O detekci změn v panelových datech

Marie Hušková and coauthors

spoluautoři: J. Antoch, L. Horváth, J. Hanousek, S. Wang

ROBUST 2018, Rybník

Outline

O detekci změn v panelových datech
Marie Hušková a spoluautoři
Outline
Introduction
N large, T fixed or moderate
Tests and estimators
Application

Outline

Outline

I. Change point detection in panel data – location model

The considered model:

$$X_{i,t} = \mu_i + \delta_i I\{t > t_0\} + \varepsilon_{i,t}, \quad 1 \le i \le N, 1 \le t \le T, \qquad (1)$$

$\varepsilon_{i,t}$ – error terms with $E\varepsilon_{i,t} = 0$ for all i and t plus additional properties

$t_0 = \lfloor T\theta_0 \rfloor, \theta_0 \in [0, 1)$ – the time of change, unknown

μ_i and $\delta_i = \delta_{i,N}$ — might be non-random or random

μ_i changes to $\mu_i + \delta_{i,N}$ in panel i at time t_0

no parameter of factor loading in i-th panel

Typically, both T and N are assumed to be large.

At first we consider both T and N large but then N large and T fixed or moderate.

(i) The null hypothesis

$$H_0: \quad \delta_i = 0 \quad \text{for all } 1 \le i \le N.$$

(ii)If H_0 rejected change point t_0 should be estimated

Motivation: financial econometrics (influence of crisis, influence of change of rules,...)

Tests – Chan, Horváth, Hušková (2013)

$$V_{N,T}(k) \frac{1}{N^{1/2}} \frac{T}{k(T-k)} \sum_{i=1}^{N} \left[\frac{1}{\sigma_i^2} \left(S_i(k) - \frac{k}{T} S_i(T) \right)^2 - \frac{k(T-k)}{T} \right],$$

$$S_i(k) = \sum_{j=1}^{k} X_{i,j}$$

$\max_k |V_{N,T}(k)|$ – test statistic

For large values the null hypothesis is rejected
Approximation of critical values based on limit distribution ($\max(T < N) \to \infty$)

limit distribution under H_0 - supremum of Gaussian process

Estimators

Model:

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

$\gamma_i \eta_t$ – common factor –they common factors is negligible, it is dominating represents dependence among panels (unknown)-it can influence the limit behavior of the estimator

The estimator $\widehat{t}_{N,T}$ for t_0 is defined as a maximizers of the sum of the CUSUM processes across the panels:

$$\widehat{t}_{N,T} = \text{argmax}_{1 \le t < T} \Big| \sum_{i=1}^{N} \Big(S_i(t) - \frac{k}{T} S_i(T)\Big)\Big|$$

Limit behavior of the estimator studied under several setups (the influence of common factor is negligible or it is dominating or both errors and common factors influence)

Bai (2010), Horváth et al (2017)

Outline

Detection of a change in panel data when N is large and T is fixed

Model:

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

where $\mathbf{x}_{i,t} = (x_{i,t}(1), \dots, x_{i,t}(d))^T$ random or nonrandom
$\boldsymbol{\beta}_i \in R^d$ an unknown regression vector in the i^{th} panel
The regressor in the i^{th} panel changes from $\boldsymbol{\beta}_i$ to $\boldsymbol{\beta}_i + \boldsymbol{\delta}_i$ at time t_0, where t_0 is unknown.

Possible motivation: Test for breaks in coefficients for the US mutual fund return data around the sub-prime crisis – January 2006 to February 2010 this corresponds with other studies (e.g. Dick-Nielsen, Feldhtter, and Lando, 2012). The Fama-French three factor model is used .
the used estimators:

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

$$\mathbf{Y}_{i,t} = (y_{i,1}, y_{i,2}, \dots, y_{i,t})^T$$
$$\mathbf{X}_{i,t} = (\mathbf{x}_{i,1}, \mathbf{x}_{i,2}, \dots, \mathbf{x}_{i,t})^T$$
$$\mathbf{Z}_{i,t} = \mathbf{X}_{i,t}^T \mathbf{X}_{i,t},$$

Tests and estimators

Our interest to test the null hypothesis there is no change, i.e.

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

where $\|\cdot\|$ denotes the Euclidean norm in R^d
and to estimate the change point t_0.

- Assumptions on the error terms $e_{i,t}$:

$$e_{i,t} = \varepsilon_{i,t} + \gamma_i^T \boldsymbol{\Lambda}_t$$

$\boldsymbol{\varepsilon}_i = (\varepsilon_{i,1}, \ldots, \varepsilon_{i,T})$, $i = 1 \ldots, N$,– independent zero mean random vectors and $var(\varepsilon_{i,t}) = \sigma_i^2$, $E\varepsilon_{i,t_1}\varepsilon_{i,t_2} = 0$ for all i and $1 \le t_1 < t_2 \le T$ plus some higher moments finite.

$\{\boldsymbol{\lambda}_t,\ t = 1, \ldots, T\}$ and $\{\boldsymbol{\varepsilon}_i,\ i = 1, \ldots, N\}$ are independent

γ_i – loading factors and $\boldsymbol{\Lambda}_t$ – common factors (random vectors)
– they make problem neither of these factors are known

- Assumptions on $\mathbf{x}_{it}$ – uniformly bounded and there exit the inverse matrices of $\sum_{j=1}^{t} \mathbf{x}_{i,j}\mathbf{x}_{ij}^{T}$
- $\mathbf{C}_{it}$ for all i, t are symmetric non-negative definite and uniformly bounded

Possible test statistics $1 < t < T$, T – finite:

$$U_N(t) = \sum_{i=1}^{N} \left(\widehat{\boldsymbol{\beta}}_{i,t} - \widehat{\boldsymbol{\beta}}_{i,T}\right)^{T} \mathbf{C}_{i,t} \left(\widehat{\boldsymbol{\beta}}_{i,t} - \widehat{\boldsymbol{\beta}}_{i,T}\right)$$

$\widehat{\boldsymbol{\beta}}_{i,t}$ – LSE of $\boldsymbol{\beta}_{i,t}$ based on $Y_{i,j}$, $j = 1, \dots, t$

With particular choices of $\mathbf{C}_{i,t}$ we get, e.g.,

$$U_{N,1}(t) = \sum_{i=1}^{N} \left(\widehat{\boldsymbol{\beta}}_{i,t} - \widetilde{\boldsymbol{\beta}}_{i,t}\right)^{T} \mathbf{C}_{i,t,1} \left(\widehat{\boldsymbol{\beta}}_{i,t} - \widetilde{\boldsymbol{\beta}}_{i,t}\right),$$

$$U_{N,2}(t) = \sum_{i=1}^{N} \left(\sum_{j=1}^{t} \mathbf{x}_{i,j}(Y_{i,j} - \mathbf{x}_{i,j}^{T}\widehat{\boldsymbol{\beta}}_{i,T})\right)^{T}$$

$$\times \mathbf{C}_{i,t,2} \left(\sum_{j=1}^{t} \mathbf{x}_{i,j}(Y_{i,j} - \mathbf{x}_{i,j}^{T}\widehat{\boldsymbol{\beta}}_{i,T})\right)$$

Test based on

$$\max_t |U_N(t) - E_{H_0} U_N(t)|$$

Notice that under H_0 and negligibility of common factors

$$E_{H_0} U_N(t) = \sigma_i^2 tr\{\boldsymbol{C}_{i,t}(\boldsymbol{Z}_{i,t}^{-1} - \boldsymbol{Z}_{i,T}^{-1})\}$$

$$\boldsymbol{Z}_{i,t} = \sum_{j=1}^{t} \boldsymbol{x}_{i,j}\boldsymbol{x}_{i,j}^T$$

σ_i^2 estimated by $\widehat{\sigma}_i^2$

Under null hypothesis, some assumptions and "small"common factors, i.e.

$$\lim_{N\to\infty} \frac{1}{N^{1/2}} ||\gamma_i||^2 = 0$$

then

$(U_N(t) - E_{H_0} U_N(t)) \frac{1}{\sqrt{var_{H_0}(U_N(t))}}$ has asymptotically $N(0,1)$ as $N \to \infty$

Test statistics:

$$\overline{u}_N = \max_t \frac{1}{\sqrt{N}} |U_N(t) - \sigma_i^2 tr\{\mathbf{C}_{i,t}(\mathbf{Z}_{i,t}^{-1} - \mathbf{Z}_{i,T}^{-1})\}|$$

$$\widehat{u}_N = \max_t \frac{1}{\sqrt{N}} |U_N(t) - \widehat{\sigma}_i^2 tr\{\mathbf{C}_{i,t}(\mathbf{Z}_{i,t}^{-1} - \mathbf{Z}_{i,T}^{-1})\}|$$

For large values the null hypothesis is rejected, critical values needed.

Approximation of critical values obtained through wild bootstrap:

$$\phi_{i,t} = (\widehat{\boldsymbol{\beta}}_{i,t} - \widehat{\boldsymbol{\beta}}_{i,T})^T \mathbf{C}_{i,t}(\widehat{\boldsymbol{\beta}}_{i,t} - \widehat{\boldsymbol{\beta}}_{i,T})$$

$$\phi_{i,t}^* = \xi_i(\phi_{i,t} - \overline{\phi}_t), \quad \overline{\phi}_t = \frac{1}{N}\sum_{i=1}^{N} \phi_{i,t}$$

$\xi_i, \ldots, \xi_i$ –i.i.d. r.v. with zero mean and unit variance independent on $Y_{i,t}$

$$\widehat{u}_N^* = \max_t \frac{1}{\sqrt{N}} |\sum_{i=1}^{N} \phi_{i,t}^*|$$

By simulations this is working reasonably well .

Alternatives

Choose
$\mathbf{C}_{i,t} = \mathbf{Z}_{i,t}\mathbf{G}_{i,T}\mathbf{Z}_{i,t}$ where $\mathbf{G}_{i,T}$ is semi-definite
and if

$$\frac{1}{\sqrt{N}}\sum_i \delta_i^T \left(\mathbf{Z}_{i,t_0}\mathbf{Z}_{i,T}^{-1}\widetilde{\mathbf{Z}}_{i,t_0}\mathbf{G}_{iT}\widetilde{\mathbf{Z}}_{i,t_0}\mathbf{Z}_{i,T}^{-1}\mathbf{Z}_{i,t_0}\right)\delta_i \to \infty$$

then the test is consistent.

Simulations

Linear model

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

with $d = 2, \boldsymbol{\beta} = (1,2)^T$

$\mathbf{C}_{i,t} = \mathbf{Z}_{i,t}\mathbf{Z}_{i,t}.$

$\mathbf{x}_{i,t}$ – $N(1,1)$ independent

$\mathbf{e}_i = (e_{i,1}, \ldots, e_{i,T})^T$, $i = 1, \ldots, N$ –independent random vector s normal distributions with zero mean and various dependence structures

A. Empirical versus theoretical significance level

TABLE 1.1. Simulation Result for H_0

	$T \backslash N$	σ_i^2 known N100	N200	N500	N1000	σ_i^2 estimated by $\hat{\sigma}_i^2$ N100	N200	N500	N1000
IID	*T50*	0.045	0.049	0.051	0.055	0.037	0.040	0.052	0.048
	T100	0.049	0.044	0.044	0.051	0.044	0.045	0.041	0.033
	T200	0.056	0.045	0.047	0.043	0.039	0.049	0.049	0.043
Unequal Variance	*T50*	0.042	0.041	0.039	0.047	0.032	0.036	0.032	0.043
	T100	0.039	0.028	0.046	0.045	0.039	0.041	0.030	0.040
	T200	0.040	0.041	0.039	0.033	0.037	0.035	0.051	0.039
GARCH(1,1)	*T50*	0.045	0.037	0.056	0.048	0.020	0.036	0.027	0.029
	T100	0.038	0.048	0.051	0.042	0.031	0.025	0.037	0.045
	T200	0.040	0.044	0.040	0.037	0.054	0.047	0.049	0.040

B. Empirical power

TABLE 2.1. Simulation Result for H_A with $\delta = 0.1$

		change in intercept $\boldsymbol{\delta} = (0.1, 0)^\top$ number of panels				change in slope $\boldsymbol{\delta} = (0, 0.1)^\top$ number of panels			
		N100	***N200***	***N500***	***N1000***	***N100***	***N200***	***N500***	***N1000***
IID		σ_i^2 known							
	T50	0.047	0.086	0.175	0.37	0.151	0.377	0.834	0.987
	T100	0.112	0.26	0.648	0.934	0.59	0.933	1	1
	T200	0.403	0.792	0.994	1	0.99	1	1	1
		σ_i^2 estimated by $\hat{\sigma}_i^2$							
	T50	0.042	0.071	0.172	0.35	0.139	0.35	0.765	0.981
	T100	0.113	0.264	0.625	0.904	0.569	0.914	0.999	1
	T200	0.408	0.791	0.998	1	0.986	1	1	1
Unequal Variance		σ_i^2 known							
	T50	0.043	0.068	0.161	0.339	0.123	0.308	0.787	0.972
	T100	0.097	0.235	0.598	0.901	0.559	0.903	1	1
	T200	0.402	0.736	0.993	1	0.983	1	1	1
		σ_i^2 estimated by $\hat{\sigma}_i^2$							
	T50	0.036	0.046	0.149	0.284	0.13	0.307	0.765	0.976
	T100	0.088	0.21	0.576	0.888	0.522	0.885	0.999	1
	T200	0.376	0.754	0.997	1	0.987	1	1	1
GARCH(1,1)		σ_i^2 known							
	T50	0.041	0.06	0.123	0.295	0.098	0.251	0.693	0.96
	T100	0.076	0.219	0.55	0.884	0.525	0.895	0.999	1
	T200	0.423	0.759	0.994	1	0.983	1	1	1
		σ_i^2 estimated by $\hat{\sigma}_i^2$							
	T50	0.027	0.043	0.124	0.269	0.104	0.277	0.666	0.96
	T100	0.085	0.19	0.541	0.892	0.504	0.878	1	1
	T200	0.38	0.785	0.99	1	0.988	1	1	1

C. Changes in some panels

TABLE 3.1. Simulation Result for H_A with $\delta = 0.25$ and $\upsilon = 0.5$

		change in intercept $\boldsymbol{\delta} = (0.25, 0)^\top$ number of panels				change in slope $\boldsymbol{\delta} = (0, 0.25)^\top$ number of panels			
		N100	*N200*	*N500*	*N1000*	*N100*	*N200*	*N500*	*N1000*
		sigmasq known							
	T50	0.251	0.535	0.942	1	0.884	0.998	1	1
	T100	0.789	0.985	1	1	1	1	1	1
	T200	0.999	1	1	1	1	1	1	1
IID		sigmasq unknown							
	T50	0.227	0.52	0.939	1	0.895	1	1	1
	T100	0.778	0.986	1	1	1	1	1	1
	T200	1	1	1	1	1	1	1	1
		sigmasq known							
	T50	0.24	0.518	0.928	0.999	0.876	1	1	1
	T100	0.751	0.985	1	1	1	1	1	1
	T200	1	1	1	1	1	1	1	1
nequalVariance		sigmasq unknown							
	T50	0.18	0.467	0.924	0.999	0.839	0.999	1	1
	T100	0.735	0.977	1	1	0.999	1	1	1
	T200	0.997	1	1	1	1	1	1	1
		sigmasq known							
	T50	0.164	0.44	0.9	0.996	0.831	0.991	1	1
	T100	0.736	0.973	1	1	0.999	1	1	1
	T200	1	0.999	1	1	1	1	1	1
GARCH(1,1)		sigmasq unknown							
	T50	0.143	0.378	0.879	0.995	0.81	0.988	1	1
	T100	0.712	0.972	1	1	1	1	1	1
	T200	0.998	1	1	1	1	1	1	1

Influential common factors

Assumptions:

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

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

$1 \le s, v \le z \le T$, 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 $\mathbf{Q}(s_0, v_0, z_0)$ is different from 0.

Then under H_0 and the above assumptions

$$\{\frac{1}{r_N}(U_N(t) - A_N^{(1)}(t)),\ \underline{t}_0 \le t \le T - \bar{t}_0\} \xrightarrow{\mathcal{D}} \{\xi_t^{(3)}, \underline{t}_0 \le t \le T - \bar{t}_0\},$$

where

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

Estimators of t_0

Possible estimator of t_0 for $\boldsymbol{C}_{i,t} = \boldsymbol{Z}_{i,t}\boldsymbol{G}_{i,T}\boldsymbol{Z}_{i,t}$ where $\boldsymbol{G}_{i,T}$ is semi-definite:

$$argmax\{|U_N(t) - E_0 U_N(t)|\}$$

under some additional assumptions

If

$$\frac{1}{r_N}\sum_i ||\gamma_i||^2 = O_P(1)$$

for some $r_N/N \to \infty$ and

$$\frac{1}{r_N}\sum_i \boldsymbol{\delta}_i^T \Big(\boldsymbol{Z}_{i,t_0}\boldsymbol{Z}_{i,T}^{-1}\widetilde{\boldsymbol{Z}}_{i,t_0}\boldsymbol{G}_{iT}\widetilde{\boldsymbol{Z}}_{i,t_0}\boldsymbol{Z}_{i,T}^{-1}\boldsymbol{Z}_{i,t_0}\Big)\boldsymbol{\delta}_i \to \infty$$

then the estimator is consistent, e.i. $P(\widehat{t}_N = t_0) \to 1$

$$\frac{1}{\sqrt{N}}\sum_i \delta_i^T\Big(\mathbf{Z}_{i,t_0}\mathbf{Z}_{i,T}^{-1}\widetilde{\mathbf{Z}}_{i,t}\mathbf{G}_T\widetilde{\mathbf{Z}}_{i,t}\mathbf{Z}_{i,T}^{-1}\mathbf{Z}_{i,t_0}\Big)\delta_i, \quad t \geq t_0,$$

$$\frac{1}{\sqrt{N}}\sum_i \delta_i^T\Big(\widetilde{\mathbf{Z}}_{i,t_0}\mathbf{Z}_{i,T}^{-1}\mathbf{Z}_{i,t}\mathbf{G}_T\mathbf{Z}_{i,t}\mathbf{Z}_{i,T}^{-1}\widetilde{\mathbf{Z}}_{i,t_0}\Big)\delta_i, \quad t \leq t_0.$$

The upper terms for $t \leq t_0$ are nondecreasing and for for $t \geq t_0$ are nonincreasing. So that we still need for consistency that the term for $t = t_0$ is the larger one.

Application – The capital asset pricing model(CAPM)

Test for breaks in coefficients for the US mutual fund return data around the sub-prime crisis – January 2006 to February 2010 this corresponds with other studies (e.g. Dick-Nielsen, Feldhütter, and Lando, 2012).
The Fama-French three factor model augmented with the Carhart (1997) momentum factor is defined as

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

$1 \le t \le T,\ 1 \le i \le N$
$R_{i,t} - R_t^f$ – the excess return on the mutual fund
$R_t^M - R_t^f$ – the market risk premium
R_t^{HML} – the return difference between portfolios with the highest decile of stocks and lowest decile of stocks across in terms of the ratio of book equity-to-market equity (HML);
R_t^{SMB} – 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);
$e_{i,t}$ are the random errors.

Data – selected the mutual funds that have no missing returns for the period of the subprime crisis (January 2006 to February 2010).

Categories:

Large Blend, Large Growth, Large Value,

Middle Blend, Middle Growth, Middle Value,

Small Blend, Small Growth and Small Value.

The developed test procedure applied – test statistics, critical values, change point estimators, some illustrative graphs

$$U_N(t) = \sum_{i=1}^{N} \left(\widehat{\beta}_{i,t} - \widehat{\beta}_{i,T}\right)^T \boldsymbol{C}_{i,t}\left(\widehat{\beta}_{i,t} - \widehat{\beta}_{i,T}\right)$$

$$V_N(t) = \frac{1}{\sqrt{N}} |U_N(t) - E_{H_0} U_N(t)|$$

$$V_N = \max_t V_N(t)$$

10.4. **Result of the First Interested Period: January 2006 to February 2010.**
Table 10.1 shows the test result of the first interested period.

TABLE 10.1. Test Result

Category	V_N	10% CV	5% CV	1% CV	$\hat{t}$	T	N
Large Blend	47125.28	18360.55	21964.06	28961.44	32	50	659
Large Growth	149623.8	36993.04	43448.65	56891.96	39	50	748
Large Value	48210.06	9137.873	10575.63	13238.53	38	50	528
Middle Blend	81238.26	38006.6	44298.88	56424.48	39	50	145
Middle Growth	151827.1	23034.25	26999.37	34195.55	39	50	329
Middle Value	36235.31	14909.84	17268.74	21878.15	34	50	173
Small Blend	37067.02	14893.47	17363.72	22394.08	39	50	279
Small Growth	67004.26	14741.61	16944.26	21574.51	33	50	313
Small Value	40582.5	35503.5	41432.85	52510.34	-	50	135

$V_N(t)$ and critical values
for aggressive mutual funds(upper) and large blend mutual funds(lower)

Figure 10.2 shows $V_N(t)$ and critical values for the category of aggressive mutual funds.

FIGURE 10.2. $V_N(t)$ and Critical Values

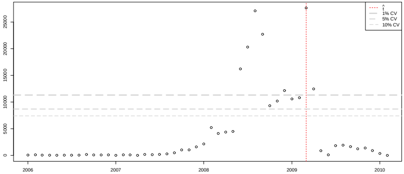


Figure 10.3 shows $V_N(t)$ and critical values for the category of large blend mutual funds.

FIGURE 10.3. $V_N(t)$ and Critical Values

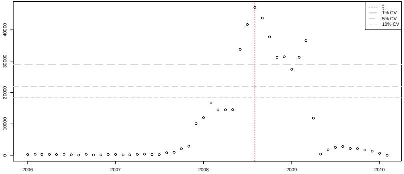


Figure 10.4 shows $V_N(t)$ and critical values for the category of large growth mutual funds.

FIGURE 10.4. $V_N(t)$ and Critical Values

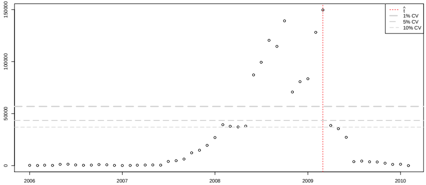


Figure 10.5 shows $V_N(t)$ and critical values for the category of large value mutual funds.

FIGURE 10.5. $V_N(t)$ and Critical Values

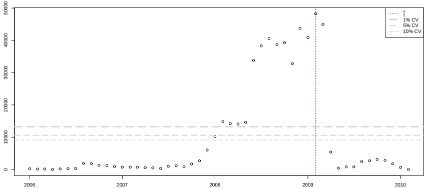


TABLE 11.1. Test on Selected Panels

Smallest beta = 0.1								Largest 1-beta = 0.1							
Category	V_N	10% CV	5% CV	1% CV	$\hat{t}$	T	N	Category	V_N	10% CV	5% CV	1% CV	$\hat{t}$	T	N
Large Blend	6239	6280	7480	9814	32	50	66	Large Blend	108517	52249	62346	80453	32	50	66
Large Growth	15016	8517	9606	11926	34	50	75	Large Growth	243320	110773	131786	171330	39	50	75
Large Value	17068	9176	10289	12922	38	50	53	Large Value	89840	13956	16102	20585	39	50	53
Middle Blend	14583	10772	12807	16281	39	50	15	Middle Blend	212368	73717	86487	111756	39	50	15
Middle Growth	13542	7949	9016	11524	39	50	33	Middle Growth	220673	28754	33202	41498	39	50	33
Middle Value	10930	15216	18017	23595	39	50	18	Middle Value	76995	30573	35787	46239	34	50	18
Small Blend	24897	10627	12556	16474	39	50	28	Small Blend	84757	24902	29248	37366	39	50	28
Small Growth	9870	9556	11200	14335	31	50	32	Small Growth	109139	24987	29106	37693	32	50	32
Small Value	16839	15461	18065	22683	39	50	14	Small Value	110660	93172	108716	144201	39	50	14
beta = 0.25								1-beta = 0.25							
Large Blend	4896	4260	4952	6519	32	50	165	Large Blend	88610	34218	40323	53865	32	50	165
Large Growth	18225	8429	9631	12261	32	50	187	Large Growth	237686	71132	84844	112372	39	50	187
Large Value	13205	6385	7188	8726	38	50	132	Large Value	87729	12399	14246	17962	38	50	132
Middle Blend	11826	10151	11796	14737	39	50	37	Middle Blend	159071	64635	75943	99273	39	50	36
Middle Growth	9716	6285	7133	8723	33	50	83	Middle Growth	227813	27627	32140	40432	39	50	82
Middle Value	8273	10685	12647	16377	39	50	44	Middle Value	67462	24941	29282	37805	34	50	43
Small Blend	20421	7444	8559	11018	39	50	70	Small Blend	82266	19579	22836	29031	39	50	70
Small Growth	9186	6573	7701	9716	31	50	79	Small Growth	109459	21709	25239	31376	32	50	78
Small Value	13796	10901	12570	15802	39	50	34	Small Value	88383	64263	75383	97039	39	50	34
beta = 0.4								1-beta = 0.4							
Large Blend	3807	3464	4025	5342	32	50	264	Large Blend	76061	27238	32496	42417	32	50	264
Large Growth	23293	7141	8292	10420	32	50	299	Large Growth	221431	57417	68238	89051	39	50	299
Large Value	10506	5336	5964	7302	38	50	211	Large Value	82159	11604	13341	16714	38	50	211
Middle Blend	8304	8391	9670	12141	39	50	58	Middle Blend	132045	54321	64841	86547	39	50	58
Middle Growth	23953	6253	7154	8903	32	50	132	Middle Growth	216500	26496	30473	39857	39	50	132
Middle Value	6604	8442	9947	12998	34	50	69	Middle Value	60437	21284	24728	31868	34	50	69
Small Blend	16667	6311	7308	9364	39	50	112	Small Blend	72680	20515	23864	30195	39	50	112
Small Growth	7951	5836	6714	8365	31	50	125	Small Growth	101291	19367	22299	28391	32	50	125
Small Value	11372	9000	10333	13161	39	50	54	Small Value	73969	52567	61312	79992	39	50	54

Figure 11.1. Large Blend

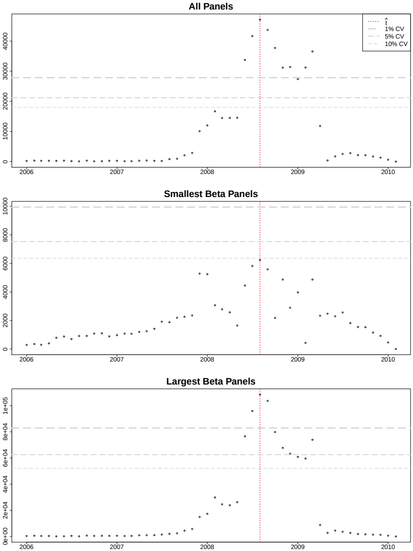
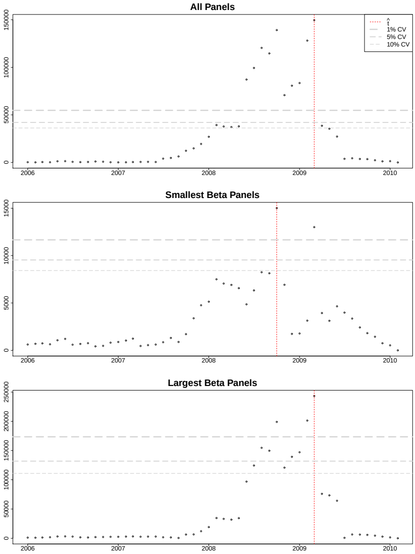


FIGURE 11.3. Large Growth



%centering %includegraphics[height=8cm]fig8.pdf %labelfig:1 %endfigure

FIGURE 11.5. Large Value

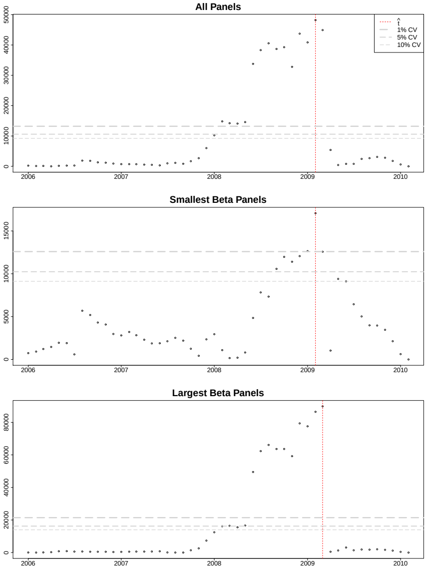
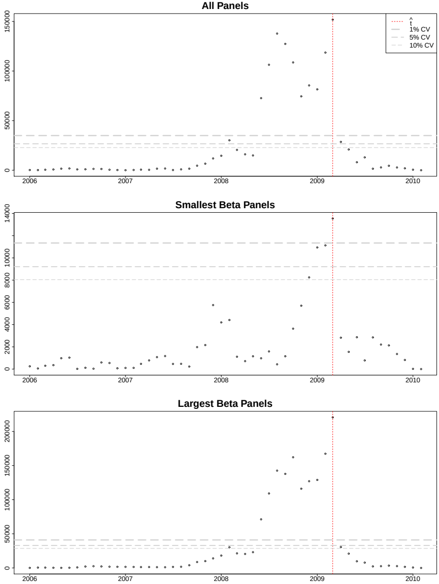


Figure 11.9. Middle Growth



References

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

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

Bai, J. and Carrioni-i-Silvestre, J.L.(2009): Structural changes, common stochastic trends, and unit roots in panel data. *The Review of Economic Studies* 76, 471–501.

Baltagi, B.H., Kao, C. and Liu, L.(2016): Estimation and identification of change points in panel models with nonstationary or stationary regressors and error terms.

Darling–Erdő limit results for change–point detection in panel data. *Journal of Statistical Planning and Inference* 143, 2013, 955–970.

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

Horváth, L., Hušková, M.,Rice,G. and Wang, J.(2017): CUSUM estimator for the time of change in linear panel data models. Econometric Theory 33, 366– 412.

Antoch, J., Hanousek,J., Horváth,L.,Huskova, M. and Wang, S.(2017):STRUCTURAL BREAKS IN PANEL DATA: LARGE NUMBER OF PANELS AND SHORT LENGTH TIME SERIES, submitted

Horváth, L. and Rice, G. (2014) Extensions of some classical methods in change point analysis (with discussions). *Test.* To appear.

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

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

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

THANK YOU !!!