. SIMULATION STUDY

A simulation experiment was performed to study the *finite sample* properties of both proposed test statistics for a common change in panel means. In particular, the interest lies in the empirical *sizes* of the proposed tests under the null hypothesis and in the empirical *rejection* rate (power) under the alternative. Random samples of panel data (5000 each time) are generated from the panel changepoint model (1). The panel size is set to $T = 10$ and $T = 25$ in order to demonstrate the performance of the testing approaches in case of extremely short panels. The number of panels is considered to be $N = 50$ and $N = 200$.

The correlation structure within each panel is modeled via random vectors generated from iid, AR(1), and GARCH(1,1) sequences. The considered AR(1) process has coefficient $\phi = 0.3$. In case of GARCH(1,1) process, we use coefficients $\alpha_0 = 1$, $\alpha_1 = 0.1$, and $\beta_1 = 0.2$, which, according to [10, Example 1], gives a strictly stationary process. In all three sequences, the innovations are obtained as iid random variables from the standard normal $\mathsf{N}(0,1)$ or the Student t_5 distribution. The common loadings ξ_t's are independent having the standard Laplace distribution (the location parameter is set to 0 and the scale parameter equals 1) or the standard Cauchy distribution (the location parameter is equal 0 and the scale parameter is 1). The panel-specific factor loadings ζ_i's are chosen randomly such that they are independently and uniformly distributed on $[-0.5, 0.5]$. Simulation scenarios are produced as all possible combinations of the above mentioned settings. The covariance matrix for the asymptotic distribution from Theorem 3.1 is estimated as proposed in Subsection 3.2. To simulate the asymptotic distribution of the test statistics, 2000 multivariate random vectors are generated using the pre-estimated covariance matrix.

To access the theoretical results under H_0 numerically, Table 1 provides the empirical specificity (one minus size) of the tests for both test statistics, $\mathcal{R}_N(T)$ as well as $\mathcal{S}_N(T)$, with the significance level $\alpha = 5\%$.

T	N	innovations	factors	IID		AR(1)		GARCH(1,1)	
10	50	$\mathsf{N}(0,1)$	Laplace	.955	.956	.934	.932	.962	.968
			Cauchy	.859	.849	.902	.901	.839	.837
		t_5	Laplace	.959	.967	.933	.930	.945	.966
			Cauchy	.851	.847	.908	.901	.847	.836
	200	$\mathsf{N}(0,1)$	Laplace	.953	.954	.937	.967	.960	.939
			Cauchy	.867	.860	.926	.912	.840	.838
		t_5	Laplace	.961	.964	.934	.968	.941	.963
			Cauchy	.879	.858	.913	.905	.851	.843
25	50	$\mathsf{N}(0,1)$	Laplace	.957	.959	.933	.965	.944	.966
			Cauchy	.868	.863	.903	.899	.842	.843
		t_5	Laplace	.949	.961	.939	.967	.947	.963
			Cauchy	.867	.859	.909	.904	.856	.953
	200	$\mathsf{N}(0,1)$	Laplace	.950	.950	.938	.964	.953	.948
			Cauchy	.880	.875	.928	.914	.866	.858
		t_5	Laplace	.951	.952	.961	.966	.943	.947
			Cauchy	.889	.865	.917	.913	.863	.857

Tab. 1. Empirical specificity (1−size) of the $\mathcal{R}_T(N)$ and $\mathcal{S}_T(N)$ tests under H_0, considering the significance level of 5%.

It may be seen that both approaches are close to the theoretical value of specificity 0.95. However, slightly better performance of $\mathcal{R}_T(N)$ is noticeable—it keeps the theoretical size more firmly. As expected, the best results are achieved in case of independence within the panels, because there is no information overlap between two consecutive observations. The precision of not rejecting the null is increasing as the number of panels is getting higher and the panel is getting longer as well. The heavy tailed common factors (Cauchy-distributed) result in generally smaller specificity.

The performance of both testing procedures under H_A in terms of the empirical rejection rates is shown in Table 2, where the changepoint is set to $\tau = \lfloor T/2 \rfloor$ and the change magnitudes δ_i's are drawn independently from the uniform distribution on $[1, 3]$ for either one half of the panels (with no change for the second half) or all of them.

One can conclude that the power of both tests increases as the panel size and the number of panels increase. Moreover, higher power is obtained when a larger portion of panels is subject to have a change in their means. The test power slightly drops when switching from independent observations within the panel to dependent ones: innovations with heavier tails (i. e., t_5) yield smaller power than innovations with lighter tails. Generally, the $\mathcal{S}_T(N)$ procedure outperforms the $\mathcal{R}_T(N)$ approach in all scenarios with respect to the power of the test. The common factors with lighter tails represented by the Laplace distribution lead to higher powers.

H_A	T	N	innovations	factors	IID		AR(1)		GARCH(1,1)	
50%	10	50	$\mathsf{N}(0,1)$	Laplace	.31	.34	.30	.32	.29	.30
				Cauchy	.29	.32	.26	.29	.23	.26
			t_5	Laplace	.30	.32	.29	.31	.24	.27
				Cauchy	.28	.30	.25	.28	.20	.25
		200	$\mathsf{N}(0,1)$	Laplace	.52	.56	.49	.52	.41	.44
				Cauchy	.48	.50	.45	.48	.37	.40
			t_5	Laplace	.50	.53	.49	.51	.39	.42
				Cauchy	.46	.48	.44	.47	.35	.39
	25	50	$\mathsf{N}(0,1)$	Laplace	.69	.74	.63	.65	.60	.66
				Cauchy	.60	.61	.58	.59	.54	.56
			t_5	Laplace	.66	.69	.60	.62	.58	.59
				Cauchy	.59	.61	.56	.58	.55	.54
		200	$\mathsf{N}(0,1)$	Laplace	.79	.82	.79	.81	.75	.76
				Cauchy	.60	.61	.58	.59	.54	.56
			t_5	Laplace	.78	.80	.77	.78	.73	.75
				Cauchy	.59	.61	.56	.63	.55	.54
100%	10	50	$\mathsf{N}(0,1)$	Laplace	.79	.80	.77	.77	.70	.72
				Cauchy	.74	.75	.73	.74	.72	.73
			t_5	Laplace	.76	.77	.75	.76	.69	.68
				Cauchy	.71	.71	.70	.69	.65	.67
		200	$\mathsf{N}(0,1)$	Laplace	.95	.96	.93	.93	.89	.90
				Cauchy	.82	.83	.81	.82	.77	.78
			t_5	Laplace	.95	.96	.91	.92	.86	.88
				Cauchy	.80	.81	.79	.79	.76	.78
	25	50	$\mathsf{N}(0,1)$	Laplace	.88	.89	.87	.87	.79	.80
				Cauchy	.82	.83	.81	.80	.76	.78
			t_5	Laplace	.86	.86	.83	.85	.74	.76
				Cauchy	.79	.80	.77	.78	.73	.74
		200	$\mathsf{N}(0,1)$	Laplace	.97	.99	.96	.97	.93	.95
				Cauchy	.86	.88	.85	.87	.83	.84
			t_5	Laplace	.96	.98	.94	.95	.91	.92
				Cauchy	.83	.83	.81	.82	.79	.81

Tab. 2. Empirical sensitivity (power) of the $\mathcal{R}_T(N)$ and $\mathcal{S}_T(N)$ tests under H_A, considering the significance level of 5%.

Finally, an early changepoint is discussed very briefly. We stay with the standard normal innovations, iid observations within the panel, the common factors having the standard Laplace distribution, the factor loadings uniformly distributed on $[-0.5, 0.5]$, the size of changes δ_i being independently uniform on $[1, 3]$ in all panels, and the changepoint time to be $\tau = 3$ in case of $T = 10$ and $\tau = 5$ for $T = 25$. The empirical sensitivities of both tests for small values of τ are shown in Table 3.

T	τ	N	iid, $\mathsf{N}(0,1)$,	Laplace	T	τ	N	iid, $\mathsf{N}(0,1)$,	Laplace
10	3	50	.51	.57	25	5	50	.59	.62
		200	.84	.86			200	.91	.92

Tab. 3. Empirical sensitivity of the $\mathcal{R}_T(N)$ and $\mathcal{S}_T(N)$ tests for small values of τ under H_A, considering the significance level of 5%.

When the changepoint is not in the middle of the panel, the power of the test generally falls down. The reason for such decrease is that the left or right part of the panel possesses less observations with constant mean, which leads to a decrease of precision in the correlation estimation. Nevertheless, the $\mathcal{S}_T(N)$ test again outperforms the $\mathcal{R}_T(N)$ version and, moreover, provides solid results even for early changepoints.

5. APPLICATION IN INSURANCE INDUSTRY

As mentioned in the introduction, our primary motivation for testing the panel mean change comes from the *insurance business*. The data set is provided by the National Association of Insurance Commissioners (NAIC) database, see [13]. We concentrate on the 'Commercial auto/truck liability/medical' insurance line of business. The data collect records from $N = 157$ insurance companies (one extreme insurance company was omitted from the analysis). Each insurance company provides $T = 10$ yearly total claim amounts starting from year 1988 up to year 1997. One can consider normalizing the claim amounts by the premium received by company i in year t. That is thinking of panel data $Y_{i,t}/p_{i,t}$, where $p_{i,t}$ is the mentioned premium. This may yield a stabilization of series' variability, which corresponds to the first part of Assumption $\mathcal{A}2$. Figure 1 graphically shows series of normalized claim amounts and their logarithmic versions.

The data are considered as panel data in the way that each insurance company corresponds to one panel, which is formed by the company's yearly total claim amounts normalized by the earned premium. The length of the panel is quite short. This is very typical in insurance business, because considering longer panels may invoke incomparability between the early claim amounts and the late ones due to changing market or policies' conditions over time.

We want to test whether or not a change in the normalized claim amounts occurred in a common year, assuming that the normalized claim amounts are approximately constant in the years before and after the possible change for every insurance company. A significance level of 5% is considered. Our second ratio type test statistic gives $\mathcal{S}_{157}(10) = 10{,}544$. The asymptotic critical value is 8,698. These values mean that we do reject the null hypothesis of no change in the panel means. However, the null hypothesis is not rejected using the asymptotic test based on $\mathcal{R}_{157}(10)$, which can be explained by lower power of this test compared to the one based on $\mathcal{S}_N(T)$, see Table 4. We also try to take the decadic logarithms of claim amounts normalized by the earned premium and to consider log normalized amounts as the panel data observations. However, we reject the hypothesis of no change in the panel means (i. e., means of $\log_{10}$ normalized amounts) again, now, based on $\mathcal{S}_N(T)$ as well as on $\mathcal{R}_N(T)$.

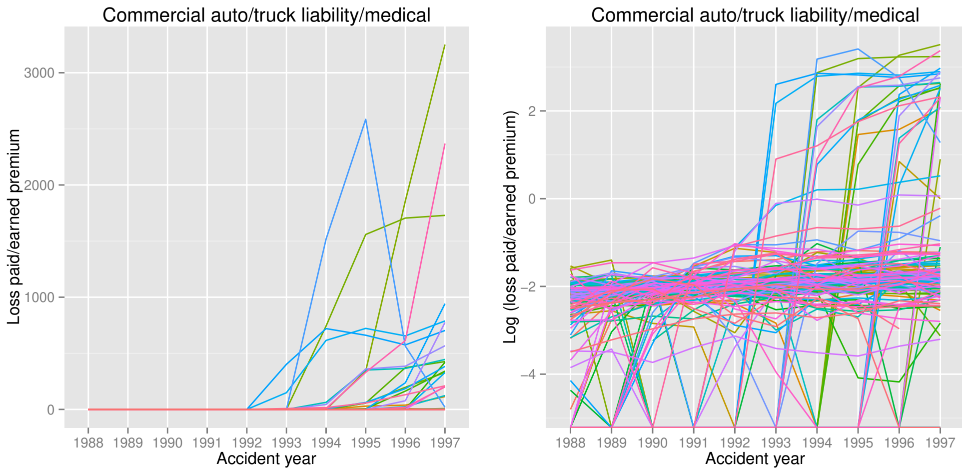

Fig. 1. Development of yearly total claim amounts normalized by the earned premium (left) together with log normalized amounts (right).

T	N	$\mathcal{R}_N(T)$	$\mathcal{S}_N(T)$	critical values	
10	157	39.9	10,544	52.4	8,698

Tab. 4. Ratio type test statistics with critical values for the 'Commercial auto/truck liability/medical' insurance.

6. CONCLUSIONS

The changepoint location problem for a *very general panel data structure* is considered in this paper: the observations within each panel are dependent, the given panels are allowed to be heteroscedastic, and even dependent among each other. The mutual dependence is modeled using some common random factors and some unknown panel specific loadings. The follow up period is allowed to be extremely short and the changepoint magnitudes may differ across the panels accounting also for a specific situation that some magnitudes are equal to zero (thus, no jump is present in such case). Another advantage of the proposed approach is that it does not require an apriori estimation of the changepoint location. We considered *two competitive ratio type test statistics* and their asymptotic properties are derived. Under the null hypothesis of no change, the test statistics weakly converge to a functional of the multivariate normal random vector with the zero mean vector and the covariance structure depending on the intra-panel covariances. Under the alternative hypothesis, both test statistics are shown to converge to infinity with the increasing number of panels and, thus, both procedures are proved to be consistent.

From the practical point of view, the general structure with heteroscedastic and possibly dependent panels with extremely short follow up periods is a lot more *realistic scenario for practical utilization* of the proposed changepoint tests than a situation with independent or even homoscedastic panels.

A simulation study illustrates that even for an *extremely small panel size* (10 observations only), both competitive test statistics perform quite well: the empirical specificity of both tests is very close to the theoretical value of one minus the significance level, but a slightly better performance is observed for the test statistic $\mathcal{R}_N(T)$ when considering some heavy tailed common factors for the mutual panel dependence. On the other hand, $\mathcal{S}_N(T)$ slightly outperforms the previous one in terms of the power, which is still comparable among various error dependence and panel dependence structures. The power increases as the number of panels gets higher. Furthermore, the sensitivity is also affected by the length of the follow up period, the proportion of panels for which a non-zero jump magnitude is observed, and the changepoint location. Longer follow up periods and higher proportions of the panels with jumps in their means imply better powers for both tests. When considering the changepoint location, then the highest power is observed for the changepoint located close to the middle of the follow up period.

The theory can be further *extended* to propose a consistent changepoint estimate, which is otherwise not needed for the tests based on the ratio type statistics. Such estimate can be used to obtain a *bootstrapped counterpart* for the asymptotic distribution of both test statistics. The ratio type test statistic allows us to omit the variance estimation and the bootstrap technique overcomes the estimation of the correlation structure. Hence, neither nuisance nor smoothing parameters are present in the whole testing process, which makes it very simple for practical use. Moreover, the whole stochastic theory behind requires relatively simple assumptions, which are not too restrictive. The whole setup can be also modified by considering a large panel size accounting also for situations with T tending to infinity. Consequently, the whole theory would lead to convergences to functionals of Gaussian processes with a covariance structure derived in a similar fashion as for fixed and small T.

A. PROOFS

Proof. [Theorem 3.1] Firstly, we show that a multivariate CLT holds for a sequence of the T-dimensional independent random vectors $\{\sigma_i[\sum_{s=1}^{1}\varepsilon_{i,s},\ldots,\sum_{s=1}^{T}\varepsilon_{i,s}]^\top\}_{i\in\mathbb{N}}$. Assumption $\mathcal{A}1$ allows us to denote $\mathsf{Var}\,[\sum_{s=1}^{1}\varepsilon_{i,s},\ldots,\sum_{s=1}^{T}\varepsilon_{i,s}]^\top=\mathbf{\Lambda}$ for all $i=1,\ldots,N$. The tth diagonal element of the covariance matrix $\mathbf{\Lambda}$ is $\mathsf{Var}\,\sum_{s=1}^{t}\varepsilon_{1,s}=r(t)$ and the upper off-diagonal element on position (t,v) is

$$\mathsf{Cov}\left(\sum_{s=1}^{t}\varepsilon_{1,s},\sum_{u=1}^{v}\varepsilon_{1,u}\right)=\mathsf{Var}\,\sum_{s=1}^{t}\varepsilon_{1,s}+\mathsf{Cov}\left(\sum_{s=1}^{t}\varepsilon_{1,s},\sum_{u=t+1}^{v}\varepsilon_{1,u}\right)=r(t)+R(t,v),$$

for $t<v$. Thus,

$$\frac{1}{\sqrt{\varsigma_N}}\sum_{i=1}^{N}\sigma_i\left[\sum_{s=1}^{1}\varepsilon_{i,s},\ldots,\sum_{s=1}^{T}\varepsilon_{i,s}\right]^\top\xrightarrow[N\to\infty]{\mathscr{D}}[X_1,\ldots,X_T]^\top\sim\mathcal{N}_T(\mathbf{0},\mathbf{\Lambda}),\tag{A.1}$$

where $\varsigma_N=\sum_{i=1}^{N}\sigma_i^2$. Indeed according to the Cramér–Wold theorem, it is sufficient to ensure that all assumptions of the one-dimensional Lyapunov CLT [3, p. 371] for triangular arrays are valid for any linear combination of the elements of the random vector $\frac{\sigma_i}{\sqrt{\varsigma_N}}[\sum_{s=1}^{1}\varepsilon_{i,s},\ldots,\sum_{s=1}^{T}\varepsilon_{i,s}]^\top$, $i\in\mathbb{N}$. For arbitrary fixed $\mathbf{b}=[b_1,\ldots,b_T]^\top\in\mathbb{R}^T$, we get

$$\sum_{i=1}^{N}\mathsf{Var}\left(\frac{\sigma_i}{\sqrt{\varsigma_N}}\mathbf{b}^\top\left[\sum_{s=1}^{1}\varepsilon_{i,s},\ldots,\sum_{s=1}^{T}\varepsilon_{i,s}\right]^\top\right)=\mathbf{b}^\top\boldsymbol{\Lambda}\mathbf{b}.$$

Moreover for some $\chi>0$, the Lyapunov's condition is satisfied, because the Jensen inequality together with Assumption $\mathcal{A}2$ give

$$\begin{aligned}&\left(\mathbf{b}^\top\boldsymbol{\Lambda}\mathbf{b}\right)^{-\frac{2+\chi}{2}}\sum_{i=1}^{N}\mathsf{E}\left|\frac{\sigma_i}{\sqrt{\varsigma_N}}\mathbf{b}^\top\left[\sum_{s=1}^{1}\varepsilon_{i,s},\ldots,\sum_{s=1}^{T}\varepsilon_{i,s}\right]^\top\right|^{2+\chi}\\ &\leqslant\left(\mathbf{b}^\top\boldsymbol{\Lambda}\mathbf{b}\right)^{-\frac{2+\chi}{2}}\varsigma_N^{-\frac{2+\chi}{2}}T^{1+\chi}\sum_{i=1}^{N}\sigma_i^{2+\chi}\sum_{t=1}^{T}|b_t|^{2+\chi}\mathsf{E}\left|\sum_{s=1}^{t}\varepsilon_{i,s}\right|^{2+\chi}\\ &\leqslant\left(\mathbf{b}^\top\boldsymbol{\Lambda}\mathbf{b}\right)^{-\frac{2+\chi}{2}}\varsigma_N^{-\frac{2+\chi}{2}}T^{1+\chi}\sum_{i=1}^{N}\sigma_i^{2+\chi}\sum_{t=1}^{T}|b_t|^{2+\chi}t^{1+\chi}\sum_{s=1}^{t}\mathsf{E}\,|\varepsilon_{i,s}|^{2+\chi}\to 0,\quad N\to\infty.\end{aligned}$$

Let us define

$$U_N(t):=\frac{1}{\sqrt{\varsigma_N}}\sum_{i=1}^{N}\sum_{s=1}^{t}(Y_{i,s}-\mu_i)\quad\text{and}\quad\widetilde{U}_N(t):=\frac{1}{\sqrt{\varsigma_N}}\sum_{i=1}^{N}\sum_{s=1}^{t}\sigma_i\varepsilon_{i,s}.$$

Under H_0 and according to (A.1), we have

$$[\widetilde{U}_N(1),\ldots,\widetilde{U}_N(T)]^\top\xrightarrow[N\to\infty]{\mathscr{D}}[X_1,\ldots,X_T]^\top.$$

Assumption $\mathcal{A}3$ yields

$$U_N(t)-\widetilde{U}_N(t)=\frac{1}{\sqrt{\varsigma_N}}\sum_{i=1}^{N}\sum_{s=1}^{t}\zeta_i\xi_t=\left(\sum_{s=1}^{t}\xi_t\right)\left(\frac{1}{\sqrt{\varsigma_N}}\sum_{i=1}^{N}\zeta_i\right)\xrightarrow[N\to\infty]{\mathsf{P}}0$$

and the Slutsky's theorem provides

$$[U_N(1),\ldots,U_N(T)]^\top\xrightarrow[N\to\infty]{\mathscr{D}}[X_1,\ldots,X_T]^\top.$$

Moreover, let us define the reverse analogue to $U_N(t)$, i. e.,

$$V_N(t):=\frac{1}{\sqrt{\varsigma_N}}\sum_{i=1}^{N}\sum_{s=t+1}^{T}(Y_{i,s}-\mu_i)=U_N(T)-U_N(t).$$

Hence,

$$\begin{aligned}U_N(s)-\frac{s}{t}U_N(t)&=\frac{1}{\sqrt{\varsigma_N}}\sum_{i=1}^{N}\left\{\sum_{r=1}^{s}\left[(Y_{i,r}-\mu_i)-\frac{1}{t}\sum_{v=1}^{t}(Y_{i,v}-\mu_i)\right]\right\}\\ &=\frac{1}{\sqrt{\varsigma_N}}\sum_{i=1}^{N}\sum_{r=1}^{s}\left(Y_{i,r}-\bar{Y}_{i,t}\right)\end{aligned}$$

and, consequently,

$$V_N(s)-\frac{T-s}{T-t}V_N(t)=\frac{1}{\sqrt{\varsigma_N}}\sum_{i=1}^{N}\left\{\sum_{r=s+1}^{T}\left[(Y_{i,r}-\mu_i)-\frac{1}{T-t}\sum_{v=t+1}^{T}(Y_{i,v}-\mu_i)\right]\right\}$$
$$=\frac{1}{\sqrt{\varsigma_N}}\sum_{i=1}^{N}\sum_{r=s+1}^{T}\left(Y_{i,r}-\widetilde{Y}_{i,t}\right).$$

Using the Cramér–Wold device, we end up with

$$\max_{t=2,\dots,T-2}\frac{\max_{s=1,\dots,t}\left|U_N(s)-\frac{s}{t}U_N(t)\right|}{\max_{s=t,\dots,T-1}\left|V_N(s)-\frac{T-s}{T-t}V_N(t)\right|}$$
$$\xrightarrow[N\to\infty]{\mathscr{D}}\max_{t=2,\dots,T-2}\frac{\max_{s=1,\dots,t}\left|X_s-\frac{s}{t}X_t\right|}{\max_{s=t,\dots,T-1}\left|(X_T-X_s)-\frac{T-s}{T-t}(X_T-X_t)\right|},$$

$$\max_{t=2,\dots,T-2}\frac{\sum_{s=1}^{t}\left[U_N(s)-\frac{s}{t}U_N(t)\right]^2}{\sum_{s=t}^{T-1}\left[V_N(s)-\frac{T-s}{T-t}V_N(t)\right]^2}$$
$$\xrightarrow[N\to\infty]{\mathscr{D}}\max_{t=2,\dots,T-2}\frac{\sum_{s=1}^{t}\left[X_s-\frac{s}{t}X_t\right]^2}{\sum_{s=t}^{T-1}\left[(X_T-X_s)-\frac{T-s}{T-t}(X_T-X_t)\right]^2}.$$

□

P r o o f. [Theorem 3.2] Let $t=\tau+1$. Firstly, we are going to focus on the numerator $\mathcal{R}_N(T)$. Due to (A.1) and Assumptions $\mathcal{A}1$–$\mathcal{A}4$, we obtain under alternative H_A

$$\frac{1}{\sqrt{\varsigma_N}}\max_{s=1,\dots,\tau+1}\left|\sum_{i=1}^{N}\left[\sum_{r=1}^{s}\left(Y_{i,r}-\bar{Y}_{i,\tau+1}\right)\right]\right|\geqslant\frac{1}{\sqrt{\varsigma_N}}\left|\sum_{i=1}^{N}\sum_{r=1}^{\tau}\left(Y_{i,r}-\bar{Y}_{i,\tau+1}\right)\right|$$
$$=\frac{1}{\sqrt{\varsigma_N}}\left|\sum_{i=1}^{N}\sum_{r=1}^{\tau}\left(\mu_i+\zeta_i\xi_r+\sigma_i\varepsilon_{i,r}-\frac{1}{\tau+1}\sum_{v=1}^{\tau+1}(\mu_i+\zeta_i\xi_v+\sigma_i\varepsilon_{i,v})-\frac{1}{\tau+1}\delta_i\right)\right|$$
$$=\frac{1}{\sqrt{\varsigma_N}}\left|\sum_{i=1}^{N}\sigma_i\sum_{r=1}^{\tau}(\varepsilon_{i,r}-\bar{\varepsilon}_{i,\tau+1})+\sum_{i=1}^{N}\sum_{r=1}^{\tau}\zeta_i\xi_r-\frac{\tau}{\tau+1}\sum_{i=1}^{N}\sum_{v=1}^{\tau+1}\zeta_i\xi_v-\frac{\tau}{\tau+1}\sum_{i=1}^{N}\delta_i\right|$$
$$=\mathcal{O}_{\mathsf{P}}(1)+o_{\mathsf{P}}(1)+\frac{\tau}{(\tau+1)\sqrt{\varsigma_N}}\left|\sum_{i=1}^{N}\delta_i\right|\xrightarrow{\mathsf{P}}\infty,\quad N\to\infty,$$

where $\bar{\varepsilon}_{i,\tau+1}=\frac{1}{\tau}\sum_{v=1}^{\tau+1}\varepsilon_{i,v}$ and $\varsigma_N=\sum_{i=1}^{N}\sigma_i^2$.

Secondly, for the numerator of $\mathcal{S}_N(T)$ it holds

$$\sum_{s=1}^{\tau+1}\left\{\frac{1}{\sqrt{\varsigma_N}}\sum_{i=1}^{N}\left[\sum_{r=1}^{s}\left(Y_{i,r}-\bar{Y}_{i,\tau+1}\right)\right]\right\}^2\geqslant\left\{\frac{1}{\sqrt{\varsigma_N}}\sum_{i=1}^{N}\left[\sum_{r=1}^{\tau+1}\left(Y_{i,r}-\bar{Y}_{i,\tau+1}\right)\right]\right\}^2$$
$$=\left[\frac{1}{\sqrt{\varsigma_N}}\sum_{i=1}^{N}\sum_{r=1}^{\tau}\left(\mu_i+\zeta_i\xi_r+\sigma_i\varepsilon_{i,r}-\frac{1}{\tau+1}\sum_{v=1}^{\tau+1}(\mu_i+\zeta_i\xi_v+\sigma_i\varepsilon_{i,v})-\frac{1}{\tau+1}\delta_i\right)\right]^2$$
$$=\frac{1}{\varsigma_N}\left[\sum_{i=1}^{N}\sigma_i\sum_{r=1}^{\tau}(\varepsilon_{i,r}-\bar{\varepsilon}_{i,\tau+1})+\sum_{i=1}^{N}\sum_{r=1}^{\tau}\zeta_i\xi_r-\frac{\tau}{\tau+1}\sum_{i=1}^{N}\sum_{v=1}^{\tau+1}\zeta_i\xi_v-\frac{\tau}{\tau+1}\sum_{i=1}^{N}\delta_i\right]^2\xrightarrow{\mathsf{P}}\infty,$$

as $N \to \infty$, again because of Assumptions $\mathcal{A}1-\mathcal{A}4$.

Finally, since there is no change after $\tau+1$ and $\tau \leqslant T-3$, then by Theorem 3.1 we have for the denominators of $\mathcal{R}_N(T)$ and $\mathcal{S}_N(T)$ the following

$$\frac{1}{\sqrt{\varsigma_N}} \max_{s=\tau+1,\ldots,T-1} \left| \sum_{i=1}^{N} \sum_{r=s+1}^{T} \left(Y_{i,r} - \widetilde{Y}_{i,\tau+1} \right) \right| \xrightarrow[N\to\infty]{\mathscr{D}} \max_{s=\tau+1,\ldots,T-1} \left| Z_s - \frac{T-s}{T-\tau} Z_{\tau+1} \right|,$$

$$\frac{1}{\sqrt{\varsigma_N}} \sum_{s=\tau+1}^{T-1} \left[\sum_{i=1}^{N} \sum_{r=s+1}^{T} \left(Y_{i,r} - \widetilde{Y}_{i,\tau+1} \right) \right]^2 \xrightarrow[N\to\infty]{\mathscr{D}} \sum_{s=\tau+1}^{T-1} \left(Z_s - \frac{T-s}{T-\tau} Z_{\tau+1} \right)^2 .$$

□

ACKNOWLEDGEMENT

Financial support through the Czech Science Foundation project GAČR No. 18-00522Y is gratefully acknowledged by Matúš Maciak. Institutional support to Barbora Peštová was provided by RVO:67985807. The research of Michal Pešta was supported by the Czech Science Foundation project GAČR No. 18-01781Y.

(Received November 27, 2017)

REFERENCES

[1] J. Bai: Common breaks in means and variances for panel data. J. Econometrics *157* (2010), 78–92. DOI:10.1016/j.jeconom.2009.10.020

[2] B. H. Baltagi, Q. Feng, and C. Kao: Estimation of heterogeneous panels with structural breaks. J. Econometrics *191* (2016), 176–195. DOI:10.1016/j.jeconom.2015.03.048

[3] P. Billingsley: Probability and Measure. Second edition. John Wiley and Sons, New York 1986.

[4] J. Chan, L. Horváth, and M. Hušková: Change-point detection in panel data. J. Statist. Plann. Inference *143* (2013), 955–970. DOI:10.1016/j.jspi.2012.11.004

[5] Z. Chen and Z. Tian: Ratio tests for variance change in nonparametric regression. Statistics *48* (2014), 1–16. DOI:10.1080/02331888.2012.708030

[6] M. Csörgő and L. Horváth: Limit Theorems in Change-Point Analysis. Wiley Publisher, Chichester 1997.

[7] L. Horváth, Z. Horváth, and M. Hušková: Ratio tests for change point detection. In: Beyond Parametrics in Interdisciplinary Research: Festschrift in Honor of Professor Pranab K. Sen (N. Balakrishnan, E. A. Peña, and M. J. Silvapulle, eds.), IMS Collections, Beachwood 2009, pp. 293–304. DOI:10.1214/193940307000000220

[8] L. Horváth and M. Hušková: Change-point detection in panel data. J. Time Series Anal. *33* (2012), 631–648. DOI:10.1111/j.1467-9892.2012.00796.x

[9] D. Kim: Estimating a common deterministic time trend break in large panels with cross section dependence. J. Econometrics *164* (2011), 310–330. DOI:10.1016/j.jeconom.2011.06.018

[10] A. M. Lindner: Stationarity, mixing, distributional properties and moments of GARCH(p,q)-processes. In: Handbook of Financial Time Series (T. G. Andersen., R. A. Davis, J. P. Kreiss, and T. Mikosch, eds.), Springer, Berlin 2009, pp. 481–496. DOI:10.1007/978-3-540-71297-8_2

[11] Y. Liu, C. Zou, and R. Zhang: Empirical likelihood ratio test for a change-point in linear regression model. Comm. Statist. Theory Methods *37* (2008), 2551–2563. DOI:10.1080/03610920802040373

[12] B. Madurkayová: Ratio type statistics for detection of changes in mean. Acta Universitatis Carolinae: Mathematica et Physicas *52* (2011), 47–58.

[13] G. G. Meyers and P. Shi: Loss Reserving Data Pulled from NAIC Schedule P. `http://www.casact.org/research/index.cfm?fa=loss_reserves_data`. [Updated September 01, 2011; Accessed June 10, 2014] (2011).

[14] M. Pesaran: Estimation and inference in large heterogeneous panels with a multifactor error structure. Econometrica *74* (2006), 967–1012. DOI:10.1111/j.1468-0262.2006.00692.x

[15] M. Pešta and Š. Hudecová: Asymptotic consistency and inconsistency of the chain ladder. Insurance Math. Econom. *51* (2012), 472–479. DOI:10.1016/j.insmatheco.2012.07.004

[16] B. Peštová and M. Pešta: Testing structural changes in panel data with small fixed panel size and bootstrap. Metrika *78* (2015), 665–689. DOI:10.1007/s00184-014-0522-8

[17] B. Peštová and M. Pešta: Change point estimation in panel data without boundary issue. Risks *5* (2017), 7. DOI:10.3390/risks5010007

[18] J. Qian and L. Su: Shrinkage estimation of regression models with multiple structural changes. Econometr. Theory *32* (2016), 1376–1433. DOI:10.1017/s0266466615000237

[19] S. De Watcher and E. Tzavalis: Detection of structural breaks in linear dynamic panel data models. Comput. Statist. Data Anal. *56* (2012), 3020–3034. DOI:10.1016/j.csda.2012.02.025

Matúš Maciak, Charles University, Faculty of Mathematics and Physics, Department of Probability and Mathematical Statistics, Sokolovská 83, 186 75 Praha 8. Czech Republic.
e-mail: matus.maciak@mff.cuni.cz

Barbora Peštová, The Czech Academy of Sciences, Institute of Computer Science, Department of Medical Informatics and Biostatistics, Pod Vodárenskou věží 271/2, 182 07 Praha 8. Czech Republic.
e-mail: pestova@cs.cas.cz

Michal Pešta, Charles University, Faculty of Mathematics and Physics, Department of Probability and Mathematical Statistics, Sokolovská 83, 186 75 Praha 8. Czech Republic.
e-mail: michal.pesta@mff.cuni.cz