Charles University in Prague

Faculty of Mathematics and Physics

COMMENTS AND ERRATA TO THE EXPERT REPORTS ON DOCTORAL THESIS

Barbora Peštová

Testing Structural Changes Using Ratio Type Statistics

Department of Probability and Mathematical Statistics

Supervised by
Prof. RNDr. Marie Hušková, DrSc.

Study program
Mathematics

Study branch
4M4 – Probability and Mathematical Statistics

Reply to Prof. RNDr. Daniela Jarušková, CSc.

Question from Chapter 6

I agree with the remark that the form of the change point estimate presented in Chapter 6 is not suitable. The assertion of Theorem 6.3 is not correct, because the corresponding proof contains a mistake—there should be an opposite sign (minus instead of plus) on 108^9 before indicator $\mathcal{I}$. Therefore, I provide here the corrected version of the whole Section 6.5.

6.5 Change point estimation

Despite the fact that the aim of the chapter is to establish testing procedures for detection of a panel mean change, it is necessary to construct a *consistent estimate* for a possible change point. There are two reasons for that: Firstly, the estimation of the covariance matrix $\mathbf{\Lambda}$ from Theorem 6.1 requires panels as vectors with elements having common mean (i.e., without a jump). Secondly, the bootstrap procedure, introduced later on, requires centered residuals to be resampled.

A consistent estimate of the change point in the panel data is proposed in Bai (2010), but under circumstances that the change occurred for sure. In our situation, we do not know whether a change occurs or not. Therefore, we modify the estimate proposed by Bai (2010) in the following way. 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 the very last time point T with probability going to one. In other words, the value of the change point estimate can be T meaning no change. This is in contrast with Bai (2010), where T is not reachable.

Let us define the estimate of τ as

$$\widehat{\tau}_N := \arg \min_{t=2,\ldots,T} \frac{1}{w(t)} \sum_{i=1}^{N} \sum_{s=1}^{t} (Y_{i,s} - \bar{Y}_{i,t})^2, \tag{6.6}$$

where $\{w(t)\}_{t=2}^{T}$ is a sequence of weights specified later on.

Now, we show the desired property of consistency for the proposed change point estimate under the following assumptions.

Assumption E1. The sequence

$$\left\{ \frac{t}{w(t)} \left(1 - \frac{r(t)}{t^2} \right) \right\}_{t=2}^{T}$$

is decreasing.

Assumption E2. There exist constants $L > 0$ and $N_0 \in \mathbb{N}$ such that

$$L < \sigma^2 \left[\frac{t}{w(t)} \left(1 - \frac{r(t)}{t^2} \right) - \frac{\tau}{w(\tau)} \left(1 - \frac{r(\tau)}{\tau^2} \right) \right] + \frac{\tau(t-\tau)}{t w(t)} \frac{1}{N} \sum_{i=1}^{N} \delta_i^2,$$

for each $t = \tau + 1, \ldots, T$ and $N \geqslant N_0$.

Assumption E3. $\lim_{N\to\infty} \frac{1}{N^2} \sum_{i=1}^{N} \delta_i^2 = 0$.

Assumption E4. $\mathsf{E}\, \varepsilon_{1,t}^4 < \infty$, $t \in \{1, \ldots, T\}$.

Theorem 6.3 (Change point estimate consistency). *Suppose that panel data $\{Y_{i,t}\}_{i,t=1}^{N,T}$ follow model (6.1). Assume that $\tau \neq 1$. Then under Assumptions P1, E1, E2, E3, and E4*

$$\lim_{N\to\infty} \mathsf{P}\left[\widehat{\tau}_N = \tau\right] = 1.$$

Proof. Let us define

$$S_N^{(i)}(t) := \frac{1}{w(t)} \sum_{s=1}^{t} (Y_{i,s} - \bar{Y}_{i,t})^2$$

and, consequently, $S_N(t) := \frac{1}{N}\sum_{i=1}^N S_N^{(i)}(t)$. Then,

$$S_N^{(i)}(t) = \begin{cases} \frac{\sigma^2}{w(t)}\sum_{s=1}^t(\varepsilon_{i,s}-\bar{\varepsilon}_{i,t})^2, & t \leqslant \tau,\\ \frac{1}{w(t)}\left[\sum_{s=1}^\tau(\sigma\varepsilon_{i,s}-\sigma\bar{\varepsilon}_{i,t}-\frac{t-\tau}{t}\delta_i)^2+\sum_{s=\tau+1}^t(\sigma\varepsilon_{i,s}-\sigma\bar{\varepsilon}_{i,t}+\frac{\tau}{t}\delta_i)^2\right], & t > \tau; \end{cases}$$

where $\bar{\varepsilon}_{i,t} = \frac{1}{t}\sum_{s=1}^t \varepsilon_{i,s}$. By the definition of the cumulative autocorrelation function, we have for $2 \leqslant t \leqslant \tau$

$$\mathsf{E}\, S_N^{(i)}(t) = \frac{\sigma^2}{w(t)}\sum_{s=1}^t \mathsf{E}\,(\varepsilon_{i,s}-\bar{\varepsilon}_{i,t})^2 = \frac{\sigma^2}{w(t)}\sum_{s=1}^t\left[1-\frac{2}{t}\sum_{r=1}^t \mathsf{E}\,\varepsilon_{i,s}\varepsilon_{i,r}+\frac{1}{t^2}r(t)\right] = \frac{\sigma^2}{w(t)}\left(t-\frac{r(t)}{t}\right).$$

In the other case when $t > \tau$, one can calculate

$$\begin{aligned}\mathsf{E}\, S_N^{(i)}(t) &= \frac{\sigma^2}{w(t)}\left(t-\frac{r(t)}{t}\right)+\frac{\tau}{w(t)}\left(\frac{t-\tau}{t}\right)^2\delta_i^2+\frac{t-\tau}{w(t)}\left(\frac{\tau}{t}\right)^2\delta_i^2\\ &= \frac{\sigma^2 t}{w(t)}\left(1-\frac{r(t)}{t^2}\right)+\frac{\tau(t-\tau)}{tw(t)}\delta_i^2.\end{aligned}$$

Realize that $S_N^{(i)}(t) - \mathsf{E}\, S_N^{(i)}(t)$ are independent with zero mean for fixed t and $i = 1,\ldots,N$. Due to Assumption E4, for $2 \leqslant t \leqslant \tau$ it holds

$$\mathsf{Var}\, S_N(t) = \frac{1}{N^2}\sum_{i=1}^N \frac{\sigma^4}{w^2(t)}\mathsf{Var}\left[\sum_{s=1}^t(\varepsilon_{i,s}-\bar{\varepsilon}_{i,t})^2\right] = \frac{1}{N}C_1(t,\sigma),$$

where $C_1(t,\sigma) > 0$ is some constant not depending on N. If $t > \tau$, then

$$\begin{aligned}\mathsf{Var}\, S_N(t) &= \frac{1}{N^2}\sum_{i=1}^N\frac{1}{w^2(t)}\mathsf{Var}\left[\sigma^2\sum_{s=1}^\tau(\varepsilon_{i,s}-\bar{\varepsilon}_{i,t})^2-2\frac{t-\tau}{t}\sigma\delta_i\sum_{s=1}^\tau(\varepsilon_{i,s}-\bar{\varepsilon}_{i,t})\right.\\ &\qquad\left.+\left(\frac{t-\tau}{t}\right)^2\delta_i^2+\sigma^2\sum_{s=\tau+1}^t(\varepsilon_{i,s}-\bar{\varepsilon}_{i,t})^2+2\frac{\tau}{t}\sigma\delta_i\sum_{s=\tau+1}^t(\varepsilon_{i,s}-\bar{\varepsilon}_{i,t})+\left(\frac{\tau}{t}\right)^2\delta_i^2\right]\\ &\leqslant \frac{1}{N}C_2(t,\tau,\sigma)+\frac{1}{N^2}C_3(t,\tau,\sigma)\sum_{i=1}^N\delta_i^2+\frac{1}{N^2}C_4(t,\tau,\sigma)\left|\sum_{i=1}^N\delta_i\right|,\end{aligned}$$

where $C_j(t,\tau,\sigma) > 0$ does not depend on N for $j = 2,3,4$.

The Chebyshev inequality provides $S_N(t) - \mathsf{E}\, S_N(t) = O_{\mathsf{P}}\left(\sqrt{\mathsf{Var}\, S_N(t)}\right)$ as $N \to \infty$. According to Assumption E3 and the Cauchy-Schwarz inequality, we have

$$\frac{1}{N^2}\left|\sum_{i=1}^N\delta_i\right| \leqslant \frac{1}{N}\sqrt{\frac{1}{N}\sum_{i=1}^N\delta_i^2} \to 0, \quad N \to \infty.$$

Since the index set $\{2,\ldots,T\}$ is finite and τ is finite as well, then

$$\max_{2\leqslant t\leqslant T}\mathsf{Var}\, S_N(t) \leqslant \frac{1}{N}K_1(\sigma)+K_2(\sigma)\frac{1}{N^2}\sum_{i=1}^N\delta_i^2+K_3(\sigma)\frac{1}{N^2}\left|\sum_{i=1}^N\delta_i\right| \leqslant \frac{1}{N}K_4(\sigma),$$

where $K_j(\sigma) > 0$ are constants not depending on N for $j = 1,2,3,4$. Thus, we also have uniform stochastic boundedness, i.e.,

$$\max_{2\leqslant t\leqslant T}|S_N(t)-\mathsf{E}\, S_N(t)| = O_{\mathsf{P}}\left(\frac{1}{\sqrt{N}}\right), \quad N \to \infty.$$

Adding and subtracting, one has

$$
\begin{aligned}
S_N(t) - S_N(\tau) &= S_N(t) - \mathsf{E}\, S_N(t) - [S_N(\tau) - \mathsf{E}\, S_N(\tau)] + \mathsf{E}\, S_N(t) - \mathsf{E}\, S_N(\tau) \\
&\geqslant -2 \max_{2\leqslant r\leqslant T} |S_N(r) - \mathsf{E}\, S_N(r)| + \mathsf{E}\, S_N(t) - \mathsf{E}\, S_N(\tau) \\
&= -2 \max_{2\leqslant r\leqslant T} |S_N(r) - \mathsf{E}\, S_N(r)| + \sigma^2 \left[\frac{t}{w(t)} \left(1 - \frac{r(t)}{t^2}\right) - \frac{\tau}{w(\tau)} \left(1 - \frac{r(\tau)}{\tau^2}\right) \right] \\
&\quad + \mathcal{I}\{t > \tau\} \frac{\tau(t-\tau)}{t w(t)} \frac{1}{N} \sum_{i=1}^N \delta_i^2 .
\end{aligned}
$$

The above inequality holds for each $t \in \{2, \ldots, T\}$ and, particularly, it holds for $\widehat{\tau}_N$. Note that $\widehat{\tau}_N = \arg\min_t S_N(t)$. Hence, $S_N(\widehat{\tau}_N) - S_N(\tau) \leqslant 0$. Therefore,

$$
\begin{aligned}
&2\sqrt{N} \max_{2\leqslant r\leqslant T} |S_N(r) - \mathsf{E}\, S_N(r)| \\
&\geqslant \sqrt{N} \left\{ \sigma^2 \left[\frac{\widehat{\tau}_N}{w(\widehat{\tau}_N)} \left(1 - \frac{r(\widehat{\tau}_N)}{\widehat{\tau}_N^2}\right) - \frac{\tau}{w(\tau)} \left(1 - \frac{r(\tau)}{\tau^2}\right) \right] + \mathcal{I}\{\widehat{\tau}_N > \tau\} \frac{\tau(\widehat{\tau}_N - \tau)}{\widehat{\tau}_N w(\widehat{\tau}_N)} \frac{1}{N} \sum_{i=1}^N \delta_i^2 \right\}. \quad (6.7)
\end{aligned}
$$

If $\widehat{\tau}_N > \tau$ almost surely for infinitely many N, then the left hand side of (6.7) is $O_{\mathsf{P}}(1)$ as $N \to \infty$, but the right hand side is unbounded because of Assumption E2. If $\widehat{\tau}_N < \tau$ almost surely for infinitely many N, then due to the monotonicity Assumption E1

$$
0 \xleftarrow[N\to\infty]{\mathsf{P}} 2 \max_{2\leqslant r\leqslant T} |S_N(r) - \mathsf{E}\, S_N(r)| \geqslant \sigma^2 \left[\frac{\widehat{\tau}_N}{w(\widehat{\tau}_N)} \left(1 - \frac{r(\widehat{\tau}_N)}{\widehat{\tau}_N^2}\right) - \frac{\tau}{w(\tau)} \left(1 - \frac{r(\tau)}{\tau^2}\right) \right] > 0,
$$

which is a contradicting conclusion. Hence, $\mathsf{P}[\widehat{\tau}_N = \tau] \to 1$ as $N \to \infty$. □

Assumption E2 assures that the values of changes have to be large enough compared to the variability of the random noise in the panels and to the strength of dependencies within the panels as well. Assumption E3 is needed to control the asymptotic boundedness of the variability of $\frac{1}{w(t)} \sum_{i=1}^N \sum_{s=1}^t (Y_{i,s} - \bar{Y}_{i,t})^2$, because a finite T cannot do that.

Assumptions E2 and E3 are satisfied for $0 < \delta \leqslant \delta_i < \Delta, \forall i$ (a common lower and upper bound for the change amount) and suitable σ, $r(t)$, and $w(t)$. The monotonicity Assumption E1 in not very restrictive at all. For example in case of independent observations within the panel (i.e., $r(t) = t$) and weight function $w(t) = t^q$, $q \geqslant 2$, this assumption is automatically fulfilled, since sequence $\{t^{1-q} - t^{-q}\}_{t=2}^T$ is decreasing. This also gives us an idea how to choose weights $w(t)$.

If one is interested in sensitivity of the change point estimate (i.e., what is the size of the change that can be estimated), let us consider the following model scenario: $T = 10$, $\tau = 5$, $\sigma = 0.1$, independent observations within the panel, and $w(t) = t^2$. Then, Assumption E2 is satisfied if $\frac{1}{N}\sum_{i=1}^N \delta_i^2 > 0.029$ for all $N \geqslant N_0$. In case of a common value of $\delta = \delta_i$ for all i, we need $\delta > \sqrt{0.029} \approx 0.170$.

Assumption E2 can be considered as too complicated. Therefore, one can replace it by the following simpler, but more restrictive assumption.

Assumption E5.

$$
\lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^N \delta_i^2 = \infty.
$$

On one hand, this assumption might be considered as too strong, because a common fixed (not depending on N) value of $\delta = \delta_i$ for all i does not fulfill Assumption E5. On the other hand, Assumption E5 is satisfied when $\delta_j^2/N \to \infty$ as $N \to \infty$ for some $j \in \mathbb{N}$ and $\delta_i = 0$ for all $i \neq j$. This stands for a situation when all the panels do not change in mean except one panel having a sufficiently large change in mean with respect to the number of panels.

Various competing consistent estimates of a possible change point can be suggested, e.g., the maximizer of $\sum_{i=1}^N \left[\sum_{s=1}^t (Y_{i,s} - \bar{Y}_{i,T})\right]^2$. To show the consistency, one needs to postulate different assumptions on the cumulative autocorrelation function and shifted cumulative correlation function compared to Theorem 6.3 and this may be rather complex.

Reply to Doc. RNDr. Zuzana Prášková, CSc.

Question from Chapter 6

I agree with the first two remarks. The above stated correction clarifies the issues.

Regarding the third remark, the coincidence of the asymptotic distributions from Theorem 6.1 and Theorem 6.6 is clear due to the last line of the proof of Theorem 6.7 and the proof of Theorems 6.1 (definitions of $U_N(t)$ and $V_N(t)$) and Theorem 6.6 (definitions of $\widehat{U}_N(t)$ and $\widehat{V}_N(t)$). Alternatively, one may derive the law of $X_s - \frac{s}{t}X_t$ from Theorem 6.1 and the law of $\mathcal{X}_s - \frac{s}{t}\mathcal{X}_t$ from Theorem 6.6 in case $T = \tau$ (i.e., under H_0). After some calculation, one concludes that these two distributions are the same.

R software was used for all computations, simulations, and data analyses. The source code is not directly provided, however algorithms are included for all resampling methods used in the thesis. These algorithms are independent on programming language. For other computations, simple and standard methods are used (e.g, when computing rejection rates) and the test statistics are evaluated directly as prescribed by mathematical formulas.

I agree that there are some typos:

- 50_6: It should be $\delta = 1$ and $\delta = 2$ instead of $\delta = 0.1$ and $\delta = 0.2$.
- There is doubled bibliography entry "Hušková et al (2008)".

References

Bai, J. (2010). Common breaks in means and variances for panel data. *Journal of Econometrics*, 157(1):78–92.