

Early warning signals of financial crises using persistent homology

Sungkyunkwan University
Department of Mathematics
Hyelin choi

The Dotcom Bubble Crash (03/10/2000)

Cause:

Irrational hype & overvaluation of Internet stocks

Impact:

78% NASDAQ collapse; ~\$5 trillion in market value erased



eBay: ~91% drop



Yahoo!: ~98% drop



Amazon: ~93% drop

Lehman Brothers Bankruptcy (09/15/2008)

Cause:

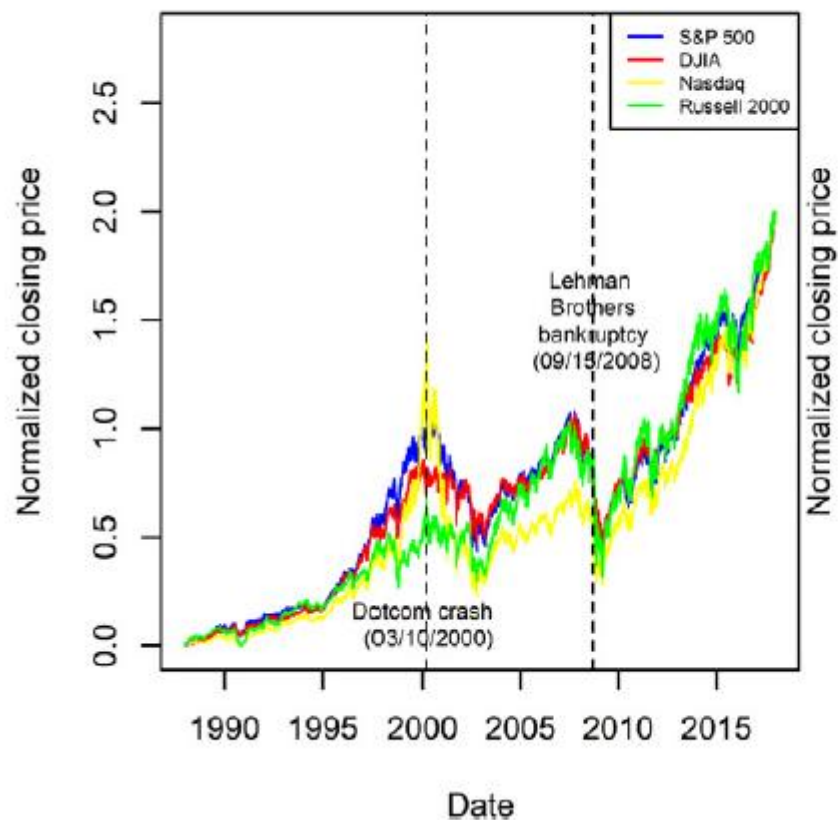
Over-leverage on toxic mortgage securities

Impact:

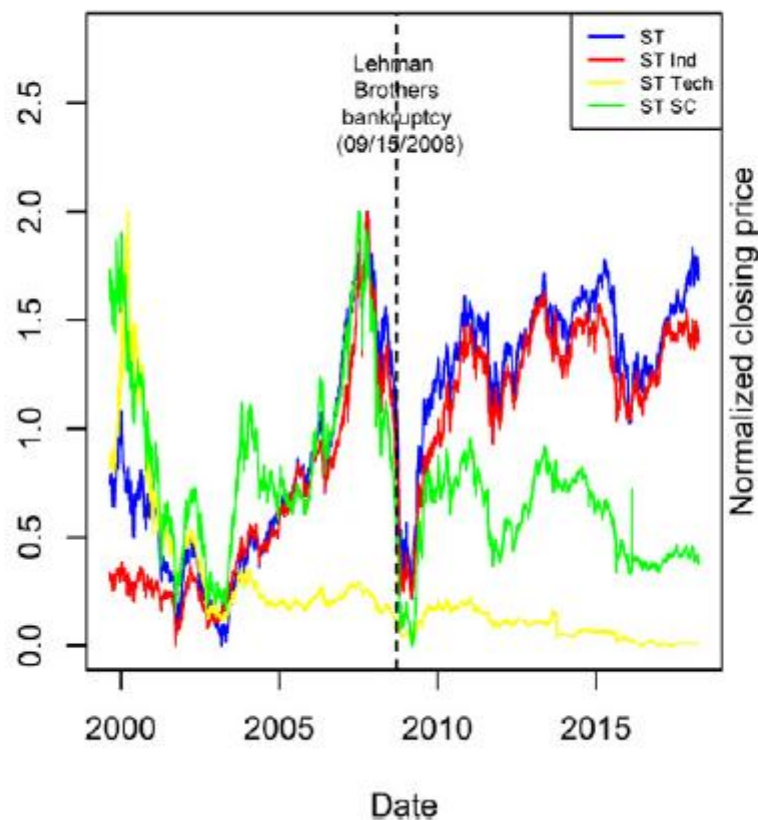
Global equities ~30% drop; credit markets froze

<https://www.youtube.com/shorts/m8d23np0oDk>

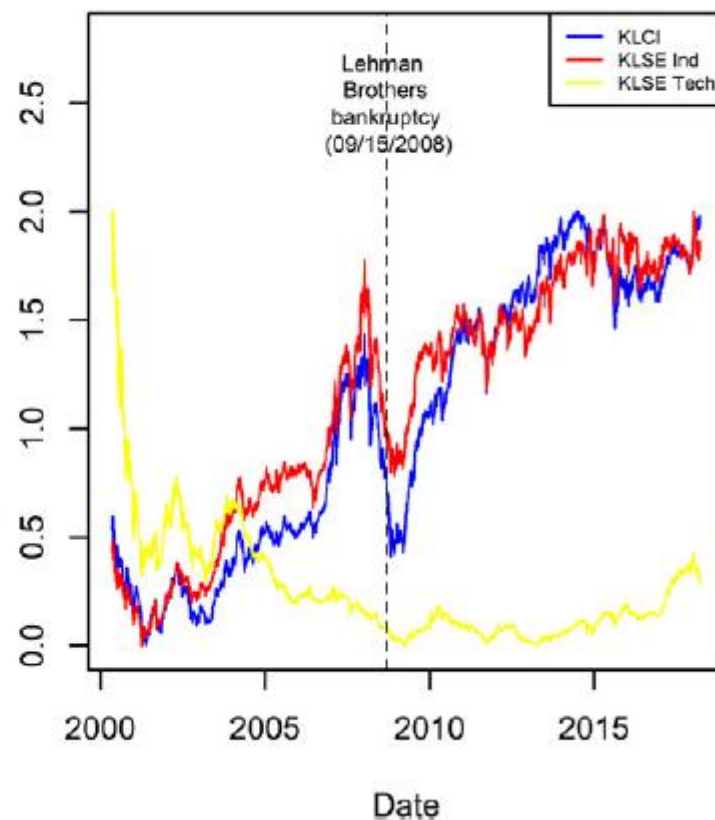
**(a) 4 Stock indices
of the US market**



**(b) 4 stock indices
of Singapore market**



**(c) 3 Stock indices
of Malaysia market**



Critical Slowing Down (CSD) Theory

- Definition:

As a system nears a tipping point, it recovers more slowly from perturbations

At tipping point, stability collapses, triggering an abrupt change

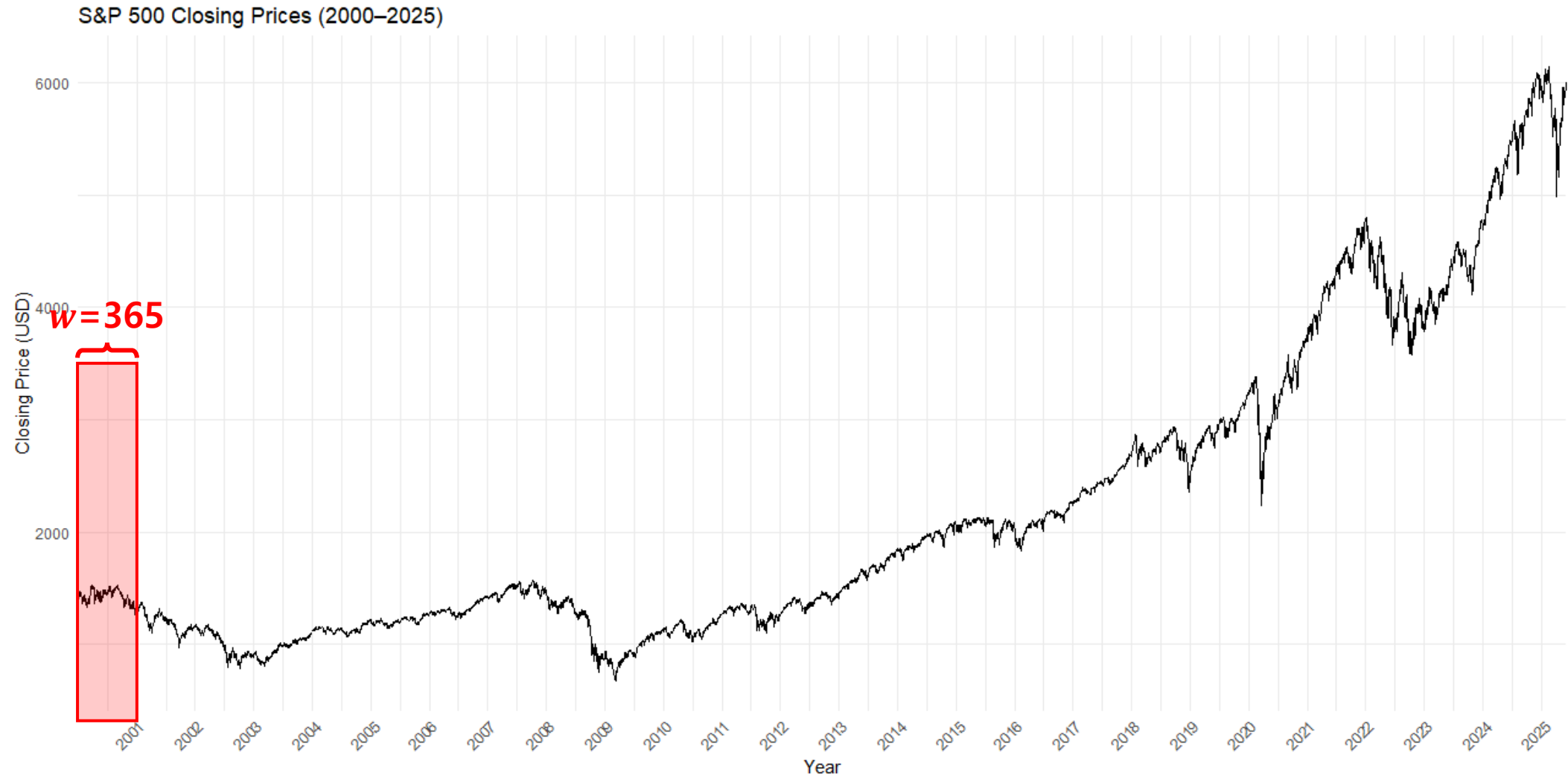
- Statistical Signs:

- Autocorrelation (ACF1) ↑
- Variance ↑
- Mean power spectrum (MPS) at low frequencies ↑

- Finance Insight:

Rising CSD metrics show loss of market resilience

Background: Sliding Window



Step 1: Data Collection

- **Dotcom Crash (03/10/2000)**
 - US
- **Lehman Brothers Bankruptcy (09/15/2008)**
 - US
 - Singapore
 - Malaysia

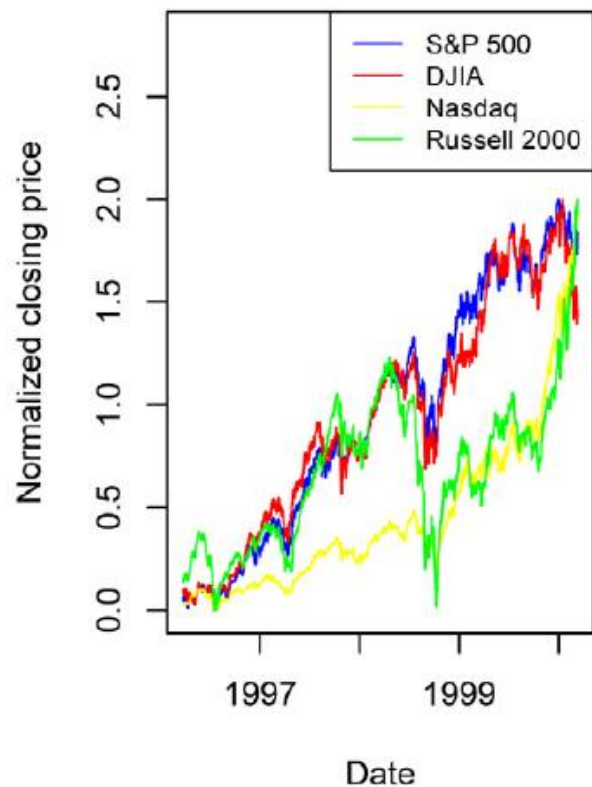
We use the 1,001 trading days prior to each event.

Step 1: Data Collection

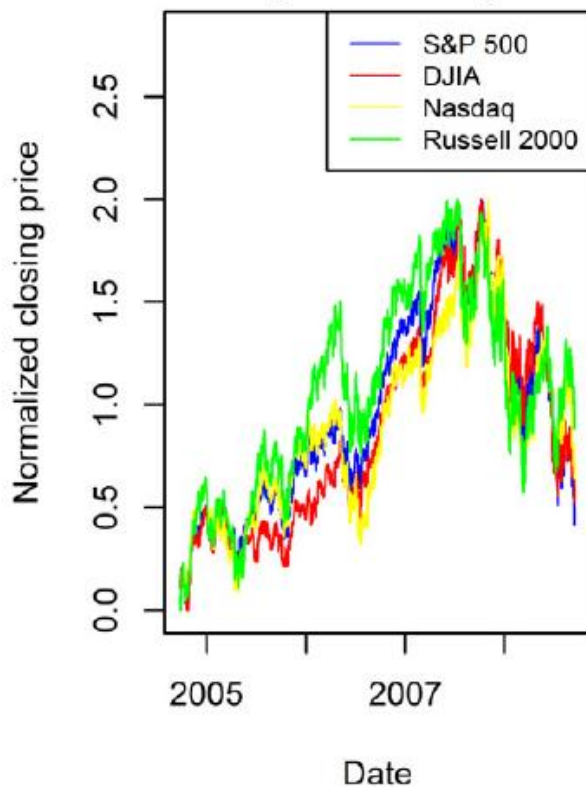
- Collect daily closing prices of d stock indices over the 1,001 trading days prior to each crash.

	1 st index (leading companies of all sectors)	2 nd index (leading companies in the industrial sector)	3 rd index (leading companies in the technology sector)	4 th index (leading small-cap companies)
US ($d = 4$)	S&P 500	DJIA	Nasdaq	Russel 2000
Singapore ($d = 4$)	ST	ST Ind	ST Tech	ST SC
Malaysia ($d = 3$)	KLCI	KLSE Ind	KLSE Tech	

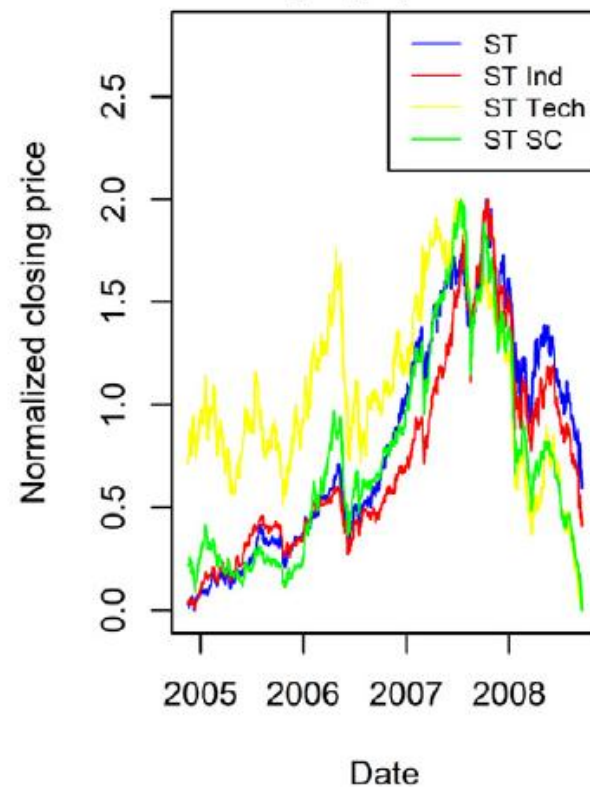
**(a) 4 US stock indices
prior to Dotcom
crash (03/10/2000)**



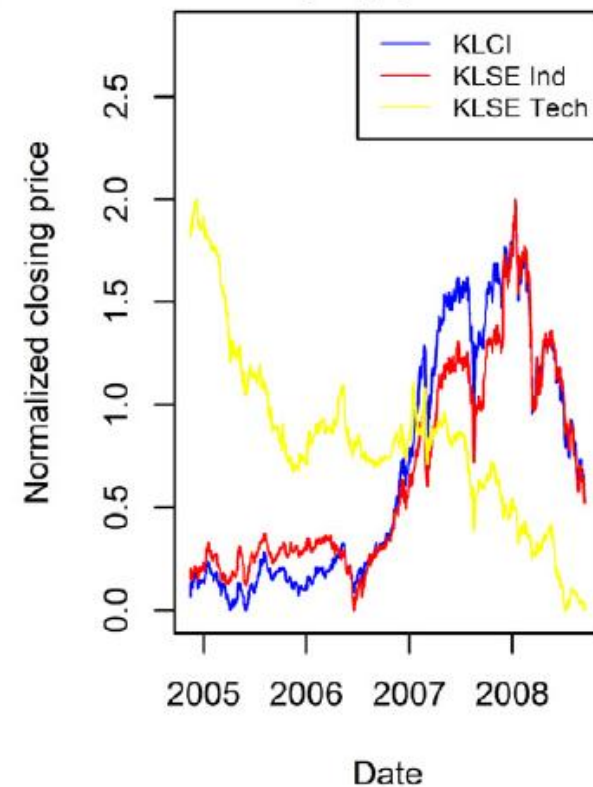
**(c) 4 US stock indices
prior to Lehman
Brothers bankruptcy
(09/15/2008)**



**(e) 4 Singapore stock
indices prior to
Lehman Brothers
bankruptcy (09/15/2008)**



**(g) 3 Malaysia stock
indices prior to
Lehman Brothers
bankruptcy (09/15/2008)**



Step 2: Log-returns Transformation

Compute log-returns

$$x_i(t) = \ln \left(\frac{P_i(t)}{P_i(t-1)} \right)$$

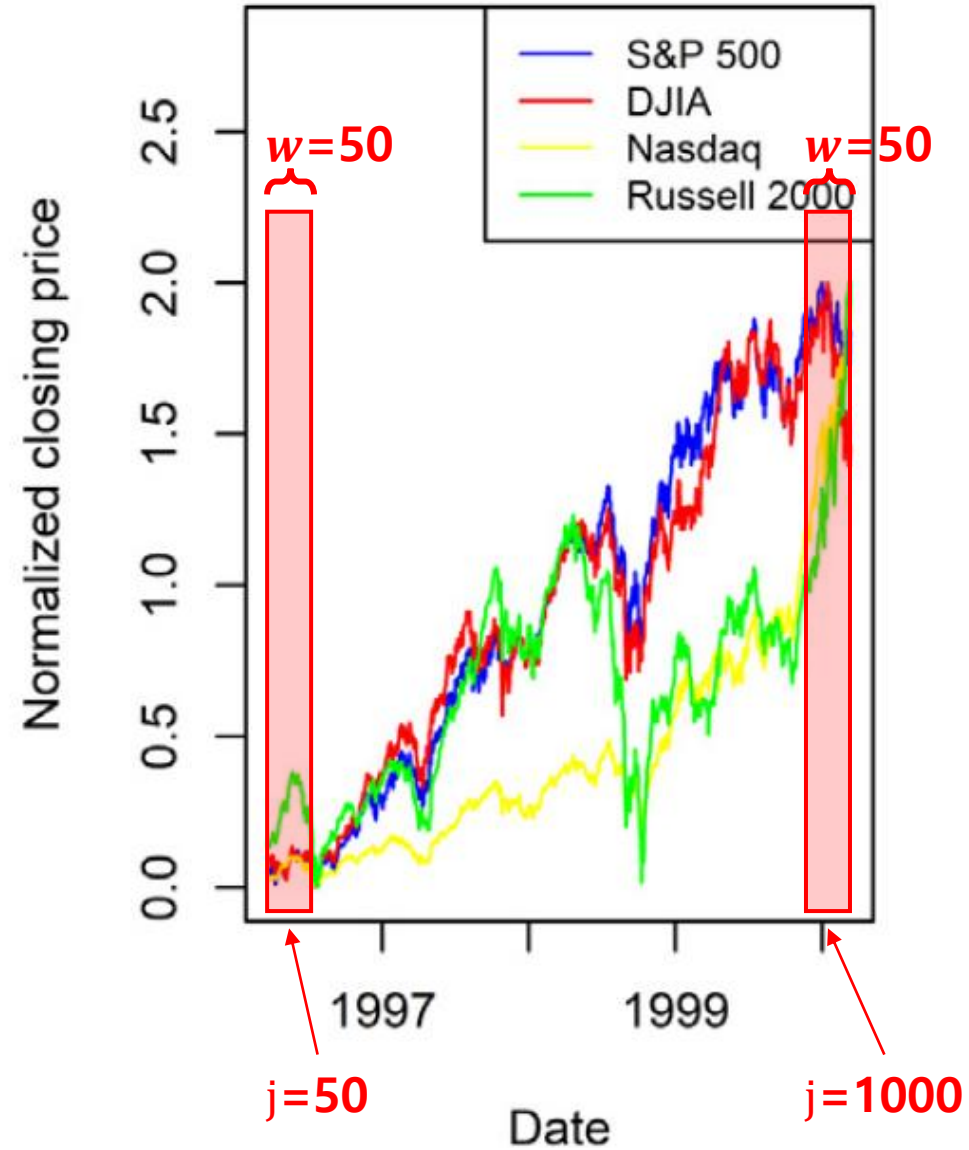
- $x_i(t)$: log-return of index i on day t
- $P_i(t)$: closing price of index i on day t

Step 3: Form Point Cloud Dataset(PCD)

Form point cloud dataset (PCD) of window size $w_1 = 50$ on day $j = 50, 51, \dots, 1000$.

$$X(j) = \begin{bmatrix} x_1(j-50+1) & x_2(j-50+1) & \dots & x_d(j-50+1) \\ x_1(j-50+2) & x_2(j-50+2) & \dots & x_d(j-50+2) \\ & \vdots & \ddots & \vdots \\ & x_1(j) & x_2(j) & \dots & x_d(j) \end{bmatrix}$$

(a) 4 US stock indices
prior to Dotcom
crash (03/10/2000)



Step 4: Vietoris-Rips Complex

- Build Vietoris-Rips complexes on log-return point cloud to capture loop structure
- Track appearance (birth) and disappearance (death) of loops across scales
- Use scale parameters

$$0 = \varepsilon_0 < \varepsilon_1 < \dots < \varepsilon_{max} = 0.05$$

Step 5: Persistence Landscape

- For each birth-death pair $(\varepsilon_b^i, \varepsilon_d^i)$, define

$$f_{(\varepsilon_b^i, \varepsilon_d^i)}(x) = \max\{0, \min\{x - \varepsilon_b^i, \varepsilon_d^i - x\}\}$$

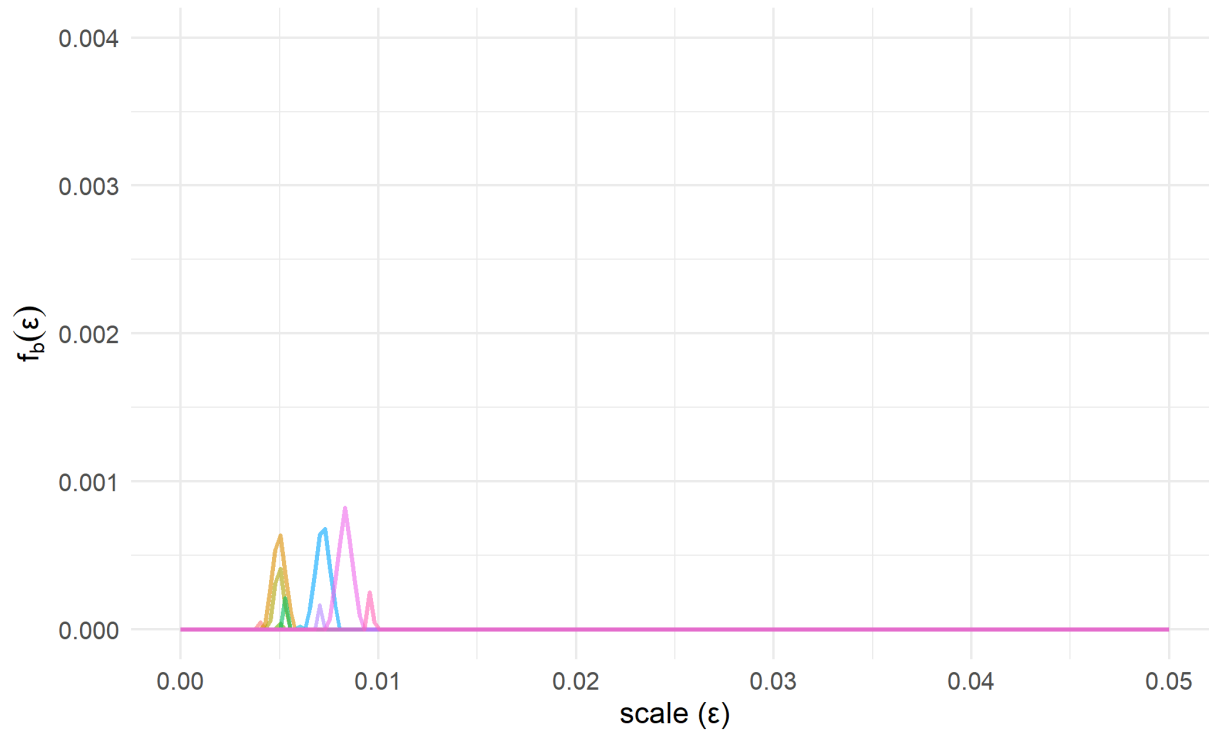
- The persistence landscape $\lambda = \{\lambda_1(x), \lambda_2(x), \dots\}$ is then

$$\lambda_k(x) = k - \max \left(\left\{ f_{(\varepsilon_b^i, \varepsilon_d^i)}(x) \mid (\varepsilon_b^i, \varepsilon_d^i) \right\}_{i=1}^n \right),$$

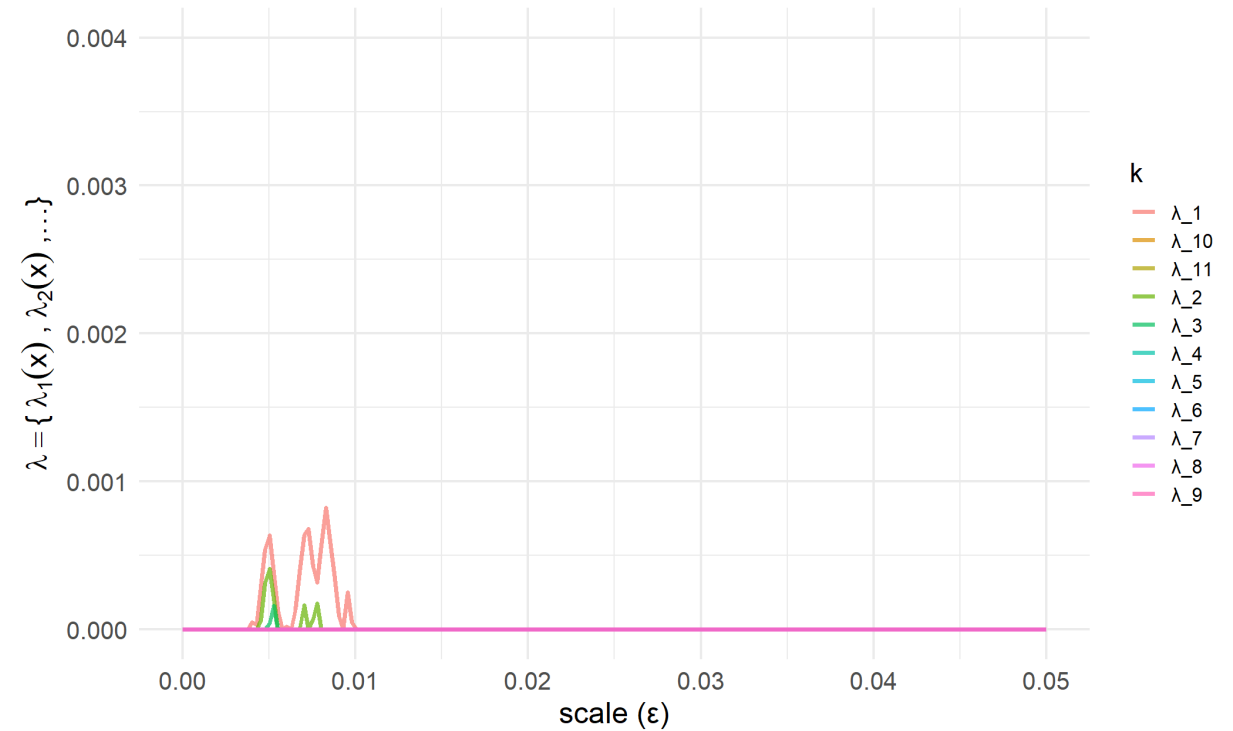
where $k - \max$ picks the k -th largest value at each x .

Persistence Landscape for $j = 50$

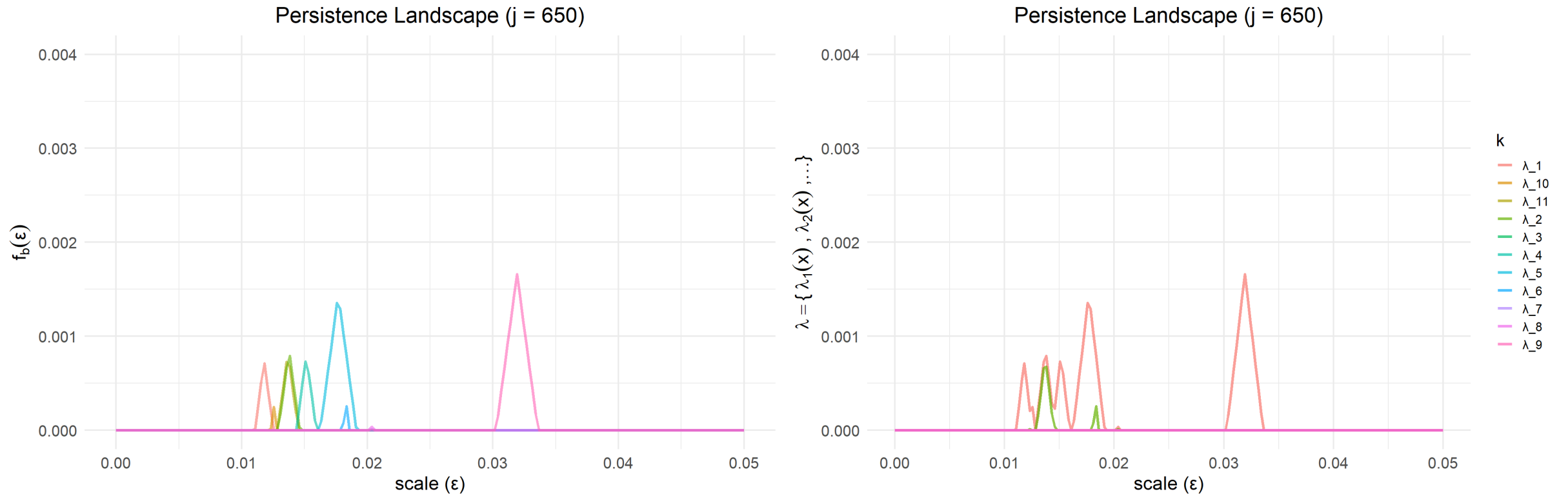
Persistence Landscape (j = 50)



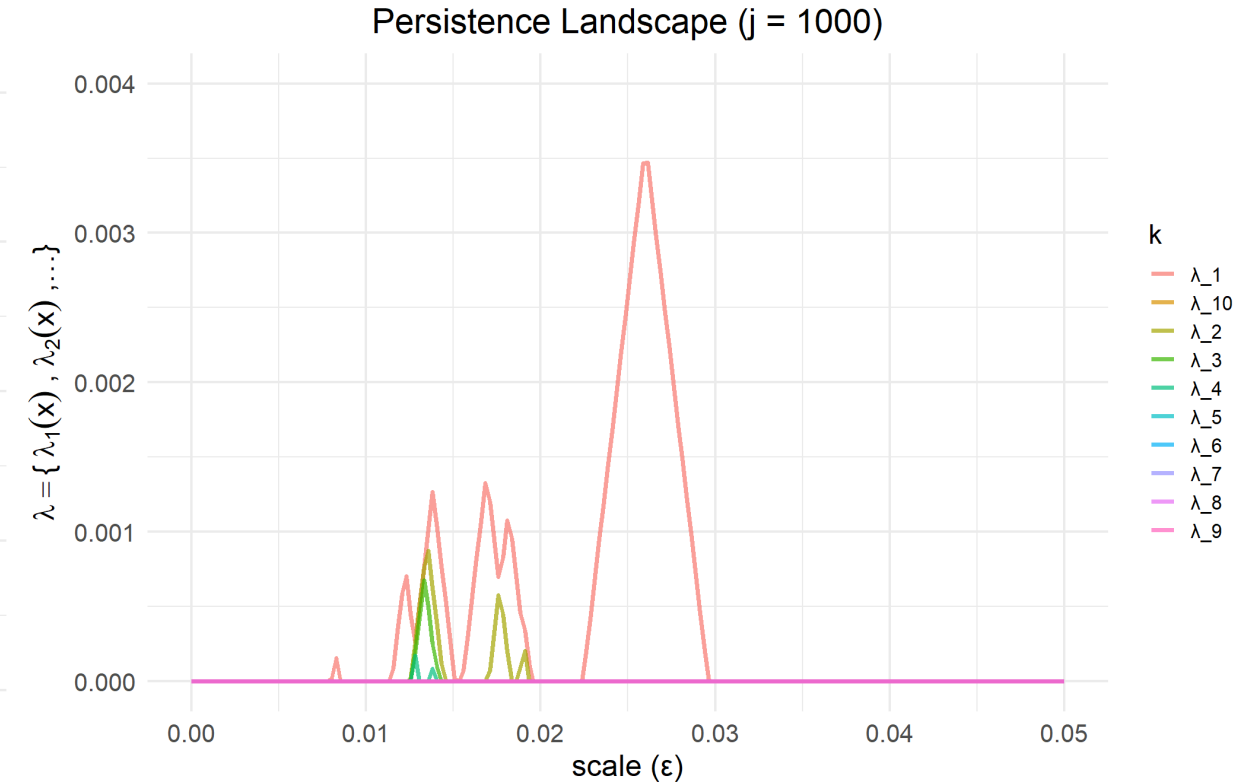
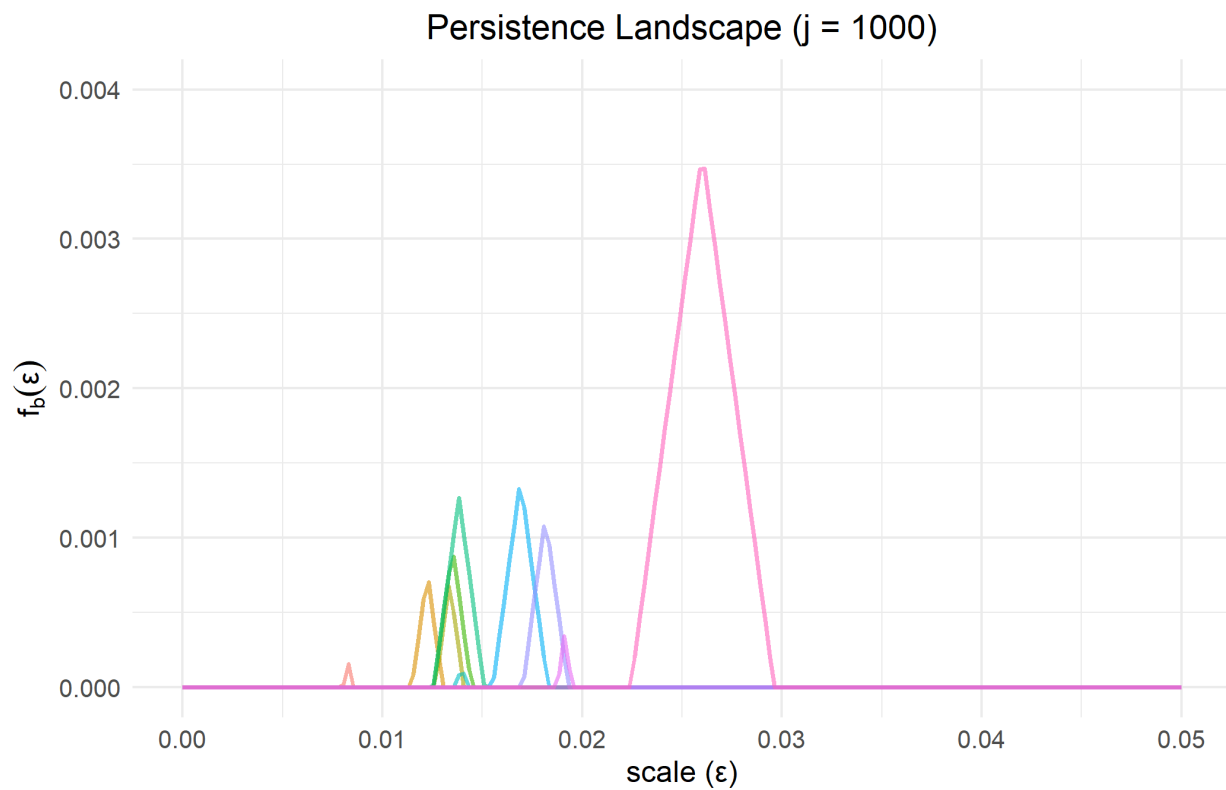
Persistence Landscape (j = 50)



Persistence Landscape for $j = 650$



Persistence Landscape for $j = 1000$

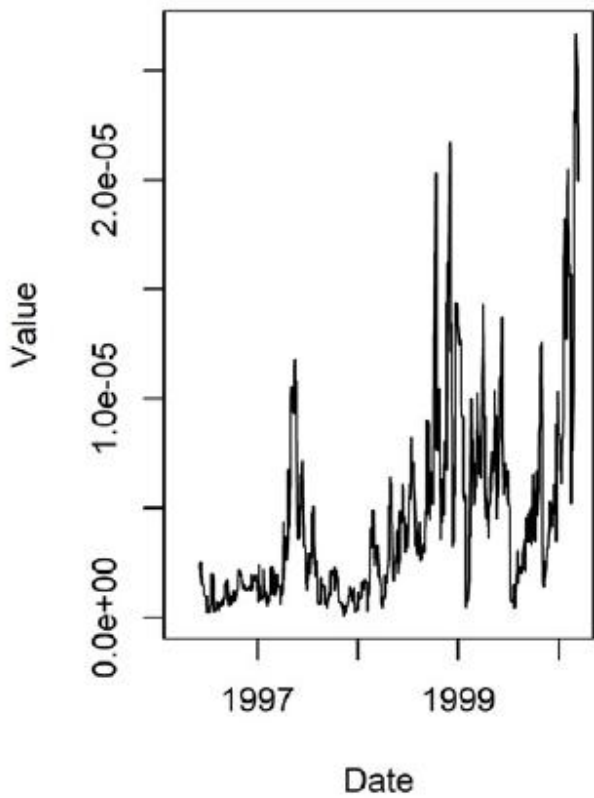


Step 6: L^1 -norm

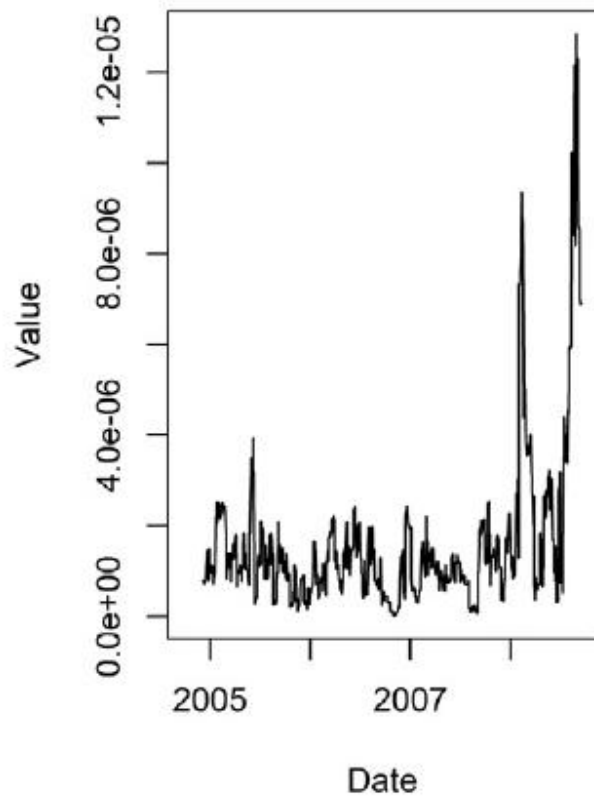
- Compute $\|\lambda\|_1 = \sum_{i=1}^{\infty} \int |\lambda_k(t)| dt$



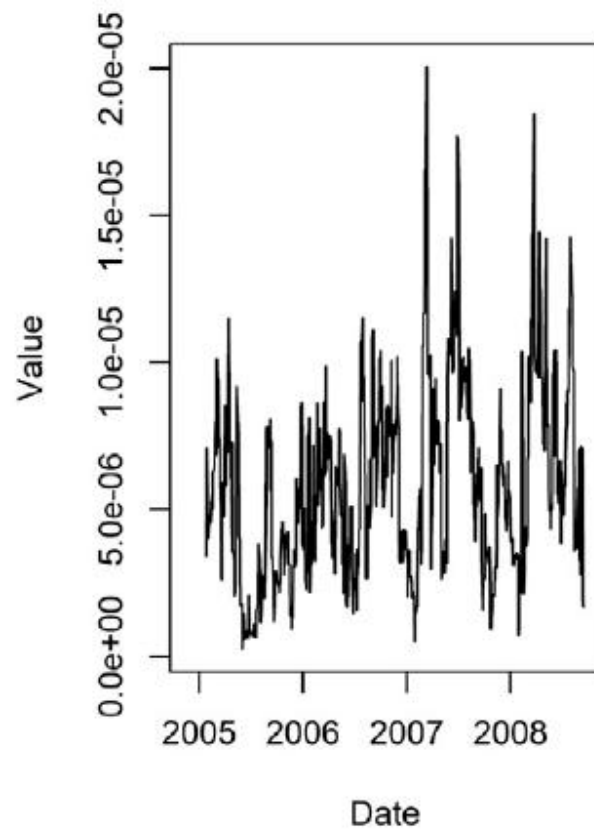
L¹-norm of US prior to
Dotcom crash



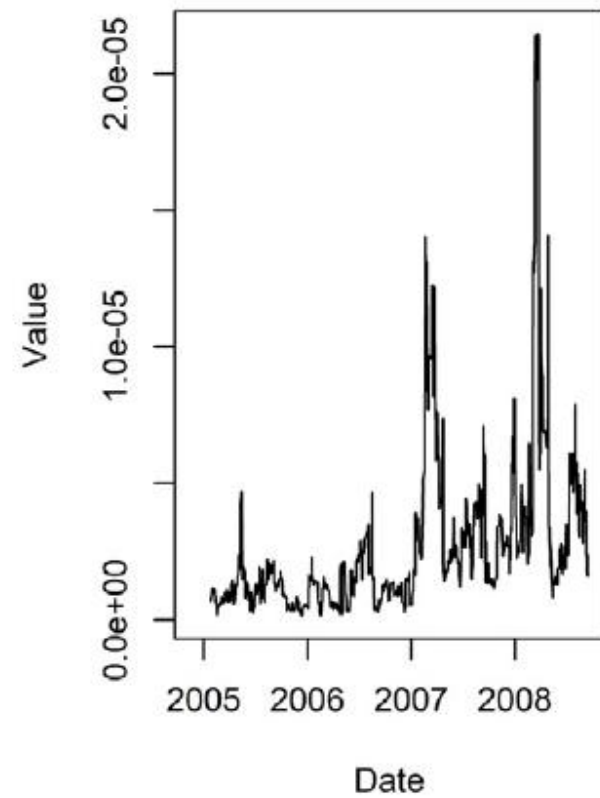
L¹-norm of US prior to
Lehman Brothers Bankruptcy



L¹-norm of Singapore prior to
Lehman Brothers Bankruptcy



L¹-norm of Malaysia prior to
Lehman Brothers Bankruptcy



Step 7: Critical Slowing Down (CSD)

Compute the following CSD Indicators using L^1 -norm using sliding window ($w_2 = 250$)

- Autocorrelation function at lag 1 (ACF1)
- Variance (VAR)
- Mean power spectrum (MPS) at low frequencies

Step 7-1: Autocorrelation function at lag 1 (ACF1)

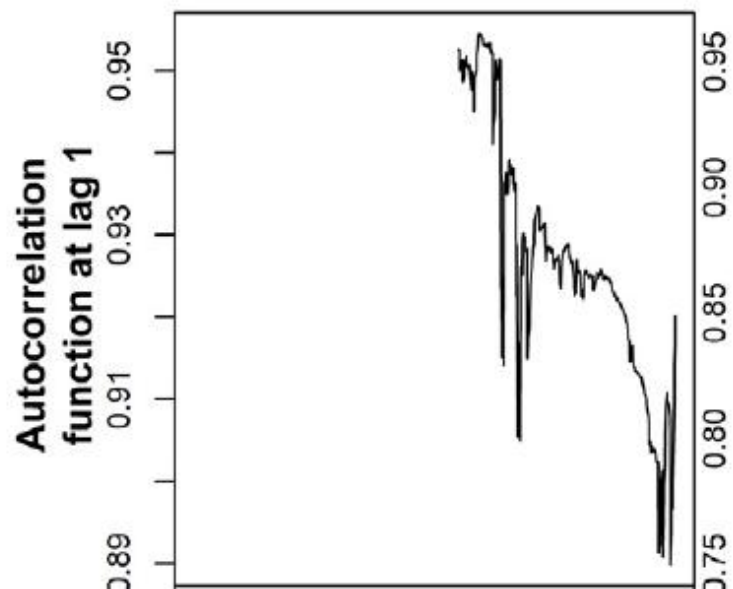
The ACF1 value at trading day l :

$$acf1(l) = \frac{\rho_1(l)}{var(l)},$$

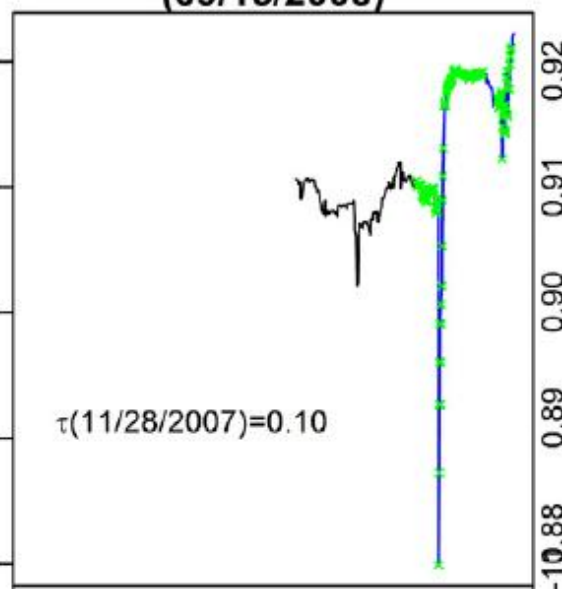
where

- $\rho_1(l) = \frac{1}{500-1} \sum_{j=l-500+1}^{l-1} (y(j) - \bar{y}(l))(y(j+1) - \bar{y}(l))$
- $\bar{y}(l) = \frac{1}{500} \sum_{j=l-500+1}^l y(j)$
- $var(l) = \frac{1}{500-1} \sum_{j=l-500+1}^l (y(j) - \bar{y}(l))^2$

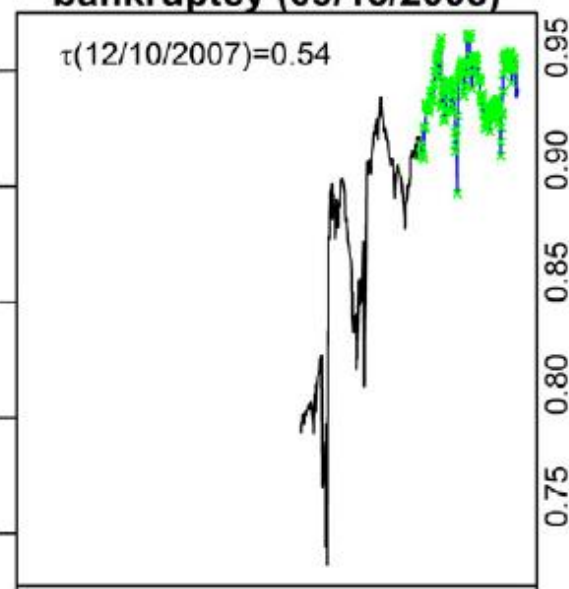
Indicators of L^1 -norms
for the US stock indices
prior to Dotcom crash
(03/10/2000)



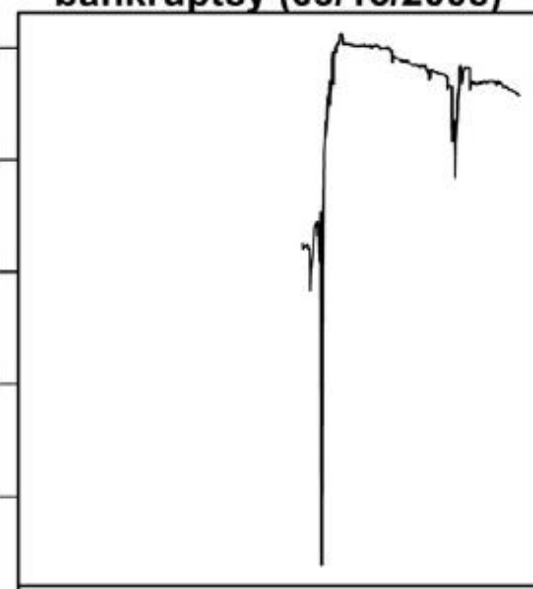
Indicators of L^1 -norms
for the US stock indices
prior to Lehman
Brothers bankruptcy
(09/15/2008)



Indicators of L^1 -norms
for Singapore stock
indices prior to
Lehman Brothers
bankruptcy (09/15/2008)



Indicators of L^1 -norms
for Malaysia stock
indices prior to
Lehman Brothers
bankruptcy (09/15/2008)



Step 7-2: Variance (VAR)

- The VAR value at trading day l :

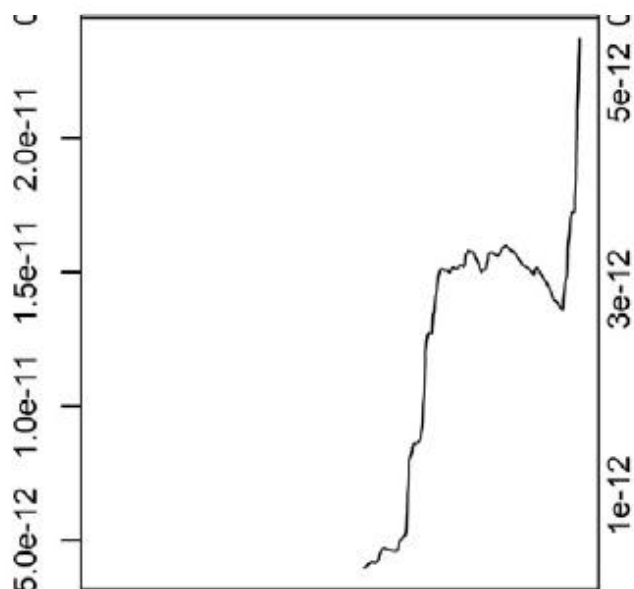
$$var(l) = \frac{1}{500 - 1} \sum_{j=l-500+1}^l (y(j) - \bar{y}(l))^2$$

where

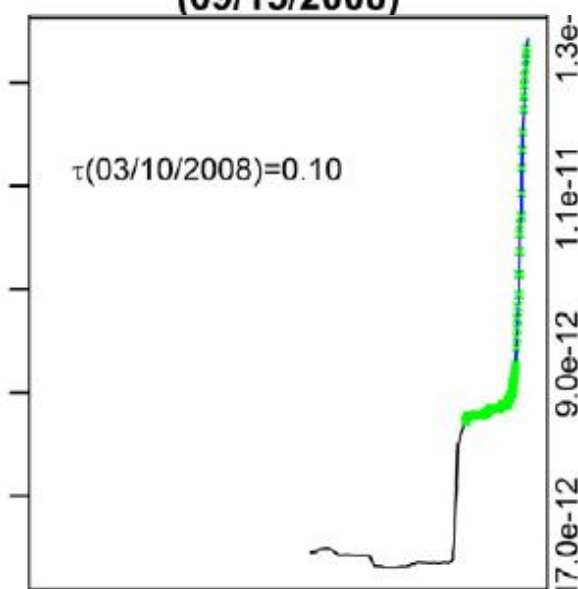
- $\bar{y}(l) = \frac{1}{500} \sum_{j=l-500+1}^l y(j)$

Variance

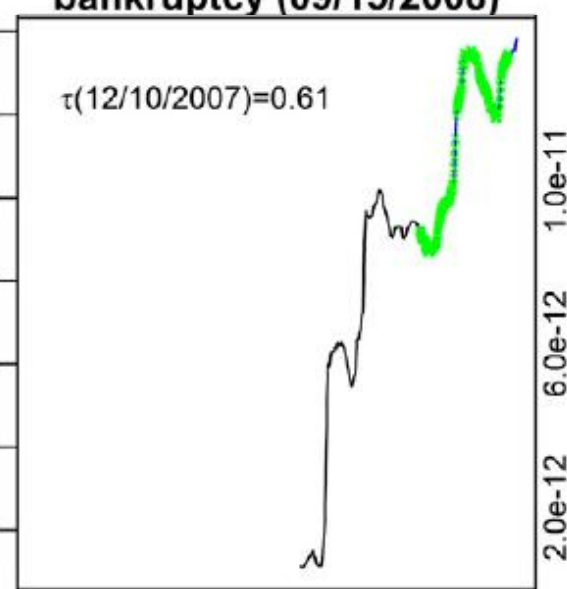
Indicators of L^1 -norms
for the US stock indices
prior to Dotcom crash
(03/10/2000)



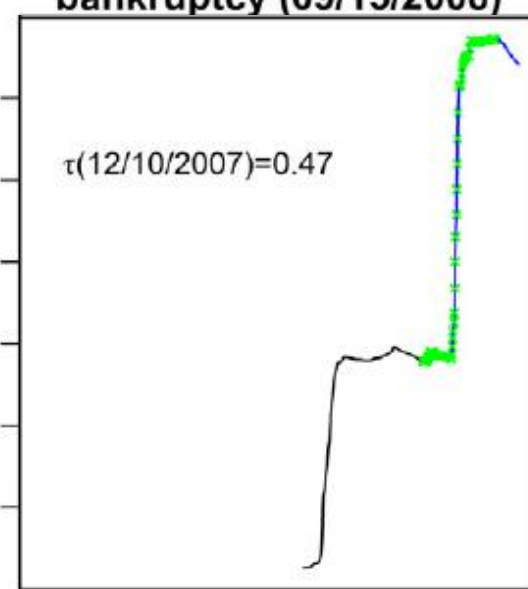
Indicators of L^1 -norms
for the US stock indices
prior to Lehman
Brothers bankruptcy
(09/15/2008)



Indicators of L^1 -norms
for Singapore stock
indices prior to
Lehman Brothers
bankruptcy (09/15/2008)



Indicators of L^1 -norms
for Malaysia stock
indices prior to
Lehman Brothers
bankruptcy (09/15/2008)



Step 7-3: Mean power spectrum (MPS) at low frequencies

- **Discrete Fourier Transform:**

$$F_k(l) = \sum_{n=l-500+1}^l y(n) e^{\frac{-2\pi i k n}{500}}$$

where $k = 1, \dots, 500$.

- **Power Spectrum:**

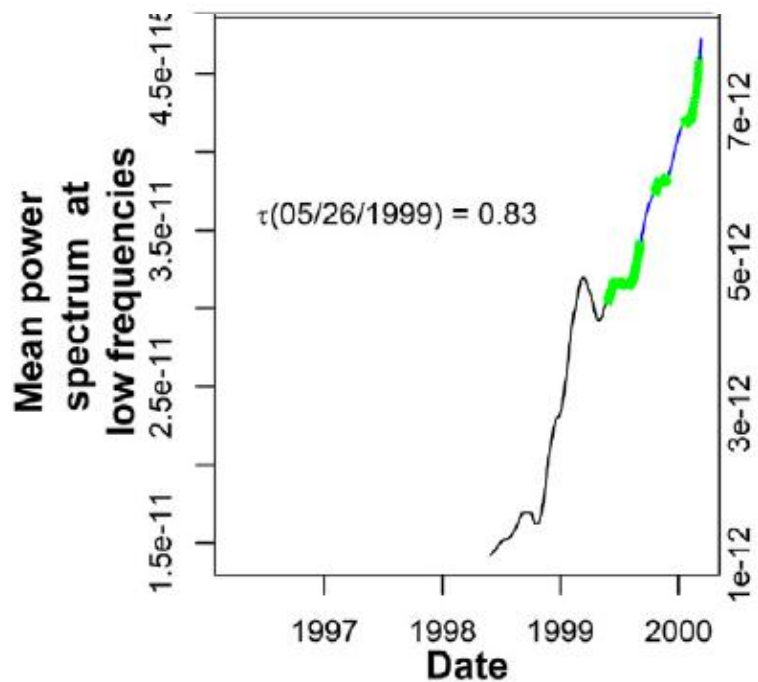
$$PS_k(l) = |F_k(l)|^2$$

Each $PS_k(l)$ is **normalized** such that its sum is equal to 1.

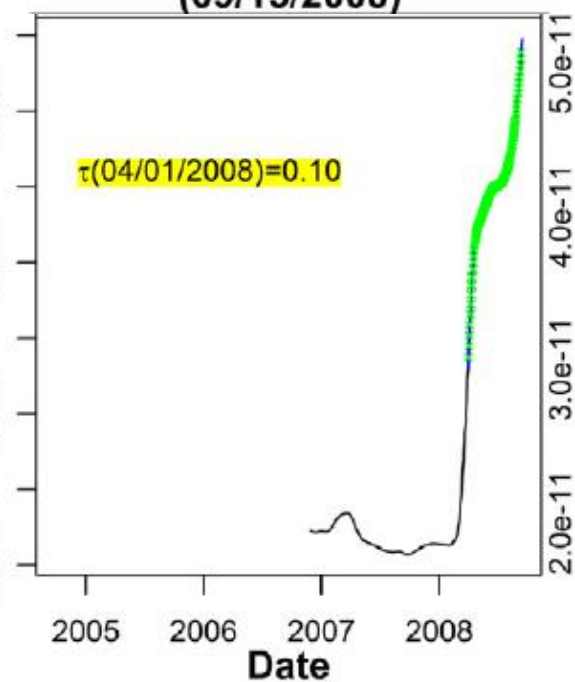
- **MPS Value:**

$$mps(l) = \frac{1}{[500/8] - 1} \sum_{k=2}^{[500/8]} PS_k(l)$$

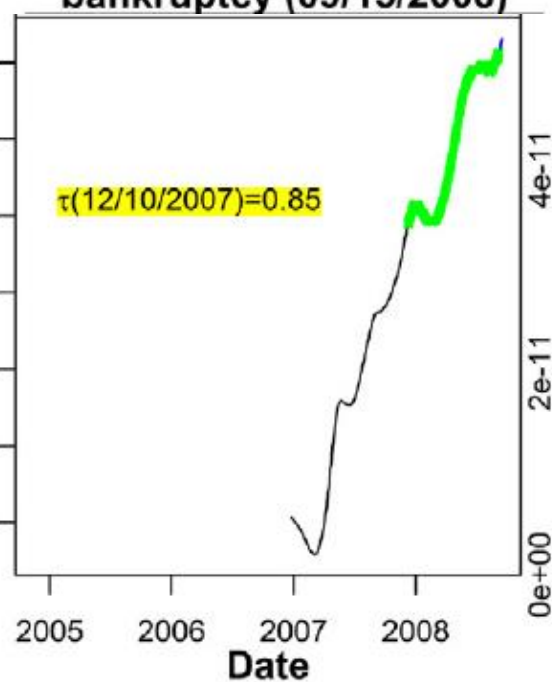
Indicators of L^1 -norms
for the US stock indices
prior to Dotcom crash
(03/10/2000)



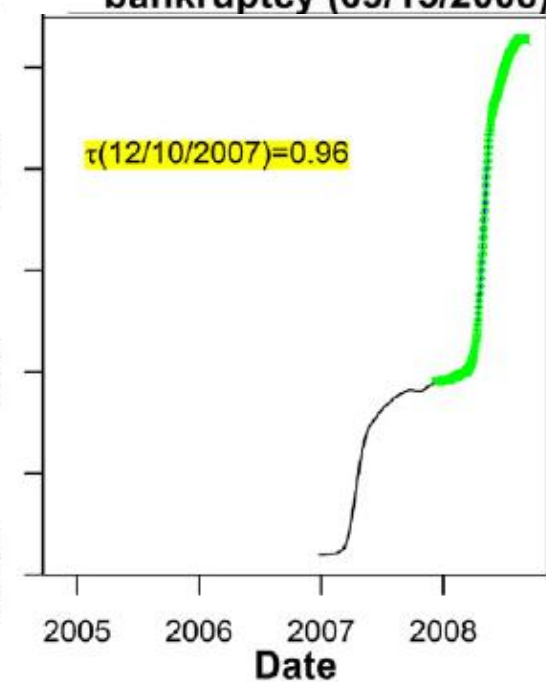
Indicators of L^1 -norms
for the US stock indices
prior to Lehman
Brothers bankruptcy
(09/15/2008)



Indicators of L^1 -norms
for Singapore stock
indices prior to
Lehman Brothers
bankruptcy (09/15/2008)



Indicators of L^1 -norms
for Malaysia stock
indices prior to
Lehman Brothers
bankruptcy (09/15/2008)



How can we confirm significant rises in CSD indicators?

- CSD Theory : ACF1, VAR, MPS values should rise before crises
However, this rising trend must be statistically significant

Significance Test: Mann-Kendall Test

- Purpose: Assess significant rising trend in CSD indicators (ACF1, VAR, MPS)
- Window size $w_3 = 250$
- Kendall's τ correlation:

$$\tau_{ACF1}(m) = \frac{S(m)}{D(m)}$$

where

- $S(m) = \sum_{p=m-250+1}^{m-1} \sum_{q=p+1}^m \text{sign}(acf1(q) - acf1(p))$
- $D(m) = \left(\frac{1}{2} (250)(249) - \frac{1}{2} \sum_{r=1}^s (\alpha_r)(\alpha_r - 1) \right)^{1/2} \left(\frac{1}{2} (250)(249) \right)^{1/2}$
- α_r is the number of points in the r th tied group

Significance Test: Mann-Kendall Test

H_0 : no monotonic trend (rising or falling) in CSD indicator

Under H_0 , we have

- $\mathbb{E}[S(m)] = 0$,
- $\sigma^2(m) = \frac{(250)(249)(505) - \sum_{r=1}^S (\alpha_r)(\alpha_r - 1)(2\alpha_r + 5)}{18}$

Transform $S(m)$ into a normally distributed Z_{obs} :

$$Z_{\text{obs}} = \begin{cases} \frac{S(m)-1}{\sigma(m)} & \text{if } S(m) > 0 \\ 0 & \text{if } S(m) = 0 \\ \frac{S(m)+1}{\sigma(m)} & \text{if } S(m) < 0. \end{cases}$$

Significance Test: Mann-Kendall Test

- Two-sided p-value:

$$p - \text{value} = 2P_{Z \sim N(0,1)}(Z > |Z_{\text{obs}}|)$$

- Criteria for Significant Rising Trend:

If $\tau_{\text{ACF1}}(m) > 0$ and $p\text{-value} < 0.05$, reject H_0 .

Structural Break Test: Chow Test

- H_0 : no structural break at time b
- H_1 : structural break at time b

- Chow F-test statistics:

$$F_{obs} = \frac{(RSS_1 - (RSS_2 + RSS_3))/k}{(RSS_2 + RSS_3)/(1000 - 549 + 1 - 2k)}$$

- Chow F-test statistics:

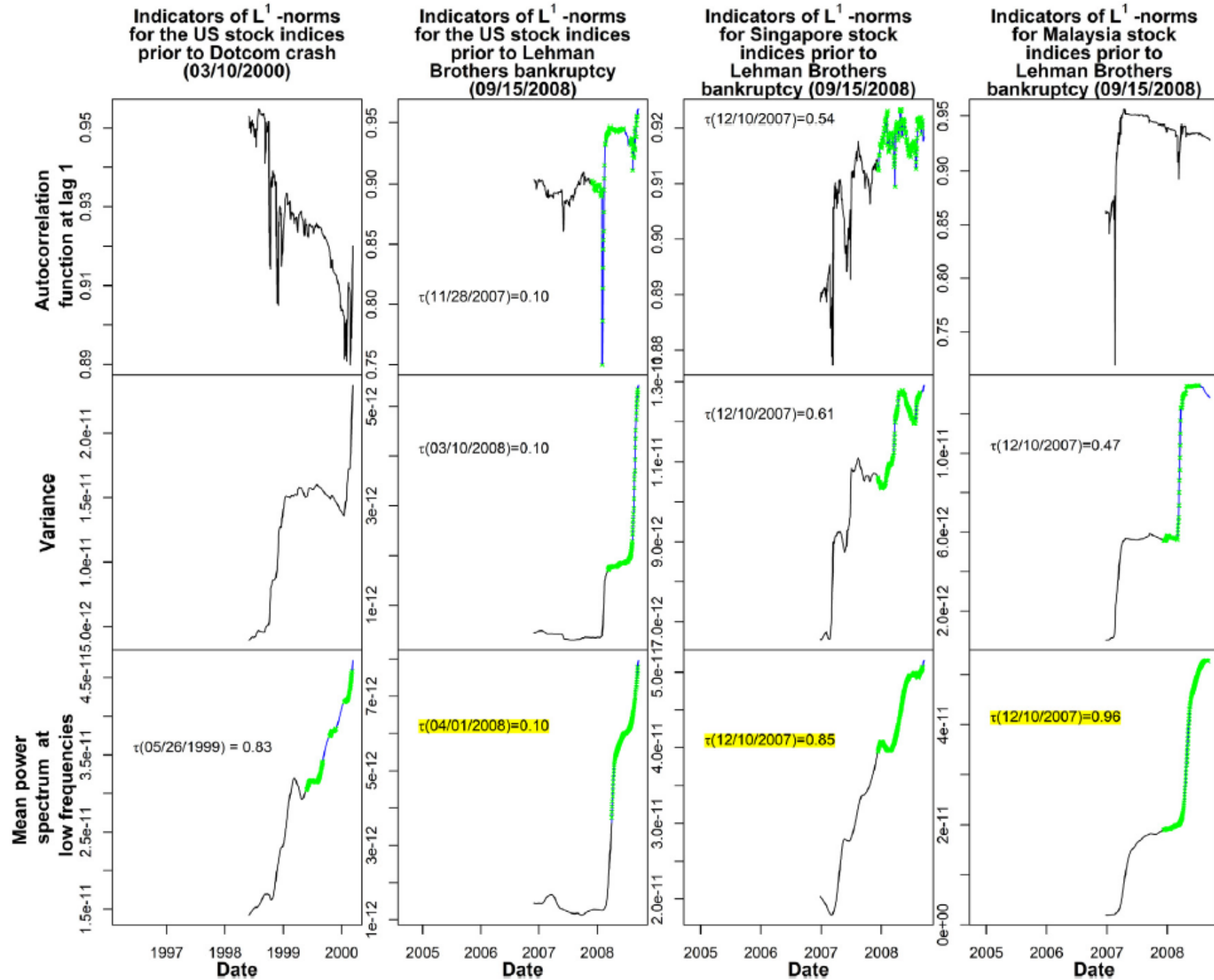
$$F_{obs} = \frac{(RSS_1 - (RSS_2 + RSS_3))/k}{(RSS_2 + RSS_3)/(1000 - 549 + 1 - 2k)}$$

where

- $RSS_1 = (acf1(l) - \hat{\beta}_0^{(1)} - \hat{\beta}_1^{(1)}l)^2$
- $RSS_2 = (acf1(l) - \hat{\beta}_0^{(2)} - \hat{\beta}_1^{(2)}l)^2$
- $RSS_3 = (acf1(l) - \hat{\beta}_0^{(3)} - \hat{\beta}_1^{(3)}l)^2$

- Criteria for Structural Break:

If $P(F > F_{obs}) < 0.05$, reject H_0 .



Thresholds for Early Warning Signals

Singapore & Malaysia

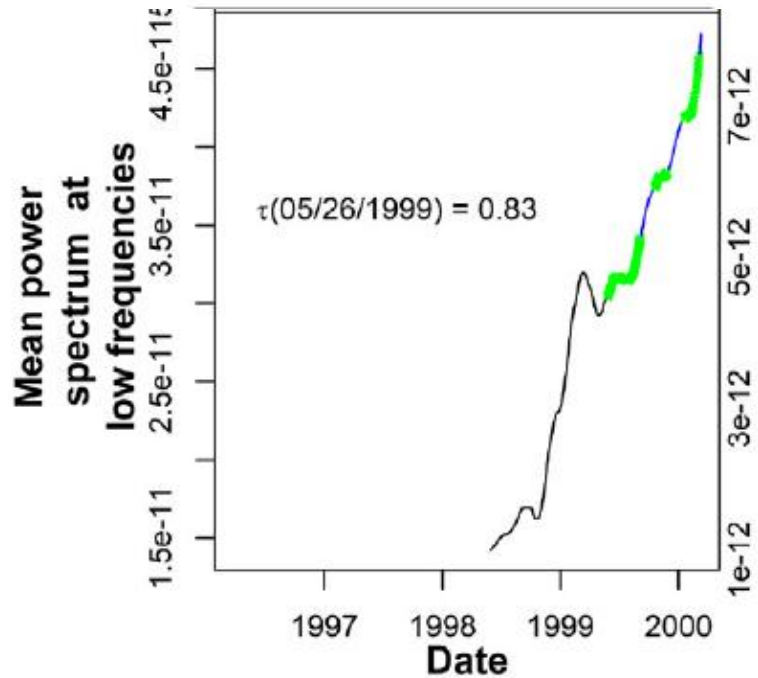
- $\{\tau\}_{m=\delta}^{1000}$: significant rising trend for ACF1, VAR, MPS
- Take threshold value T at the first breakpoint within $\{\tau\}_{m=\delta}^{1000}$

US

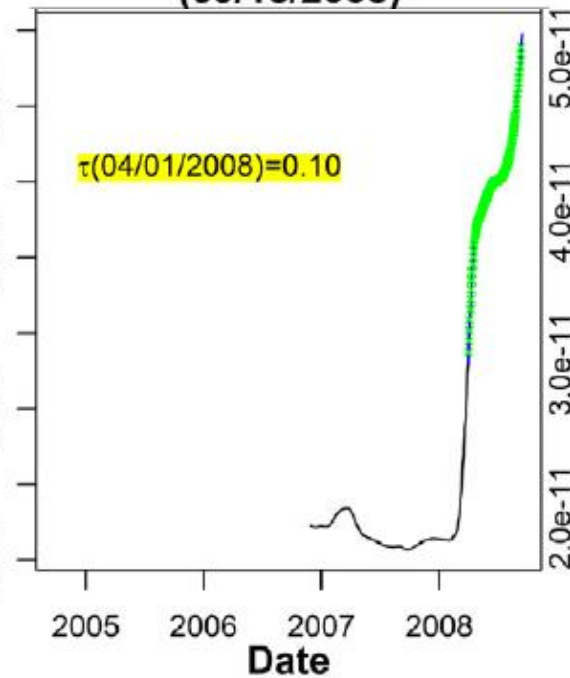
- Take threshold value $T = \min \{T_{Dotcom}, T_{Lehman}\}$

where T_{Dotcom} and T_{Lehman} are determined as in Singapore & Malaysia

Indicators of L^1 -norms
for the US stock indices
prior to Dotcom crash
(03/10/2000)

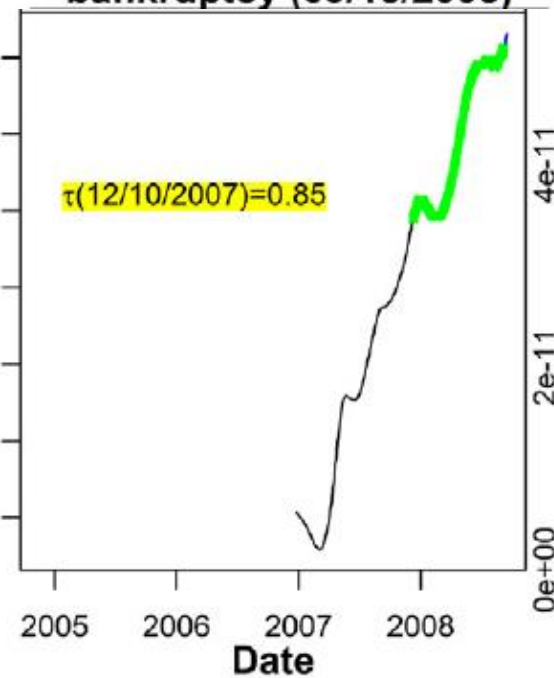


Indicators of L^1 -norms
for the US stock indices
prior to Lehman
Brothers bankruptcy
(09/15/2008)



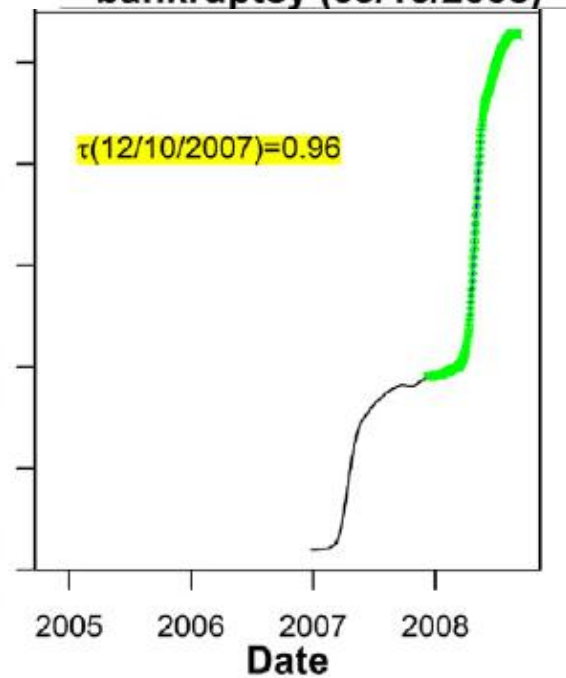
$$T_{US} = 0.10$$

Indicators of L^1 -norms
for Singapore stock
indices prior to
Lehman Brothers
bankruptcy (09/15/2008)



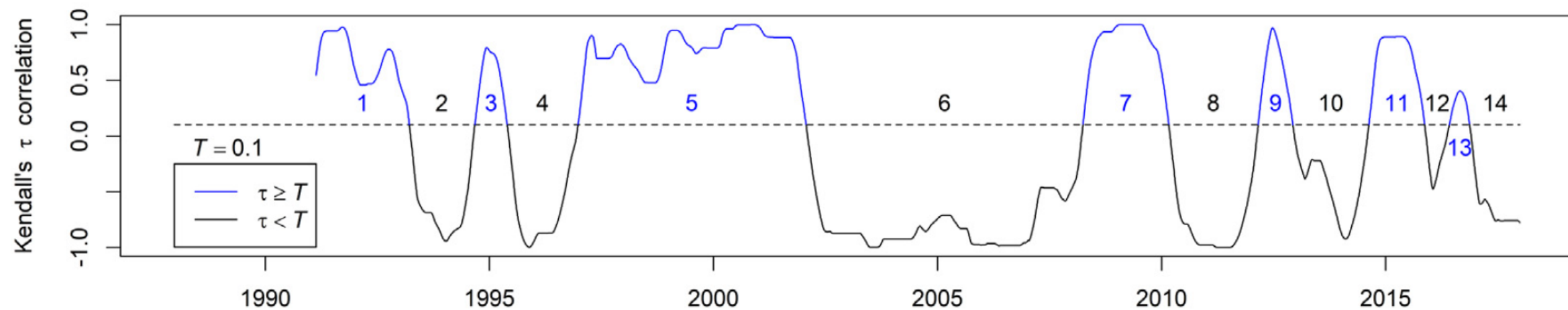
$$T_{Singapore} = 0.85 \quad T_{Malaysia} = 0.96$$

Indicators of L^1 -norms
for Malaysia stock
indices prior to
Lehman Brothers
bankruptcy (09/15/2008)



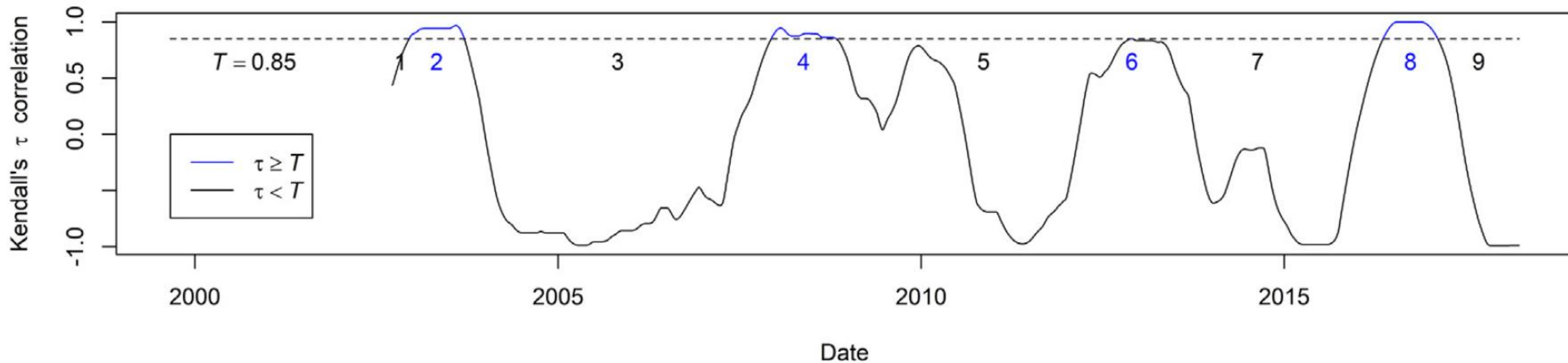
US

(a) Kendall's τ correlations of mean power spectrum at low frequencies of L^1 -norms for the US market



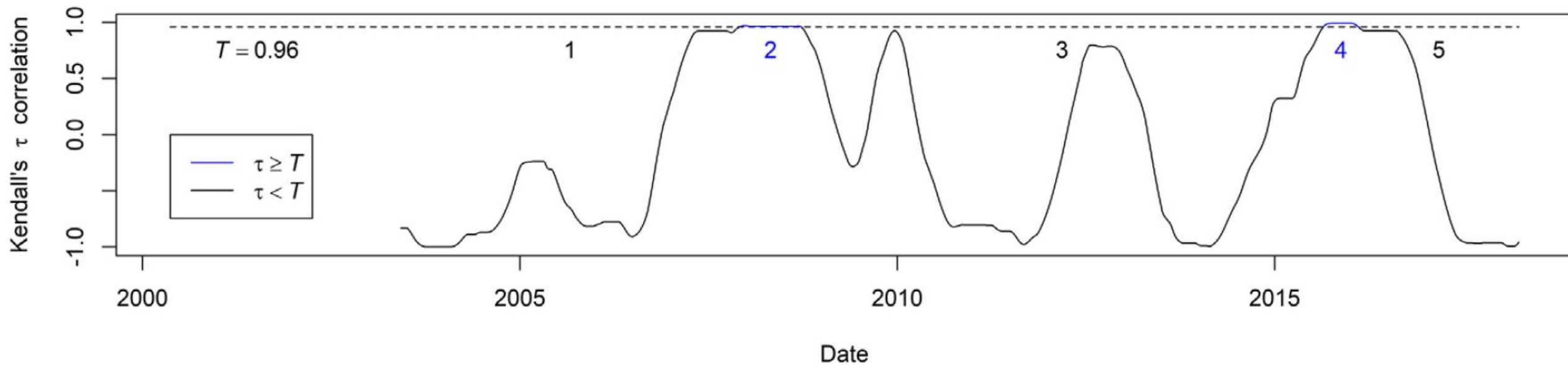
Singapore

(b) Kendall's τ correlations of mean power spectrum at low frequencies of L^1 -norms for Singapore market



Malaysia

(c) Kendall's τ correlations of mean power spectrum at low frequencies of L^1 -norms for Malaysia market



Financial Crises

US	Singapore	Malaysia
<ol style="list-style-type: none"> 1. The global mini-crash caused by the 1997 Asian economic crisis (10/27/1997) 2. Dotcom crash (03/10/2000) 3. September 11 attack (09/11/2001) 4. 2002 stock market downturn (10/09/2002) 5. Chinese stock bubble of 2007 (02/27/2007) 6. US bear market of 2007 (10/11/2007) 7. Lehman Brothers Bankruptcy (09/15/2008) 8. 2009 Dubai debt standstill (11/27/2009) 9. 2010 flash crash (03/06/2010) 10. 2015-2016 Chinese stock market crash (06/12/2015) 11. 2015-2016 US stock market sell-off (08/18/2015) 	<ol style="list-style-type: none"> 1. Chinese stock bubble of 2007 (02/27/2007) 2. US bear market of 2007 (10/11/2007) 3. Lehman Brothers Bankruptcy (09/15/2008) 4. 2009 Dubai debt standstill (11/27/2009) 5. 2015-2016 Chinese stock market crash (06/12/2015) 	<ol style="list-style-type: none"> 1. US bear market of 2007 (10/11/2007) 2. Lehman Brothers Bankruptcy (09/15/2008) 3. 2009 Dubai debt standstill (11/27/2009)

Event Classification

- True Positive (EWS): Signal & Crisis
- False Alarm (FA): Signal & No Crisis
- False Negative (FN): No Signal & Crisis
- True Negative (TN): No Signal & No Crisis

Performance Metrics

Let

- A = # of True Positive (EWS)
- B = # of False Alarm (FA)
- C = # of False Negative (FN)
- D = # of True Negative (TN)
- Probability of Successful Anticipation:
$$\frac{A + D}{A + B + C + D} \times 100\%$$
- Probability of Erroneous Anticipation:
$$\frac{B + C}{A + B + C + D} \times 100\%$$

MPS Results on PL L1-Norm

Market	Score name	Method
		MPS of the L^1 -norms
US	Probability of successful anticipation (%)	60
	Probability of erroneous anticipation (%)	40
Singapore	Probability of successful anticipation (%)	30
	Probability of erroneous anticipation (%)	70
Malaysia	Probability of successful anticipation (%)	40
	Probability of erroneous anticipation (%)	60

Residual Time Series

- Use only the first stock index

	1 st index (leading companies of all sectors)	2 nd index (leading companies in the industrial sector)	3 rd index (leading companies in the technology sector)	4 th index (leading small-cap companies)
US ($d = 4$)	S&P 500	DJIA	Nasdaq	Russel 2000
Singapore ($d = 4$)	ST	ST Ind	ST Tech	ST SC
Malaysia ($d = 3$)	KLCI	KLSE Ind	KLSE Tech	

Residual Time Series

- $\{x_i(t)\}_{t=1}^{1000}$: closing prices before financial crisis

- **Smoothed index:**

$$y_i(t) = \frac{\sum_{s=1}^{1000} K\left(\frac{t-s}{h}\right) x_i(s)}{\sum_{s=1}^{1000} K\left(\frac{t-s}{h}\right)}, \quad K(u) = \frac{1}{\sqrt{2\pi}} e^{-u^2/2}, \quad h = 25$$

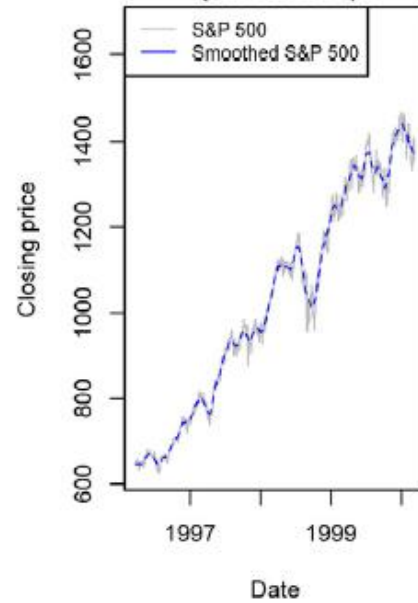
- t : target trading day
- s : index over all days in window
- h : bandwidth (25 days)
- $K(u)$: Gaussian kernel weighting function

- **Residual time series:**

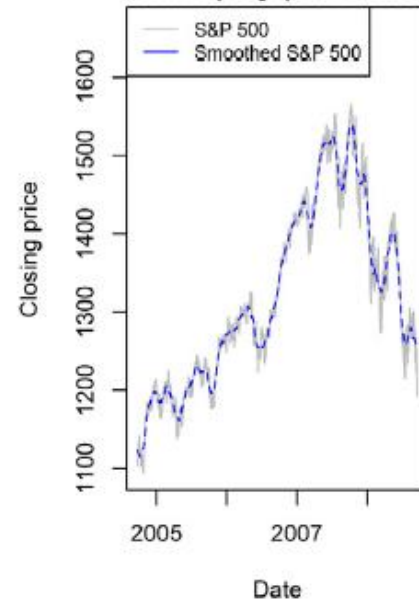
$$res_i(t) = x_i(t) - y_i(t)$$

Res

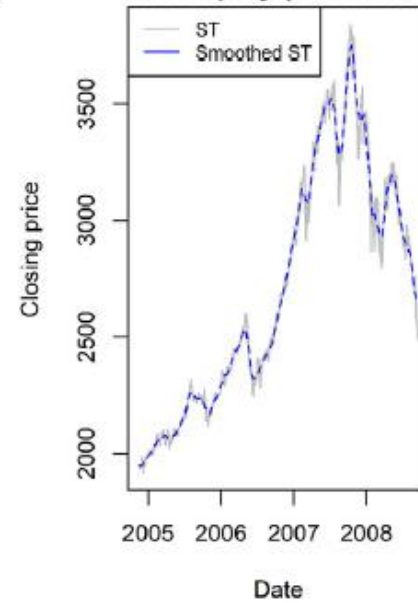
(a) S&P 500 of the US prior to Dotcom crash (03/10/2000)



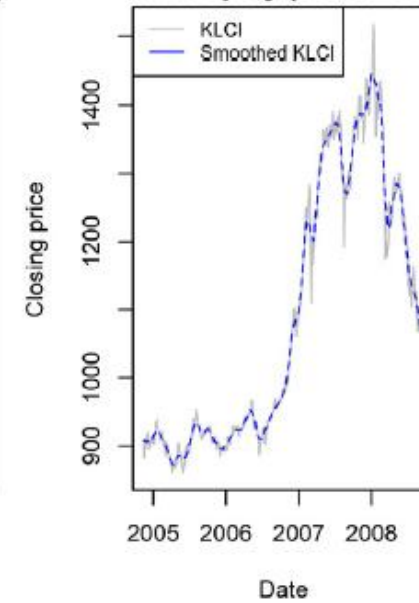
(c) S&P 500 of the US prior to Lehman Brothers bankruptcy (09/15/2008)



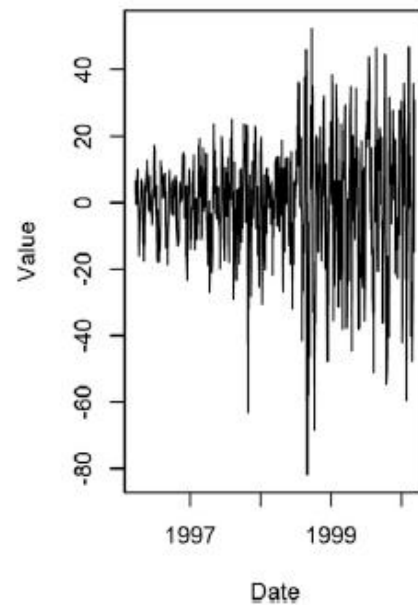
(e) ST of Singapore prior to Lehman Brothers bankruptcy (09/15/2008)



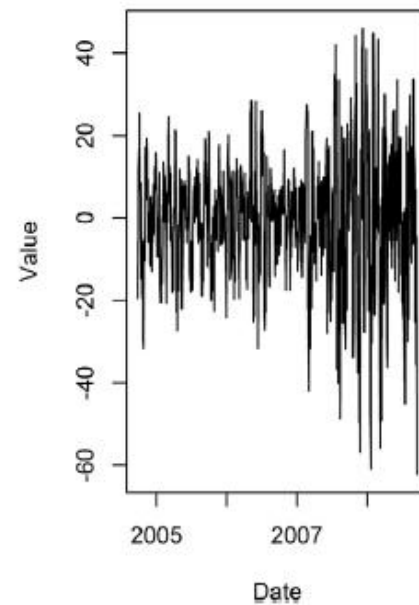
(g) KLCI of Malaysia prior to Lehman Brothers bankruptcy (09/15/2008)



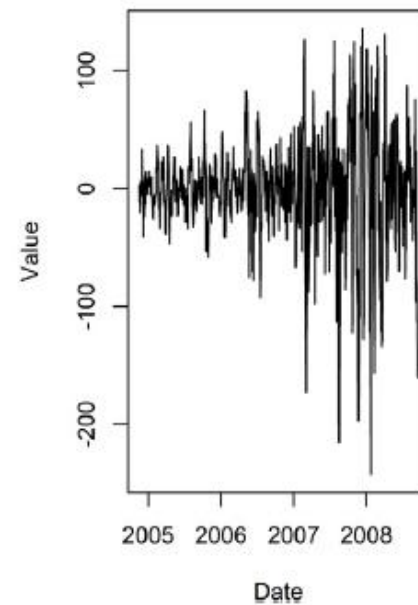
(b) Residuals from (a)



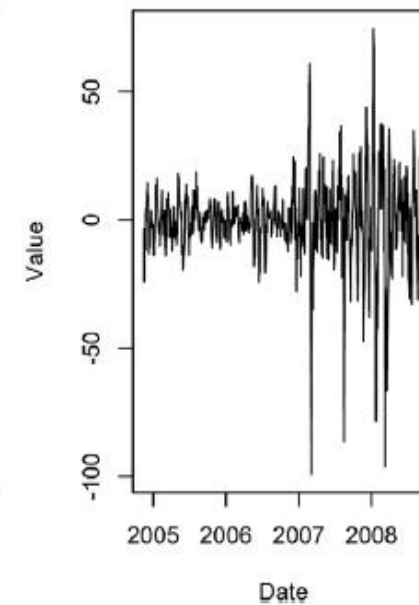
(d) Residuals from (c)



(f) Residuals from (e)



(h) Residuals from (g)



Residual Time Series EWS Workflow

1. Compute Residuals $\{res_i(t)\}_{t=1}^{1000}$ for $i = 1$

2. Compute CSD indicators

ACF1, Variance, MPS on $\{res_i(t)\}_{t=1}^{1000}$ with $w_2 = 500$

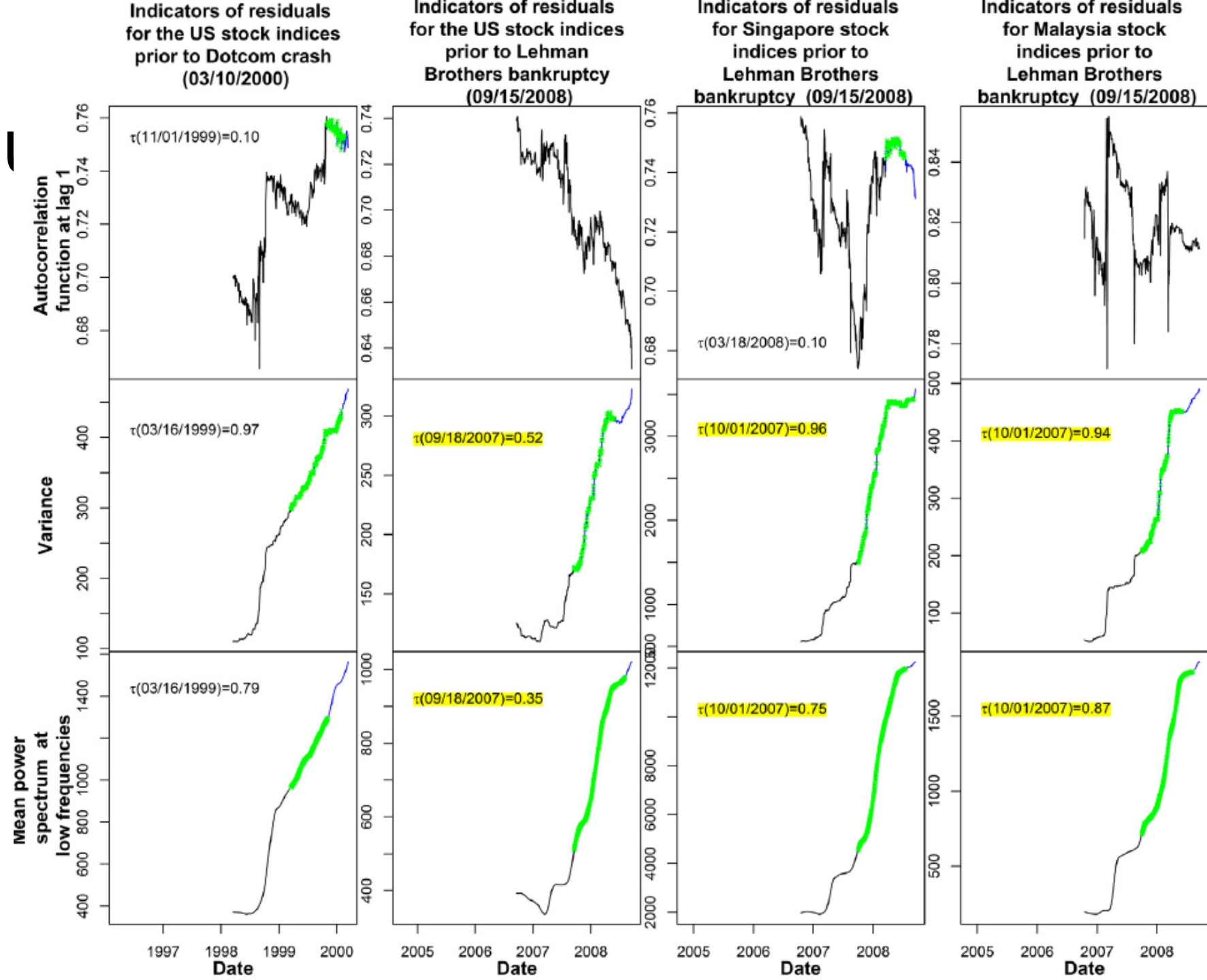
3. Significance Test

Mann-Kendall test with $w_3 = 250$

4. Structural Break Test (Chow test)

5. Threshold Application

Resi



Overall Results

Table 12

Summary for the classification scores obtained of the L^1 -norms and the residuals.

Market	Score name	Method		
		MPS of the L^1 -norms	VAR of the residuals	MPS of the residuals
US	Probability of successful anticipation (%)	60	55	64
	Probability of erroneous anticipation (%)	40	45	36
Singapore	Probability of successful anticipation (%)	30	20	38
	Probability of erroneous anticipation (%)	70	80	62
Malaysia	Probability of successful anticipation (%)	40	50	67
	Probability of erroneous anticipation (%)	60	50	33

Overall Results

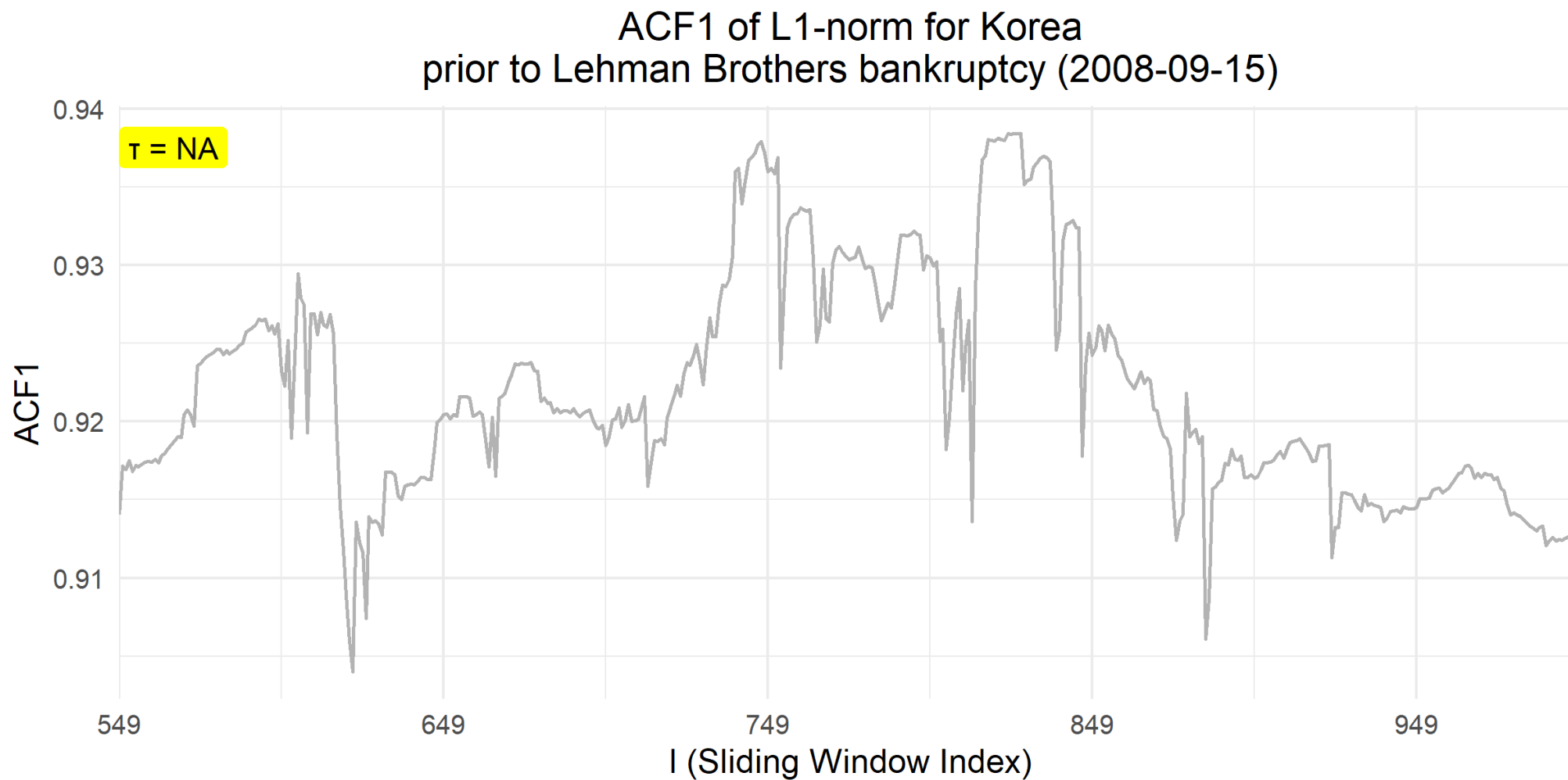
- The US's very low threshold yielded more EWS detections, whereas higher thresholds in Singapore and Malaysia led to fewer signals.
- Malaysia's weaker L1-norm performance likely stems from using only 3 indices versus 4 in the US and Singapore.
- Inappropriate window sizes for Singapore and Malaysia further degraded results, underscoring parameter sensitivity.

Extending to the Korean Market

	1 st index (leading companies of all sectors)	2 nd index (leading companies in the industrial sector)	3 rd index (leading companies in the technology sector)	4 th index (leading small-cap companies)
US ($d = 4$)	S&P 500	DJIA	Nasdaq	Russel 2000
Singapore ($d = 4$)	ST	ST Ind	ST Tech	ST SC
Malaysia ($d = 3$)	KLCI	KLSE Ind	KLSE Tech	
Korea ($d = 4$)	KOSPI 200	KOSPI Manufacturing	KOSPI Electronics	Kosdaq

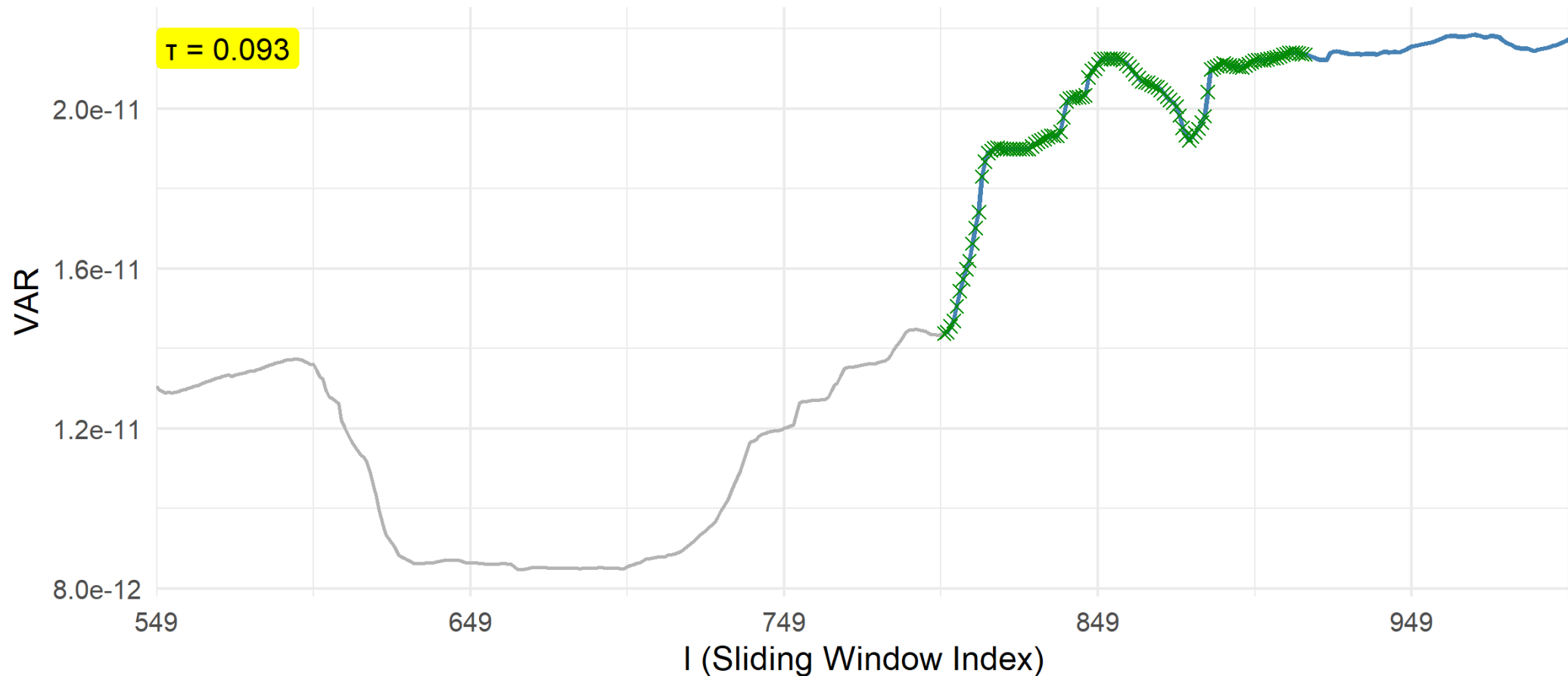
Key Financial Crises Analyzed: Lehman Brothers Bankruptcy (09/15/2008)

ACF1 on L1-norm - Korean Market



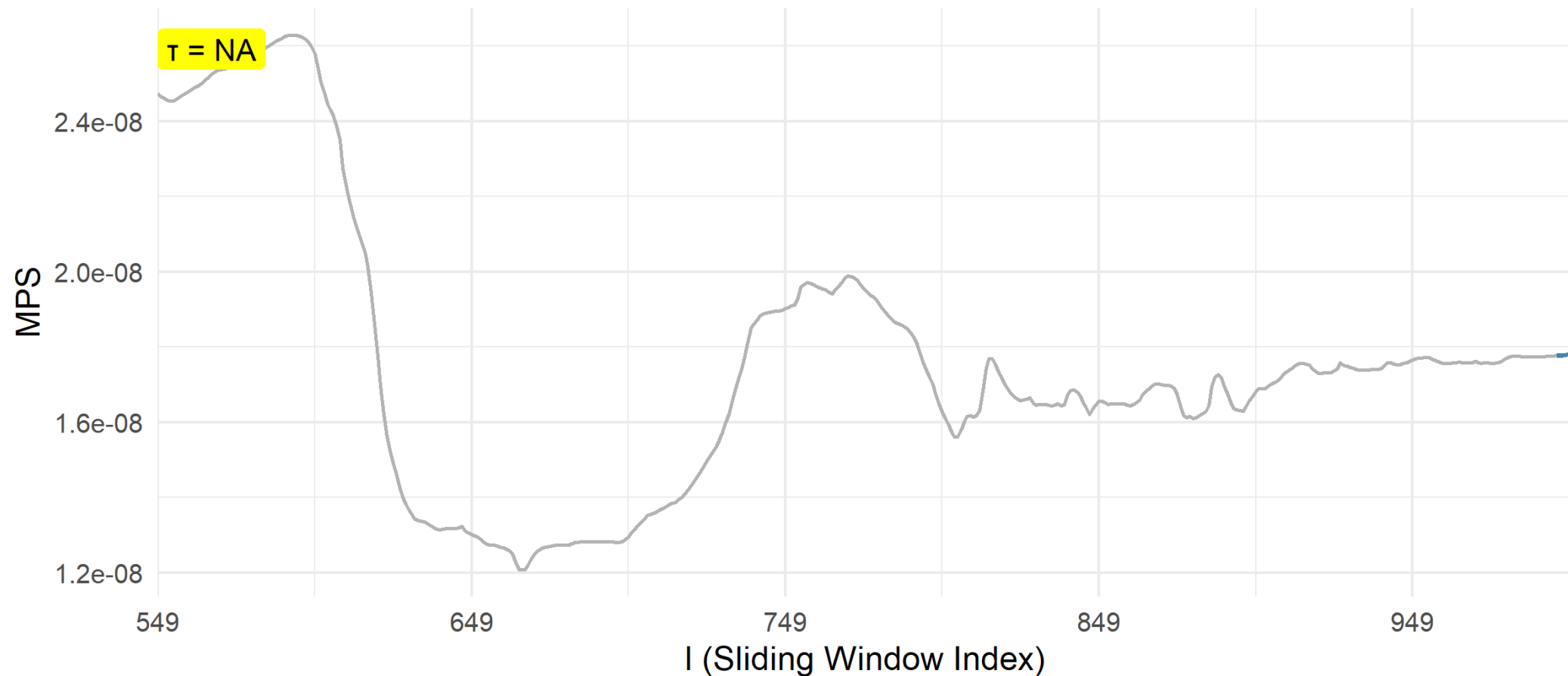
VAR on L1-norm - Korean Market

VAR of L1-norm for Korea
prior to Lehman Brothers bankruptcy (2008-09-15)



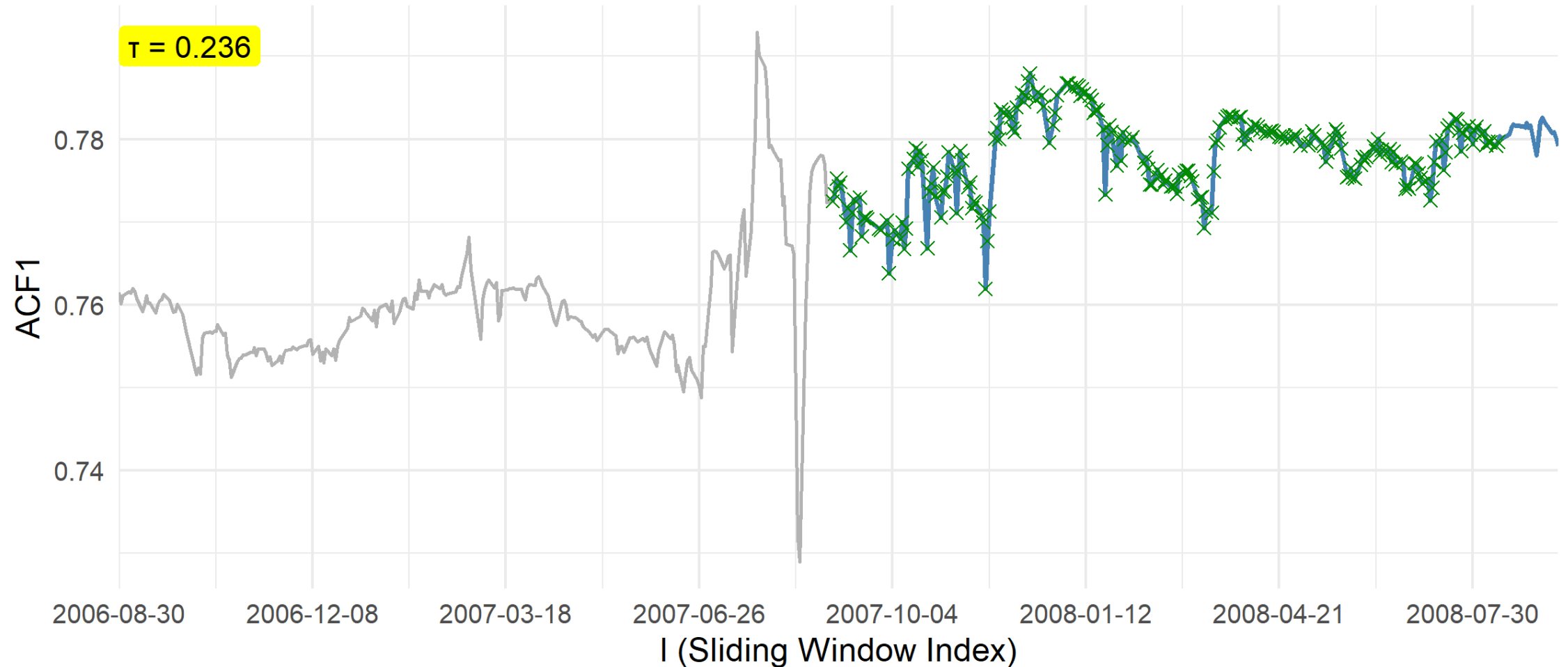
MPS on L1-norm - Korean Market

MPS of L1-norm for Korea
prior to Lehman Brothers bankruptcy (2008-09-15)



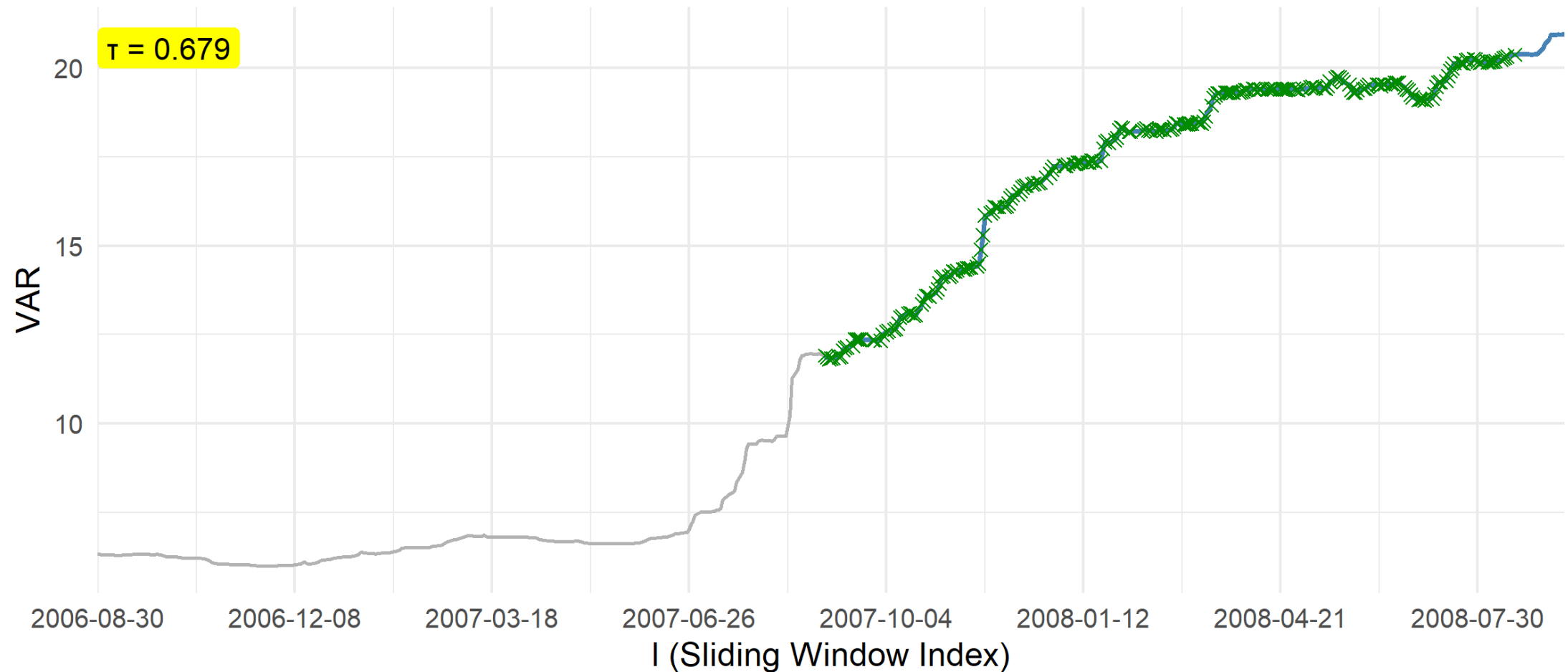
ACF1 on Residual – Korea Market

ACF1 of Residuals for Korea
prior to Lehman Brothers bankruptcy (2008-09-15)



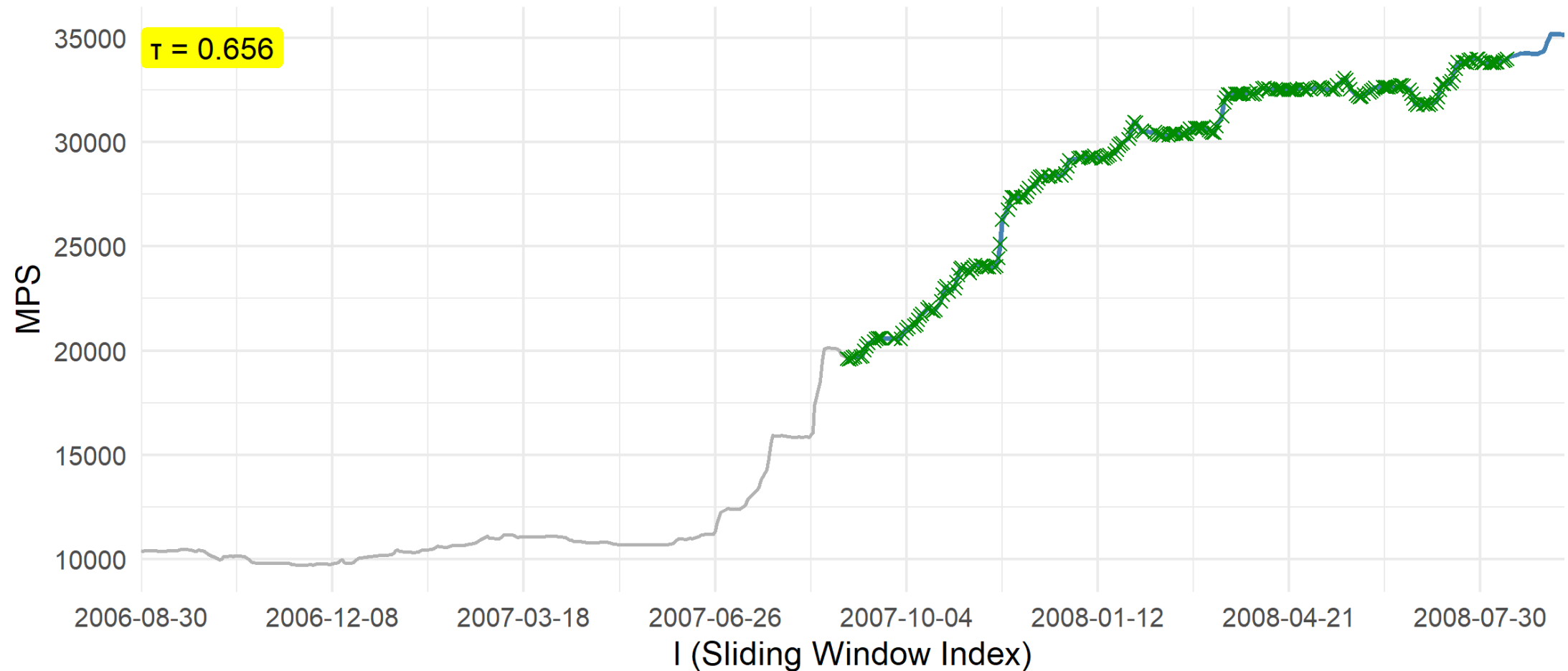
VAR on Residual – Korea Market

VAR of Residuals for Korea
prior to Lehman Brothers bankruptcy (2008-09-15)



MPS on Residual – Korea Market

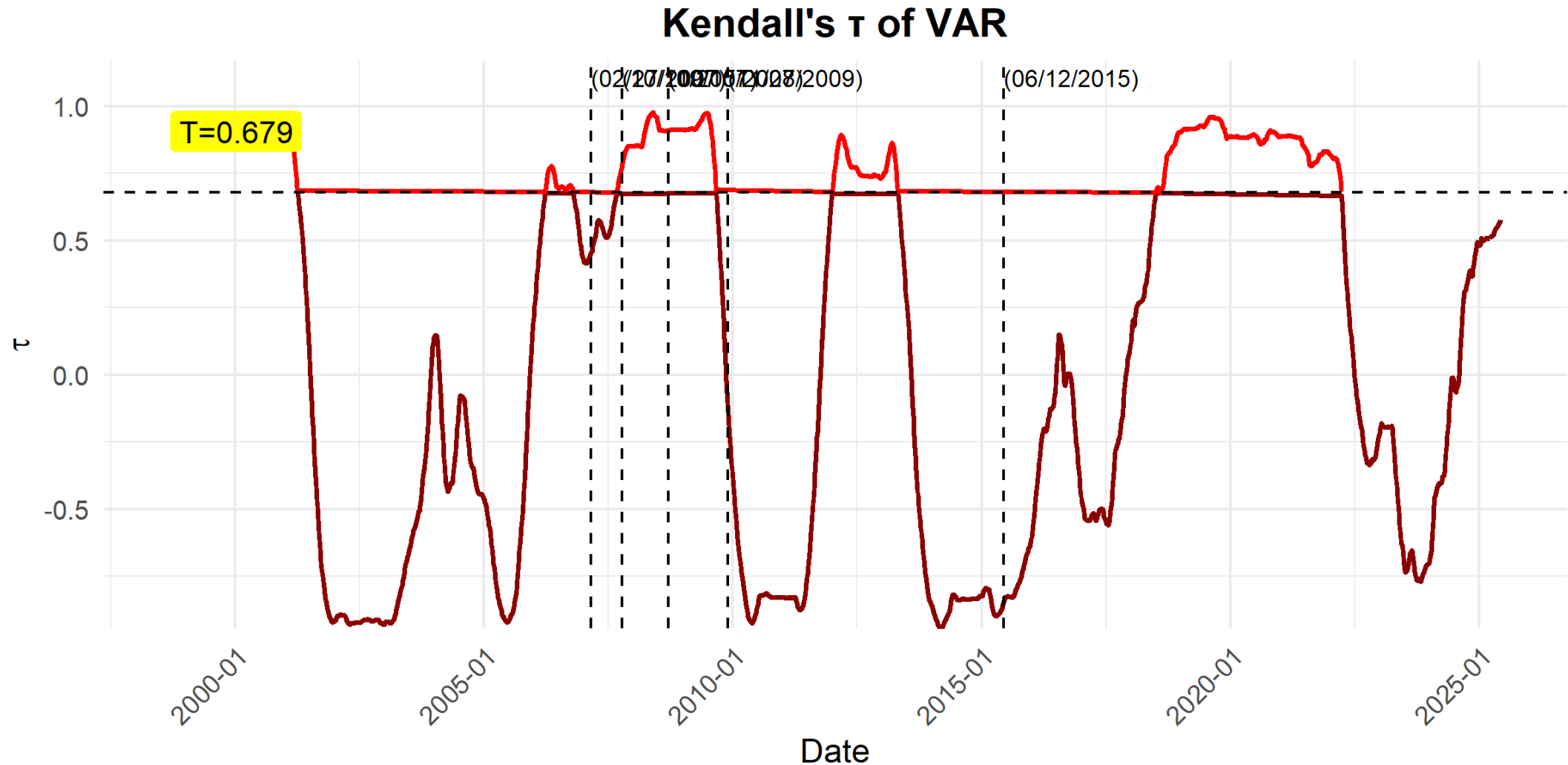
MPS of Residuals for Korea
prior to Lehman Brothers bankruptcy (2008-09-15)



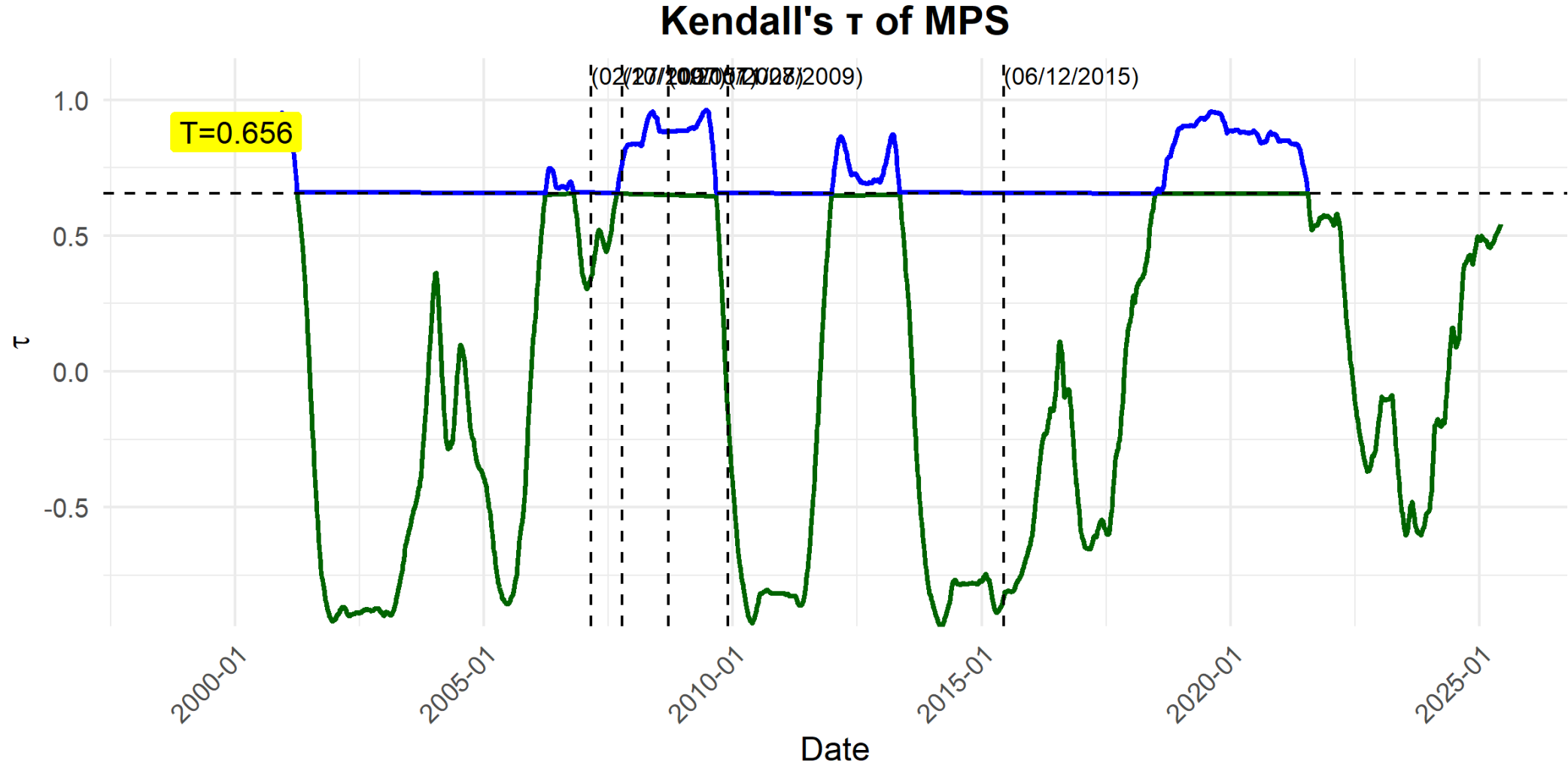
Financial Crisis

Korea
<ol style="list-style-type: none">1. Chinese stock bubble of 2007 (02/27/2007)2. US bear market of 2007 (10/11/2007)3. Lehman Brothers Bankruptcy (09/15/2008)4. 2009 Dubai debt standstill (11/27/2009)5. 2015-2016 Chinese stock market crash (06/12/2015)

VAR on Residual – Korea Market



MPS on Residual – Korea Market



Results for Korea

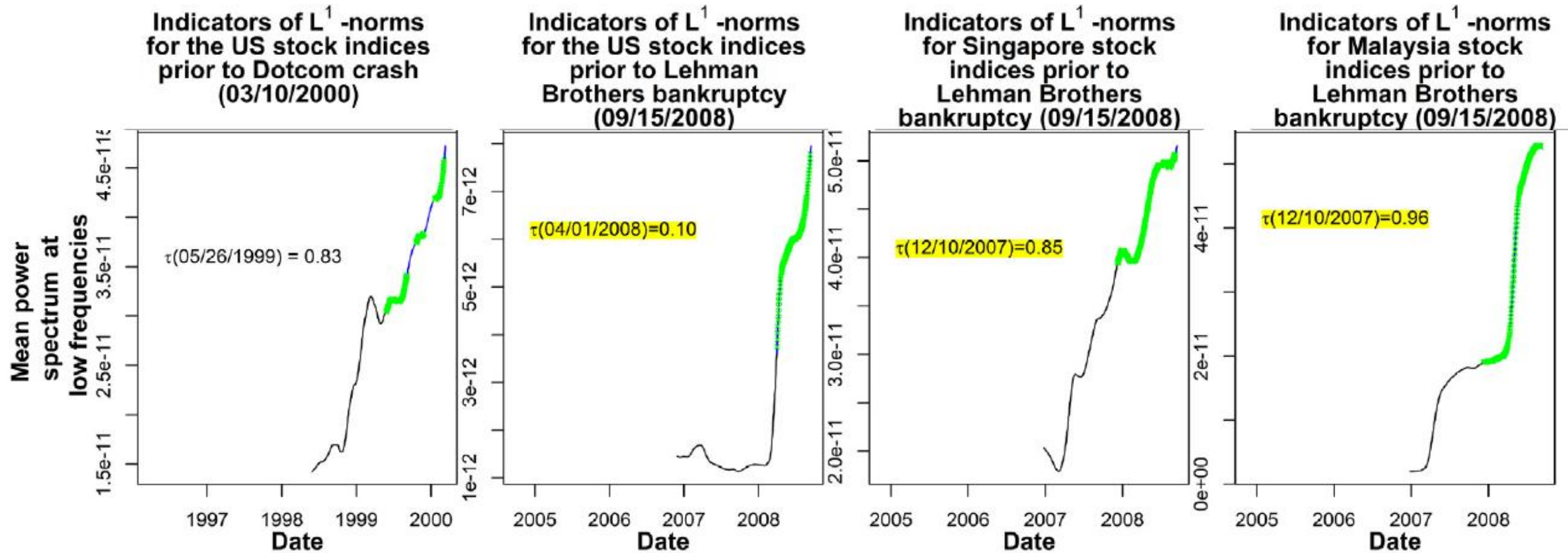
Table 12

Summary for the classification scores obtained of the L^1 -norms and the residuals.

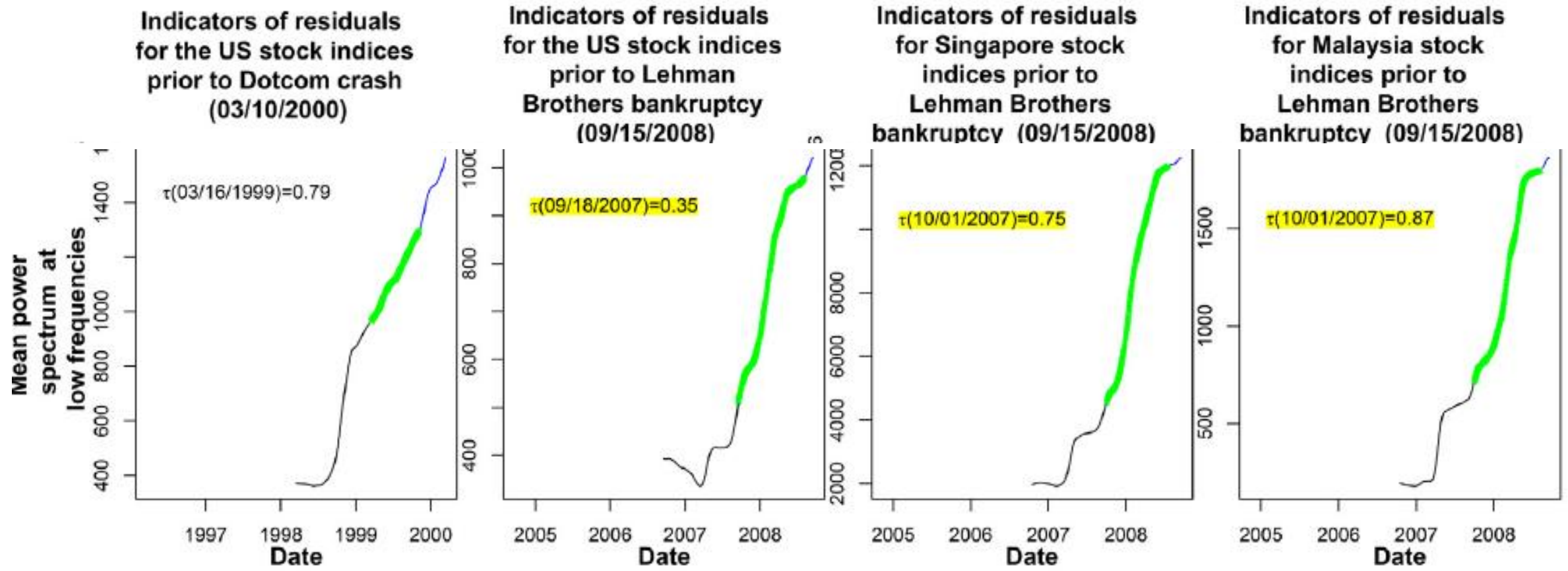
Market	Score name	Method		
		MPS of the L^1 -norms	VAR of the residuals	MPS of the residuals
US	Probability of successful anticipation (%)	60	55	64
	Probability of erroneous anticipation (%)	40	45	36
Singapore	Probability of successful anticipation (%)	30	20	38
	Probability of erroneous anticipation (%)	70	80	62
Malaysia	Probability of successful anticipation (%)	40	50	67
	Probability of erroneous anticipation (%)	60	50	33
Korea	Probability of successful anticipation (%)		54	54
	Probability of erroneous anticipation (%)		46	46

Discussions

MPS on L1-norms



MPS on Residuals



Step 7-3: Mean power spectrum (MPS) at low frequencies

- **Discrete Fourier Transform:**

$$F_k(l) = \sum_{n=l-500+1}^l y(n) e^{\frac{-2\pi i k n}{500}}$$

where $k = 1, \dots, 500$.

- **Power Spectrum:**

$$PS_k(l) = |F_k(l)|^2$$

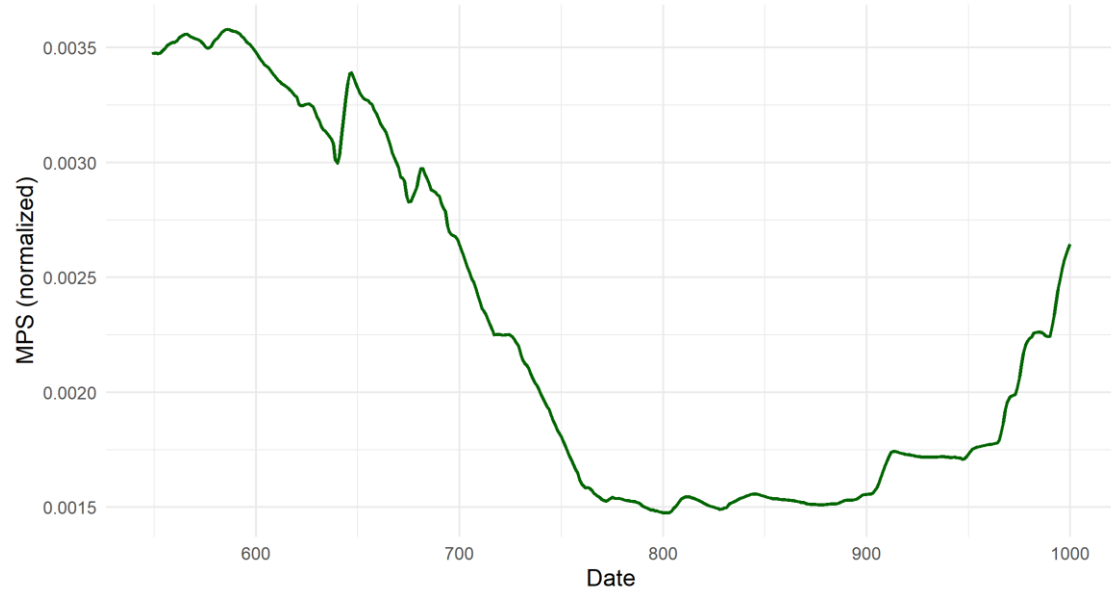
Each $PS_k(l)$ is **normalized** such that its sum is equal to 1.

- **MPS Value:**

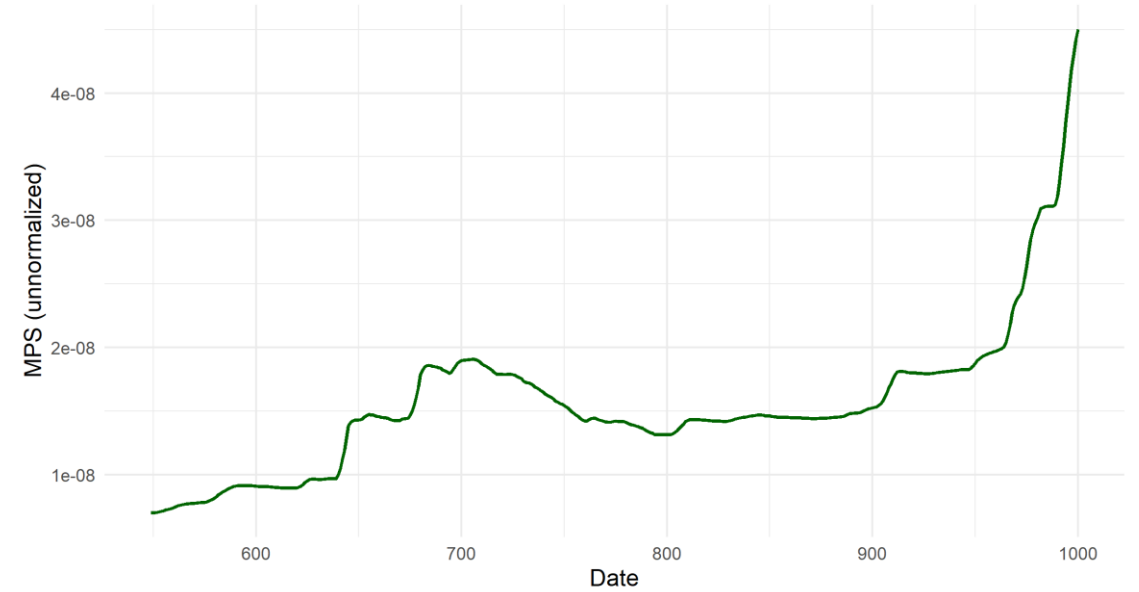
$$mps(l) = \frac{1}{[500/8] - 1} \sum_{k=2}^{[500/8]} PS_k(l)$$

MPS on L1-norms for US

**Mean Power Spectrum (Low Frequencies)
of L1-norms for the US stock indices
prior to Dotcom crash (03/10/2000)**



**Mean Power Spectrum (Low Frequencies)
of L1-norms for the US stock indices
prior to Dotcom crash (03/10/2000)**



Conclusions

- Persistent Homology + CSD Indicators:
 - Vietoris-Rips complex → Persistence landscape → ACF1, VAR, MPS
- PH-Based L1-Norm Time Series:
 - Strong upward trend before Dotcom & Lehman crises
- Top Methods Identified:
 - MPS on residuals (best overall)
 - MPS on L1-norms (close second)
 - VAR on residuals

Thank you