Sichuan University

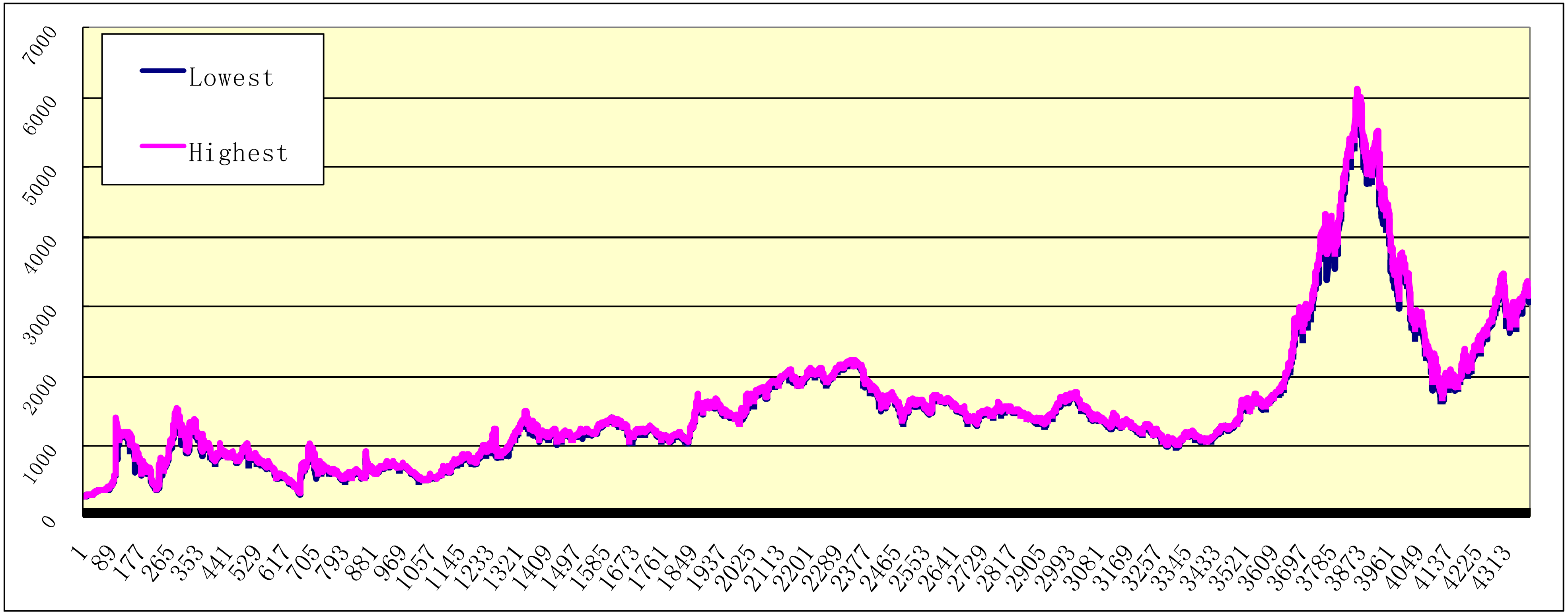
One Century-old Famous University

1896-2011




上证指数原始区间观测值序列

An original observations of the interval time series from SCI

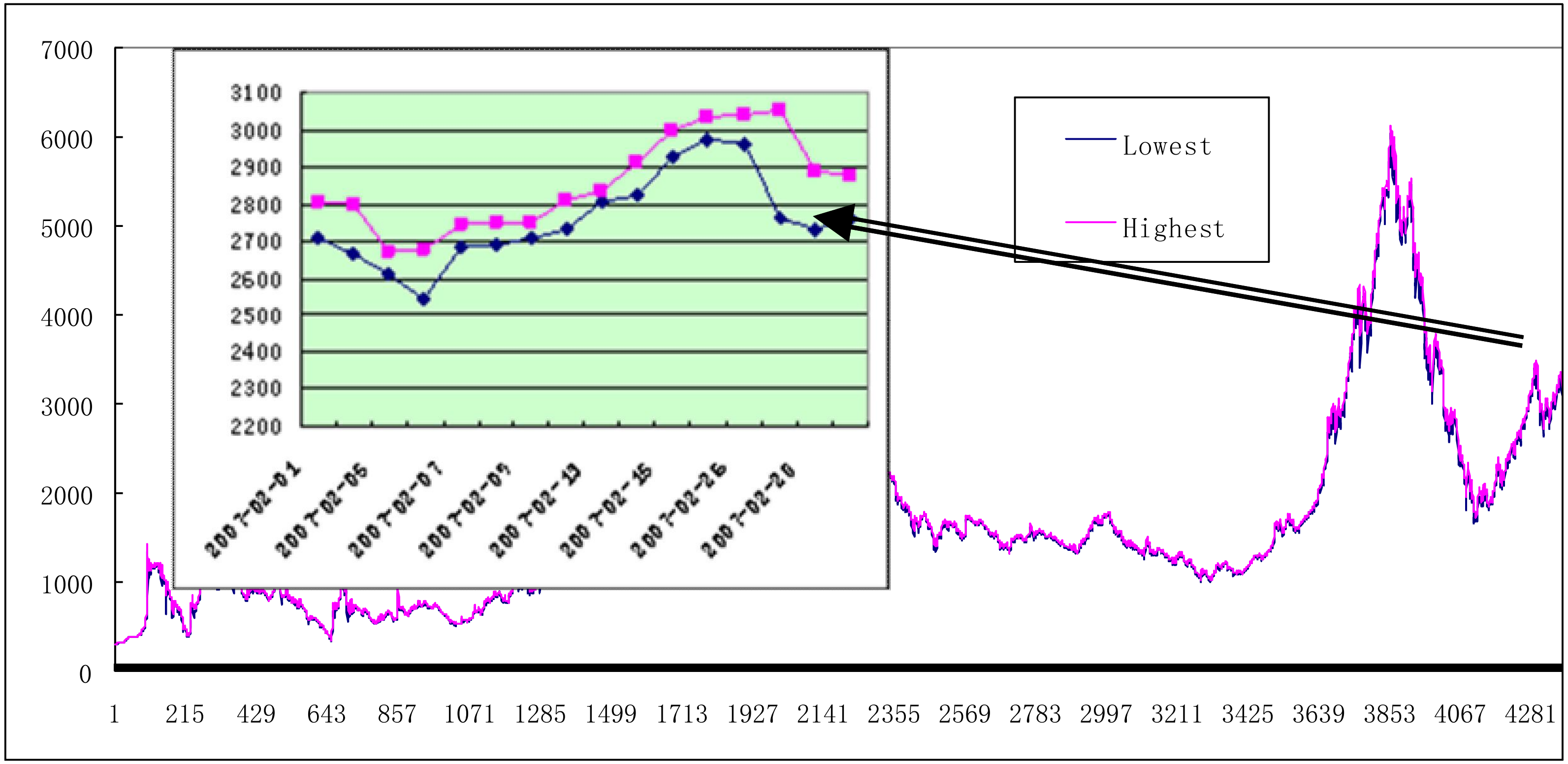


Shanghai Composite Index (SCI)（1992.1.2 -- 2009.12.01）

[lowest, highest]

上证指数原始区间观测值序列

An original observation of the interval time series from SCI



Shanghai Composite Index (SCI)（1992.1.2 -- 2009.12.01）

[lowest, highest]

Study on the nonlinear structure of the interval finance time series

Zhuyu LI Taiji WANG

Mathematical College, Sichuan University

December 16-19, 2011, Taipei

- **Is it possible that we can observe the volatility of the financial market from another perspective?**

- **What we could get from a study on the structure of the dynamic moving from the Interval-Valued Financial Time Series (IVFTS) data**

* Supported by the fund of National Philosophy and Social Science

四川大學
Sichuan University

Motivation:

- ✓In the financial market, the observations are always observed in a random way, with a fuzzy status. Such observations usually are observed in a interval-valued time series.

- ✓Based on the interval-valued time observations, we would like to try to study the dynamic moving characters, by a fuzzy view.

- ✓We will try to set up a model, estimate unknown parameters, and try to fit the model with some evaluation tools, especially on the nonlinear feature.

Explore:

- 金融区间序列分析的基础（描述方法）
- **The basic structure of the financial IVTS (How to describe)**
- 模型构建（分析平台）
- **Trying to do the Modeling (How to analyze)**
- 未知参数估计（估计方法）
- **Estimating unknown parameters in the model (How to estimate)**
- 模型拟合评价（适用标准）
- **Trying to get the way to Fit and evaluate models (How to evaluate)**
- 预测效果 (模型比较）
- **Comparing the results of the prediction on the different models (Which is better)**

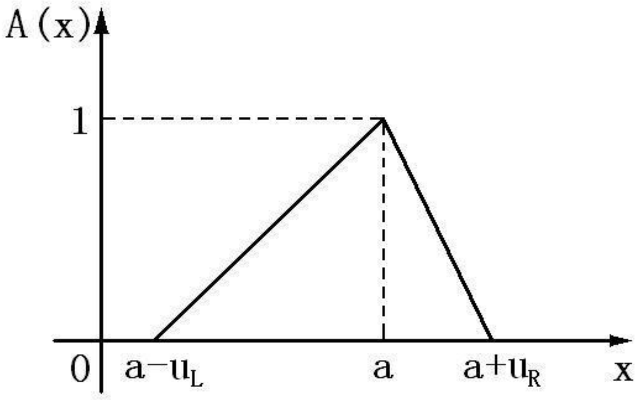
We need some concept first

区间金融时间序列、区间金融收益率序列及其平稳性

- ✓ **Interval-valued financial time series (IVFTS)**
- ✓ **Interval-valued financial yield series (IVFYS) and its stability**

A triangular fuzzy number:

$$A(x)=\begin{cases}\dfrac{x-a+u_L}{u_L} & a-u_L\le x\le a\\[2ex] \dfrac{-x+a+u_R}{u_R} & a<x\le a+u_R\\[2ex] 0 & x<a-u_L \ \text{ or } \ x>a+u_R\end{cases}$$



A triangular fuzzy data : $A=\left(a,u_L,u_R\right)$

a Is the centre of the triangular fuzzy data

u_L,u_R the left and right spread, respectively

$\left[a-u_L,a+u_R\right]$ a bounded support set of A

A fuzzy distance between the symmetrical triangular fuzzy data:

$\tilde{A}_1, \tilde{A}_2 \in F\left(R\right)$ is the symmetrical triangular fuzzy data.

$\tilde{A}_i = (a_i, u_i), i = 1, 2$

A fuzzy distance between $\tilde{A}_1$ and $\tilde{A}_2$ is:

$$d\left(\tilde{A}_1, \tilde{A}_2\right) = \sqrt{(a_1 - a_2)^2 + (u_1 - u_2)^2}$$

Interval-valued financial time series **(IVFTS)**

an interval-valued observation time series:

$$\left\{[x_t^L, x_t^U], t = 1, \ldots, N\right\} \qquad (1)$$

In which x_t^L x_t^U the lowest and highest points at time *t*, respectively. We can get

$$y_t = \frac{x_t^L + x_t^U}{2} \qquad z_t = \frac{x_t^U - x_t^L}{2}$$

Base on the definition of the symmetrical triangular fuzzy data, (1) is equal to :

$$\left\{(y_t, z_t), t = 1, \ldots, N\right\} \qquad (2)$$

(2): a form of the symmetrical triangular fuzzy series,
(1): its bounded support set.

Definition 1: a interval-valued financial yield series (IVFYS) $\tilde{r}_t$ ***区间金融收益率序列* (IVFYS)**

$$\{\tilde{r}_t\} = \left\{ \left[\ln \frac{x_t^L}{x_{t-1}^U}, \ln \frac{x_t^U}{x_{t-1}^L} \right] \middle| t = 2, 3, \ldots N \right\} \quad (3)$$

Down limit: $\ln \dfrac{x_t^L}{x_{t-1}^U}$ the lowest yield of a financial asset at time t.

Up limit: $\ln \dfrac{x_t^U}{x_{t-1}^L}$ the highest yield of a financial asset at time t.

Base on the symmetrical triangular fuzzy data, we can get

a centre series: $$\{c_t\} = \{\frac{1}{2}[\ln \frac{x_t^L}{x_{t-1}^U} + \ln \frac{x_t^U}{x_{t-1}^L}], t = 2, \dots N\}$$

反映了金融资产收益率的集中趋势

reflects the trend of the centre series based on a IVFYS

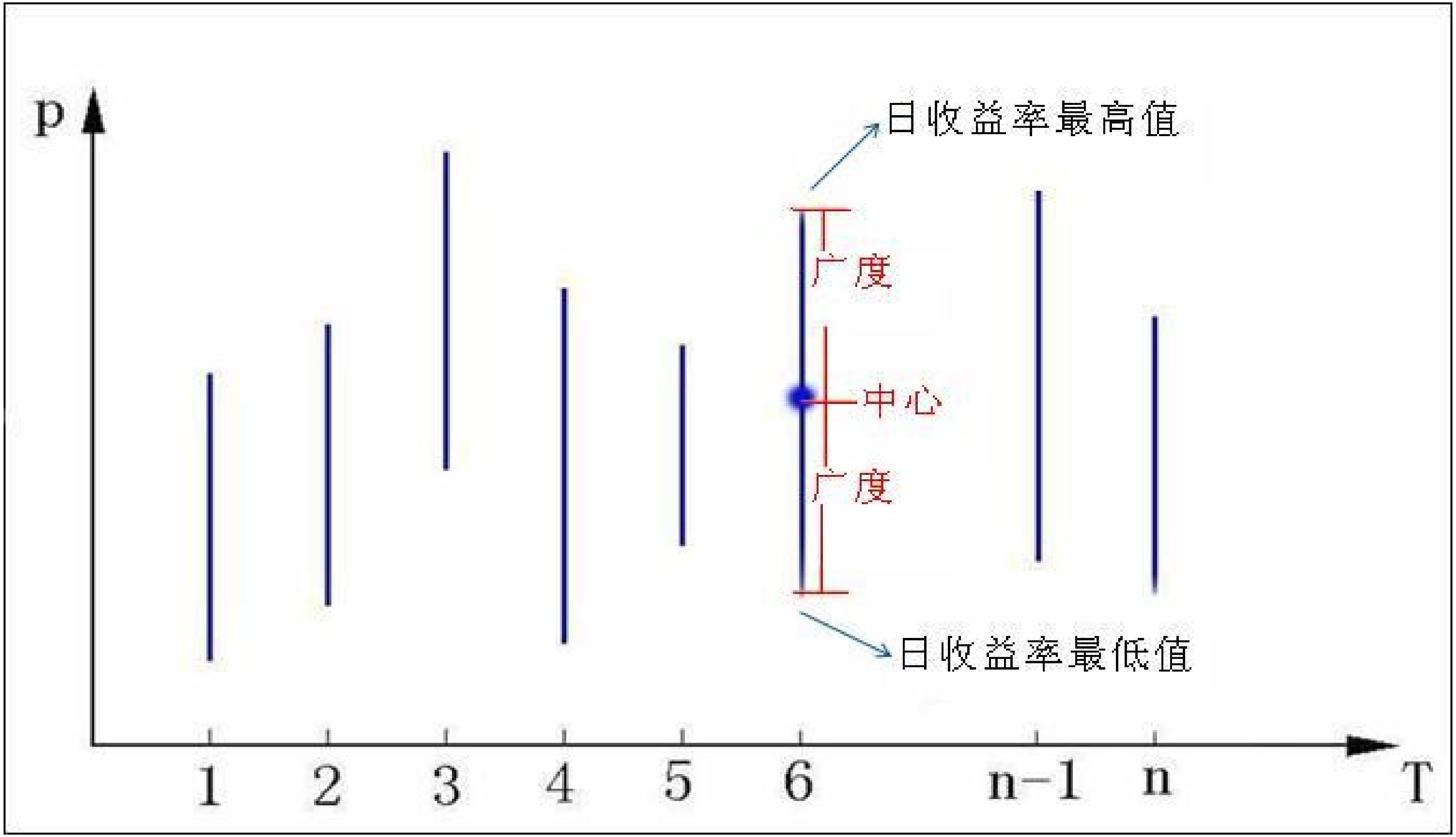
a spread series: $$\{u_t\} = \{\frac{1}{2}[\ln \frac{x_t^U}{x_{t-1}^L} - \ln \frac{x_t^L}{x_{t-1}^U}], t = 2, \dots N\}$$

反映了金融资产收益率非随机波动性的大小.

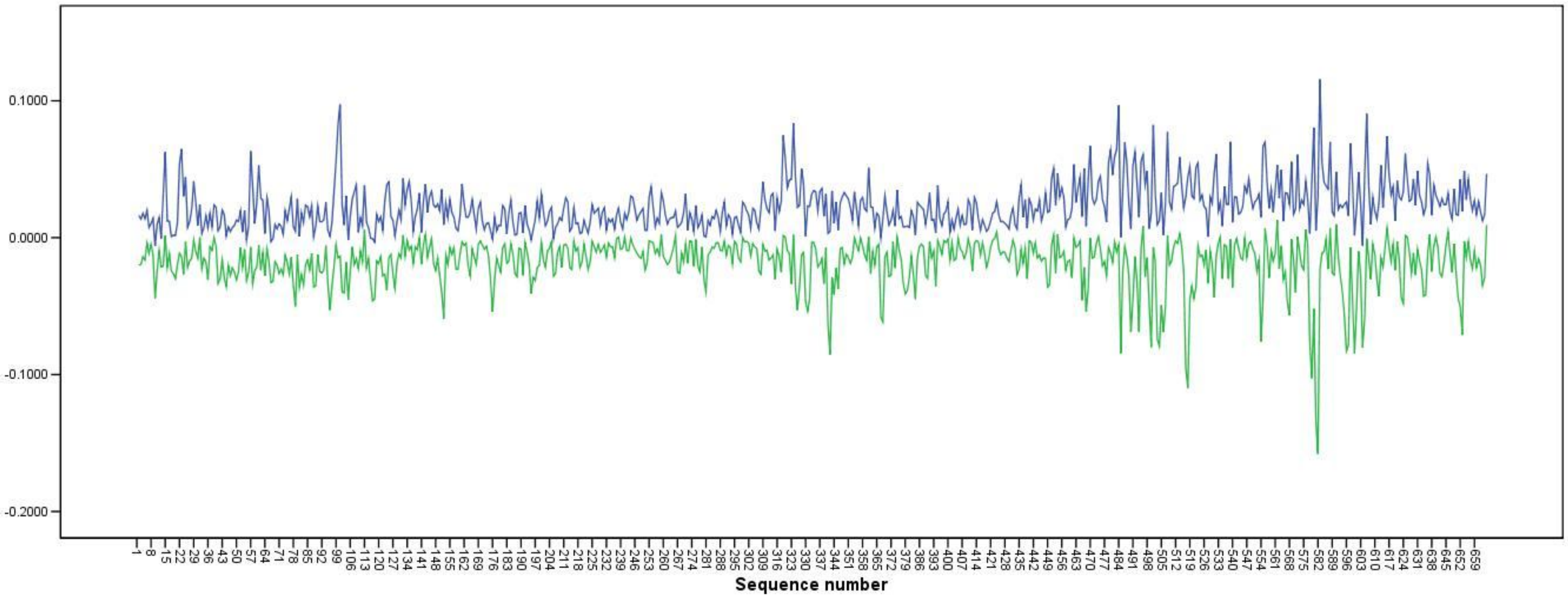
reflects the unrandom volatility of a IVFYS

(3) is equal to $$\boxed{\tilde{r}_t = \left[\ln \frac{x_t^L}{x_{t-1}^U}, \ln \frac{x_t^U}{x_{t-1}^L}\right] = (c_t, u_t),} \qquad t = 2,3,\dots\dots N$$

Centre-Spread Structure of IVFYS



A diagrammatic presentation of a IVFYS in Definition 1



a IVFYS of Shanghai Stock Market （2005.1.4-2007.9.28）

- it is easy to point out that the changes of the centre series on this bin is in (-0.2,0.2). Based on this view, we may infer: the stability of the centre series can tell the stability of the original fuzzy series, conditionally.

Definition 2:

a conditional stability of a IVFYS

Let

$$\{\tilde{r}_t\} = \{\left[\ln \frac{x_t^L}{x_{t-1}^U}, \ln \frac{x_t^U}{x_{t-1}^L}\right], t = 2, 3, \ldots N\}$$

be a IVFYS, when its centre series $\{c_t\}$ is stable, then we say

$\{\tilde{r}_t, t = 2, \ldots, N\}$ is a conditional stable IVFYS.

➢ **Modeling of IVFYS**

- Fuzzy Auto-Regression FAR(p)
- Fuzzy Double Linear Regression FDLR(p,q)
- **Explore** of a nonlinear structure

Fuzzy Auto-Regression FAR(*p*)

LI Zhuyu, ZHANG Cheng (2008), **Structure Analysis of Regression Model with Fuzzy Data**, *STATISTICAL RESEARCH*, 25 (8), 74-78.

LI Zhuyu, ZHANG Cheng, WANG Taiji (2010), **Studying on the Interval Financial Time Series and Evaluating on the Forecast**, *Journal of Applied Statistics and Management*, Vol.29 (1)： 129-136.

FAR(p)模型

$$\tilde{r}_t = A_1\tilde{r}_{t-1} + A_2\tilde{r}_{t-2} + \cdots + A_p\tilde{r}_{t-p} + \tilde{\varepsilon}_t$$

estimation: $A_i = (a_i, b_i)'$ are the unknown parameters, let its

FLP estimation be $\hat{A}_1, \hat{A}_2, \ldots, \hat{A}_p$.

Then, from the centre series, and the spread series, there is a FLP

estimation of $\tilde{r}_t$ as

$$\hat{\tilde{r}}_t = \left(\sum_{i=1}^{p} \hat{A}_i c_{t-i}, \sum_{i=1}^{p} \left|\hat{A}_i\right| u_{t-i} \right)$$

We work on FAR(*1*) with FLP

Fuzzy Double Linear Regression **FDLR(*p*,*q*)**

LI Zhuyu, LIU Weiyi, WANG Taiji (2009), **Fuzzy Bilinear Regression of Yield Series,** *STATISTICAL RESEARCH*, 26(2): 68-73.

T.J.WANG， W.Y. LIU and Z.Y. LI (2009), **Fuzzy Double Linear Regression of Financial Assets Yield**, <Cutting-Edge Research Topics on Multiple Criteria Decision Making>, pp59-62, Springer.

Model FDLR(*p, q*)

$$\begin{cases} c_t = \alpha_0 + \alpha_1 c_{t-1} + \cdots + \alpha_p c_{t-p} + \varepsilon_t \\ u_t = \beta_0 + \beta_1 u_{t-1} + \cdots + \beta_q u_{t-q} + \gamma c_t + e_t \end{cases}$$

$$t = k+1, \ldots, N, \qquad k = \max(p, q)$$

,

$\alpha_0, \alpha_1, \ldots, \alpha_p, \beta_0, \beta_1, \ldots, \beta_q, \gamma$ are parameters need be estimated.
Model FDLR(*p,q*) is the improved one from FAR(*p*). It tells the relationship between the central trend of the interval financial yield series and its *p* lags at time t, as well as the relationship of its volatile trend (the spread), the *q* lags and the central trend.

Let $\hat{\alpha}_0, \hat{\alpha}_1, \ldots, \hat{\alpha}_p, \hat{\beta}_0, \hat{\beta}_1, \ldots, \hat{\beta}_q, \hat{\gamma}$ be an estimation of the regression coefficients. We can get the following fuzzy regression model:

$$\begin{cases} \hat{c}_t = \hat{\alpha}_0 + \hat{\alpha}_1 c_{t-1} + \cdots + \hat{\alpha}_p c_{t-p} \\ \hat{u}_t = \hat{\beta}_0 + \hat{\beta}_1 u_{t-1} + \cdots + \hat{\beta}_q u_{t-q} + \hat{\gamma} c_t \end{cases}$$

(the data out of the sample could be replaced by the fitting one.)

$\hat{\tilde{r}}_t = (\hat{c}_t, \hat{u}_t) = [\hat{c}_t - \hat{u}_t, \hat{c}_t + \hat{u}_t]$: the estimation of $\tilde{r}_t$,

or the fitting value or forecast in the other words.

$\longrightarrow$ **Estimation Method**

Fuzzy Least Square Estimation (FLSE):

Theorem: Let $\mathbf{1}=\mathbf{1}_{\mathbf{N-1}}=(1,1,\ldots,1)'$, $\mathbf{c}=(c_{k+1},c_{k+2},\ldots,c_N)'$,

$\mathbf{u}=(u_{k+1},u_{k+2},\ldots,u_N)'$, $\mathbf{c}_{-i}=(c_{k+1-i},c_{k+2-i}\ldots,c_{N-i})', i=1,2,\ldots,p,$

$\mathbf{u}_{-j}=(u_{k+1-j},u_{k+2-j}\ldots,u_{N-j})', j=1,2,\ldots,q,$ $\mathbf{X}=(\mathbf{1},\mathbf{c}_{-1},\mathbf{c}_{-2},\ldots,\mathbf{c}_{-p}),$

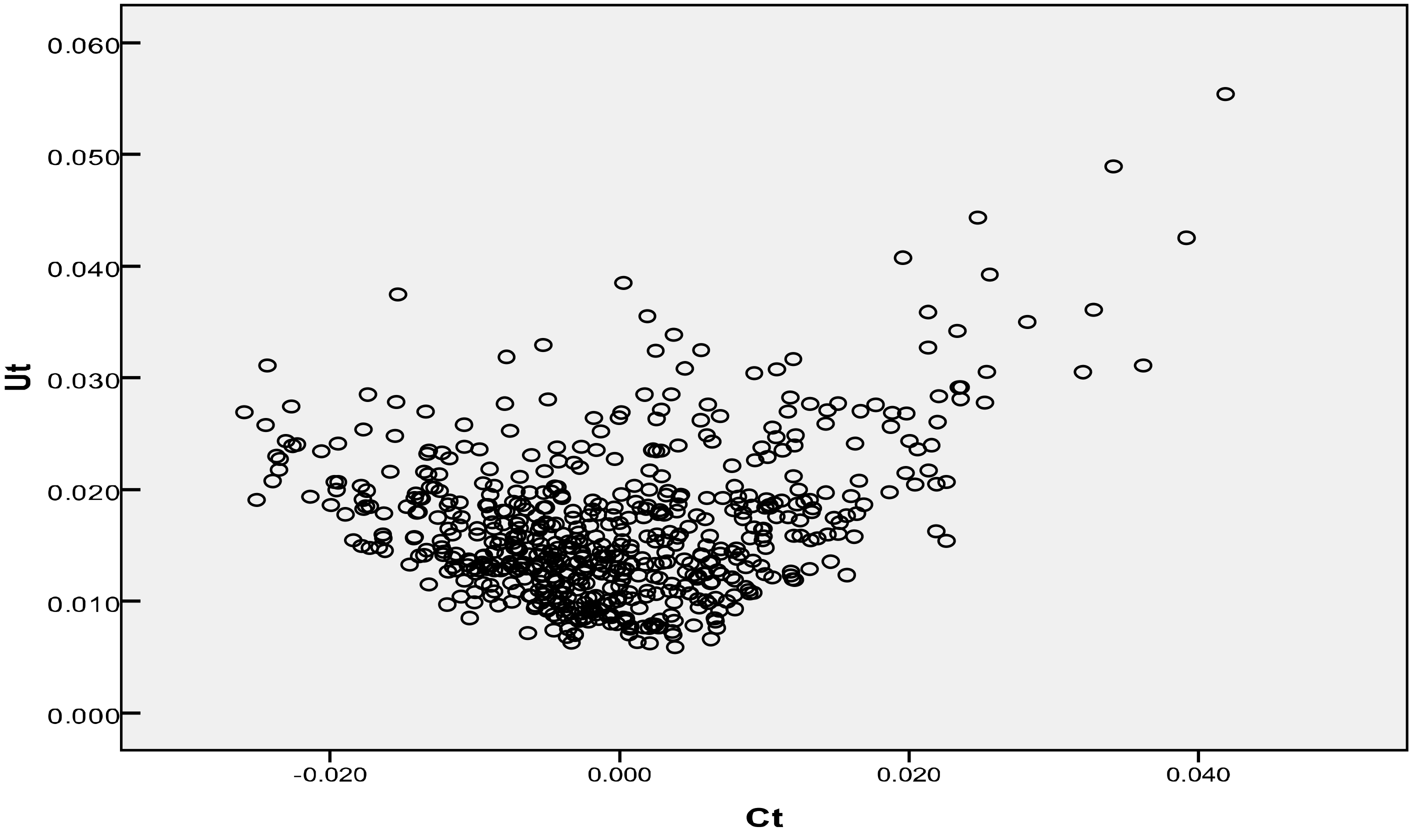
$\mathbf{Y}=(\mathbf{1},\mathbf{u}_{-1},\mathbf{u}_{-2},\ldots,\mathbf{u}_{-q},\mathbf{c}),$ $\boldsymbol{\alpha}=(\alpha_0,\alpha_1,\ldots,\alpha_p)',$

$\boldsymbol{\beta}=(\beta_0,\beta_1,\ldots,\beta_q,\gamma)'.$

Assume $(X'X)^{-1}$, $(Y'Y)^{-1}$ be existed, then estimators in model FDLR (p,q) are:

$$\boxed{\begin{cases}\hat{\boldsymbol{\alpha}}=(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{c}\\ \hat{\boldsymbol{\beta}}=(\mathbf{Y}'\mathbf{Y})^{-1}\mathbf{Y}'\mathbf{u}\end{cases}}$$

What the real relationship between the Centre and the Spread vary with time? Further explore



The Centre－Spread series at time t

Central-Spread Mixed Quadratic Regression Model C-SMQRM(1,1)

$$\begin{cases} c_t = \alpha_0 + \alpha_1 c_{t-1} + \cdots + \alpha_p c_{t-p} + \varepsilon_t \\ u_t = \beta_0 + \beta_1 u_{t-1} + \cdots + \beta_q u_{t-q} + \gamma c_t^2 + e_t \end{cases}$$

$\longrightarrow$ **FLSE**

$$\begin{cases} c_t = \hat{\alpha}_0 + \hat{\alpha}_1 c_{t-1} + \cdots + \hat{\alpha}_p c_{t-p} \\ u_t = \hat{\beta}_0 + \hat{\beta}_1 u_{t-1} + \cdots + \hat{\beta}_q u_{t-q} + \hat{\gamma} \hat{c}_t^2 \end{cases}$$

More   # Partial Nonparametric structure

$$\begin{cases} c_t = \alpha_0 + \alpha_1 c_{t-1} + \varepsilon_t \\ u_t = \beta_0 + \beta_1 u_{t-1} + m(c_t) + e_t \end{cases}, t = 2, \ldots, N.$$

The 1st one still reflects the trend of the centre series based on a IVFYS;

$\beta_0 + \beta_1 u_{t-1}$ and $m(c_t)$ reflect the trend of the unrandom volatility of a IVFYS

Linear part can be recognized as the main exchange, and the nonparametric part
Can be treaded as the no expectation one. Re-Data mining!

Evaluation of the fitting effect: Method-1

Mean Squared Error (MSE):

$$MSE = \frac{1}{N}\sum_{t=1}^{N}[(c_t - \hat{c}_t)^2 + (u_t - \hat{u}_t)^2]$$

Mean Absolute Error (MAE):

$$MAE = \frac{1}{N}\sum_{t=1}^{N}[|c_t - \hat{c}_t| + |u_t - \hat{u}_t|]$$

Evaluation of the fitting effect: Method-2

Spread Mean Square Error (SMSE)

$$SMSE = N^{-1} \sum_{t=1}^{N} (u_t - \hat{u}_t)^2$$

- Empirical analysis

Let $p = 1$, $q = 1$.

We use the original **IVFYS** data from Shanghai Stock Market to fit **FDLR(p , q)** , **C-SMQR(1,1), the Partial model.**

Original observation period be in

（2003.01.03—2005.07.22）(614 days)

沪市指数原始区间观测值序列

Shanghai Composite Index (SCI) （2003.01.03——2005.07.22）

[lowest, highest]

The test of stationary (Runs Test)

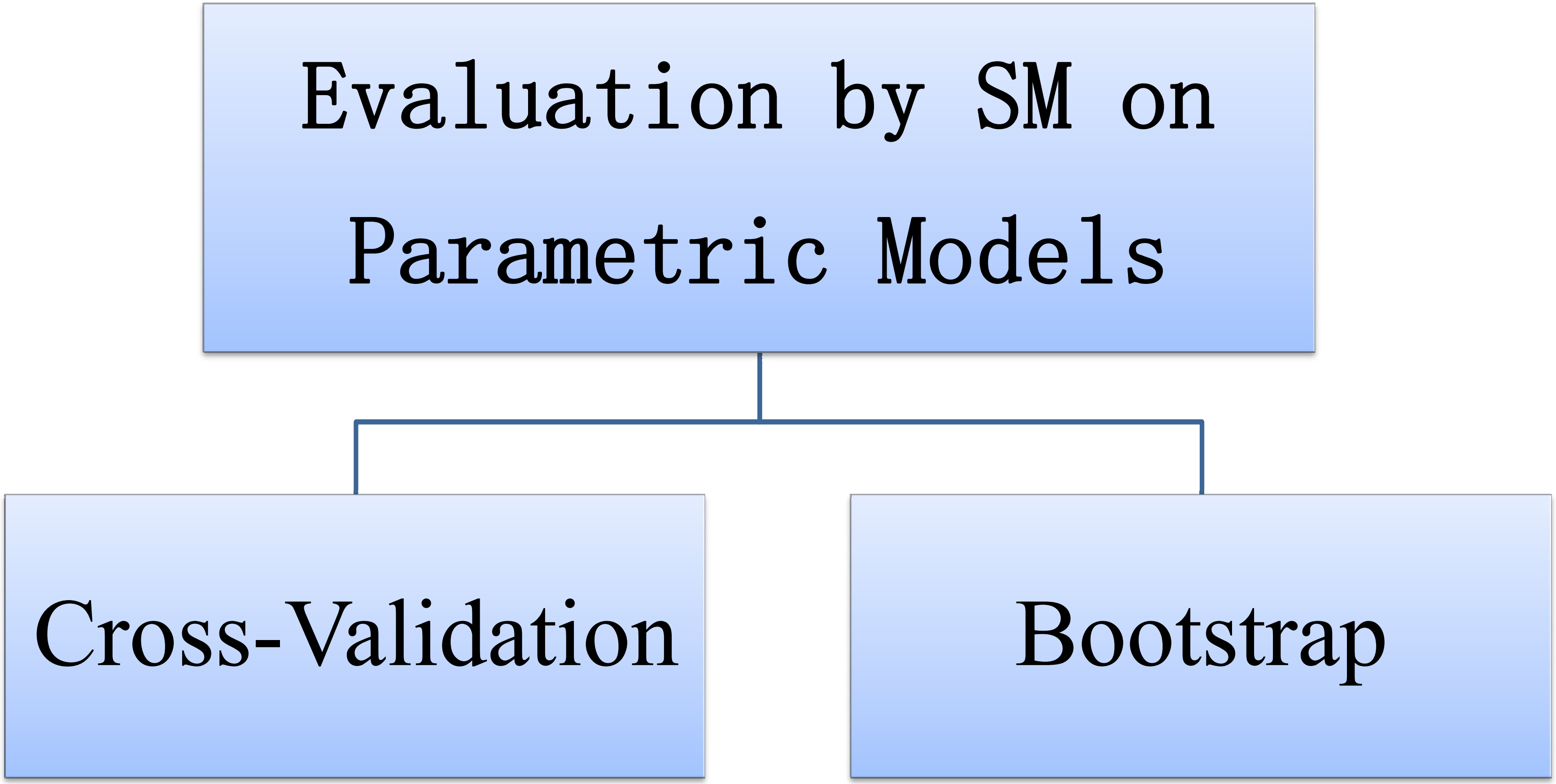
IVFYS - the centre series	
test value	-3.87E-4
cases<test value	332
cases>=test value	282
total cases	614
number of runs	285
Z	−1. 705
Asymp.sig.(2-tailed)	0. 088

* the test value is the mean of the centre series of IVFYS

Evaluation of three models by SMSE

$$SMSE = N^{-1}\sum_{t=1}^{N}(\tilde{u}_t - \hat{\hat{u}}_t)^2$$

2003. 1. 3—2005. 7. 22	FDLR（1, 1）	C-SMQRM（1, 1）	Partial（LP）	Partial（**LOW ESS**）
SMSE	2. 12E-05	1. 57E-05	10. 2 E-05	2. 32 E-05


Evaluation by SM on Parametric Models
Cross-Validation
Bootstrap

Estimated Results based on CV (Leave-one-out)

	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\gamma}$	SRR
FDLR(1,1)	0.0052	0.6906	0.0890	0.0294
C-SMQR（1,1）	0.0055	0.5738	14.48	0.0309

Evaluation on Bootstrap for FDLR（1,1）

	Original	Average of Resampling	Sample estimate Deviation	Standard Error	Coefficient of variation (σ/μ)
10	5.185E-03	5.487E-03	3.020E-04	4.030E-04	7.345E-02
	6.906E-01	6.669E-01	-2.365E-02	2.002E-02	3.002E-02
	8.897E-02	8.753E-02	-1.441E-03	2.240E-02	2.559E-01
50	5.185E-03	5.183E-03	-1.700E-06	5.550E-04	1.071E-01
	6.906E-01	6.897E-01	-9.423E-04	2.873E-02	4.166E-02
	8.897E-02	8.789E-02	-1.075E-03	1.606E-02	1.827E-01
100	5.185E-03	5.151E-03	-3.417E-05	5.470E-04	**1.062E-01**
	6.906E-01	6.919E-01	1.310E-03	3.125E-02	**4.516E-02**
	8.897E-02	8.848E-02	-4.888E-04	1.723E-02	**1.947E-01**
1000	5.185E-03	5.179E-03	-5.574E-06	5.260E-04	1.016E-01
	6.906E-01	6.906E-01	4.905E-05	2.860E-02	4.141E-02
	8.897E-02	8.971E-02	7.456E-04	1.845E-02	2.057E-01

Evaluation on Bootstrap for C-SMQRM（1,1）

	Original	Average of Resampling	Sample estimate Deviation	Standard Error	Coefficient of variation (σ/μ)
10	5.488E-03	5.350E-03	-1.380E-04	3.010E-04	5.626E-02
	5.738E-01	5.803E-01	6.594E-03	1.135E-02	1.956E-02
	1.448E+01	1.486E+01	3.783E-01	7.453E-01	5.015E-02
50	5.488E-03	5.421E-03	-6.702E-05	4.190E-04	7.729E-02
	5.738E-01	5.775E-01	3.784E-03	2.387E-02	4.133E-02
	1.448E+01	1.437E+01	-1.122E-01	8.312E-01	5.784E-02
100	5.488E-03	5.530E-03	4.172E-05	4.380E-04	**7.921E-02**
	5.738E-01	5.706E-01	-3.169E-03	2.639E-02	**4.624E-02**
	1.448E+01	1.454E+01	5.402E-02	8.260E-01	**5.682E-02**
1000	5.488E-03	5.493E-03	4.695E-06	4.330E-04	7.883E-02
	5.738E-01	5.736E-01	-1.196E-04	2.554E-02	4.453E-02
	1.448E+01	1.448E+01	-4.459E-03	9.206E-01	6.359E-02

Compare by Coefficient of Variation on RS100

Bootstrap

Coefficient of variation (SD/MEAN)	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\gamma}$
FDLR (1 , 1)	**1.062E-01**	**4.516E-02**	**1.947E-01**
C-SMQRM (1,1)	**7.921E-02**	**4.624E-02**	**5.682E-02**

Conclusion：

Through the Statistical Simulation evaluation method, a nonlinear structure Central-Spread Mixed Quadratic Regression model C-SMQR(1,1) better

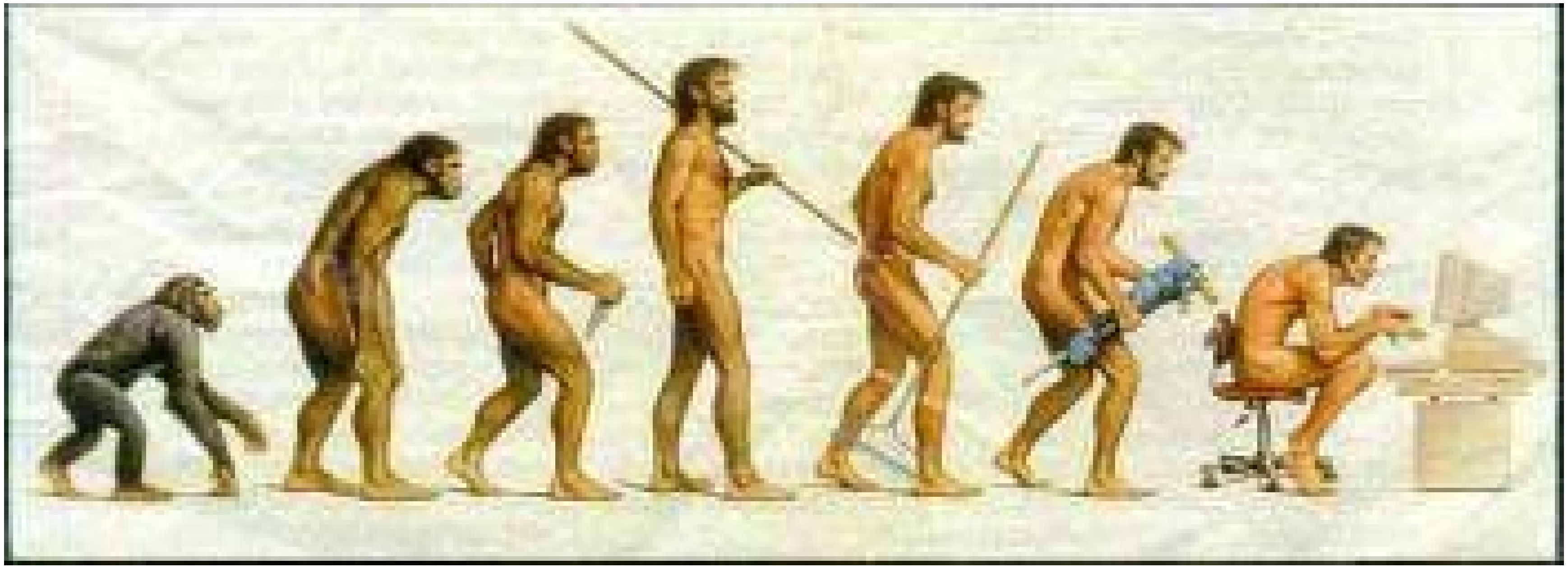
Still, compare the two simulation way, we need do more on the explore to find a more suitable model

Note:

- We propose that the dynamic change of the interval-valued time series from the financial market can be use to analyze the uncertain which is different from the traditional one on the yield study, we call it as IVFYS which show a uncertain fuzziness.

- This study is just from another view to consider the problems. To replace the traditional research of randomness on uncertain is NOT OUR thoughts from the beginning.

- **Research** on **Interval-valued financial time sequence** has give us another view to watch the financial market, it gives us a hint that people may make some new views and models that may give more information than the traditional way.
- **Still,** there are **MANY** problems in our research, e.g. the order determination; the error term, and so on… .




Thank you for your
Attention !