

Maths and Stats

New tools of the trade
Singapore August 2010

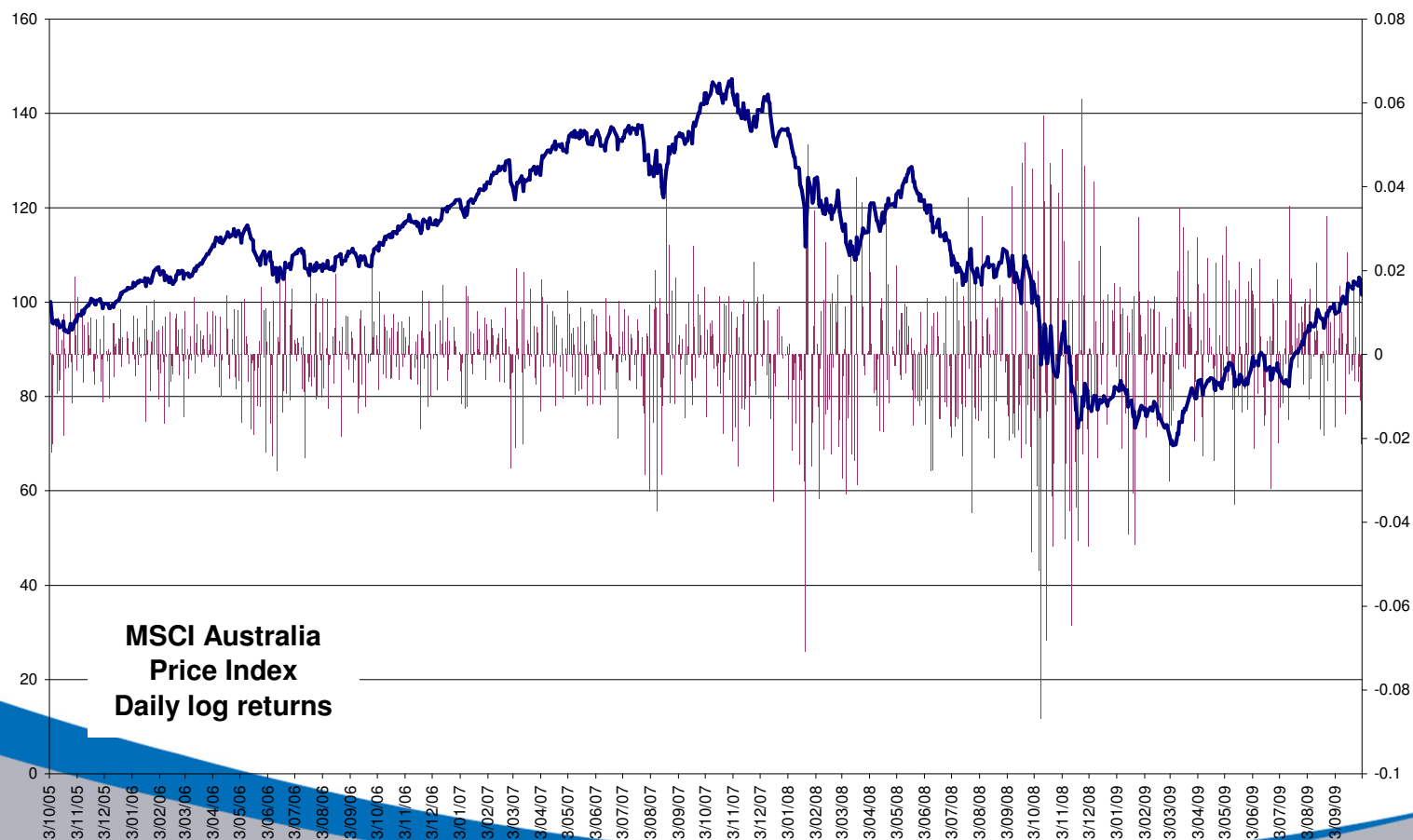


Order

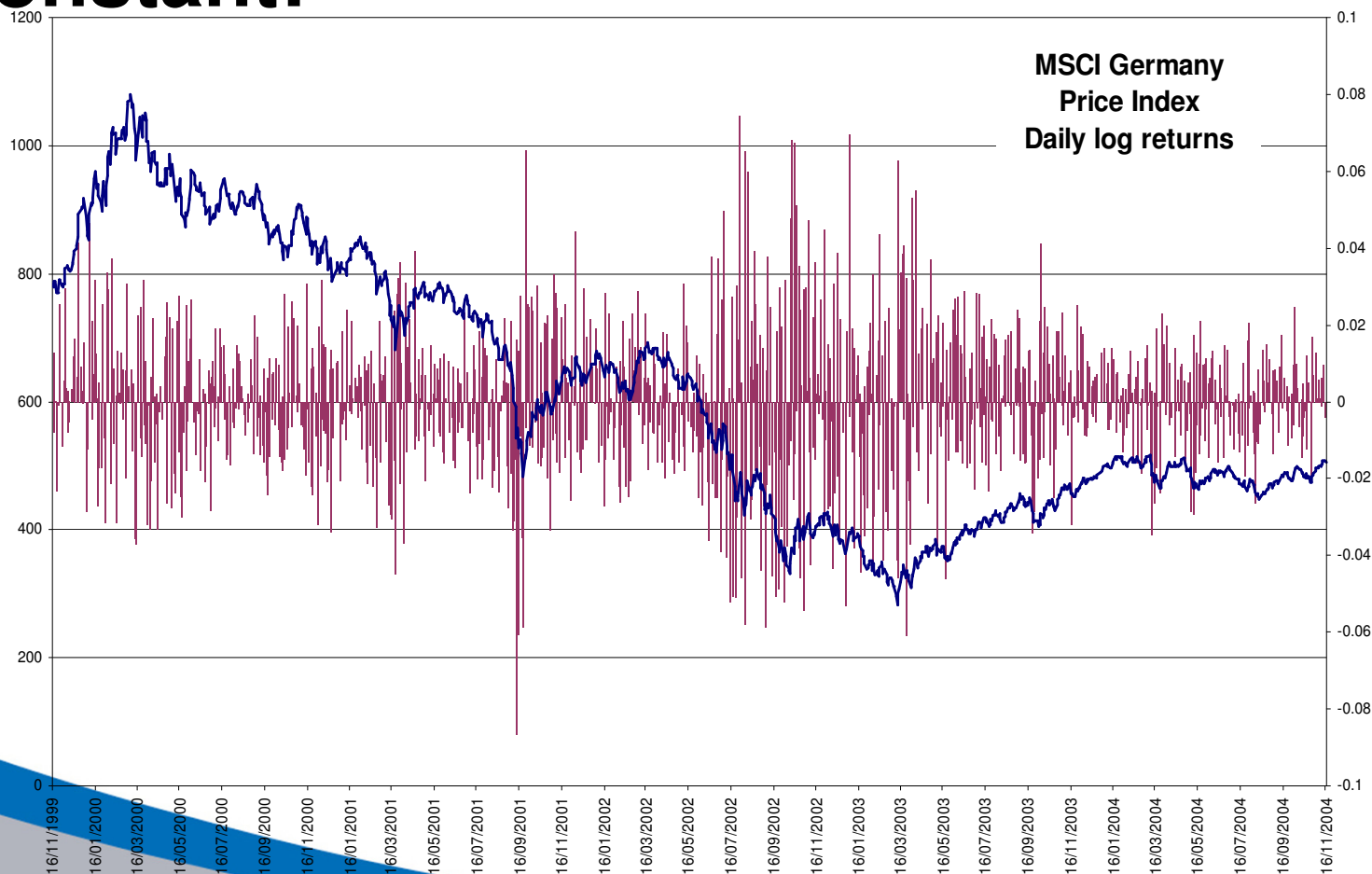
- Modelling Volatility
- Extreme Value Theory
- Statistical Bootstrapping
- Stochastic Calculus
- Copulas
- Simulation with all this



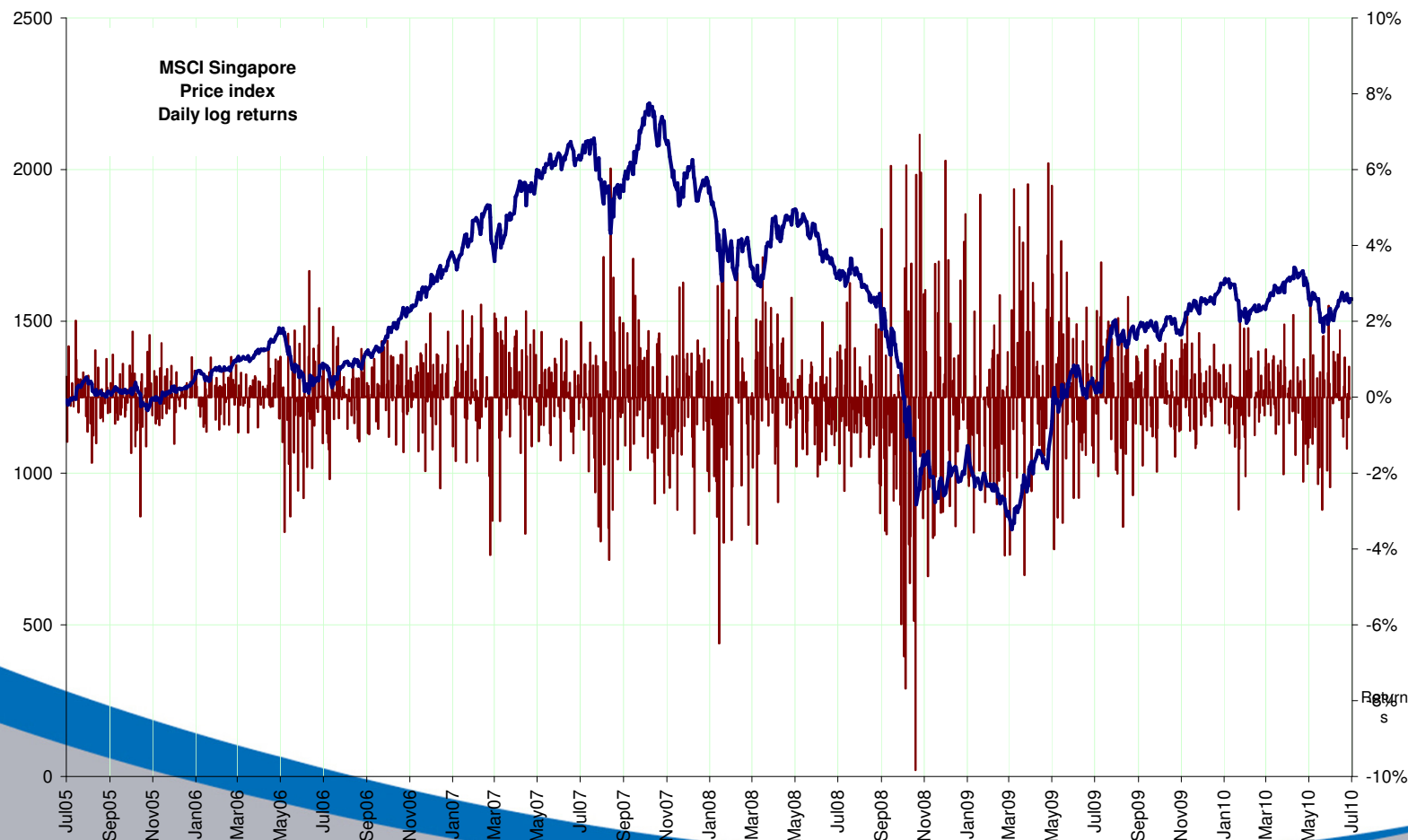
Heteroscedasticity – volatility is not constant!



Heteroscedasticity – volatility is not constant!



Heteroscedasticity – volatility is not constant!



Heteroscedasticity

- Volatility is not constant
- Volatility is predictable
 - market movement x is not predictable
 - x^2 is predictable
- High volatility occurs in heatwaves or clusters, typically around a mean value
- Low volatility periods may be long lasting
- How to use forecastable volatility for VaR?

Heteroscedasticity

- Stochastic volatility:
 - Stationary models, where the volatility in any period is a perturbation of a basic level of volatility

$$\ln \sigma_t^2 = \ln \sigma^2 + S\varepsilon_t$$

- where σ is the underlying volatility and S is the volatility of volatility
 - Log auto-regressive models

$$\ln \sigma_t^2 = a + b.\ln \sigma_{t-1}^2 + S\varepsilon_t$$

Volatility model

- Regime-switching models

$$\sigma_t = \sigma(s_t)$$

- where s_t is one of n states of the world - usually a high volatility state and a low volatility state.

$$\underline{s}_{t+1} = A \underline{s}_t$$

Volatility model

- GARCH(p,q) models
 - all α and β are positive with total sum less than 1.0

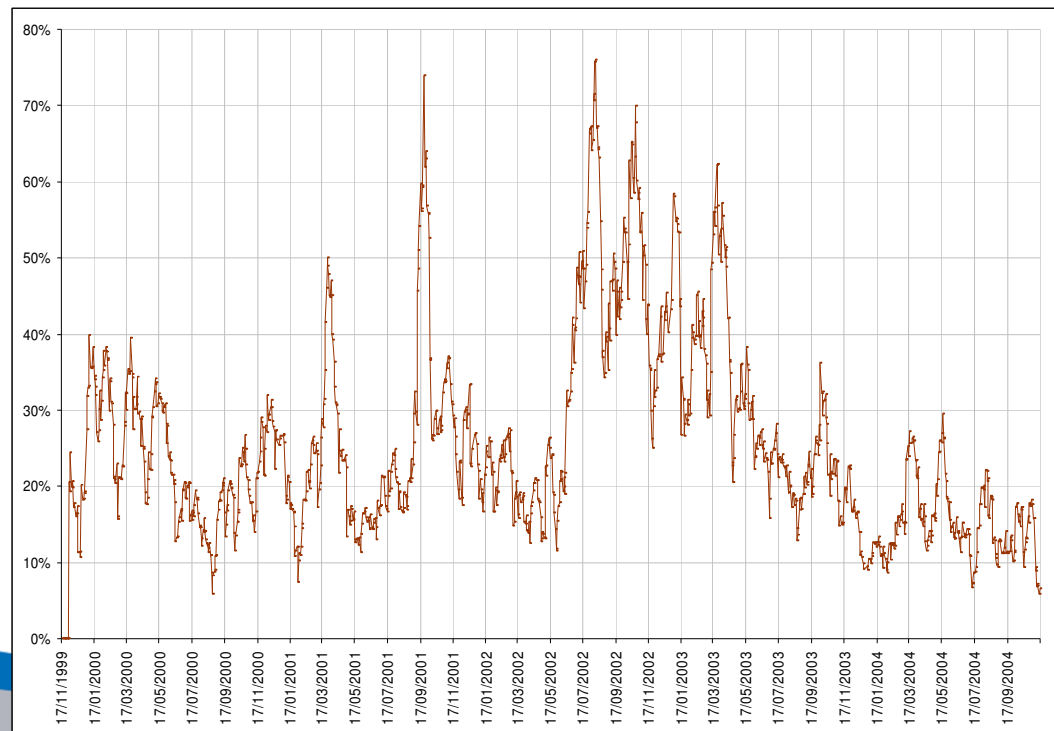
$$\sigma_t^2 = \alpha_0 + \sum_{j=1}^p \alpha_j X_{t-j}^2 + \sum_{k=1}^q \beta_k \sigma_{t-k}^2$$

Estimation methods

- Remember why we want to estimate volatility!
- We don't care what portfolio has been in the past
- We want to **forecast** volatility

Estimation methods - MA

- Fixed window
 - Moving average methods
 - $MA_n(x^2)$ – assume expected value is zero
 - Plateau effects



Estimation methods - EWMA

- Exponential window
 - Moving average methods with exponentially declining weights
 - $\sigma^2(t) = \lambda\sigma^2(t-1) + (1-\lambda)r_t^2$
 - r is today's market move
 - $\sigma^2(t) = \sum_n \lambda^{n-1}(1-\lambda)r_{t-n+1}^2$
 - No data drops out of averaging, more recent data gets greater weight

Vol – slider fixed and EWMA.xls

[spreadsheets\Vol - slider fixed and EWMA.xls](#)

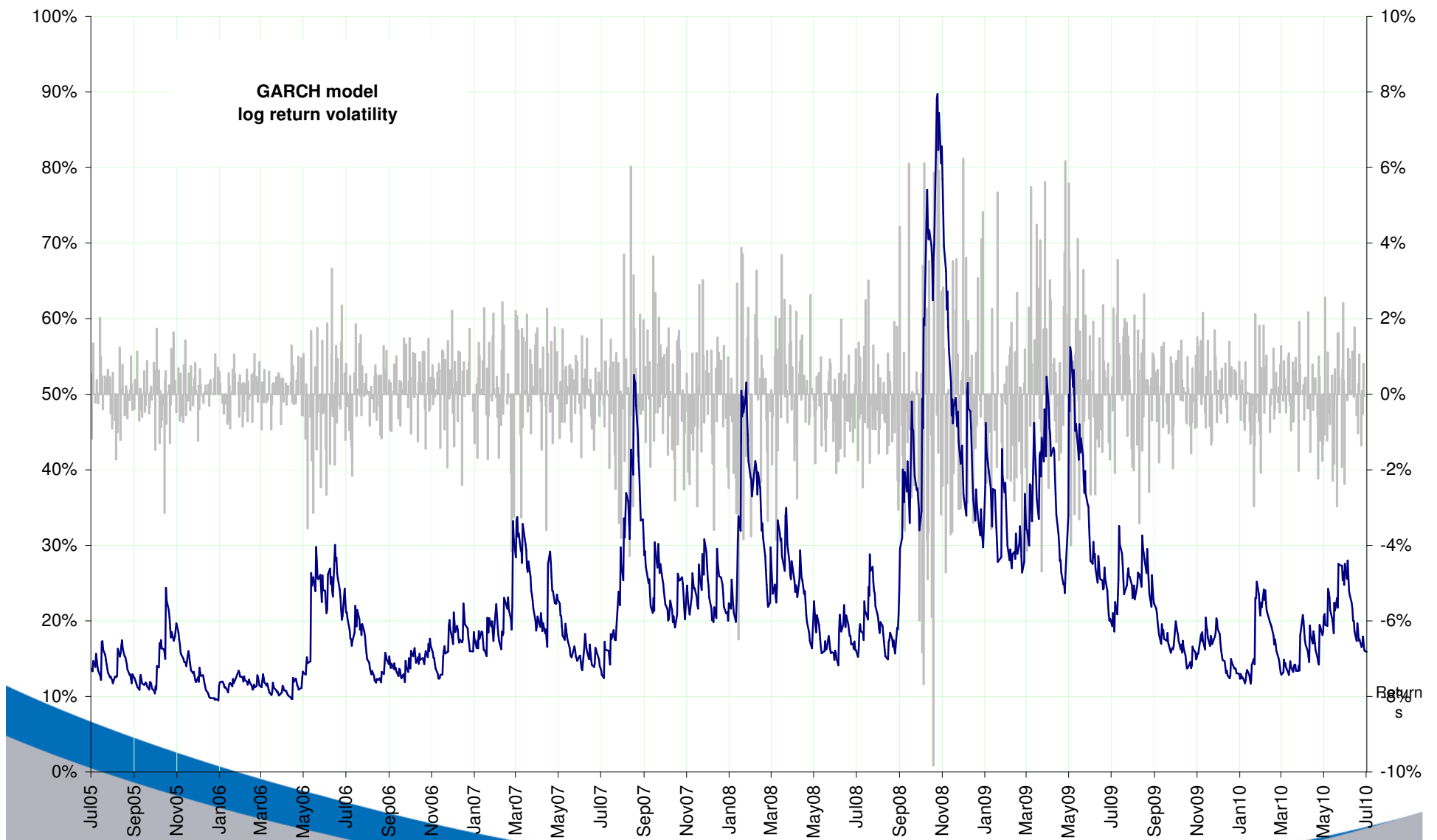


Estimation methods – GARCH

- Generalised
Auto-
Regressive
Conditional
Heteroscedasticity

GARCH

- A simple, commonly used GARCH model is:
- $\sigma^2_n = \lambda \sigma^2_m + \kappa \sigma^2_{n-1} + (1-\lambda-\kappa)u^2_{n-1}$
- Tomorrow's estimate of variance is a weighted average of:
 - a mean variance value
 - today's estimate of variance
 - today's market move
- More complicated models available



Parameter estimation

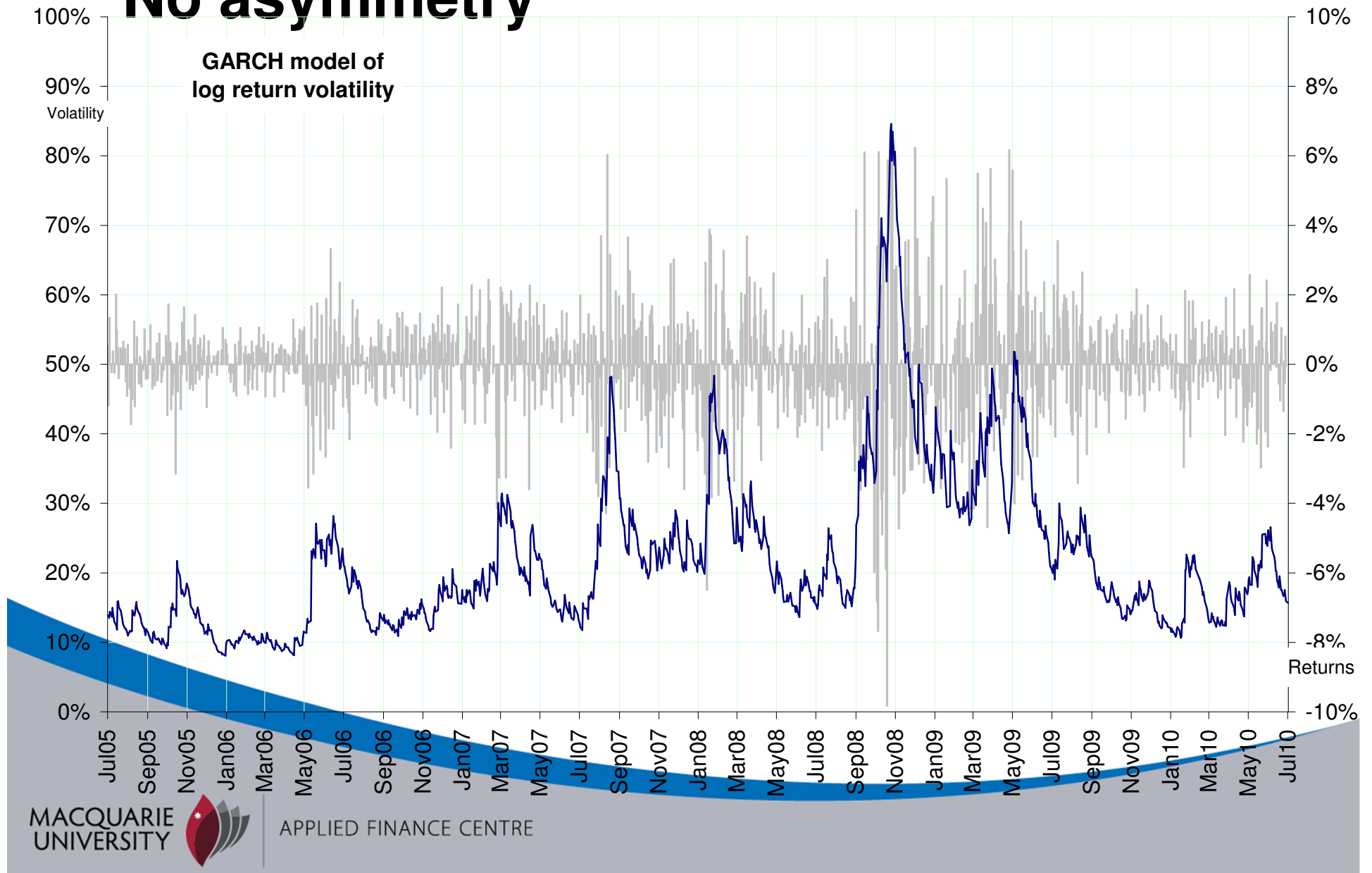
- Compare forecast volatility with actual volatility
- Non-parametric
 - If model is correct then u^2/σ^2 is roughly chi-squared with 1 df
 - Sum of scaled volatilities should equal number of observations
 - Sum of squares should be 3* observations
- Parametric
 - Log returns are Normal so can use Maximum likelihood
 - Log-likelihood = $-0.5\ln(\sigma^2) - 0.5u^2/\sigma^2$

Asymmetric volatility

- $\sigma_n^2 = \lambda \sigma_m^2 + K \sigma_{n-1}^2 + (1-\lambda-K) I_{n-1} u_{n-1}^2$
- the Indicator function I_{n-1} has the value 1 if the return u_{n-1} is negative and 0 otherwise
 - This means the expected value of the indicator function is about 0.5, allowing easy comparison of parameter values with the previous expression
- $\sigma_n^2 = \lambda \sigma_m^2 + K \sigma_{n-1}^2 + (1-\lambda-K)(w I_{n-1} + 1-w) u_{n-1}^2$
 - Allows the extent of asymmetry to be determined by a parameter w
- [spreadsheets\vol - GARCH Singapore asymmetric.xls](#)

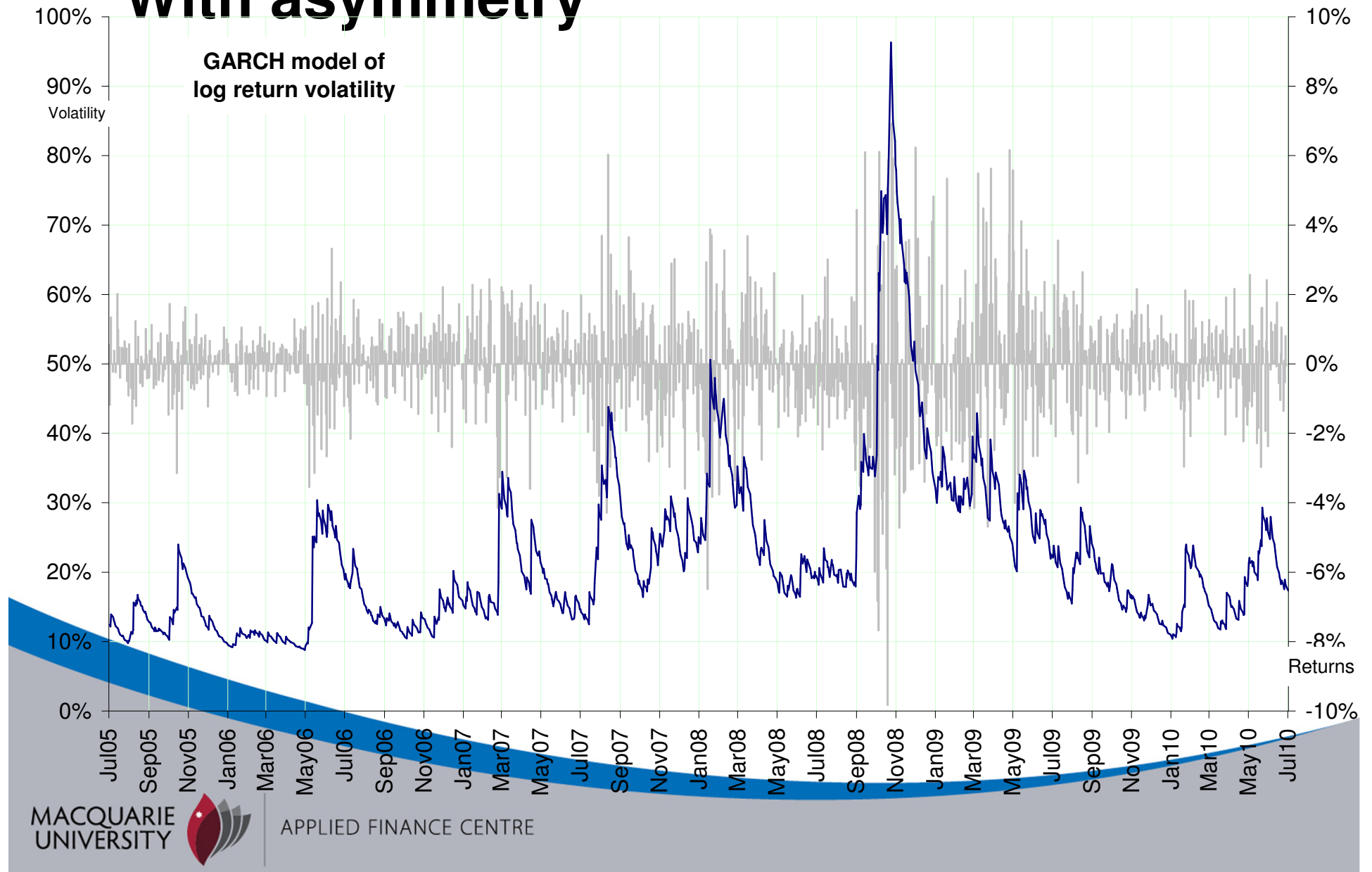
No asymmetry

GARCH model of
log return volatility

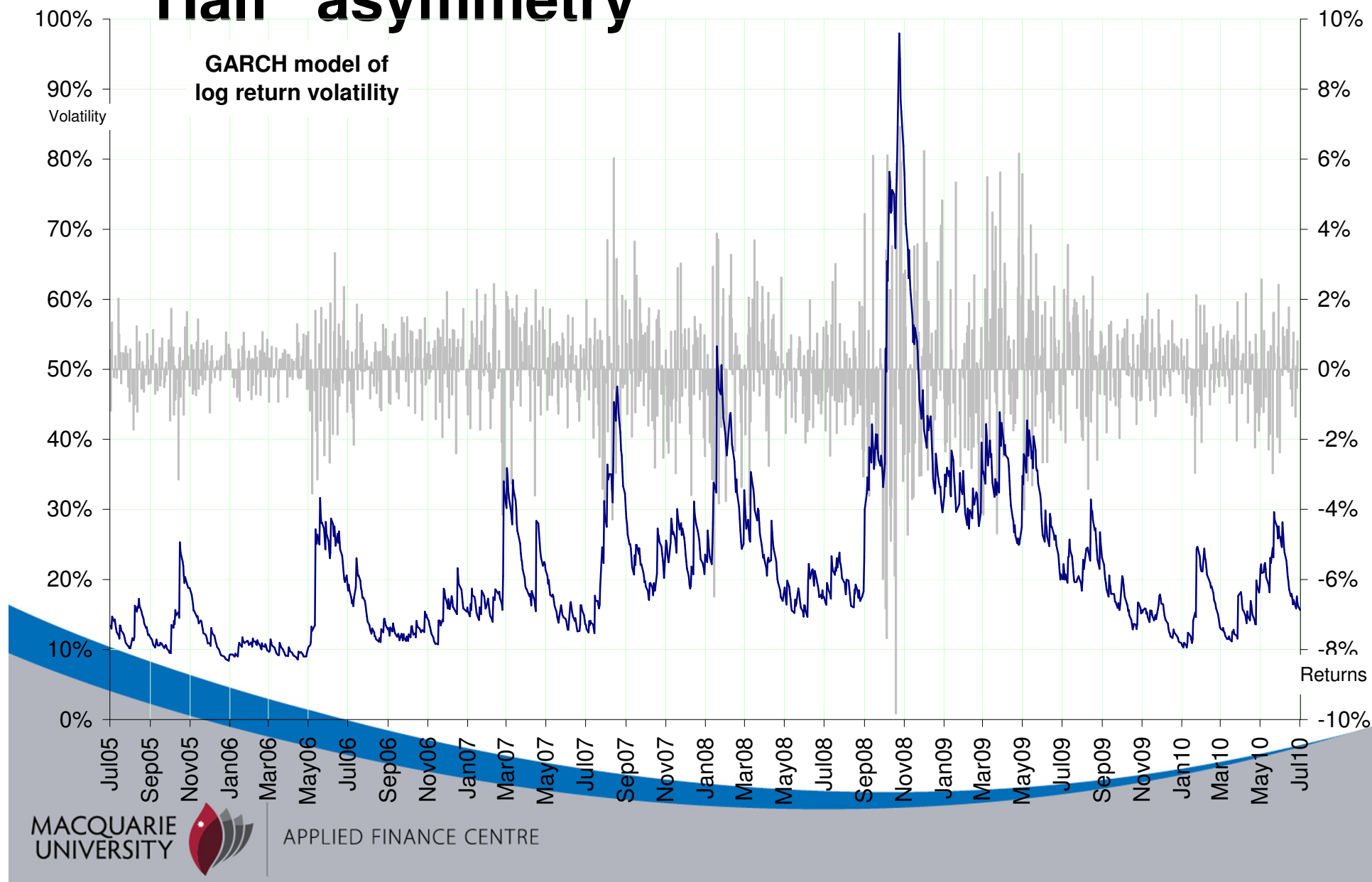


With asymmetry

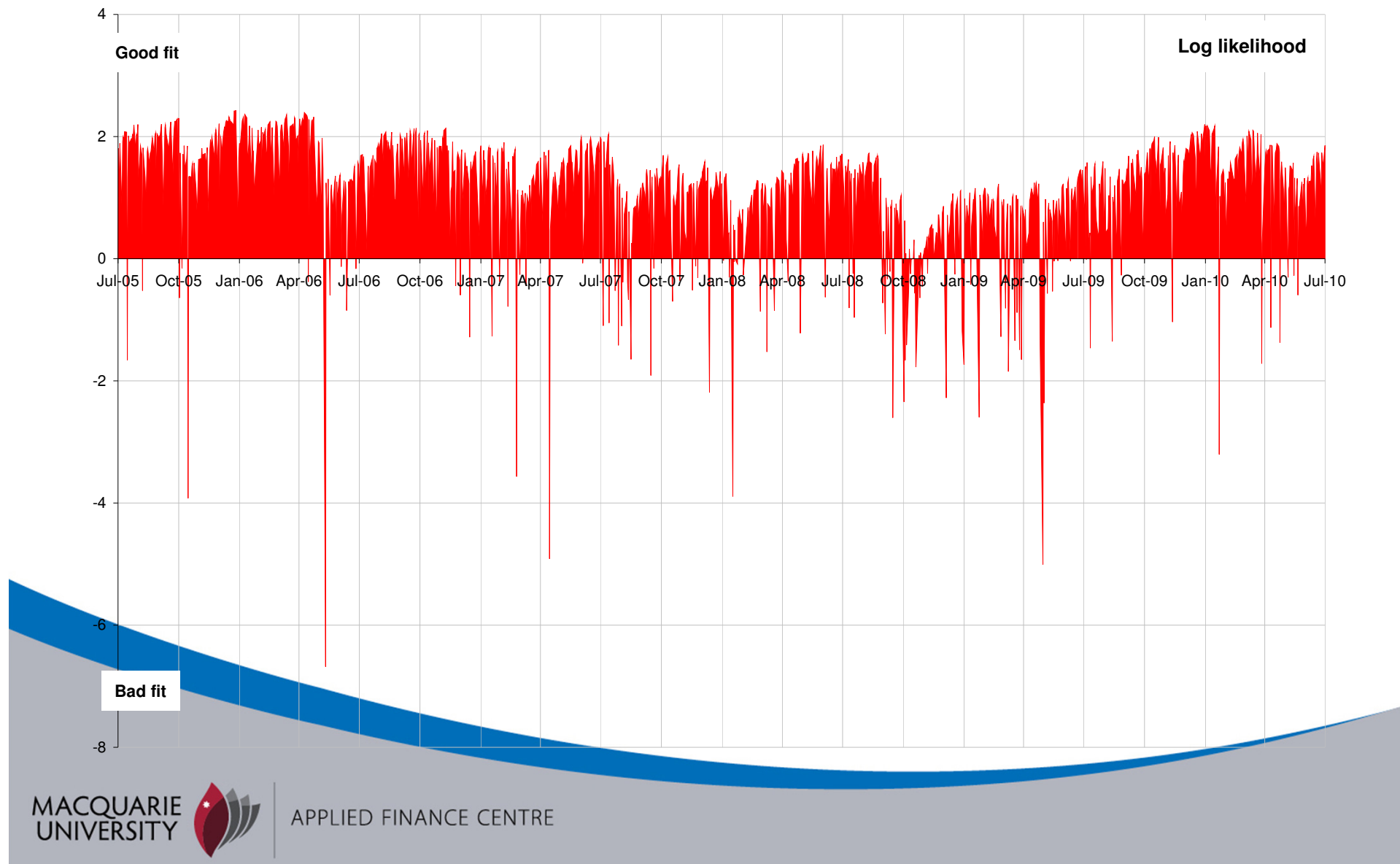
GARCH model of
log return volatility



“Half” asymmetry



Values of log likelihood



Spreadsheets

- [spreadsheets\vol - GARCH Germany.xls](#)
- [spreadsheets\vol - GARCH Australia.xls](#)
- [spreadsheets\vol - GARCH Singapore MLE.xls](#)

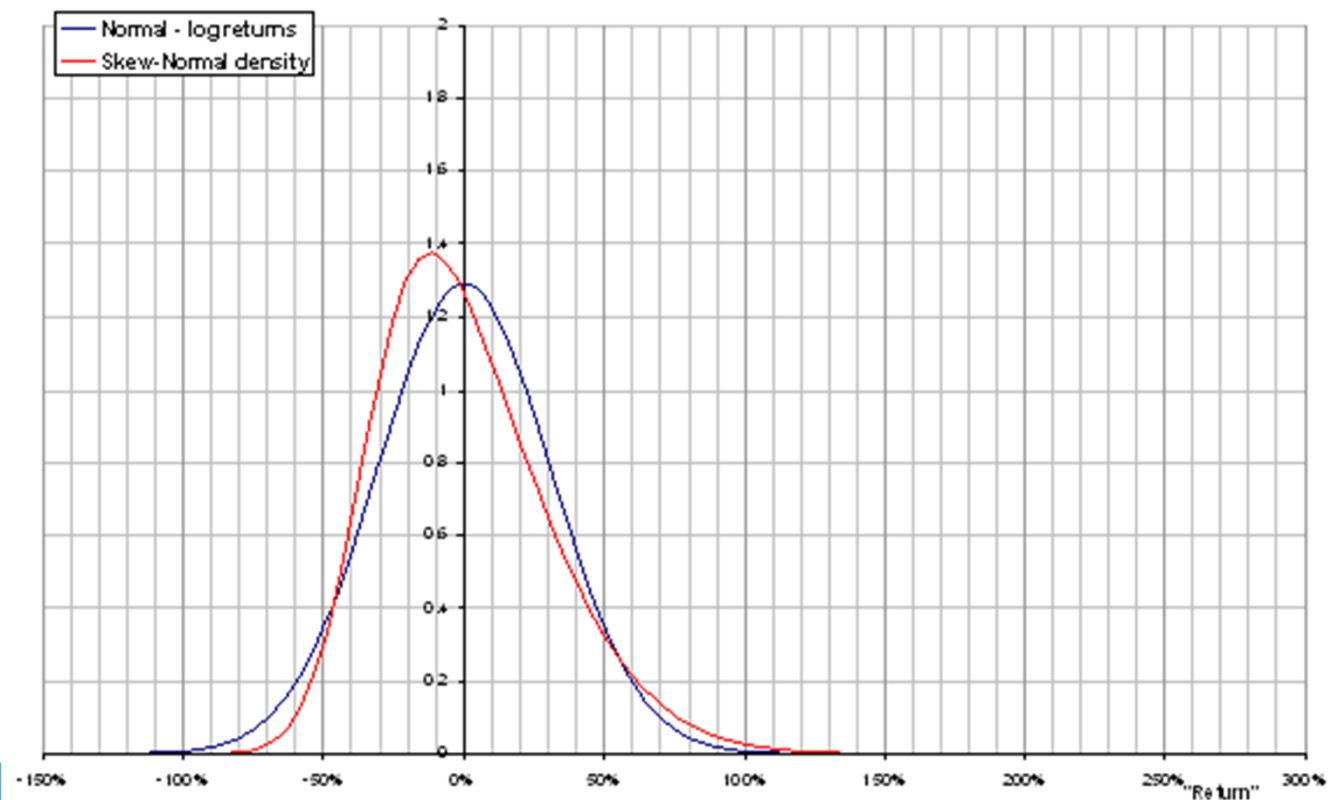
Conditional volatility

- What is the distribution of random variable z if $\sigma_t z$ is the standard deviation of log returns for day $t+1$?
 - Standardise log returns by dividing by volatility estimate
 - Examine distribution of new variables
- New data is much closer to Normality

	Raw log returns	Standardised returns
Count	1273	1273
Max	7.45%	3.57
75%	0.99%	0.64
50%	0.00%	0.00
25%	-1.08%	-0.68
Min	-8.67%	-4.24
Average	-0.03%	-0.03
Std dev	1.81%	1.00
skew	-0.02	-0.10
Excess kurtosis	1.71	0.14

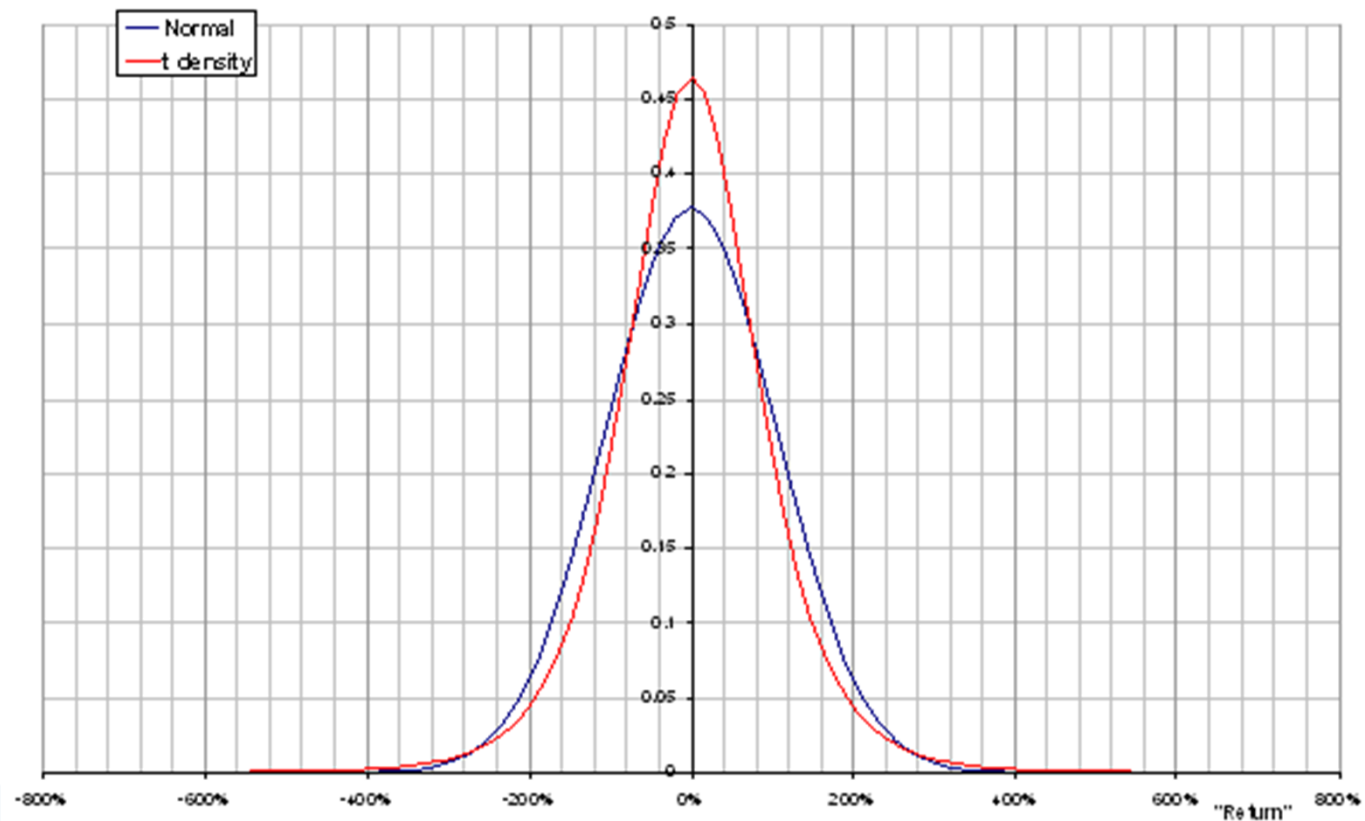
Skewness

- $S = \sum z^3 / n$



Kurtosis

- $K = (\sum z^4 / n) - 3$



Is the data Normally distributed?

- Jaque-Bera statistic is intuitive
 - Squares of *standardised differences from Normal values* for skew and volatility are asymptotically chi-squared
 - Sum of these two squares will be $\text{Chisq}(2)$
- $J = n[S^2/6 + K^2 /24]$
- Earlier German data had:
 - 1273 data points
 - skew = -0.10
 - excess kurtosis = 0.14
 - JB = 2.94, not significant

Is conditional volatility important?

UBS Annual Report 2008



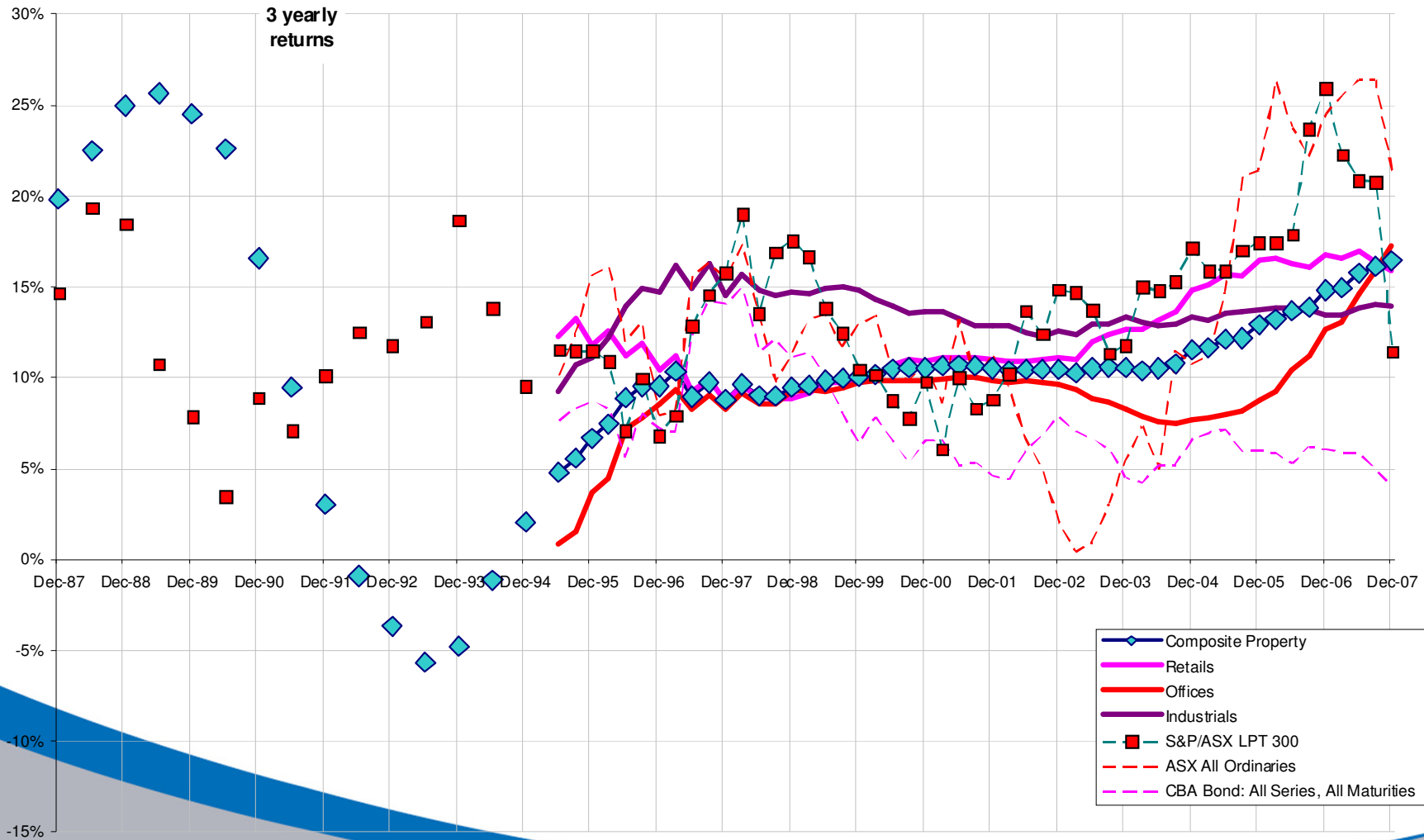
Good fit?

- Need to go through history and fit parameters on an incremental basis
 - Data for e.g. 11 Sep 2001, shouldn't be fitted using data after that date
- Beware of statistical tests
 - Have you removed all look-ahead bias?
 - Are subtle structures present in the data that could bias estimates of capital and risk

Illiquid Assets

- How can we estimate the volatility of an illiquid asset class?
- Measure over longer periods of time!
- Measure return over length of time similar to minimum holding period
- Can use overlapping periods

Direct Property Returns



	Compound return	Mean return	Standard deviation rolling 3 years	Standard deviation quarterly periods	Minimum return over 1 year	Minimum return over 3 years
Composite Property	11.0%	11.6%	10.5%	3.7%	-8.6%	-5.6%
Retails	13.7%	13.9%	6.2%	2.6%	6.9%	8.7%
Offices	9.5%	10.3%	12.9%	4.6%	-13.9%	-11.8%
Industrials	12.0%	12.4%	8.4%	3.2%	-10.0%	-4.0%
S&P/ASX LPT 300	12.9%	13.1%	7.7%	9.5%	-8.4%	3.5%
ASX All Ordinaries	14.3%	15.0%	12.1%	11.5%	-17.5%	-0.5%
CBA Bond: All Series, All Maturities	9.8%	10.0%	7.4%	5.1%	-6.8%	4.0%

What's wrong with overlaps?

- Many people assume that volatility cannot be calculated for data from overlapping time periods e.g. using rolling annual returns based on monthly data as the basic unit in calculating volatility.
- This is most probably due to people misremembering their basic course in statistics.

What's wrong with overlaps? (ii)

- What have people misremembered?
- The injunction was to not use the usual standard errors to judge the significance of the statistics.
- The estimation of the mean and standard deviation is, in most cases, unaffected by the rolling periods, but care must be taken in assessing their significance.

Volatility for overlapping periods

- The formula for the unbiased volatility where we measure volatility using rolling sums of length m is
- $\sigma^2 = \text{var}(\text{rolling } m\text{-period returns}) / m * (1 - m/(n - m + 1) + (m^2 - 1)/(3(n + m - 1)))$
- Where:
- n is the number of rolling periods we have in the sample;
- m is the number of periods in each rolling period;
- $\text{var}(\cdot)$ is the standard population estimate of the variance of the n rolling periods – sum of squared deviations around the mean divided by n .
- We require that $m < n/2$.

Non-overlapping periods

- Don't we have much less data to estimate from?
- No, it's just in a different form
- Example:
 - Yearly data usually based on years ending December
 - Can also use 12 month periods ending January, February, March ...
 - This gives 12 estimates of volatility based on different end dates
 - Each estimate is equally valid (under usual assumptions)

Non-overlapping periods

Month in which the rolling year ends	Australian Equities	Commodity futures cash backed	Bonds and cash	US equities in AUD
0	25.64%	13.07%	3.87%	21.13%
1	22.78%	13.34%	3.87%	18.60%
2	22.30%	13.47%	3.90%	21.08%
3	22.31%	13.07%	3.86%	20.47%
4	20.63%	14.18%	3.83%	19.37%
5	20.71%	16.23%	3.85%	18.94%
6	19.26%	13.00%	3.90%	20.57%
7	17.51%	13.30%	3.93%	20.00%
8	17.62%	13.21%	3.85%	20.85%
9	18.78%	15.01%	3.83%	21.06%
10	21.65%	15.10%	3.84%	20.97%
11	21.95%	12.96%	3.87%	21.40%





Extreme Value Theory

EVT and CLT

- CLT tells us what happens in the centre
 - Averages become Normally distributed
- EVT tells us what happens at the extremes
- Two versions of EVT
 - Extreme values of blocks of data come from one of three distributions
 - Extreme values of a single sample follow Generalised Pareto Distribution (GPD)

EVT – a quick intro

- Tails of distribution of returns are Generalised Pareto Distribution (GPD)
- $F(x) \cong 1 - (1 + kx/s)^{-1/k}$
- x is sufficiently large and positive, it is presumed to be larger than some threshold u
- Strictly:
- $F(x) = 1 - (1 + k(x-u)/s)^{-1/k}$

EVT

- $s > 0$ and is a scale parameter – the larger the value of s , then the more spread out are the values of x
- k is a shape parameter – as k decreases ($1/k$ increases), the tail of the distribution becomes skinnier
- $1+k(x-u)/s > 0$
- No standard notation available
 - Incredibly confusing
 - Even worse: are you looking at big positive values or big negative values

EVT PDF

- $$f(x) = \frac{1}{s} \left(1 + \frac{k(x-u)}{s} \right)^{-(k+1)/k}$$

EVT

- If U is uniform variate
- $x = u + s(U^k - 1)/k$
- x is GPD
- Mean = $u + s/(1-k)$
- Variance = $s^2 / (1-k)^2(1-2k)$
- The mean is infinite if $k > 1$
- Standard deviation is infinite if $k > 0.5$

Shape parameter

- Hill estimate is used for the shape parameter:
- $k = \Sigma \ln(x/u) / n$
- where:
- n is the number of observations where x is more extreme than u ,
 - summation is only over those extreme values

Scale parameter

- $s = k(\Sigma(x-u)^k / n)^{1/k}$
- The shape parameter k governs how fast the tail decays – like $x^{-1/k}$
- The scale parameter s gives us a measure of the typical size of the exceedances
- Both are important in the determination of the VaR

Play time

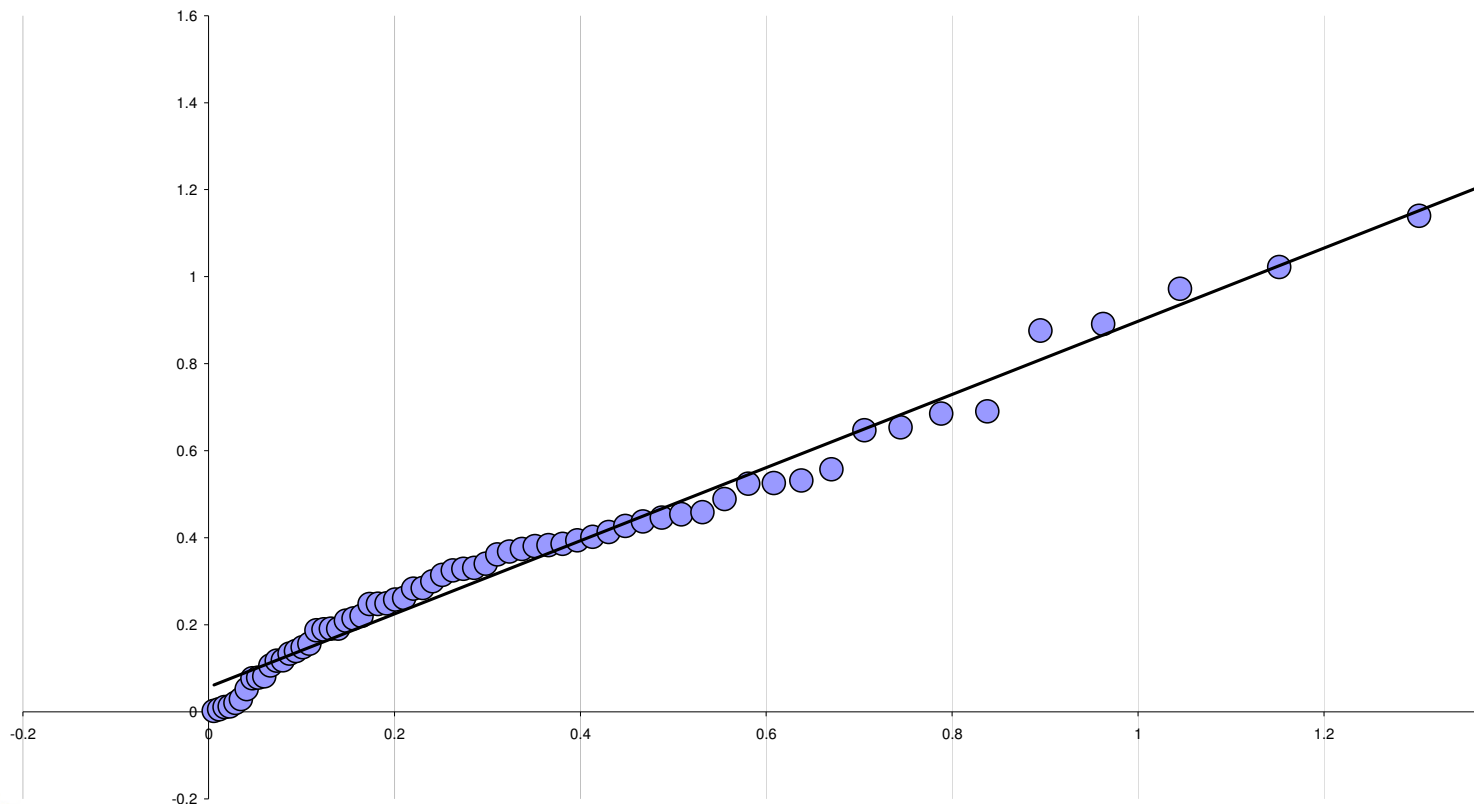
- Play with the spreadsheet
 - [spreadsheets\EVT Generalised pareto distribution and density.xls](#)
 - to see shape

Quantile Plot (Q Plot or QQ Plot)

- If GPD is good, log relative rank versus log exceedance should be a straight line.
- Q Plot drawn to enable judgement
- We may only have a few observations in the tail, and so not much statistical power – eyeballing may be the best test.
- $Prob(X > rank) = rank / \text{Number of obs} \cong cx^{-1/k}$
- $\log(rank) \sim -1/k \log(x)$

Singapore stocks, last 5 years

QQ plot – bad returns



Play time

- Play with the spreadsheet
 - [spreadsheets\EVT - estimation.xls](#) to see analyses
- What happens when you change the threshold?
- What does the tail look like on the upside?
- What if we look at monthly data?

Value at Risk

- Let p be the probability that an observation is in the tail.
 - probability of value more extreme than u is p
- the percentile at probability level q is
- $VaR_q = u + T_q$
- $T_q = (s/k) [(q/p)^{-k} - 1]$
- For good estimates need $q \ll p$.

Expected shortfall

- For estimating expected shortfall the formula is
- $ES_q = (u + T_q)/(1-k) + (s - ku)/(1-k)$
- Need to be careful to get back to percentage returns if we have done analysis on log returns



Statistical Bootstrapping

Problems

- How do we compute the distribution of a statistic when we have only have one sample of data?
- What if the statistic is so complicated that its confidence intervals and standard error are difficult, if not impossible, to otherwise specify or compute?
- What if the statistic itself is likely to be badly biased?
- Bootstrapping is a particular form of a *resampling* procedure.

Mean-variance algorithm

- Very sensitive to data sample
- Likely to be biased
- Almost impossible to give description of distribution of statistic for realistic problems
- [spreadsheets\Bootstrapping - block.xls](#)

Bootstrapped values

	<i>Aus</i>	<i>Canada</i>	<i>Germany</i>	<i>Japan</i>	<i>UK</i>	<i>US</i>	<i>Singapore</i>
<i>max</i>	68.50%	73.90%	69.07%	71.04%	83.99%	62.18%	74.93%
<i>75%ile</i>	27.23%	32.04%	26.59%	27.58%	12.70%	0.00%	34.12%
<i>median</i>	3.09%	2.04%	8.20%	10.42%	0.00%	0.00%	18.67%
<i>25%ile</i>	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	1.00%
<i>min</i>	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%

<i>average</i>	14.50%	16.84%	14.73%	15.44%	9.60%	7.11%	21.79%
----------------	--------	--------	--------	--------	-------	-------	--------

Structure in the data?

- What if there is some structure in the data?
- For return time series we could have:
 - heteroscedasticity
 - autocorrelation

Block bootstrapping

12 month

	<i>Aus</i>	<i>Canada</i>	<i>Germany</i>	<i>Japan</i>	<i>UK</i>	<i>US</i>	<i>Singapore</i>
<i>max</i>	78.35%	83.17%	81.68%	71.63%	50.60%	71.46%	74.79%
<i>75%ile</i>	27.63%	30.77%	25.99%	24.08%	13.57%	0.00%	35.18%
<i>median</i>	10.04%	1.54%	1.44%	9.01%	0.00%	0.00%	19.82%
<i>25%ile</i>	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	4.03%
<i>min</i>	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%

<i>average</i>	16.36%	16.98%	13.82%	15.99%	8.68%	6.17%	22.01%
----------------	--------	--------	--------	--------	-------	-------	--------

36 month

	<i>Aus</i>	<i>Canada</i>	<i>Germany</i>	<i>Japan</i>	<i>UK</i>	<i>US</i>	<i>Singapore</i>
<i>max</i>	79.39%	69.69%	87.44%	71.87%	46.58%	80.24%	98.56%
<i>75%ile</i>	24.30%	26.58%	19.74%	35.17%	9.31%	0.27%	39.66%
<i>median</i>	2.43%	7.20%	3.33%	17.69%	0.00%	0.00%	23.41%
<i>25%ile</i>	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	5.77%
<i>min</i>	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%

<i>average</i>	12.91%	14.57%	11.86%	20.35%	6.54%	7.22%	26.56%
----------------	--------	--------	--------	--------	-------	-------	--------

Block bootstrapping

72 month

	<i>Aus</i>	<i>Canada</i>	<i>Germany</i>	<i>Japan</i>	<i>UK</i>	<i>US</i>	<i>Singapore</i>
<i>max</i>	67.88%	83.61%	46.44%	100.00%	47.15%	84.18%	66.69%
<i>75%ile</i>	20.89%	21.71%	17.56%	36.65%	14.25%	18.48%	35.53%
<i>median</i>	2.98%	5.02%	1.89%	14.02%	0.00%	0.00%	19.62%
<i>25%ile</i>	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	1.89%
<i>min</i>	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%

<i>average</i>	11.94%	15.29%	10.08%	21.05%	7.44%	12.21%	21.99%
----------------	--------	--------	--------	--------	-------	--------	--------

120 month

	<i>Aus</i>	<i>Canada</i>	<i>Germany</i>	<i>Japan</i>	<i>UK</i>	<i>US</i>	<i>Singapore</i>
<i>max</i>	57.58%	85.69%	30.67%	86.24%	58.48%	90.57%	77.32%
<i>75%ile</i>	10.42%	33.15%	9.12%	49.86%	7.39%	0.00%	34.74%
<i>median</i>	3.09%	10.06%	0.00%	18.81%	0.00%	0.00%	25.32%
<i>25%ile</i>	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	11.96%
<i>min</i>	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%

<i>average</i>	9.04%	18.89%	5.35%	27.18%	6.87%	8.71%	23.97%
----------------	-------	--------	-------	--------	-------	-------	--------

Bootstrapping residuals

- y is a function of variables x
 - E.g. a model of average claim payments (PPCI) in a run-off triangle:

$$y = f(x)$$

- How accurate are the estimates of the parameters and how accurate is the estimate of outstanding claims from this model? Is there some bias to the raw estimate?
- Residuals can be calculated

$$r_k = \hat{y}_k - y_k$$

- These may be standardised in some way if they have different natural sizes – say late development payments are much more volatile than early development claims.

Bootstrapping estimates

- A new hypothetical data set can now be constructed from the values $\hat{y}$ combined with random values from the set of values of r .
- A new model can then be estimated from this new data set and outstanding claims and other values of interest calculated.
- Repeat this step a number of times to give a sufficient indication of the uncertainty in our estimates and to have a reasonable estimate of the outstanding claims.

Outstanding claims

- Worked example
- [spreadsheets\bootstrapping - taylor 2000 data.xls](#)





Stochastic Calculus

And its travelling companions

Ito's Lemma

- Suppose that some variable y is a function $f(s, t)$ of s and t , and that s follows a continuous random process that can be described by

$$ds(t) = \alpha dt + \sigma dz(t)$$

- Here α is the instantaneous expected rate of change in s and σ is its instantaneous standard deviation. Then y will also follow a random process — one induced by the randomness of s .
- Ito's lemma relates the characteristics of the y process to the s process. Specifically, y will follow

$$dy(t) = (\alpha f_s + 1/2 \sigma^2 f_{ss} + f_t) dt + \sigma f_s dz(t)$$

- where subscripts denote partial derivatives.

Understanding the Black-Scholes Model

- The logic of the Black-Scholes-Merton argument is as follows:
- Find a mathematical expression for changes in value of asset.
- $dS = S\mu dt + S\sigma dz$ (2)
- The random component dz has a standard Normal distribution:
- a mean of $0dt$,
- a standard deviation of $1 * \sqrt{dt}$.

- Simplistically what we've said is that the stock price changes over time in line with:
- its expected continuous rate of return, μ ,
- its percentage volatility, σ .

- Find an expression for changes in the value of the call given what's happening to the asset.

Change in C	=	dC
	how fast the option price changes as S changes	$\partial C / \partial S$ which we'll call Δ
times	the change in S	dS
plus	the change in how the option price changes as S changes	$\partial \Delta / \partial S = \partial^2 C / \partial S^2$ which we'll call Γ
times	half the square of the change in S	$\frac{1}{2} dS^2$
plus	change in option price over time	$\theta = \partial C / \partial t$
times	the change in time	dt
plus	lots of terms where we multiply <i>dt</i> and <i>dS</i> terms	safely ignore them

WOW!

- $dC = \Delta dS + \frac{1}{2}\Gamma(dS)^2 + \theta dt$
- Now, noting that dS^2 can be ignored except for dz^2 , which equals dt
- $dC = \Delta dS + \theta dt + \frac{1}{2}\Gamma S^2 \sigma^2 dt$ (3)
- $\Delta = \partial C / \partial S$
- $\theta = \partial C / \partial t$
- $\Gamma = \partial^2 C / \partial S^2 = \partial \Delta / \partial S$

- Rearranging gives an expression for changes in a portfolio value (long C, short S) that doesn't have any random component:
- It's a risk-free portfolio. Amazingly(?), the expected return of the asset also falls out.
- $dC - \Delta dS = \theta dt + \frac{1}{2} \Gamma S^2 \sigma^2 dt$
- Return on the risk-free portfolio must be the risk-free rate.
- $dC - \Delta dS = (C - \Delta S) r dt$

- Construct a partial differential equation.
- $(C - \Delta S) r dt = \theta dt + \frac{1}{2} \Gamma S^2 \sigma^2 dt$
- $\partial C / \partial t = r C - r S \partial C / \partial S - \frac{1}{2} S^2 \sigma^2 \partial^2 C / \partial S^2$
- The solution to this partial differential equation is the Black-Scholes formula.

Risk Neutral valuation

- This is a mathematical trick!
- *No arbitrage exists
if and only if*
- *the current asset prices are the present value of a
probability weighted average of the future state prices*

Example - 1

- Assume you have n assets with current prices $p_1 \dots p_n$
- AMP = 5
- BHP = 12
- CSR = 3

Example (2)

- At the next time interval there are s possible states (or scenarios) of the world and the i th asset has price s_{ij} in the j th state. The number of states is greater than or equal to the number of assets.

- Prices A, B, C

- Good 9, 19, 5

- Average 5, 13, 3

- Bad 2, 4, 2

- The number of states is greater than or equal to the number of assets. Otherwise some of the assets are redundant and are identical to a portfolio of the other assets. Binomial option pricing is an example of this.

Example (3)

- Arbitrage exists if you have a current portfolio which can be put in place for zero cost or less (i.e. you get paid) and has future values in each state which are all positive or zero. Or vice versa.
- In other words, something for nothing and with no downside risk.

Example (4)

- The risk-neutral probabilities are 22.2%, 66.7% and 11.1% for the Good, Average and Bad scenarios;
- the risk-neutral rate of return is 11.1%.

Insights

- Why use the term risk-neutral?
- Under the calculated probabilities all assets and all portfolios earn the one rate of return. It doesn't matter what the possible outcomes of the asset are, no matter how dangerous or how good, the average return is always the same.
- So this return is *neutral* to whatever risk the asset has.
- Note we are not saying anything about the risk-aversion of an investor, in fact in all this work we haven't assumed we have any investors; we've only talked about asset markets.

Insights (2)

- Notice how arbitrary the risk-neutral probabilities are.
- They are derived only from the possible states of the world and the current prices and the assumption of no arbitrage.
- They are not dependent at all on economic fundamentals or common sense or anything else; just the prices in the future states and at present.

Insights (3)

- If we have many more states than assets then there may be many possible solutions for the risk-neutral probabilities and risk-neutral returns.
- This further emphasises the arbitrary nature of the probabilities.
- If we want the risk-neutral probabilities to have some particular form, such as a Lognormal distribution, then there may be only one solution.

Insights (4)

- If there is an asset which pays off the same amount in all states then its return has to be the risk-neutral return; we call it the *risk-free asset* because there is no uncertainty in its return, and this return is the *risk-free rate of return*.
 - But note that we do not have to assume a risk-free asset exists.
- There is no requirement that the risk-neutral return be positive.
- There is no restriction that current prices be positive or that the payoff in each state is positive.

Insights (5)

- The common technique of *scenario analysis*, where a number of possible future scenarios are postulated along with a probability of occurrence for each state, is only guaranteed to be arbitrage free if the current asset prices can be calculated as a risk-neutral probability weighted average of the scenarios.
- As it is easy to calculate a risk-neutral set of probabilities then we should do so whenever attempting complicated scenario analysis.
- Not being able to find a risk-neutral solution indicates an inappropriate set of scenarios. This is especially important if options are included amongst the assets.

Risk Neutral Valuation Revisited

- The price of the option doesn't have any term for return of the stock.
- Therefore we can pretend we are in a risk-neutral world.
- Risk-neutral valuation can be summarized as follows:
- Assume that the expected return on the underlying asset is the risk-free interest rate.
- Calculate the expected payoff from the option at maturity.
- Discount the expected payoff at the risk-free rate.
- Note when we say “expected return” we are talking about horizon returns, not log returns

Simulation

- Consider a stock currently priced at \$10.00 with continuous volatility of 30% pa. Simulate the value of a call option over this stock with a strike price of \$10, a term of 20 days, where the risk-free rate is 5.00% pa continuous.
- *Tricky bit:*
- Continuous rate of return for stocks in risk-neutral world does not equal the continuous risk free rate.
- Expected returns are equal in the real world (they are arithmetic averages)
- Need to adjust expected log return for stocks by their volatility.

	Log returns	Percentage Return	Per period		
Mean	0.50%	5.13%	0.0020%	Option maturity (yrs)	<i>0.080000</i>
Volatility	<i>30.00%</i>	32.26%	1.90%	Strike	<i>\$10.00</i>
Risk free	<i>5.00%</i>	5.13%	0.02%	Period (yrs)	0.004000

Period	Random variate	Normal variate	Log return	Price movement	Price at end of period
1	0.368075	-0.336957	-0.006373	0.993647	<i>10.00</i> 9.94
2	0.987242	2.233498	0.042398	1.043309	10.37
3	0.500185	0.000463	0.000029	1.000029	10.37
4	0.766132	0.726168	0.013798	1.013894	10.51

Average	\$0.35	\$10.03
Std dev	\$0.52	\$0.84
Std error	\$0.05	\$0.08
PV of average	\$0.35	
Data Table		
Simulation number	Option price	End price
	\$0.20	10.20
1	\$0.1341	10.13
2	\$0.4653	10.47
3	\$0.1820	10.18
4	\$0.0000	9.93
5	\$0.0000	9.22
6	\$0.4604	10.46
7	\$0.3589	10.36

To and from Lognormal and Normal

- Assume that the price relative X of S has a Log-Normal distribution

$$E(X) = 1+m$$

$$\text{var } X = s^2$$

- So S has an expected return of m and volatility s
- If we want to calculate the same parameters of the return r of a portfolio of stocks, where stock i has weight w_i :

$$E(r) = \sum w_i m_i$$

$$\text{Var}(r) = \sum \sum w_i w_j s_{ij}$$

- Where s_{ij} is the covariance between the returns of stock i and j .

LogNormal to Normal

- Then $Y = \ln(X)$ is the log return of S and has the following parameters:
- $\text{var } Y = s'^2 = \ln(1 + (s^2 / (1+m)^2))$
- $EY = m' = \ln(1+m) - s'^2 / 2$
- Y has a Normal distribution

Log returns

- The formulas above for portfolio return and variance cannot be applied to log returns, they give the wrong answers. For example, a portfolio of two equally weighted stocks each with log returns of 10% p.a. and volatilities of 20% p.a. **does not** have a log return of 10% p.a.
- If we want to do portfolio arithmetic then we must (i) use the formulas below to convert from expected log-returns to expected returns, (ii) do the portfolio arithmetic and then (iii) convert back to log-returns.

Normal to Lognormal

- The formulas to move from Y to $X=\exp(Y)$ are:
- $E(X) = 1+m = \exp(m' + s'^2/2)$
- $\text{var } X = s^2 = (\exp(s'^2) - 1) * (1+m)^2$

This is a warning!

- Note that the volatility term slips into the calculation of the means and vice versa.
- These equations are given as a warning.
- In too many cases in practice this relationship is ignored or forgotten.
- In some of the spreadsheets these are used where we make an assumption about expected returns in the “real world”, and need to convert that to the appropriate value in the “log-world”.



Copulas

Diversification

- Covariance
- Correlation
- Estimation
- Difficulties of forecasting
- Copulas (implications for risk measurement)

Multivariate distributions

- What should be looked for?
 - Recap and introduction of dependence properties
 - Dependence
 - Statistical independence
 - Co-movement
 - Correlation measures
 - Tail dependence

Multivariate distributions

- Recap and introduction of dependence properties
- Dependence
- Statistical independence
- Co-movement
- Correlation measures
- Tail dependence

Multivariate distributions

- More than one variable of interest
- Need to specify dependence of one distribution on the others

Example

- Multivariate Normal distribution
 - μ , and $\mathbf{A}$ are the vector of means and the covariance matrix

$$f() \propto \int \exp\left(-\frac{1}{2}(x-\mu)^T A^{-1}(x-\mu)\right) dx$$

Example: Multivariate Normal

- Covariance matrix A has the correlations within it that define the relationships between the variables
- If X and Y are from a Multivariate Normal

then

- X+Y has a Normal distribution
 - mean = $\mu_X + \mu_Y$
 - variance = $\sigma_X^2 + \sigma_Y^2 + 2\sigma_X\sigma_Y\rho_{XY}$
- variance(X | Y) = $\sigma_X^2(1-\rho^2)$

Very nice properties
and implicitly assumed
in many cases,
but are they always
applicable?

Dependence

- We can have dependence by other structures
- Y may be a simple or complicated function of X

$$Y = f(X)$$

- Most cases it is only worthwhile to invoke a simple functional relationship for market variables
 - How do we measure dependence? ...

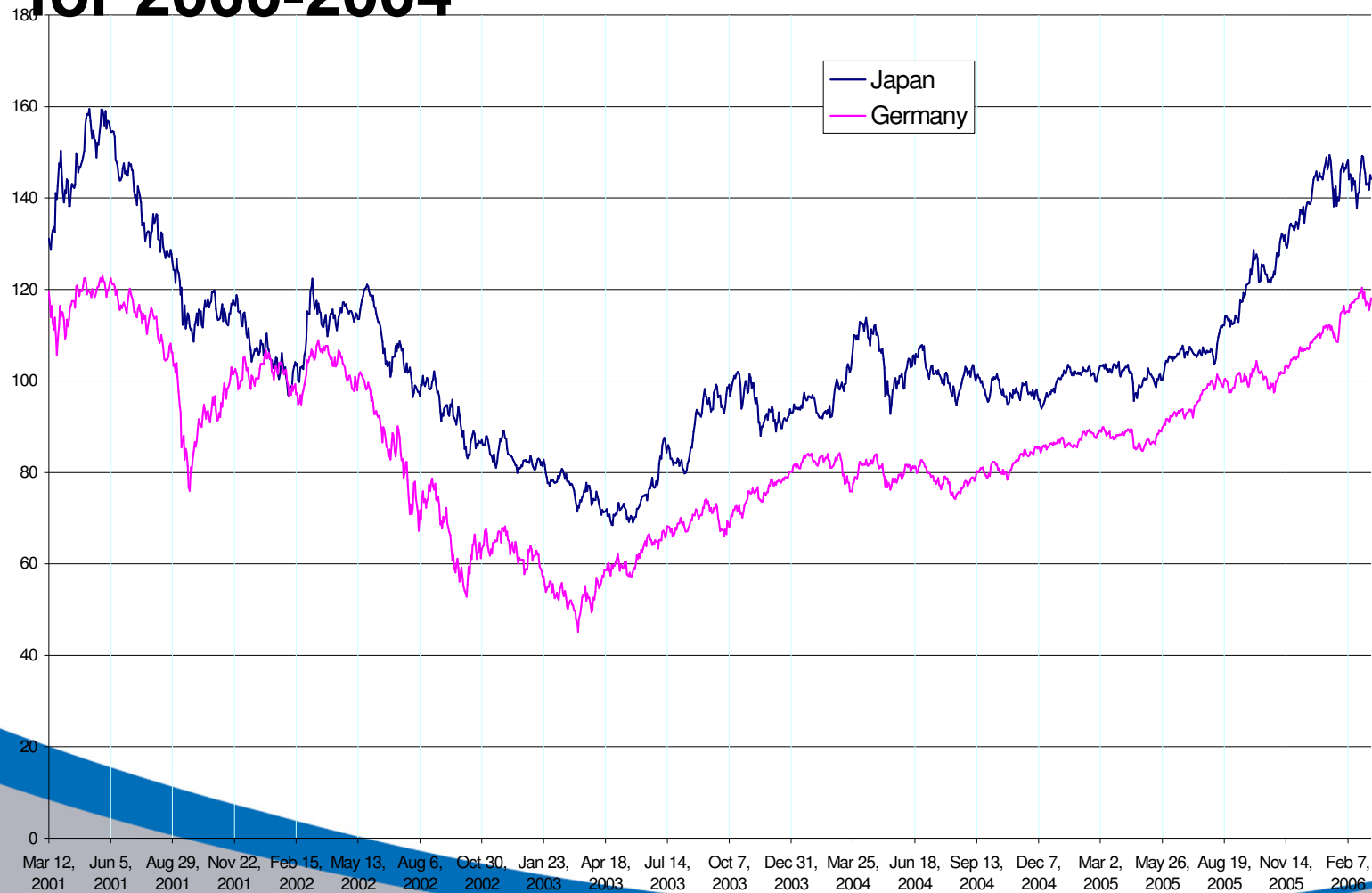
Correlation

- **Pearson** – product-moment correlation
 - $\rho = [E(XY) - E(X)E(Y)] / \sigma_X \sigma_Y$
- **Spearman** – rank correlation
 - correlation coefficient of the ranks of X and Y
- **Kendall - tau**
 - $A_{ij} = \text{sign}(X_j - X_i)\text{sign}(Y_j - Y_i)$
 - $T = \sum_{i < j} A_{ij} / (n(n-1)/2)$

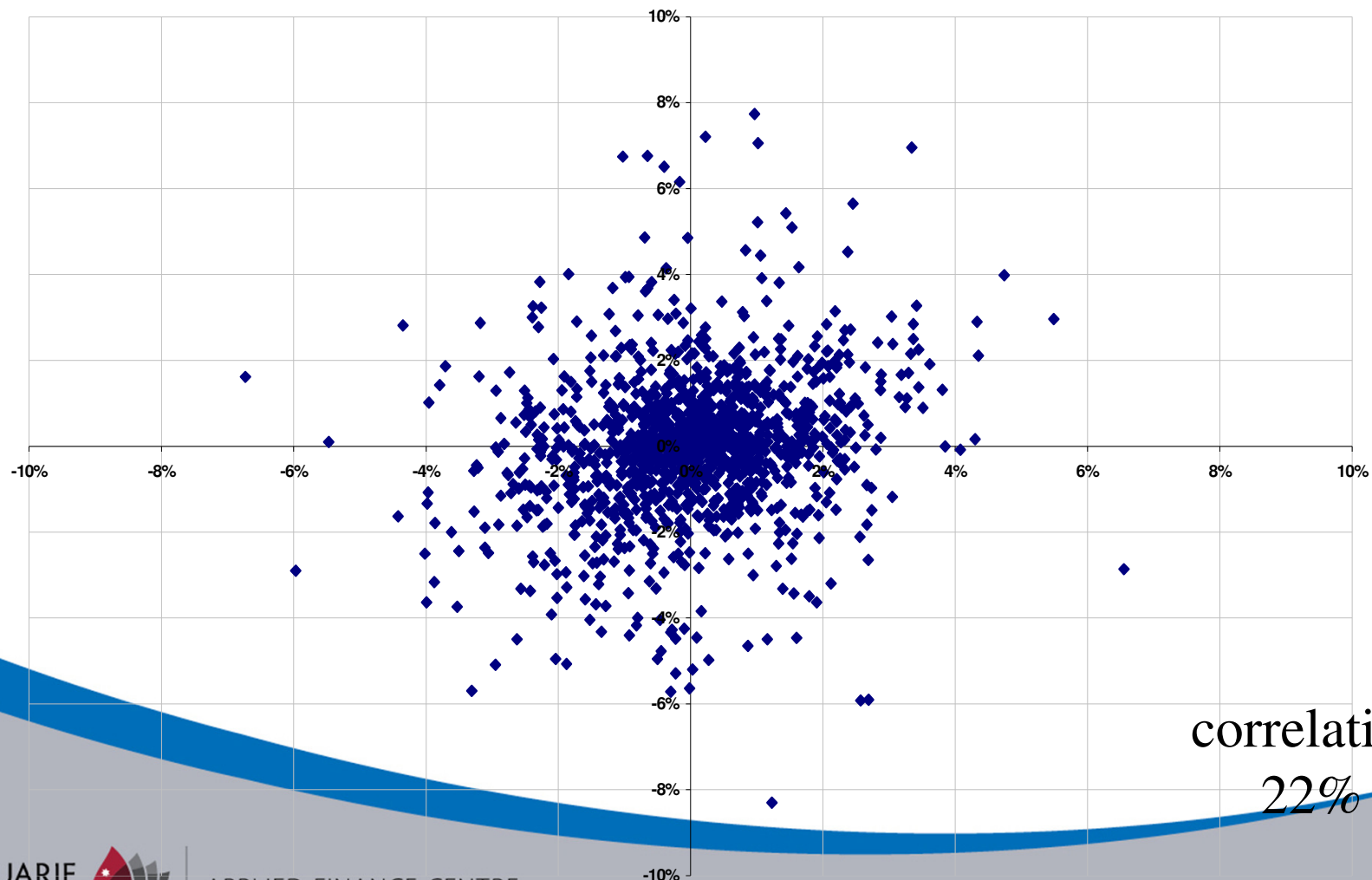
Product moment

- A natural second moment for any distribution
- $E(X^2)$, $E(Y^2)$, $E(XY)$
- Defining parameter for any elliptical distribution
 - Embodied in covariance matrix for Multivariate Normal
 - This is a major motivator for use of this coefficient

Let's look at Japanese and German returns for 2000-2004



Scatter chart of daily log returns



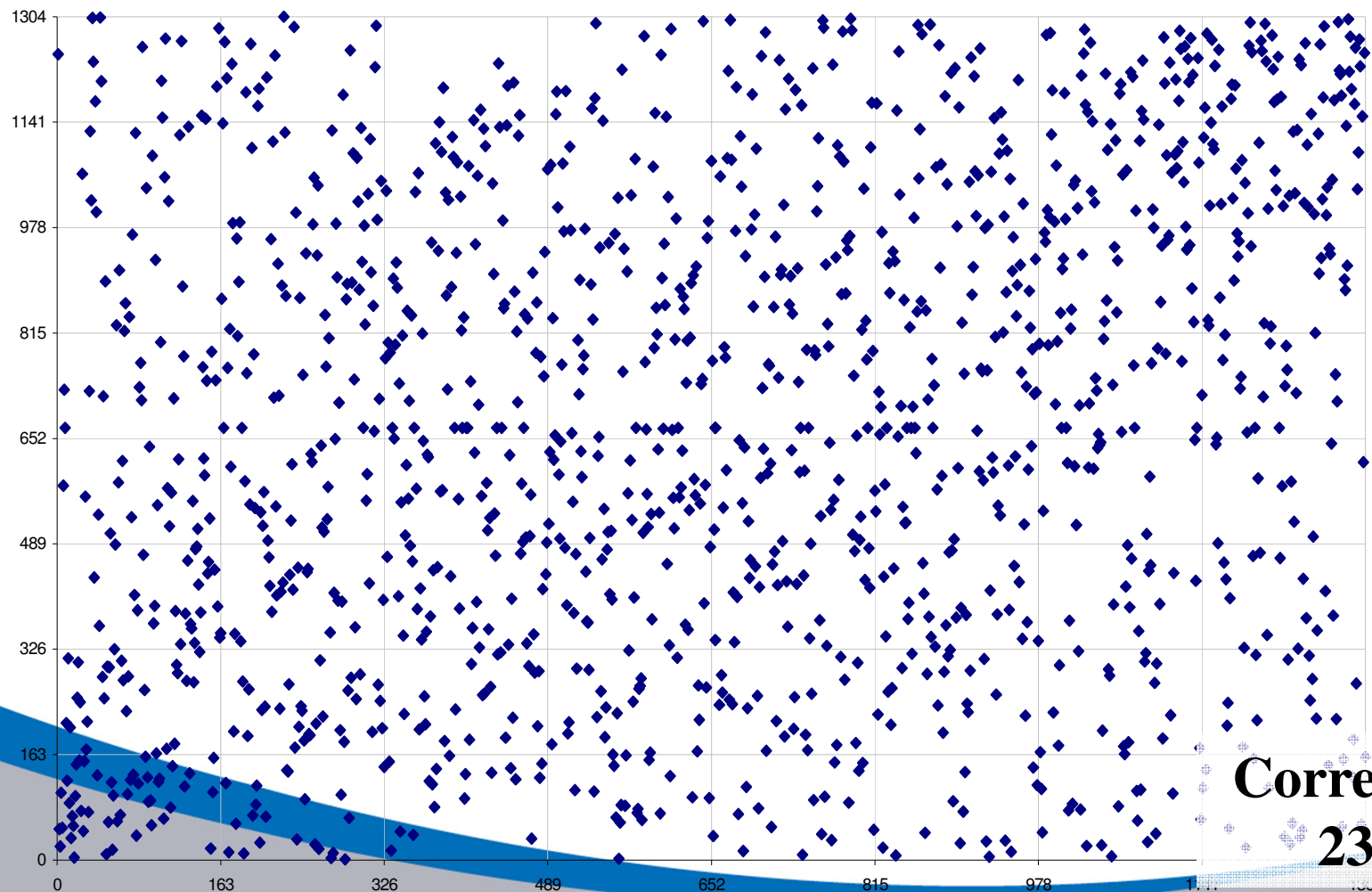
correlation
22%

Pearson's product-moment

- Problems?
 - Influenced by extreme values
- Not most suitable for asymmetric distributions or fat-tails
 - Similar problems to standard deviation

Rank correlation

- Correlation of ranks
- Intuitive, easy to calculate
- No good statistical backing
 - Except in certain cases
- Robust



Correlation
23%

Kendall's tau

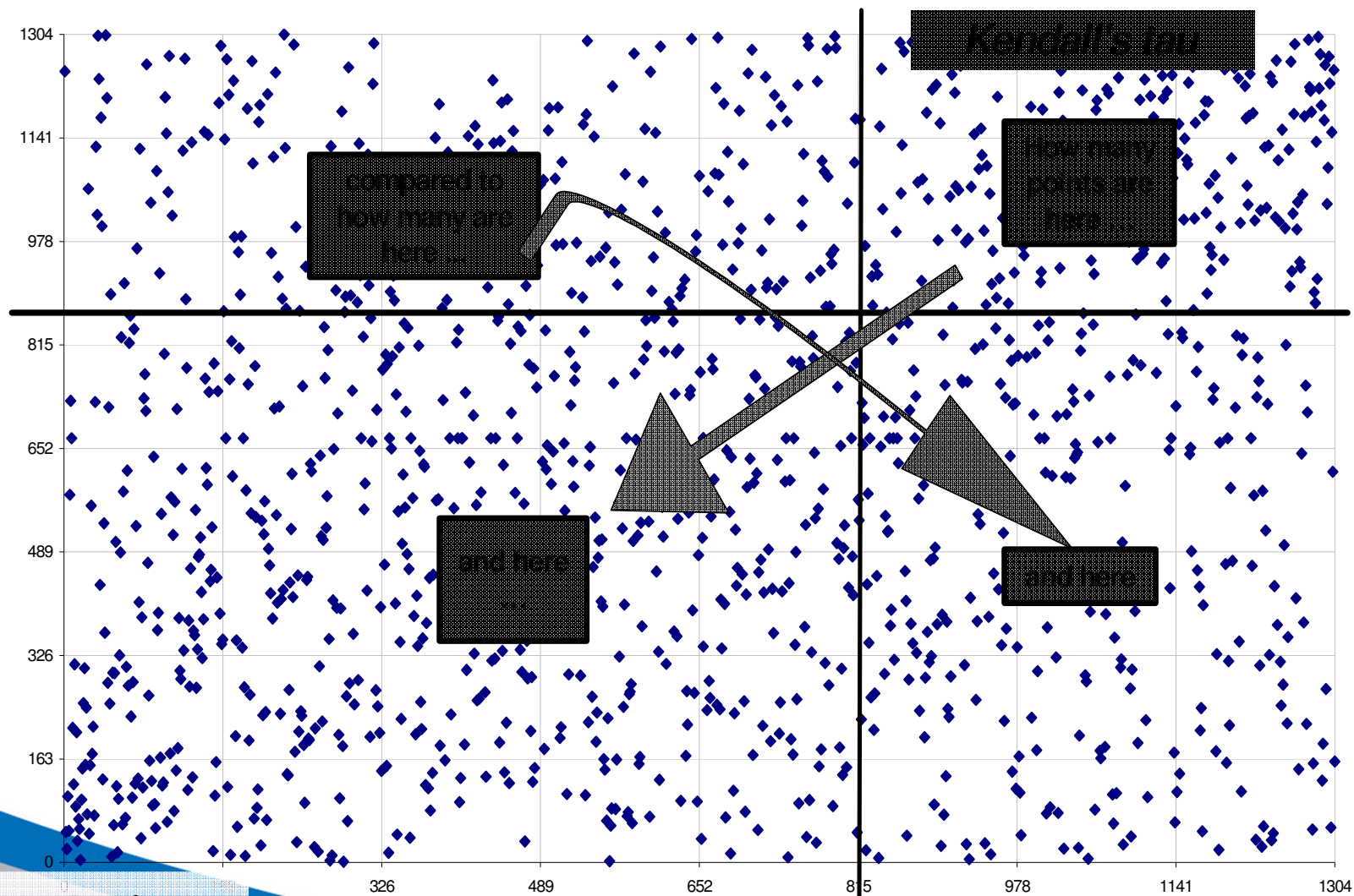
$$A_{ij} = \text{sign}(X_j - X_i) \text{sign}(Y_j - Y_i)$$

$$T = \sum_{i < j} A_{ij} / (n(n-1)/2)$$

- Robust
- Interpretable
 - compared to rank and product-moment correlation

Kendall tau – interpretation

- 2 pairs (X_i, Y_i) and (X_j, Y_j)
- What proportion of pairs are *concordant*?
 - $X_i > X_j$ implies $Y_i > Y_j$ and vice versa
- Proportion = $(1+\tau)/2$



correlation

16%

averaged across all points

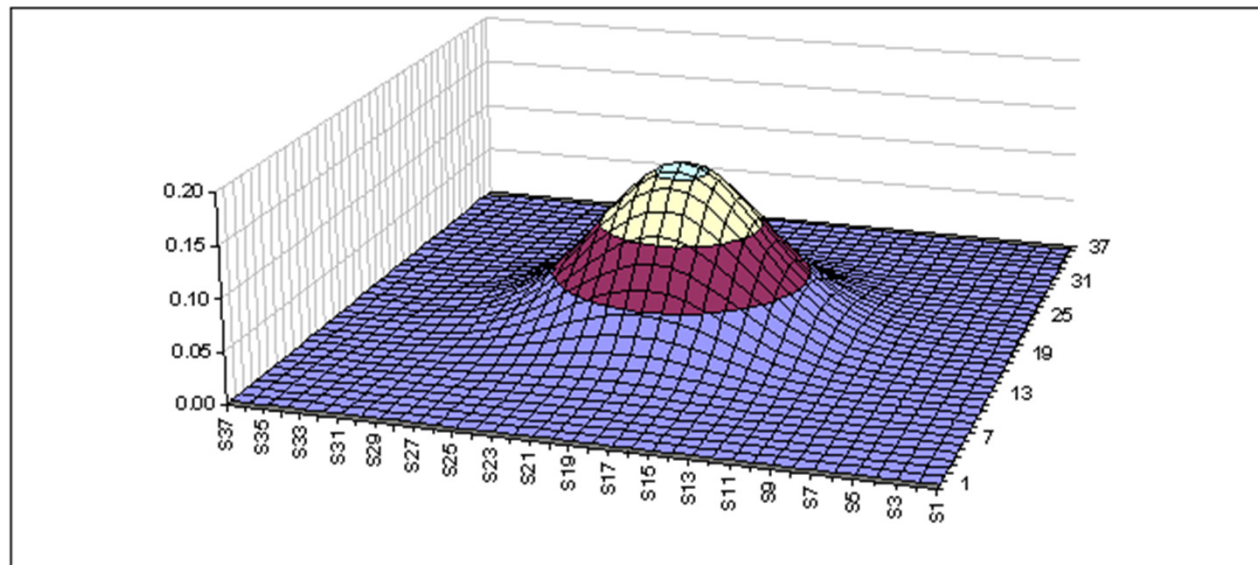
How are you measuring correlation?

What's the big deal?
Let's look at copulas

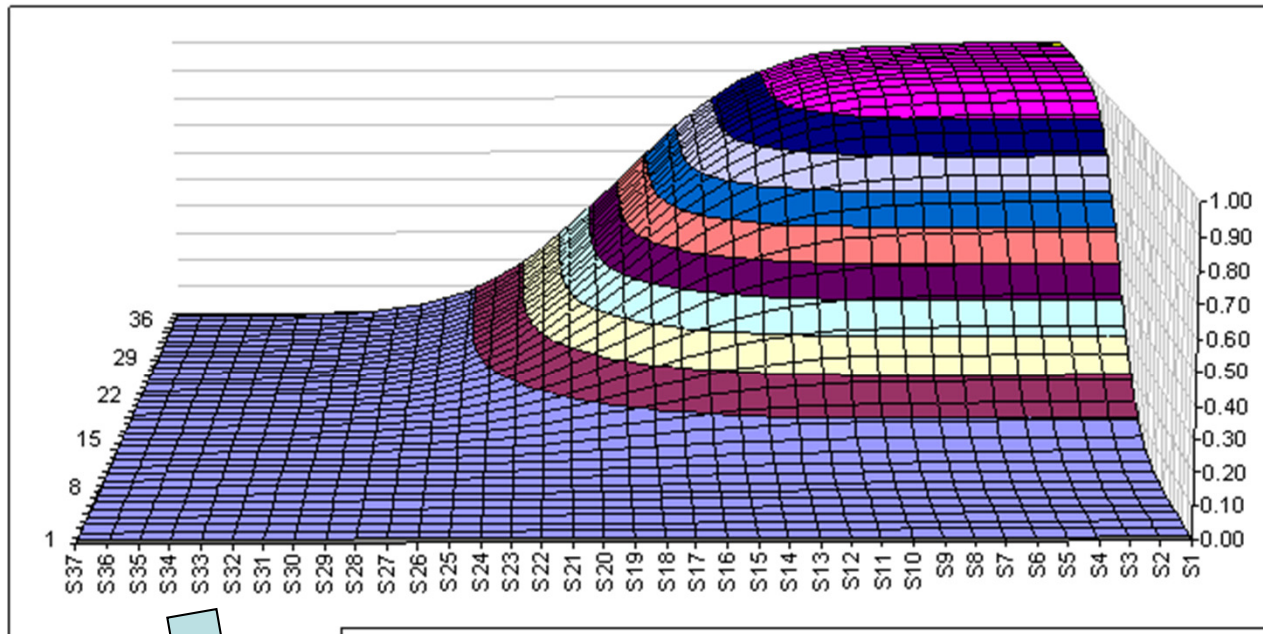
Multivariate densities

- Play with the spreadsheet

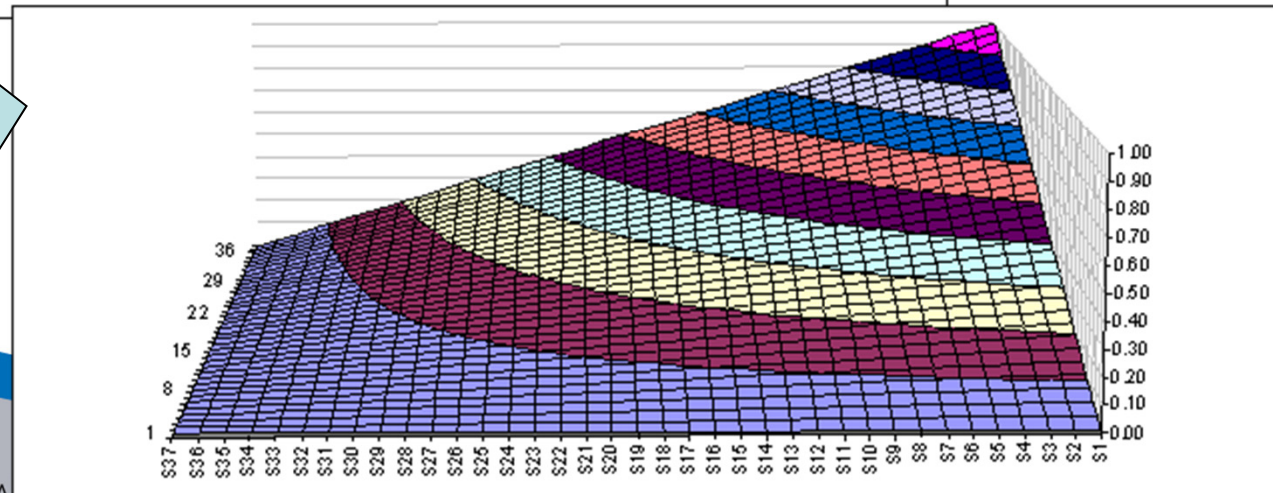
Cop - bivariate Normal.xls



Multivariate distribution



Rescale
axes



Sklar's Theorem

- Given a multivariate cumulative distribution $M(x, y, \dots, z)$
- Each variable has a marginal (cumulative) distribution $F(x), G(y), \dots H(z)$ etc

then

- $M(x, y, \dots, z) = \mathbf{C}(F(x), G(y), \dots, H(z))$ for some function $C(\cdot)$
- C is a copula

Example (i)

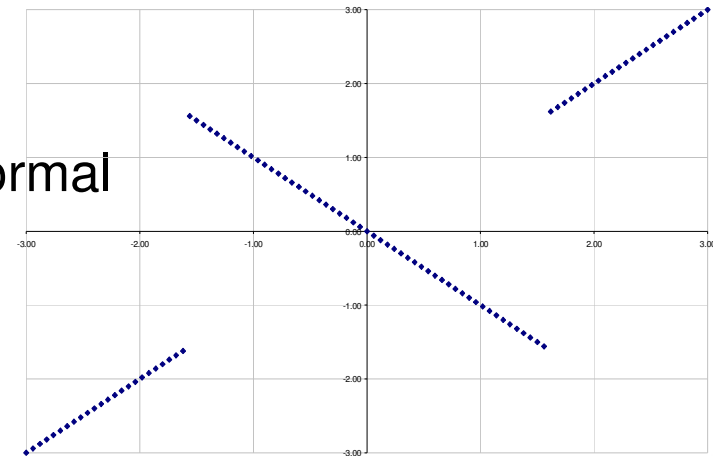
- $M(.)$ is multivariate Normal
- Each variable has a marginal Normal distribution
- Copula is called Gaussian copula

Example (ii)

- X has standard Normal marginal distribution
- Y has standard Normal marginal distribution
- X and Y are statistically independent i.e. $E(XY) = E(X)E(Y)$
- Distribution of $X+Y$ depends on copula

- Case 1:
 - Gaussian copula $\rightarrow X+Y$ is Normal

- Case 2:
 - $X = Y$ if $|X| > \pi/2$
 - $X = -Y$ otherwise
 - $X+Y$ is not Normal



Copulas

- Appropriate measure of correlation is Kendall's tau

Limits on copulas

$$\max(0, \sum x_i - n + 1) < C(x_i) < \min(x_i)$$

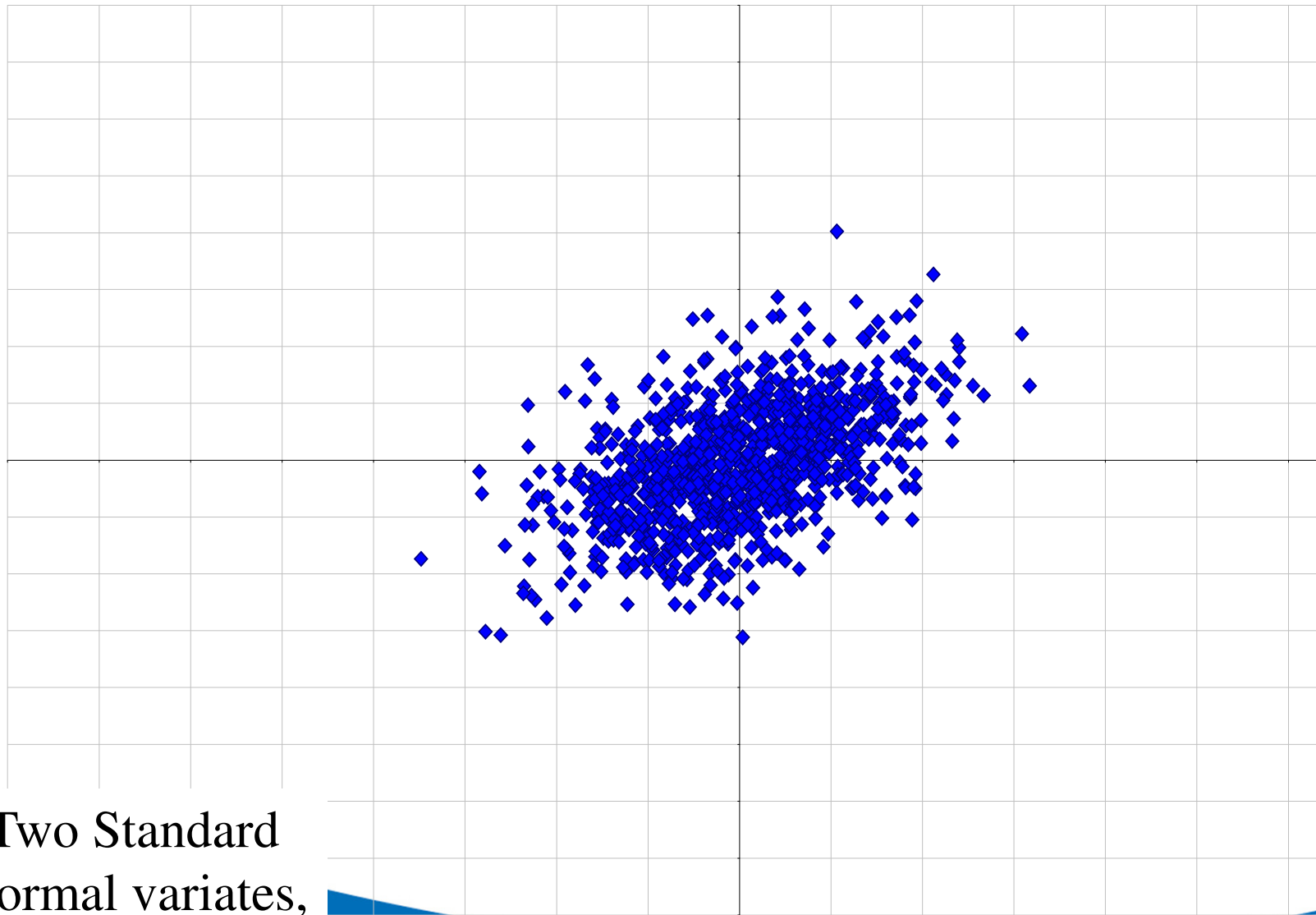
- Limits can be seen on spreadsheet with correlation = -1 and +1

Are you considering
the appropriate copula to use?
Or haven't you thought about it?

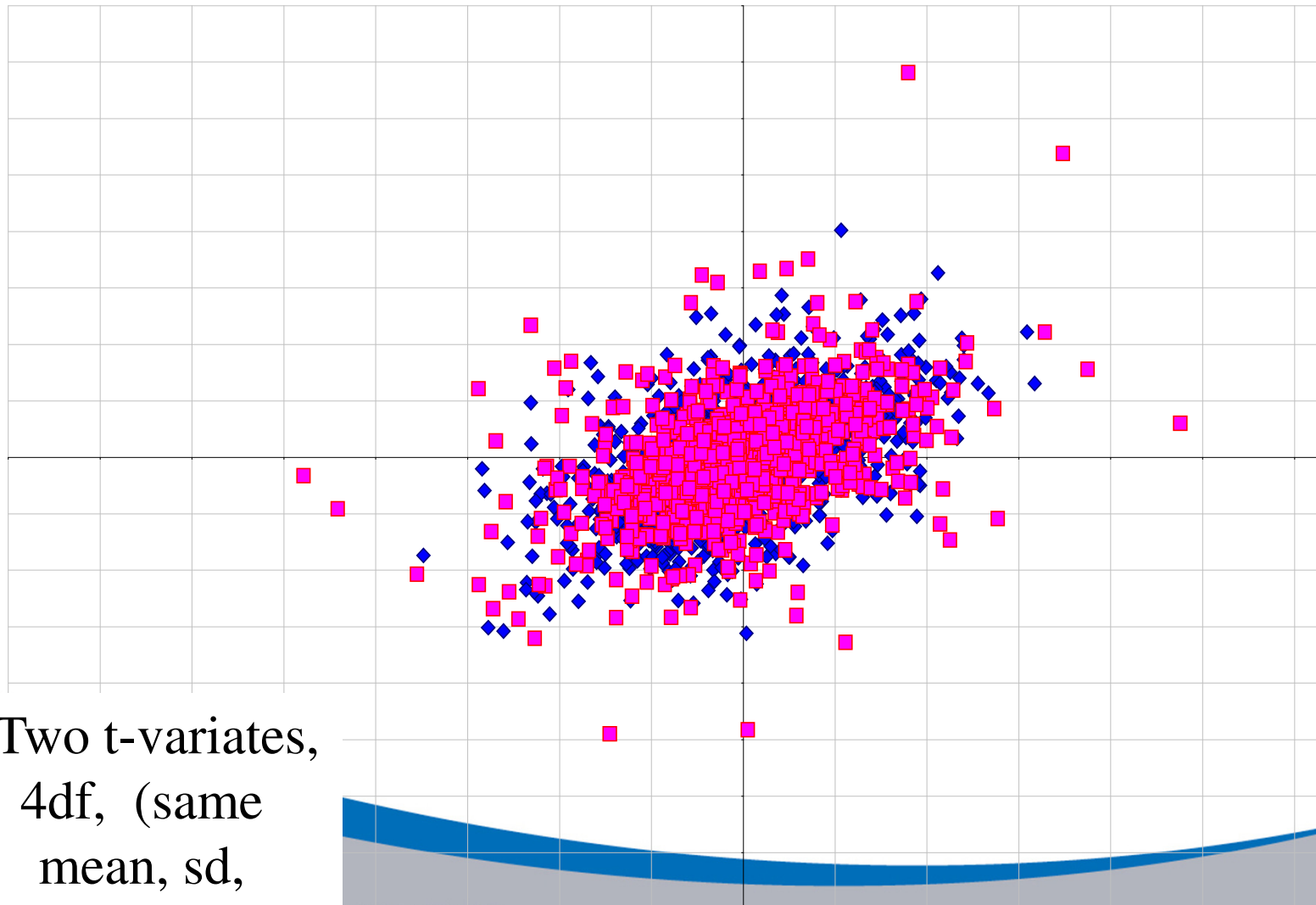
Where might copulas be important? ...

Tail dependence

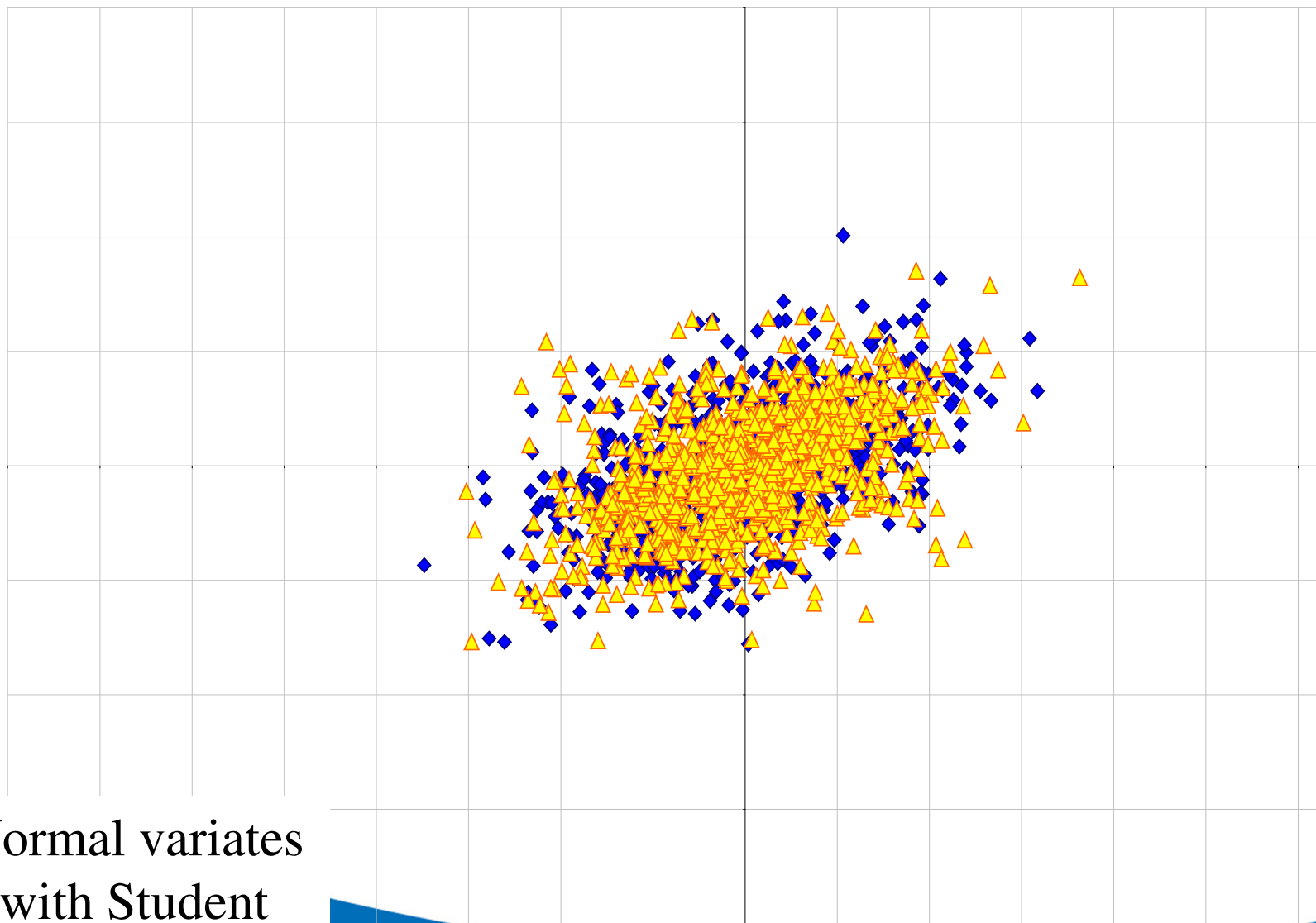
- Definition:
- $\lim_{u \rightarrow 1} P(Y > G^{-1}(u) \mid X > F^{-1}(u)) = \lambda$
 - As we get further into the tail of one distribution we know something about the likelihood of being in the tail of the second distribution
- Multivariate Normal $\rho < 1 \Rightarrow \lambda = 0$ – no tail dependence
- Multivariate t $\rho > -1 \Rightarrow \lambda > 0$ - tail dependence in all cases
- Fits stylised facts of bear markets
- Tail dependence is a copula property



Two Standard
Normal variates,
correlation =



+Two t-variates,
4df, (same
mean, sd,
correlation)



Normal variates
with Student
copula

Gaussian vs t-copula

- The spreadsheet *cop – Normal t copula 2.xls* enables you to examine the differences between a Gaussian and t copula

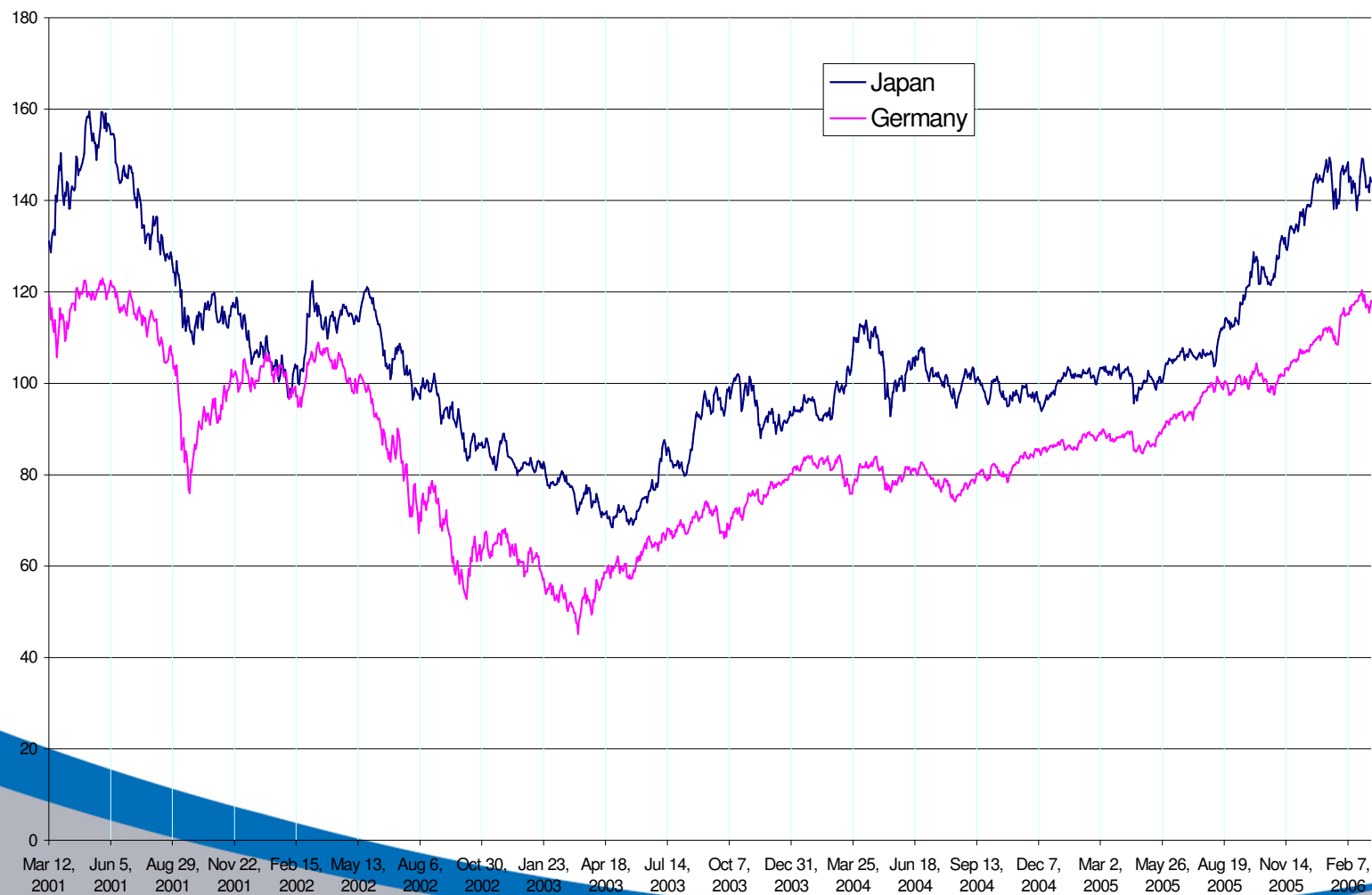
The Formula that Killed Wall Street ?

Correlation - Longer time periods

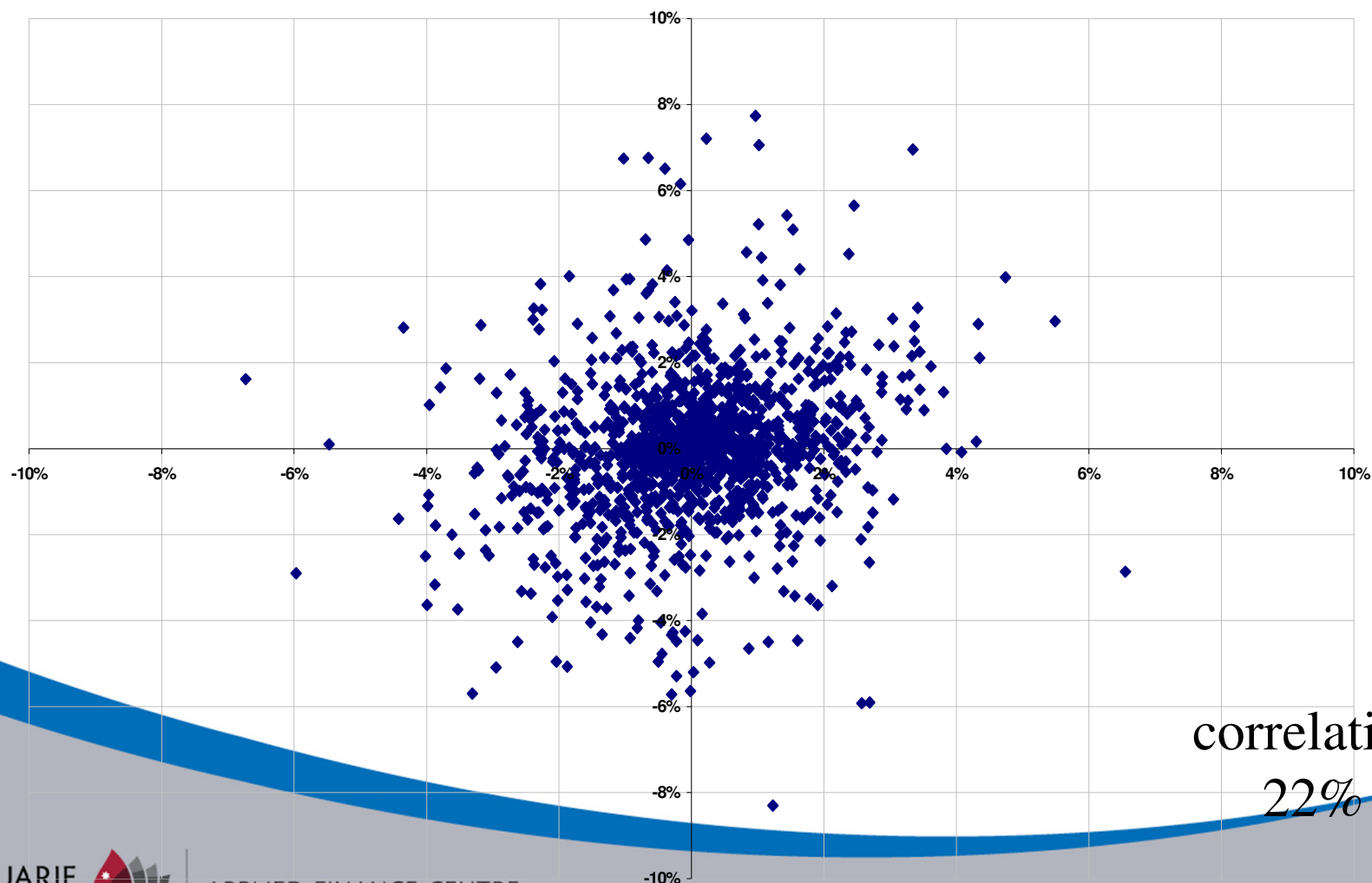
- Volatility depends on the holding period
- Does correlation depend too?

Relationship of long and short intervals

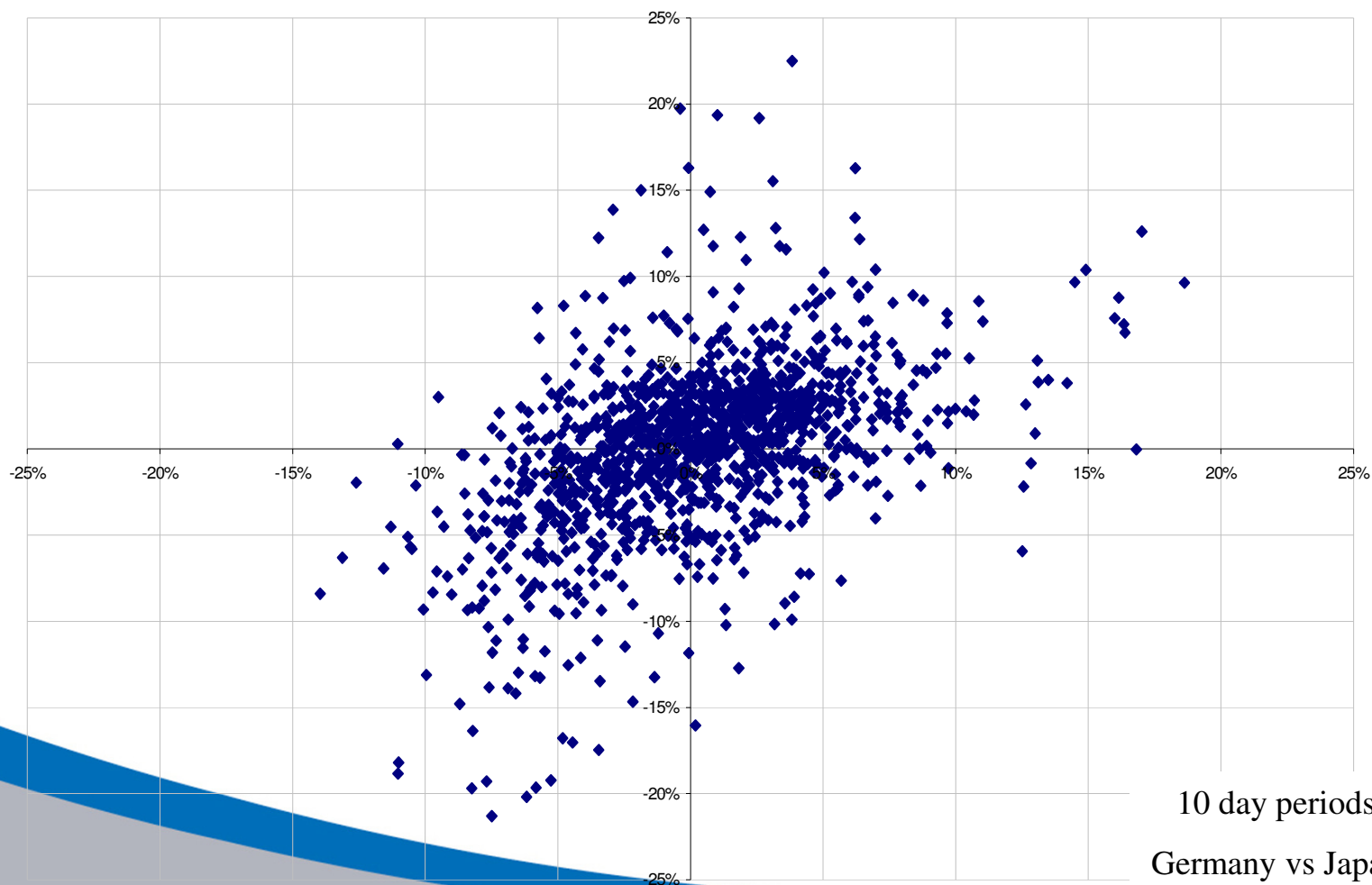
- Correlations (Pearson) change as we change the period of measurement
 - 1 day 22%
 - 10 day 50%
 - 62 day 62%
-
- Choose the interval of measurement to be similar to the interval of risk estimation and capital requirement



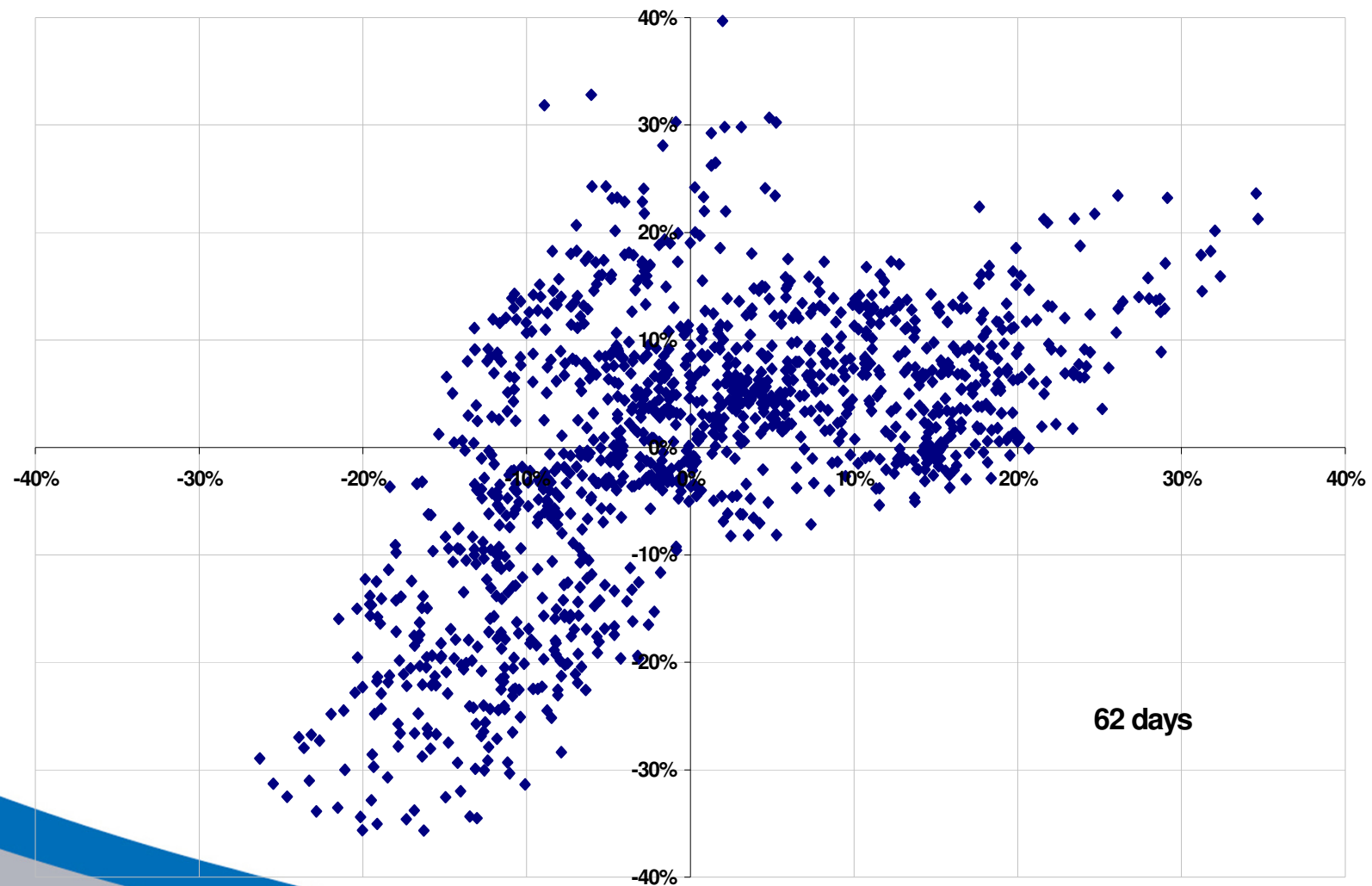
Scatter chart of daily log returns



correlation
22%



10 day periods
Germany vs Japan

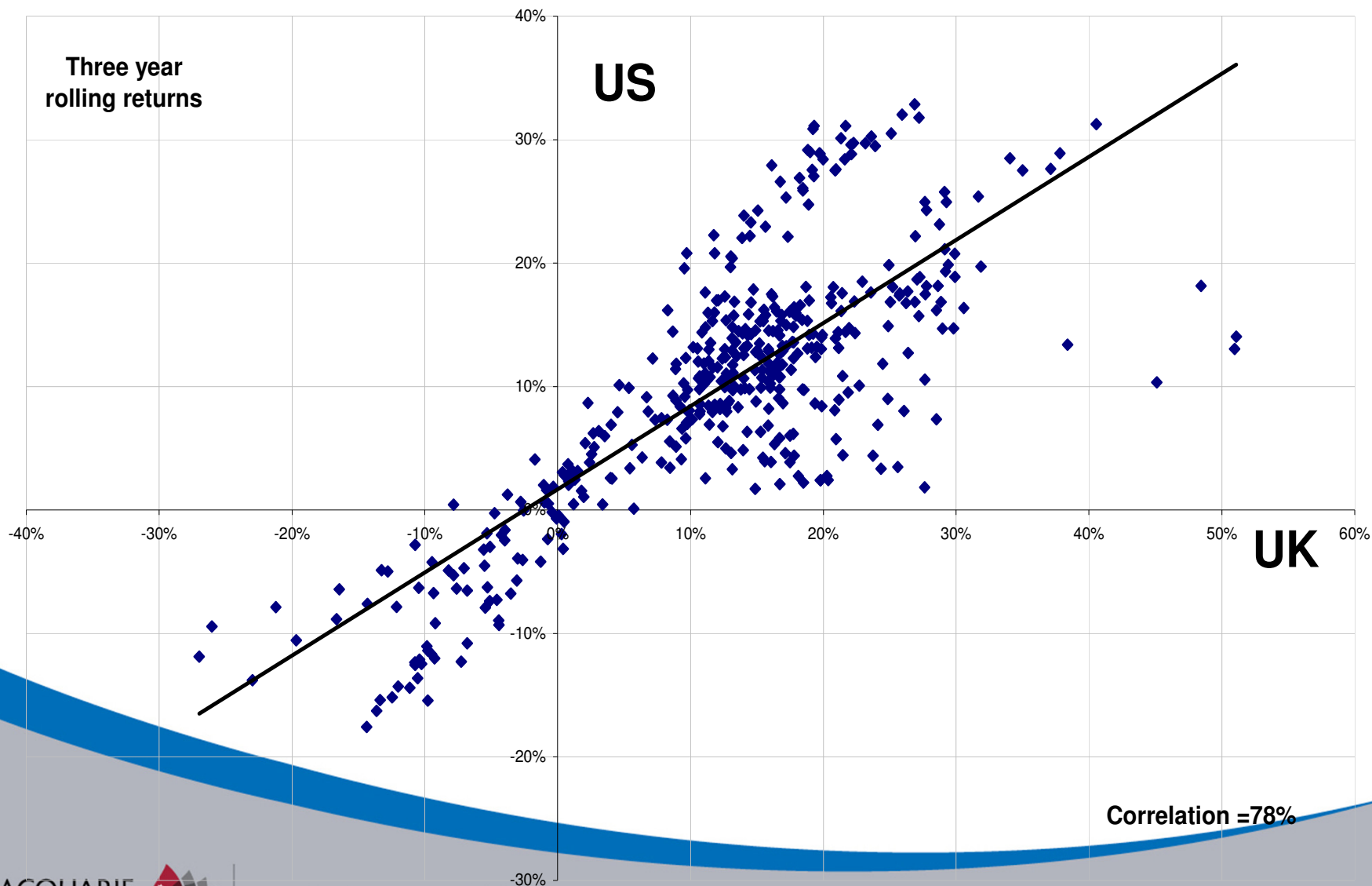


Three year
rolling returns

US

UK

Correlation = 78%



Simulations

- Monte Carlo methods
- Generation of random numbers
- Sobol sequences.
- Variance reduction.
- Importance sampling.

Generation of random numbers

- Pseudo-random numbers
 - Can be generated in a deterministic sequence from an initial seed
 - Allows replication of some results when model parameters are changed
 - Using the same random number sequence helps reduce one source of difference
 - Concentrate on effect of model parameters
- The numbers aren't random!
 - Some non-random effects can occur
 - Sequence eventually recurs
 - Poor dimensional distribution
 - Correlation of successive values

Don't
worry !

Variance reduction techniques

- Antithetic variates
 - Whenever we generate one random number e.g. 0.2, simulate the same variable using $1-0.2$ as well.
 - Only works where the result wanted, say X , is correlated with the random variate (a common case)

Variance reduction techniques (ii)

- Control variables
 - If there are variables we can calculate from the MC for which we know the values (expected return of an asset should equal the input value) then we can adjust the value of X .
 - Simplistically, find a regression equation of X as a function of F , and plug the correct value of F into the regression.

Variance reduction techniques (iii)

- Importance sampling
 - Choose random variables from a part of the distribution that have a large impact on the variable X .
 - Correct for having a non-random sample
- Quasi-random numbers
 - There are some sequences of numbers which appear random but “fill up” the $[0-1]^n$ space leaving no gaps.
 - Sobol sequences are an example

Monte Carlo in Regulation

- Summarised recently in Nicholls and Skinner
– IAAust Biennial Convention 2007,
Christchurch

Firstly, regulators and rating agencies now consider capital modelling, when incorporated into a holistic enterprise risk management (ERM) framework, essential for a well run insurance company.

...

For larger companies, or where the interactions of risks is more complicated, simulation based stochastic capital models become vital.



Putting it together



Frequency and severity for premium calculation

- Compound Poisson is textbook solution
- What needs to be considered with new techniques?
- Can we model extreme tail of claims distribution using EVT?
- Is Poisson rate constant? Maybe model rate like volatility?

Copula?

- Usual assumption is independence
- How sensitive is solution to other copulas, even if we keep correlation = 0?
- Interaction of asset distribution and claim distribution

Bootstrapping

- How sensitive is estimate of Poisson parameter to the sample?
- What about estimation of claim distribution, especially tail?
- Asset return distribution has vaguely known parameters in it e.g. ERP

Plenty to think about

- Questions
- Discussion