An Elasticity Similarity Distance & Application

Prepared by: Kelvin Li,
Metrics & Analytics
CIBC

Last Modified: Sep 12, 2009

Content

- Objective
- Inference Of Elasticity Similarity Distance
- Definition Distance (ESD)
- Interpretation Of ESD
- Application In Discrete Time Scenario
- Examples: Why ESD Is Better
- Application Of ESD In Stock Market Pair Trading

Objective

- A similarity measure is defined in terms of curve elasticity that are more appropriate for clustering time-course trajectories

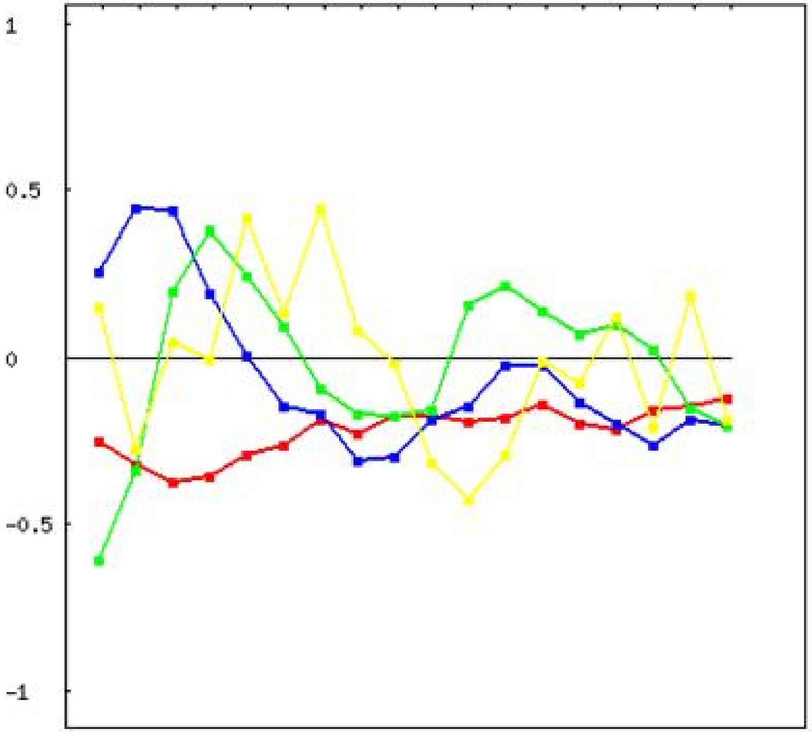
- An estimation method is proposed to handle the discrete-time nature of the observations

- Computational experiment is carried out on real stock market data

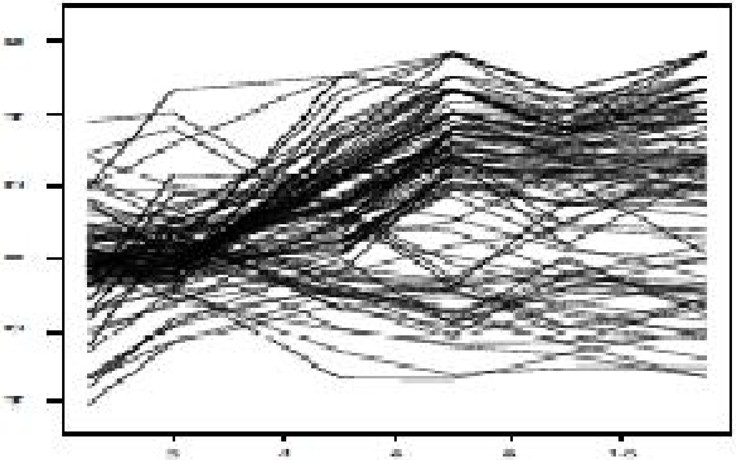
Application Areas

- The profile of different subjects may follow different trajectories, as measured at a number of sampling times, such as customer life cycle curves, to be employed in market segmentation
- The similarity among gene expression profiles over time course implies the similar transcriptional functions
- The similarity among the price or volume curves of stocks and bonds indicates the common characteristics of different companies
- The distance between survival (or hazard) curves of different drugs may demonstrate the same treatment performance

Application Areas



Life value of customers over tenure



Gene clusters over time course

Inference of Elasticity Similarity Distance

We visualize a subject profile over the time-course as the following smooth function of time :

$$g_k(t) = \mu_k(t) + e_k(t)$$

Where $\mu_k(t)$ represents actual profile pattern over the time course.

The function $e_k(t)$ is error curve caused by experimental or random errors

Applying the stochastic capital asset price equation (omit violability caused by external random effect, such as unexpected news)

$$\frac{d(\mu_k(t))}{dt} = C_k \mu_k(t)$$

The solution to the model has form:

$$x_k(t) = Q_k(t)\exp(\lambda_k t)$$

Where, $Q_k(t)$ is a matrix whose elements polynomial functions are of time. If a system is stable then $Q_k(t)$ Should be a constant C

Definition of Elasticity Similarity Distance

$$\frac{d(\mu_k(t))}{dt} = C_k \mu_k(t) \quad \text{OR} \quad \frac{d(\log(\mu_k(t)))}{dt} = C_k$$

Where, we regard C_k the relative rate of change of response values. It can be treated as an index that reflects curve's behavior over time course.

The similarity measure called elasticity similarity distance (ESD) is defined:

$$\rho = \frac{1}{T_1 - T_0}[\int_{T_0}^{T_1} | \frac{d[\log(\mu_k(t))]}{dt} - \frac{d[\log(\mu_j(t))]}{dt} | dt]$$

Characteristics and Benefit of ESD

The ESD is not affected by point intervention. For example, assuming any curve can be expressed by:

$$\mu_k(t) = \hat{\mu}_k(t) I(t)$$

Where, $\hat{\mu}_k(t)$ represents the time course profile induced by the subject's self-behavior, and $I(t)$ is the external intervention factor curve that affects $\hat{\mu}_k(t)$ over time course.

Obviously, the ESD between two curves is not affected by any external intervention effect. This also means that If ESD between $g(t)$ and $f(t)$ is 0, then $g(t)/f(t) = C$

By comparison, the distance defined by Pearson correlation coefficient will be distorted by external intervention factor or outlier effect.

Application In Discrete Time Scenario

In practice, the time is often discrete, then we can discretize the ESD from the integral form

$$\hat{\rho} = \sum_{S=1} (\Delta_s - \Delta_{s-1})[E(\frac{d \log(f(t))}{dt}) - E(\frac{d \log(f(t))}{dt})]$$

Where $\Delta_s = t_{s+1} - t_s$ is the time step size.

The formulae can be approximated by the average slope difference across the time points

$$ESD = \sum_{t=1}^{T} \omega_t \left| \frac{\Delta f_t}{f(t)} - \frac{\Delta g_t}{g(t)} \right| \approx \sum_{t=1}^{T} \omega_t \left| \frac{d[\log(f(t))]}{dt} - \frac{d[\log(g(t))]}{dt} \right| \approx \sum_{t=1}^{T} \omega_t \left| \log[(f(t+1)*g(t))/(f(t)*g(t+1))] \right|$$

Where, Δf_t and Δg_t are the changes of curves at time t. When time step sizes are equal (e.g. 1) , it simply reduces to

$$ESD = \sum_{t=1}^{T} \left| \log[(f(t+1)*g(t))/(f(t)*g(t+1))] \right|$$

Application of ESD

If market is efficient, there does not exist arbitrage opportunity. However, statistical arbitrage is likely if you know two or more stocks normally move together become temporarily uncorrelated

The **pair trading** system was deveolped by JPMorgan Chase. The idea is that competitors in the same sector were correlated in their day-to-day price movements. When the correlation broke down, i.e. one stock traded up while the other traded down, you can long one stock and short another, betting that the "spread" between the two would eventually converge. Some typical examples of potentially correlated pairs:

Wal-Mart (WMT) and Target Corporation (TGT)

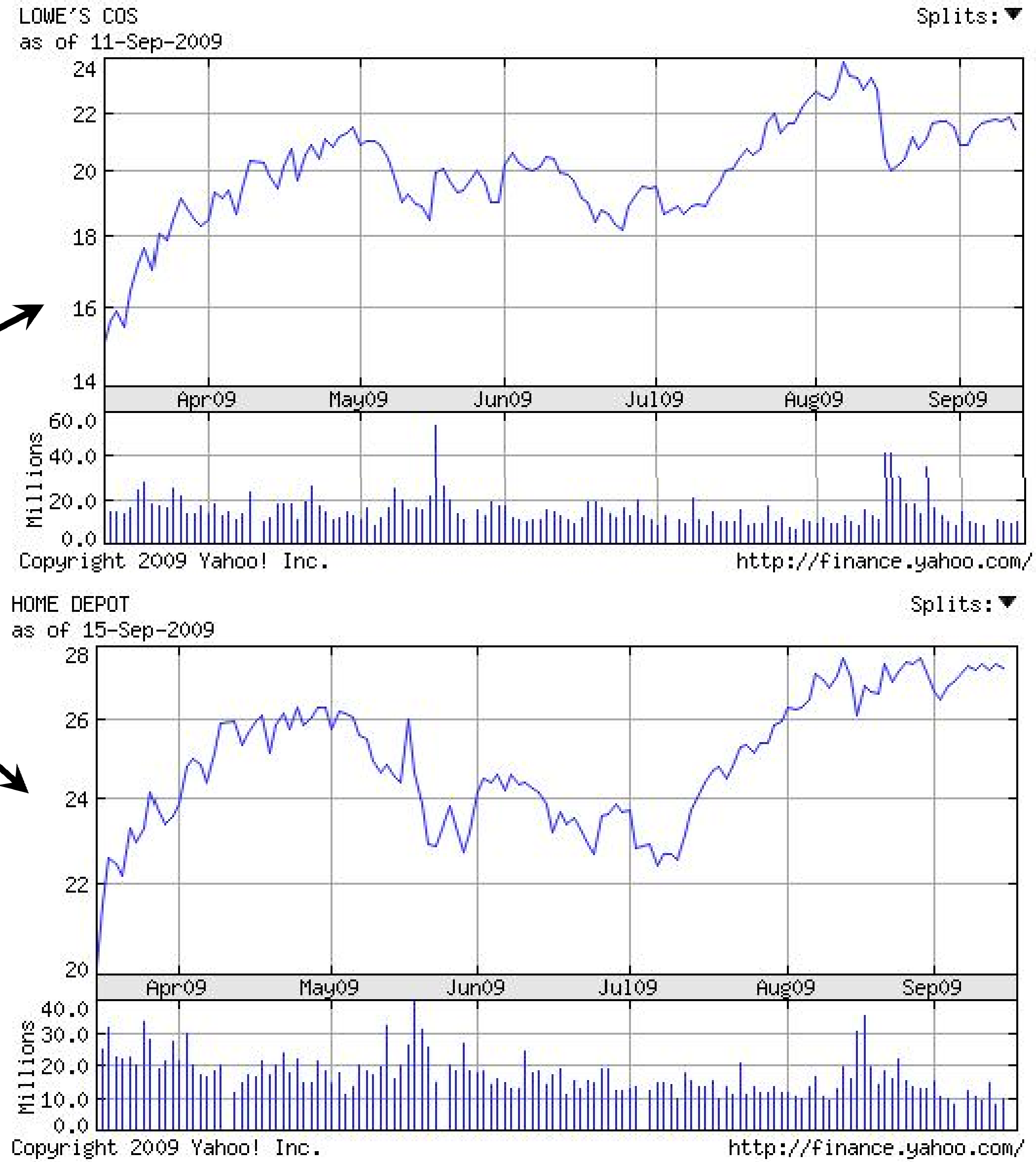
Home depot (HD) and Lows Companies (LOW)

Coca Cola (KO) and Pepsico (PEP)

The problem is, when picking pairs, whether Pearson Correlation distance is the 'best' measure to describe the distance between paired curves? Look at the following examples

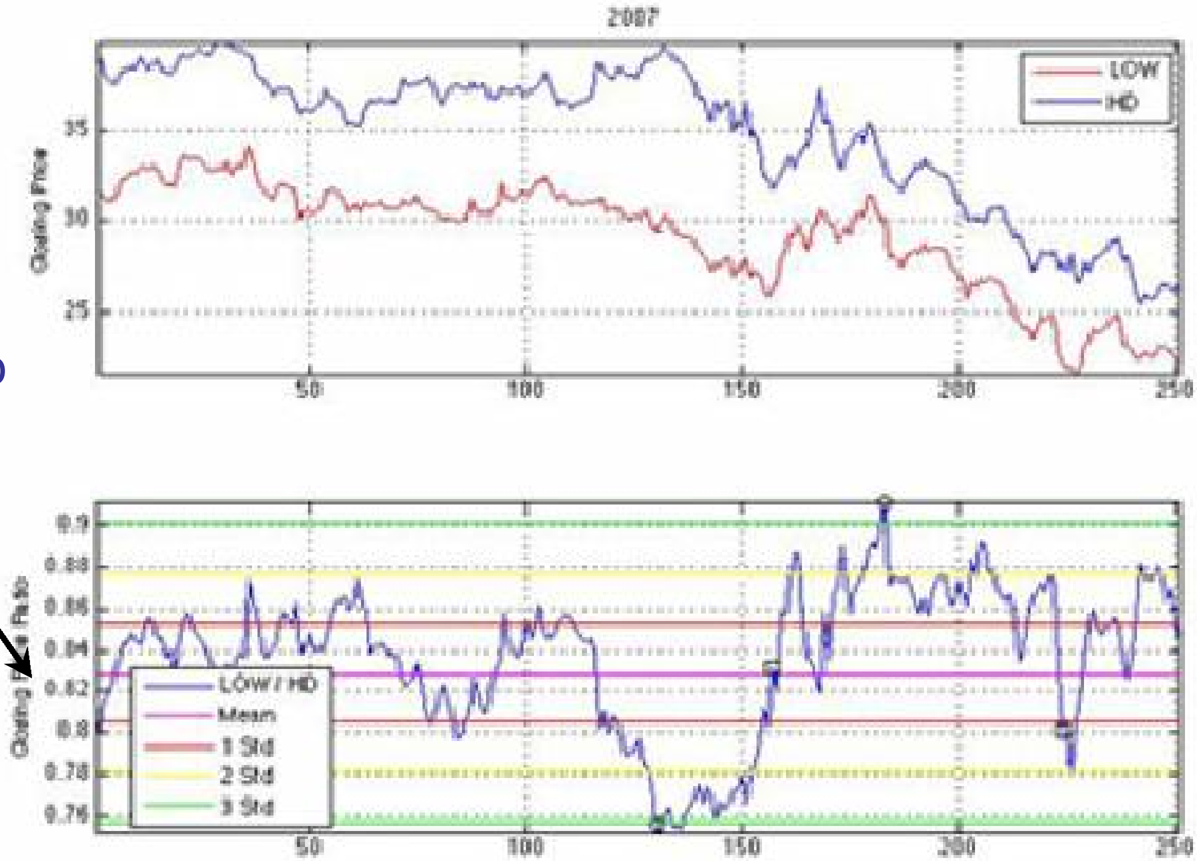
Pair Example

How to define distance?
LOW and HD



Pair Example

Price ratio



Example: Why ESD Is Better

```
data price0;
 input p @@;
 price1=2.2*p+58;
 price2=p;
 drop p;
cards;
1 2 3 4 5 6 7 8 9 10
;
run;

proc corr data=price0;
 var price1 price2;
run;

data price0;
 set price0;
 ratio_1=round(dif(price1)/(price1-
 dif(price1)),0.01);
 ratio_2=round(dif(price2)/(price2-
 dif(price2)),0.01);
run;
```

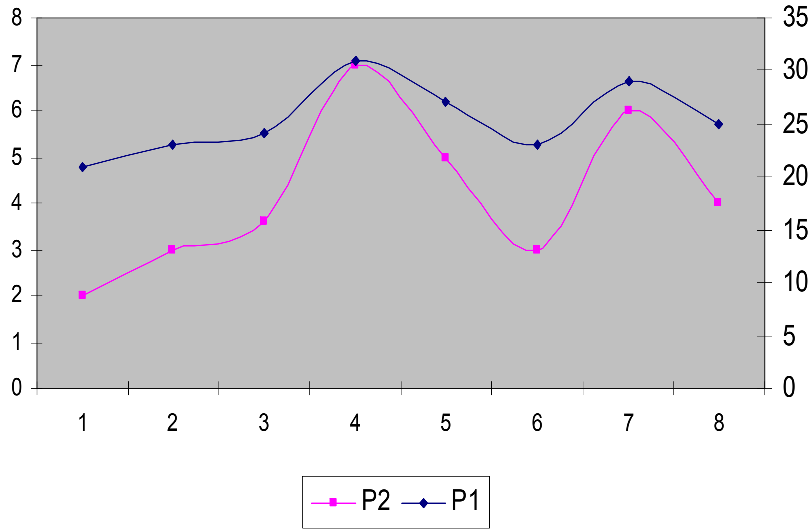
price1	price2	ratio_1	ratio_2
60.2	1	.	.
62.4	2	0.04	1
64.6	3	0.04	0.5
66.8	4	0.03	0.33
69	5	0.03	0.25
71.2	6	0.03	0.2
73.4	7	0.03	0.17
75.6	8	0.03	0.14
77.8	9	0.03	0.13
80	10	0.03	0.11

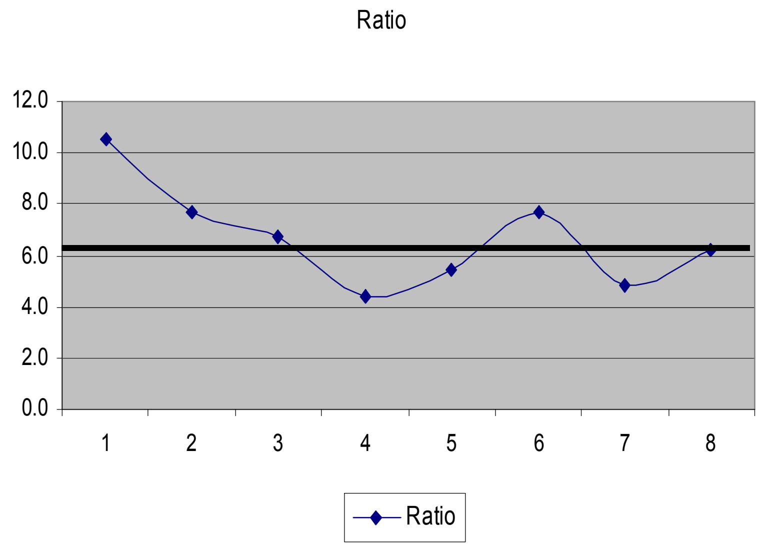
Pearson Correlation Coefficients, N = 10
Prob > |r| under H0: Rho=0

	price1	price2
price1	1.00000	1.00000 <.0001
price2	1.00000 <.0001	1.00000

Even though the Pearson Correlation distance is 0, the changing ratio can have a big discrepancy

Example: Why ESD Is Better





Even though the Pearson Correlation distance is 0, the changing ratio can have a big discrepancy

Example: Why ESD Is Better

```
data price1;
 input p intervention;
 price1=3.5*p+86; price2=p;
 price_adj1=price1*intervention;
 price_adj2=price2*intervention;
 drop p price1-price2;
cards;
1 1
2 1
3 12
4 1
5 1
6 3
7 2
8 1
9 7
10 1
;
run;
proc corr data=price1;
 var price_adj1 price_adj2;
run;
data price1;
 set price1;
 ratio_1=round(dif(price_adj1)/(price_adj1-dif(price_adj1)),0.01);
 ratio_2=round(dif(price_adj2)/(price_adj2-dif(price_adj2)),0.01);
run;
```

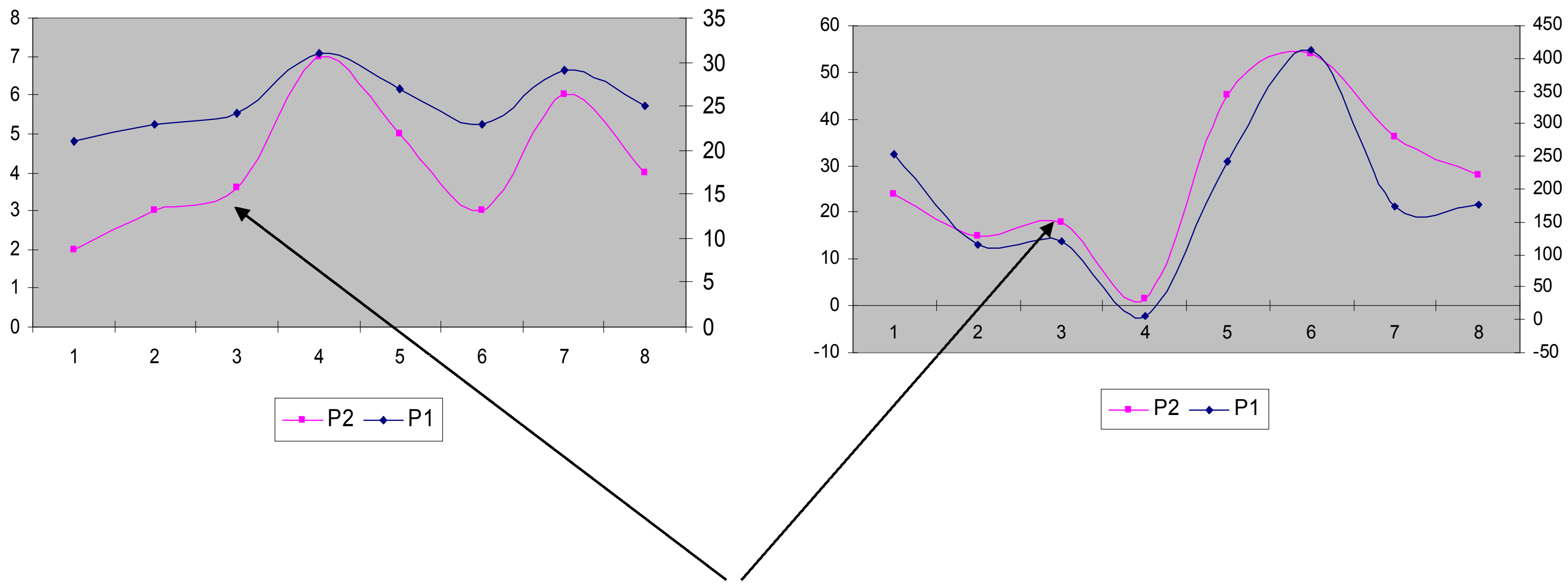
intervention	price_adj1	price_adj2	ratio_1	ratio_2
1	89.5	1	.	.
1	93	2	0.04	1
12	1158	36	11.45	17
1	100	4	-0.91	-0.89
1	103.5	5	0.04	0.25
3	321	18	2.1	2.6
2	221	14	-0.31	-0.22
1	114	8	-0.48	-0.43
7	822.5	63	6.21	6.88
1	121	10	-0.85	-0.84

Pearson Correlation Coefficients, N = 10
Prob > |r| under H0: Rho=0

	price_adj1	price_adj2
price_adj1	1.00000	0.84585 0.0020
price_adj2	0.84585 0.0020	1.00000

if there are common intervention effects, Pearson Correlation distance can distort the actual similarity between two curves

Example: Why ESD Is Better



Under some intervention event, Pearson Correlation distance is changed because data order becomes reverse, even though the changing ratio remain unchanged

Example: Why ESD Is Better

```
data price2;
  input p intervention;
  price_adj1=p*5*intervention;
  price_adj2=p*intervention;
 drop p;
cards;
1 1
2 1
3 2
4 11
5 1
6 3
7 1
8 1
9 7
10 1
;
run;

data price2;
 set price2;
 ratio_1=round(dif(price_adj1)/(price_adj1-dif(price_adj1)),0.01);
 ratio_2=round(dif(price_adj2)/(price_adj2-dif(price_adj2)),0.01);
run;
```

intervention	price_adj1	price_adj2	ratio_1	ratio_2
1	5	1	.	.
1	10	2	1	1
2	30	6	2	2
11	220	44	6.33	6.33
1	25	5	-0.89	-0.89
3	90	18	2.6	2.6
1	35	7	-0.61	-0.61
1	40	8	0.14	0.14
7	315	63	6.88	6.88
1	50	10	-0.84	-0.84

ESD will not change the similarity between two curves if there are common intervention factors

Example: Why ESD Is Better

```
data price3;
 set price2;
 r1=log(price_adj1);
 r2=log(price_adj2);
 ratio_1=round(dif(r1),0.01);
 ratio_2=round(dif(r2),0.01);
 drop r1-r2;
run;
```

intervention	price_adj1	price_adj2	ratio_1	ratio_2
1	5	1	.	.
1	10	2	0.69	0.69
2	30	6	1.1	1.1
11	220	44	1.99	1.99
1	25	5	-2.17	-2.17
3	90	18	1.28	1.28
1	35	7	-0.94	-0.94
1	40	8	0.13	0.13
7	315	63	2.06	2.06
1	50	10	-1.84	-1.84

Therefore, the ESD can really reflects the simialarity between among time-course trajectories

Example: Find Good Pairs

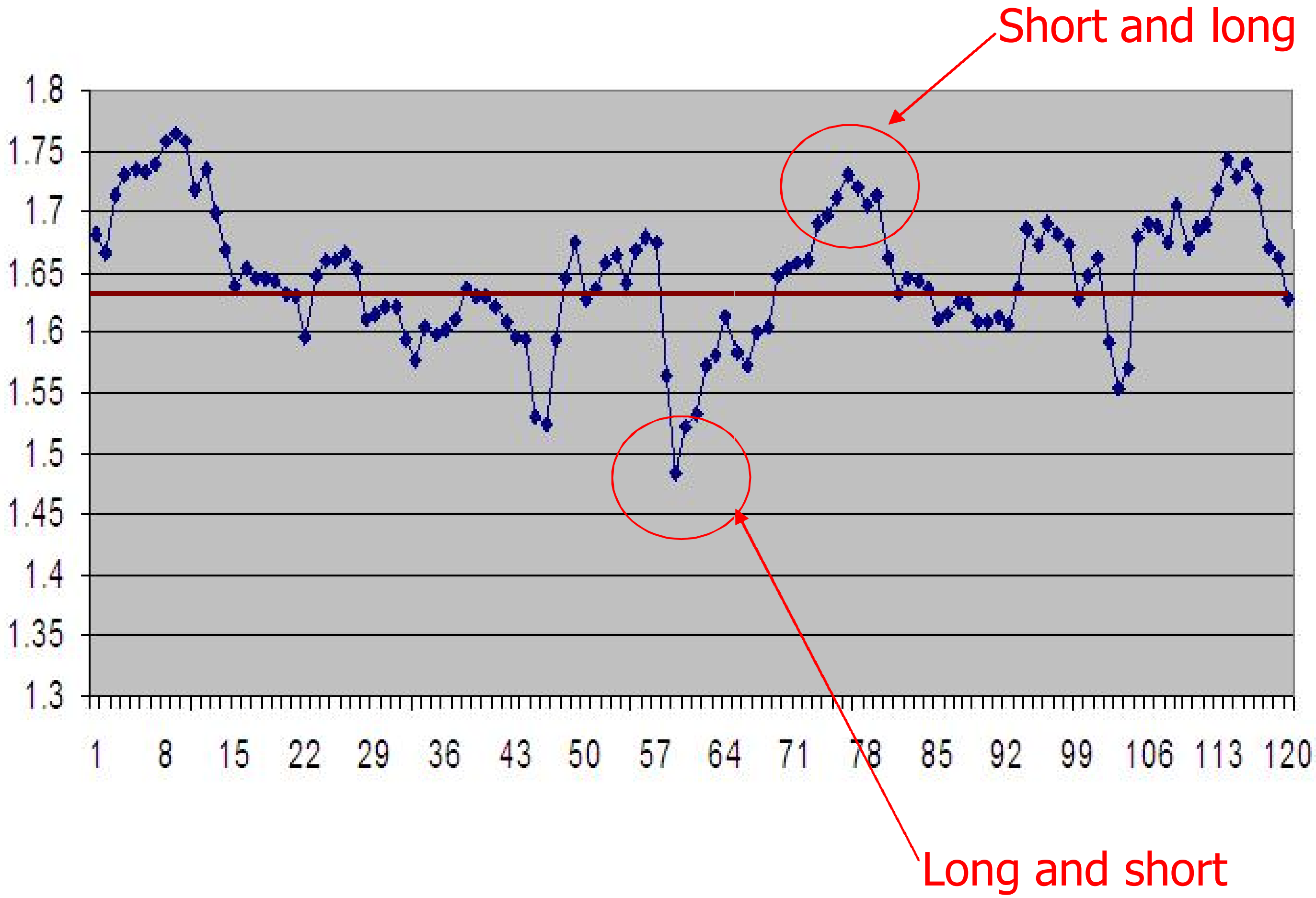
- Applying SAS Macro and real stock price data, we can find the following 'good' pairs in terms of ESD

HD	LOW
KO	PEP
XOM	CVX
WMT	TGT
GOOG	BIDU
XTO	UPL
ORCL	IBM

Example: Pair Spread

Implement Pair Trading

The ratio of the the stock price is no longer random walk but still contains non-constant variance. Using the time series GARCH model or neural net we can predict future ratio



Thank You!