Risk and Portfolio Management NYU Spring Semester 2013

Marco Avellaneda

1. Equities and Equity Derivatives

The Equities Universe

- Single-stock shares
- ETFs
- Single share futures
- CFDs
- Index Futures
- Options on single-names
- Options on ETFs
- Options on Indices

- Variance Swaps (Index)
- Dividend Swaps (SN, Index)
- Long-dated options

Risk Factors

- Share price, Index price (spot factors)
- Implied volatility (surface factors)
- Dividend Yield (curve factor)
- Interest Rate (curve factor)

These are the main factors which affect the price changes of Equities and Equity Derivatives.

The market risk of a portfolio of Equity / Derivatives is determined by extreme changes in these parameters

What do we actually know about the statistics of stock returns?

Stock returns exhibit <u>heavy tails:</u>

-- Small moves are more frequent (likely) than predicted by Gaussian PDF

-- Large moves are more likely than predicted by Gaussian distribution

<u>Validation of this statement:</u>

-- Consider a large cross-section of the US stock market (~3000 stocks)

-- Data consists of 1 year worth of data on ~ 3000 stocks

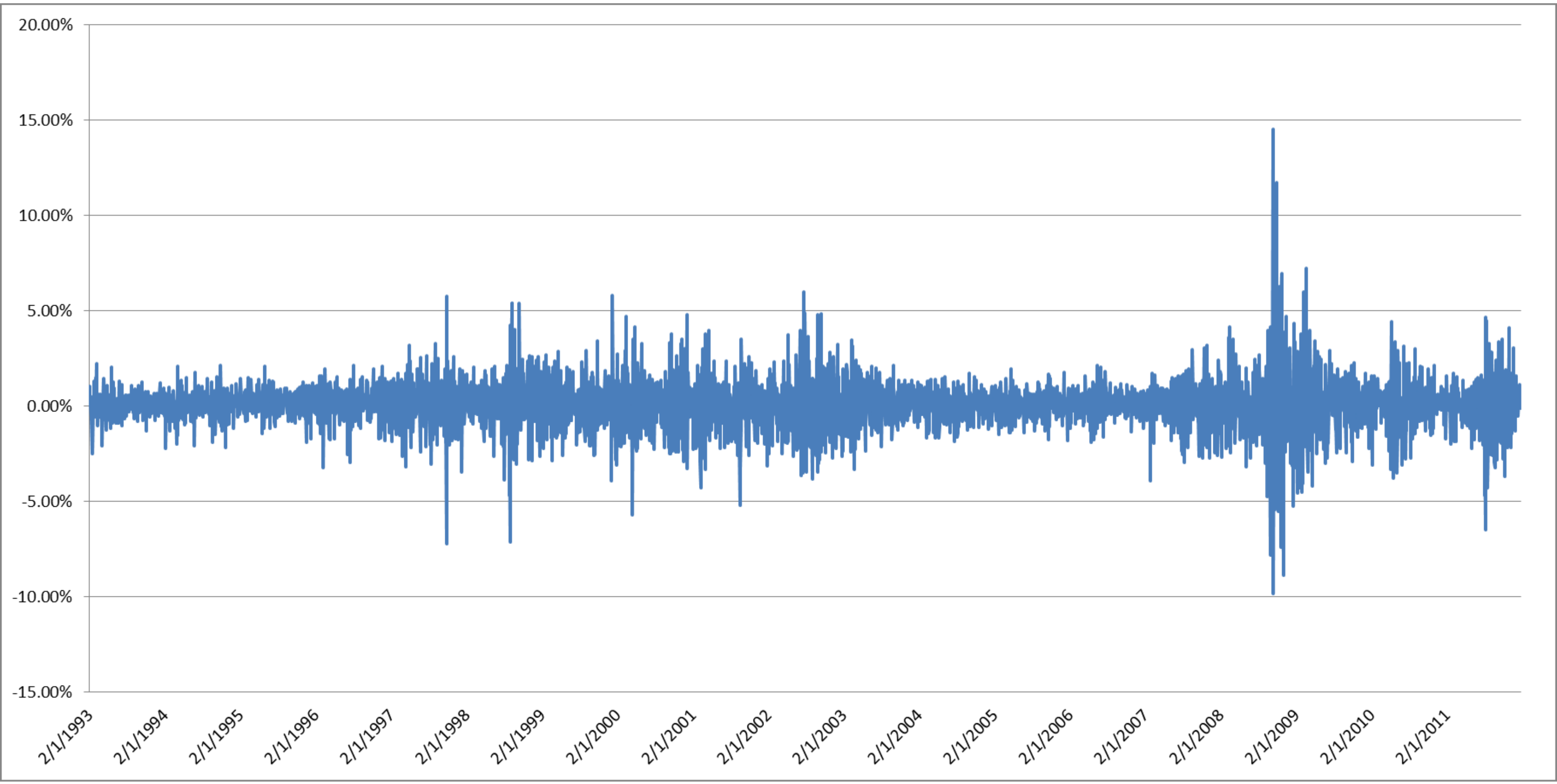
-- Standardized stock returns over T days (e.g. T=1) and fit to various probability distributions

Stylized facts of asset returns

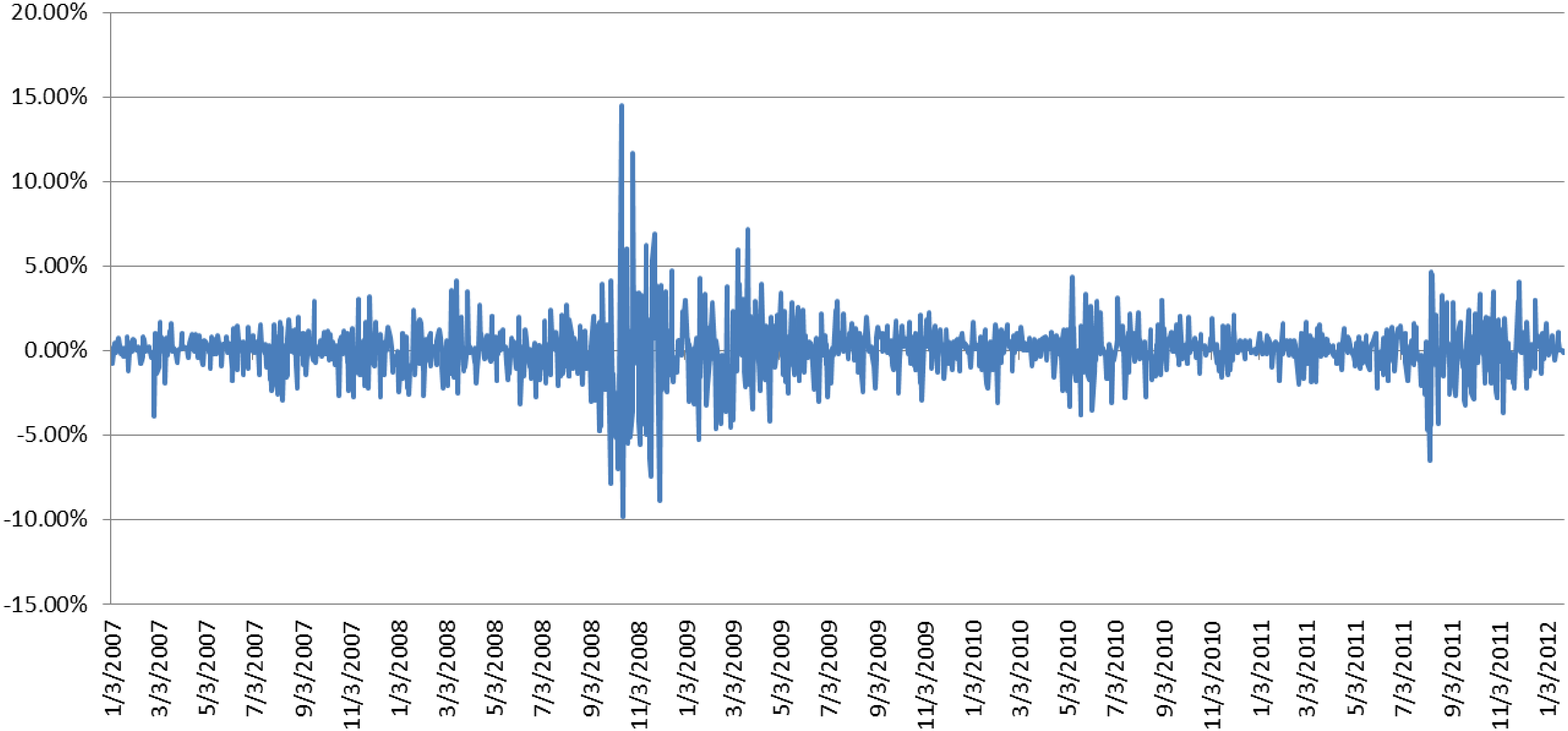
Rama Cont (2000): Empirical properties of asset returns: stylized facts and statistical issues, *Quantitative Finance.*

- Absence of serial autocorrelations
- Heavy tails
- Gain/loss assymetry
- Aggregational Gaussianity
- Intermittency (tendency to have spikes)
- Volatility Clustering
- Conditional heavy tails
- Slow decay of correlations in absolute returns
- Leverage effect
- Volume/volatility correlation
- Assymetry in time scales, for volatility estimation

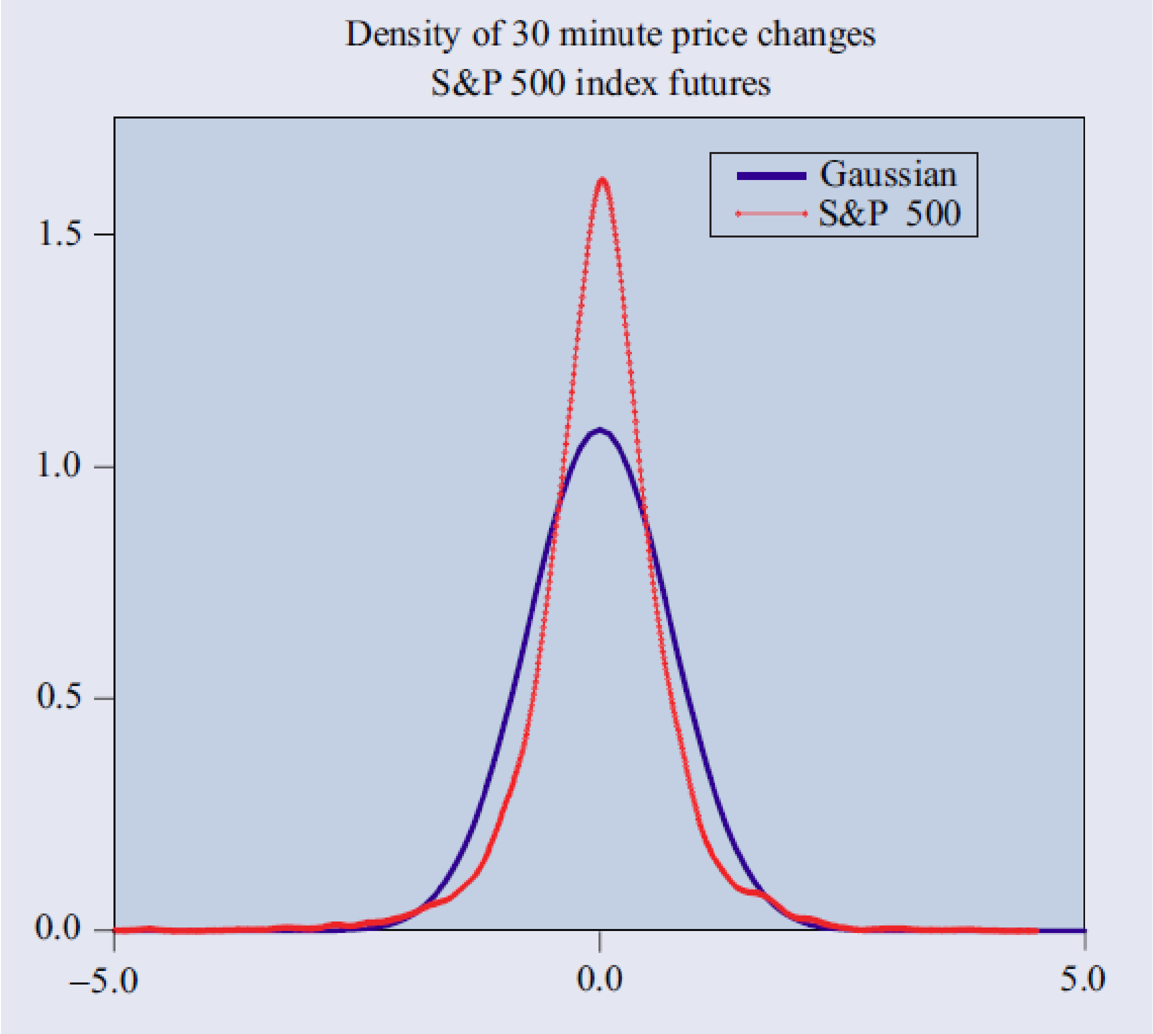
Returns of S&P 500
Feb 2, 1993 to Jan 26, 2012



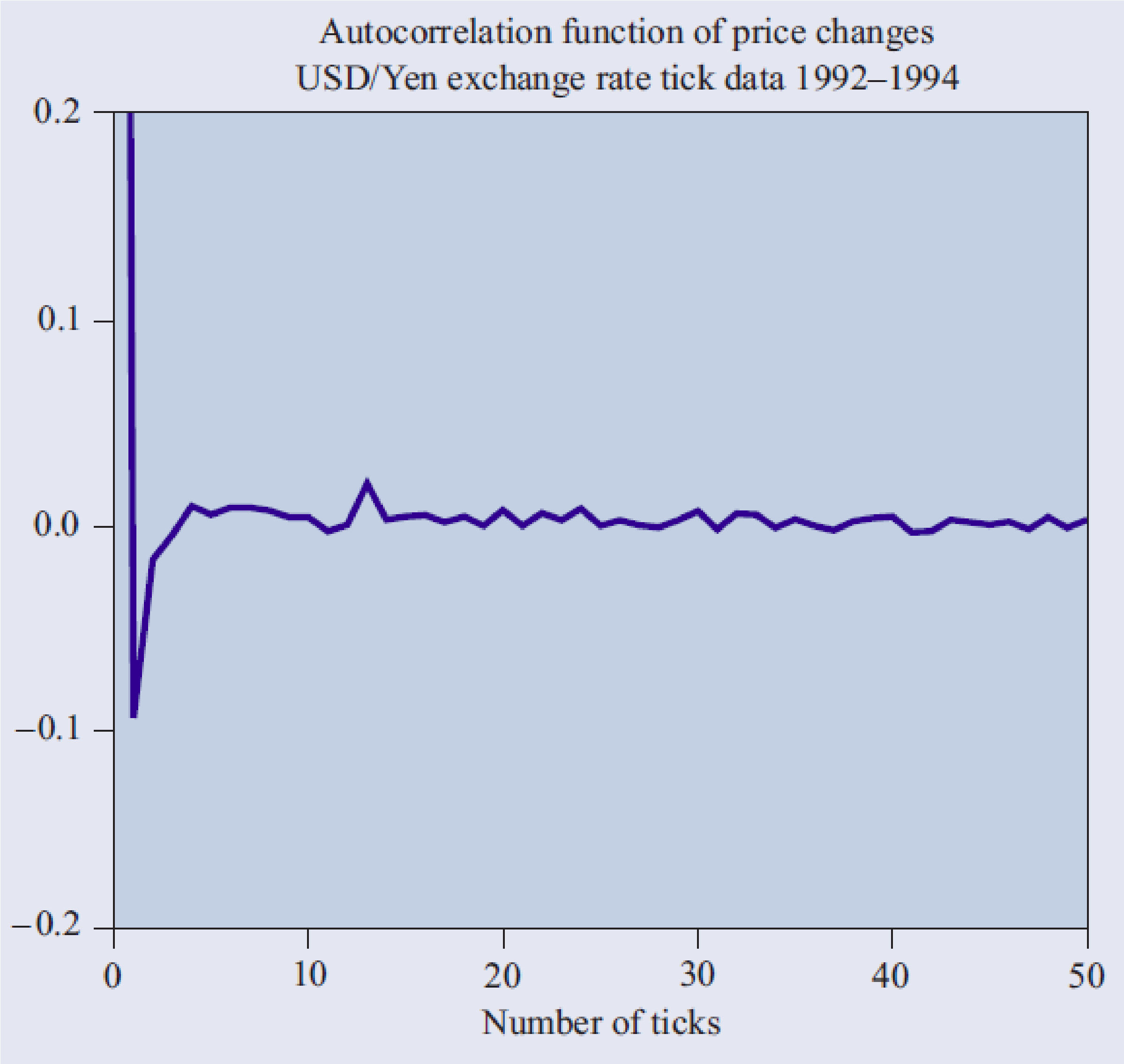
SPY Returns, since Jan 3,2007



Normalized probability density is non-Gaussian



Autocorrelation function USD/JPY



Heavy Tails for Individual Stocks

Stock returns exhibit heavy tails:

-- Small moves are more frequent (likely) than predicted by Gaussian PDF

-- Large moves are more likely than predicted by Gaussian distribution

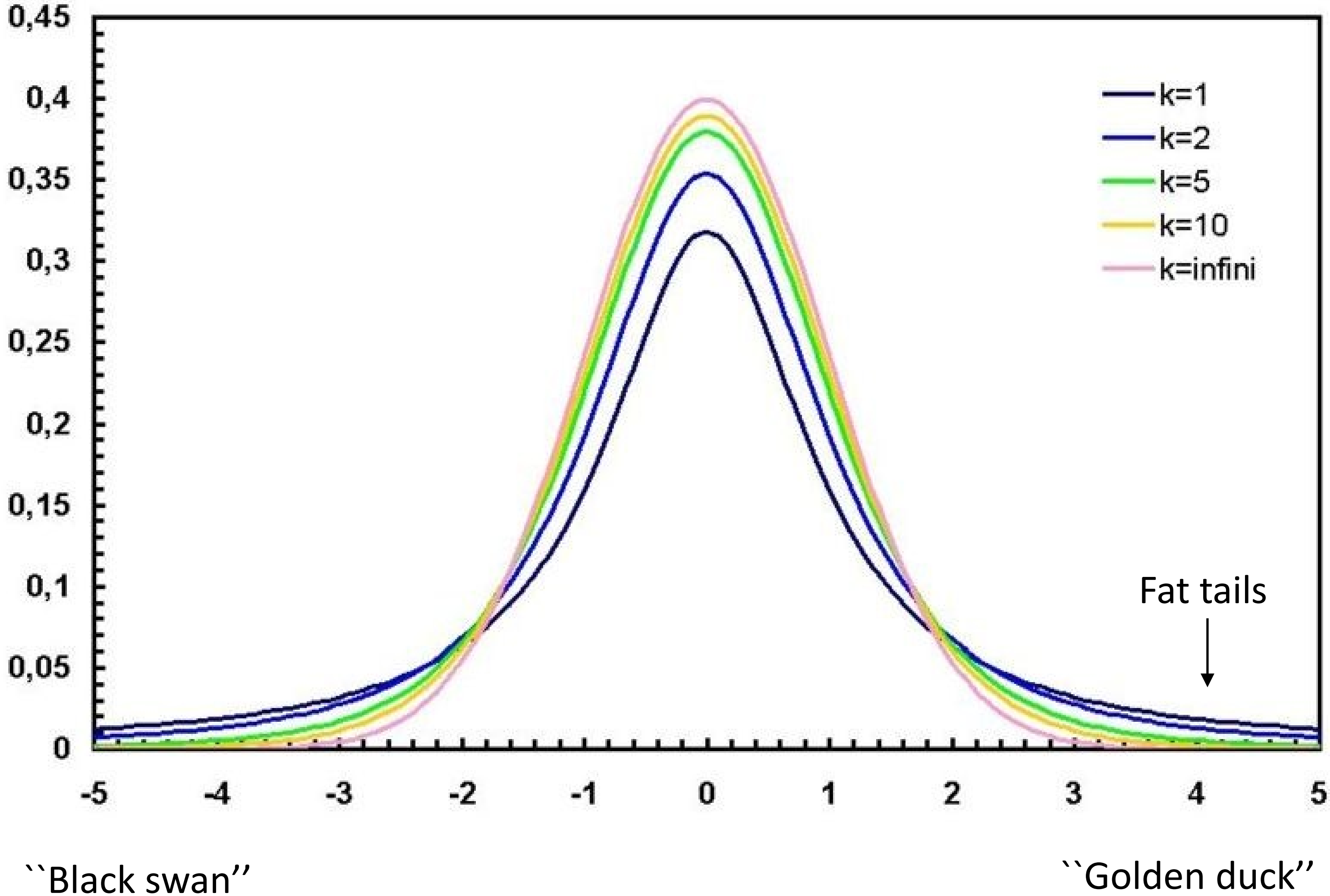
Validation of this statement:

-- Consider a large cross-section of the US stock market (~3000 stocks)

-- Data consists of 1 year worth of data on ~ 3000 stocks

-- Standardized stock returns over T days (e.g. T=1) and fit to various probability distributions

The Student-t family of distributions



Statistical model with fat tails to account for extreme moves

$$f(x) = \frac{C}{\left(1 + \frac{x^2}{k}\right)^{\frac{1+k}{2}}}$$

Student: Power-law tails

$$P\{X > r\} \sim 1/r^k$$

Gauss: Exponential tails
(Gauss ~ k=infinity)

$$g(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

As a general rule, Gaussian or normal distributions are not suitable for financial data due to fat tails

Quantile-to-quantile(QQ) plots

Generate a sample from the unknown distribution and sort it in increasing order

$$X_1, X_2, \ldots, X_N$$

Generate a vector of a known distribution (e.g. Student t)

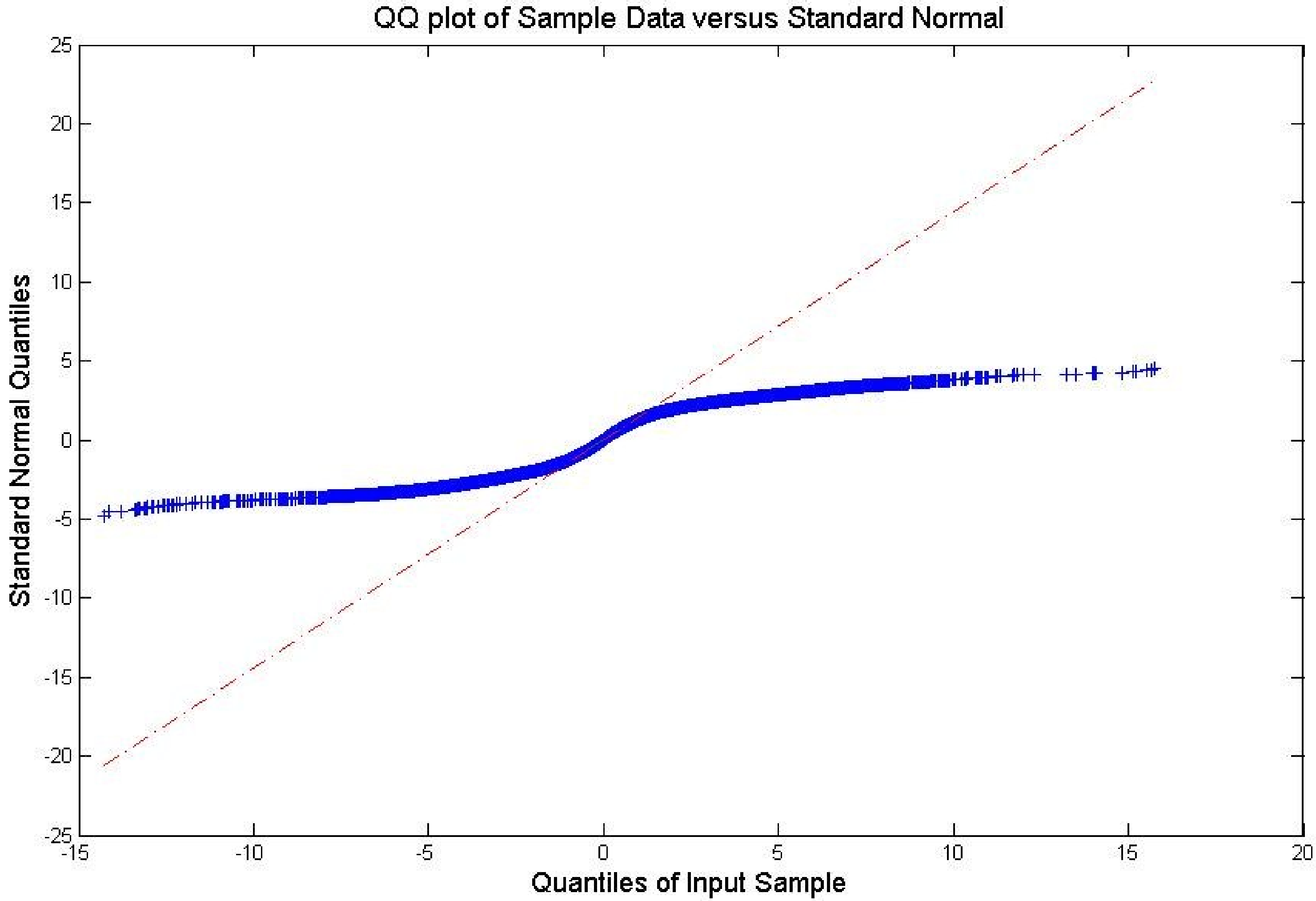
$$Y_1 = F_\alpha^{-1}\left(\frac{1}{N}\right), \ldots, Y_k = F_\alpha^{-1}\left(\frac{k}{N}\right), \ldots, Y_N = F_\alpha^{-1}(1) = \infty$$

Draw an X-Y plot of the ``data'' $\left(X_k, Y_k\right)_{k=1}^{N-1}$

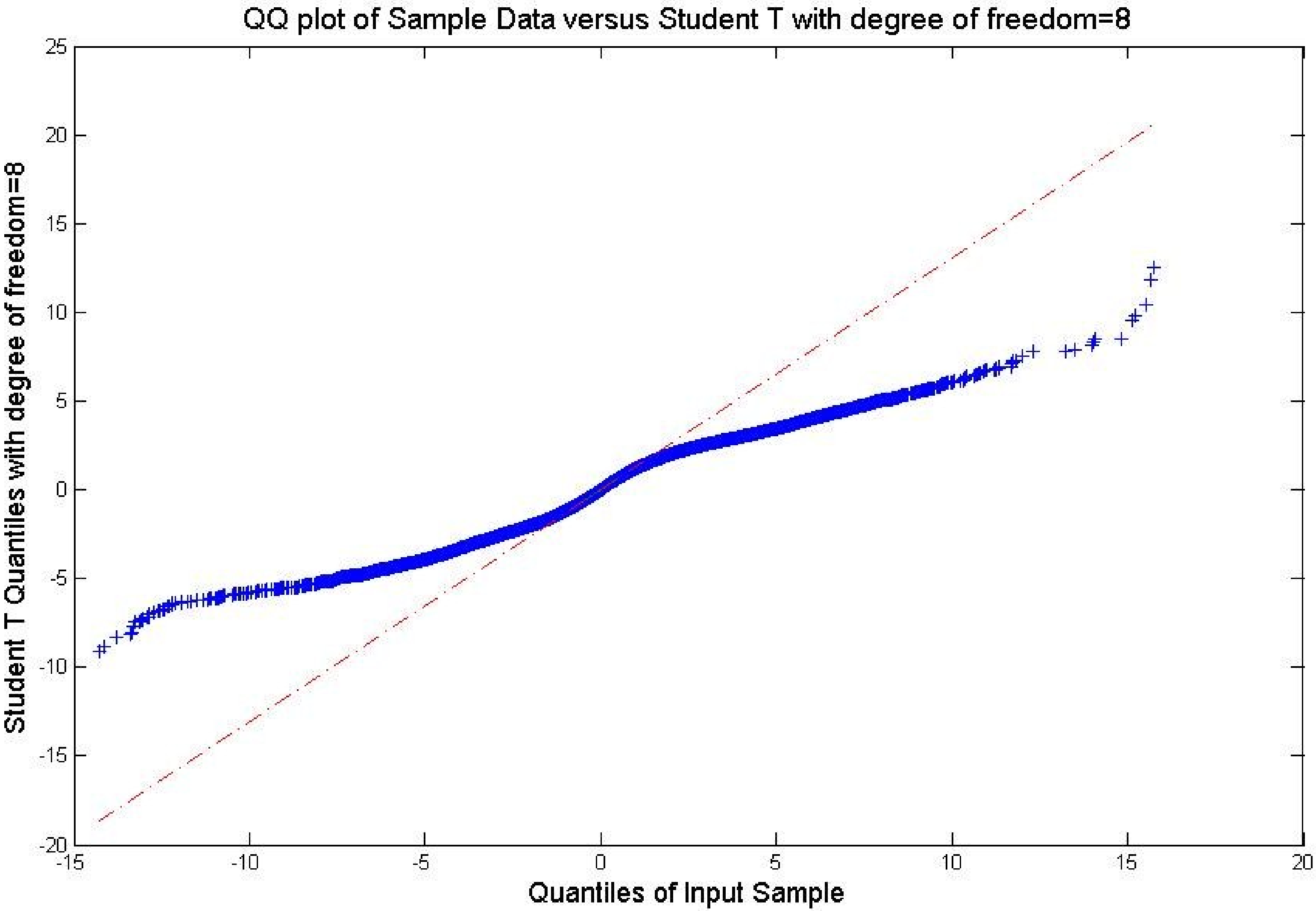
If the sample is drawn from the known distribution, then the points fall approximately on the X=Y line.

This is a good technique to fit tails.

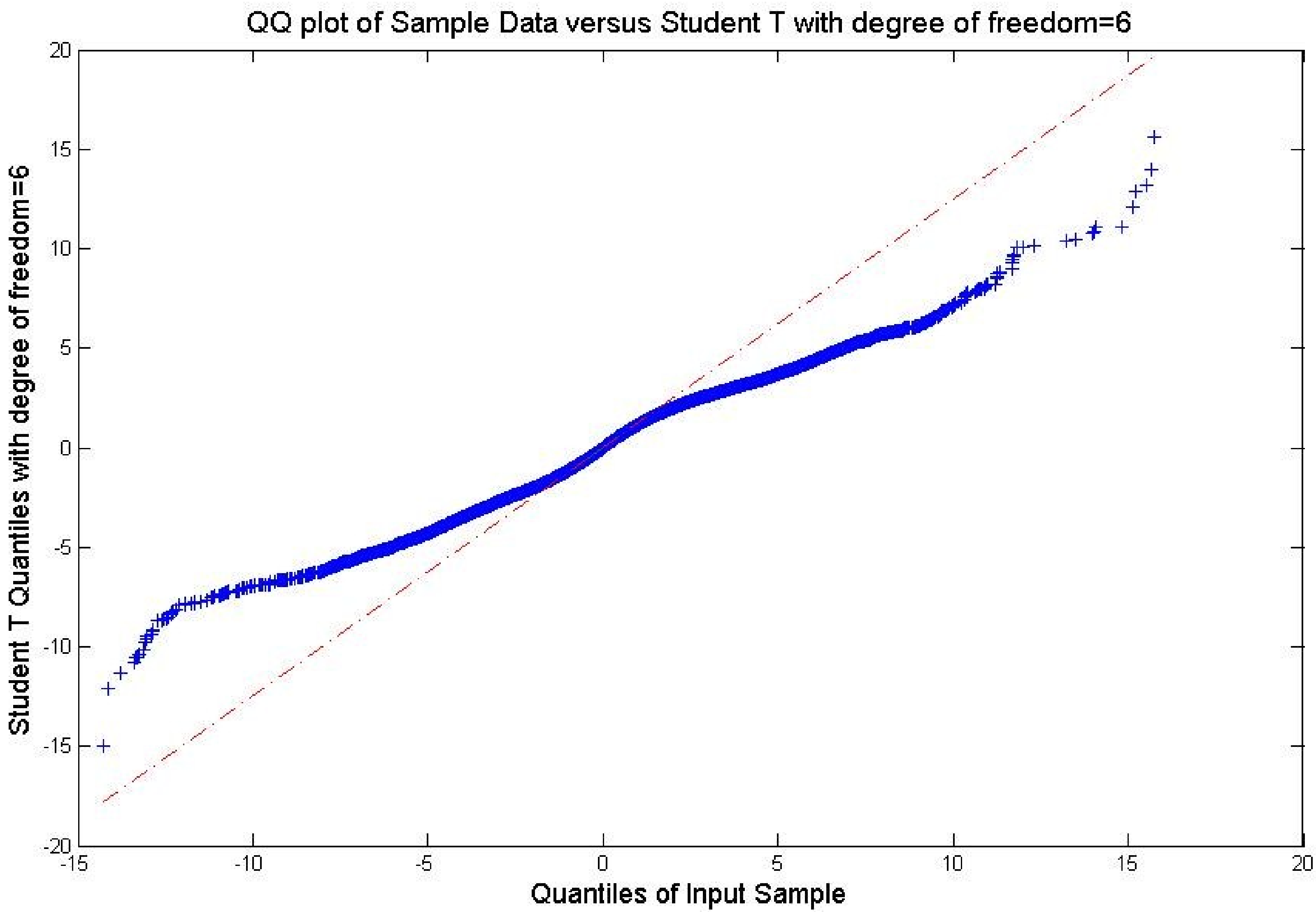
Tails (extreme values) for standardized two-day stock returns: Gaussian fit



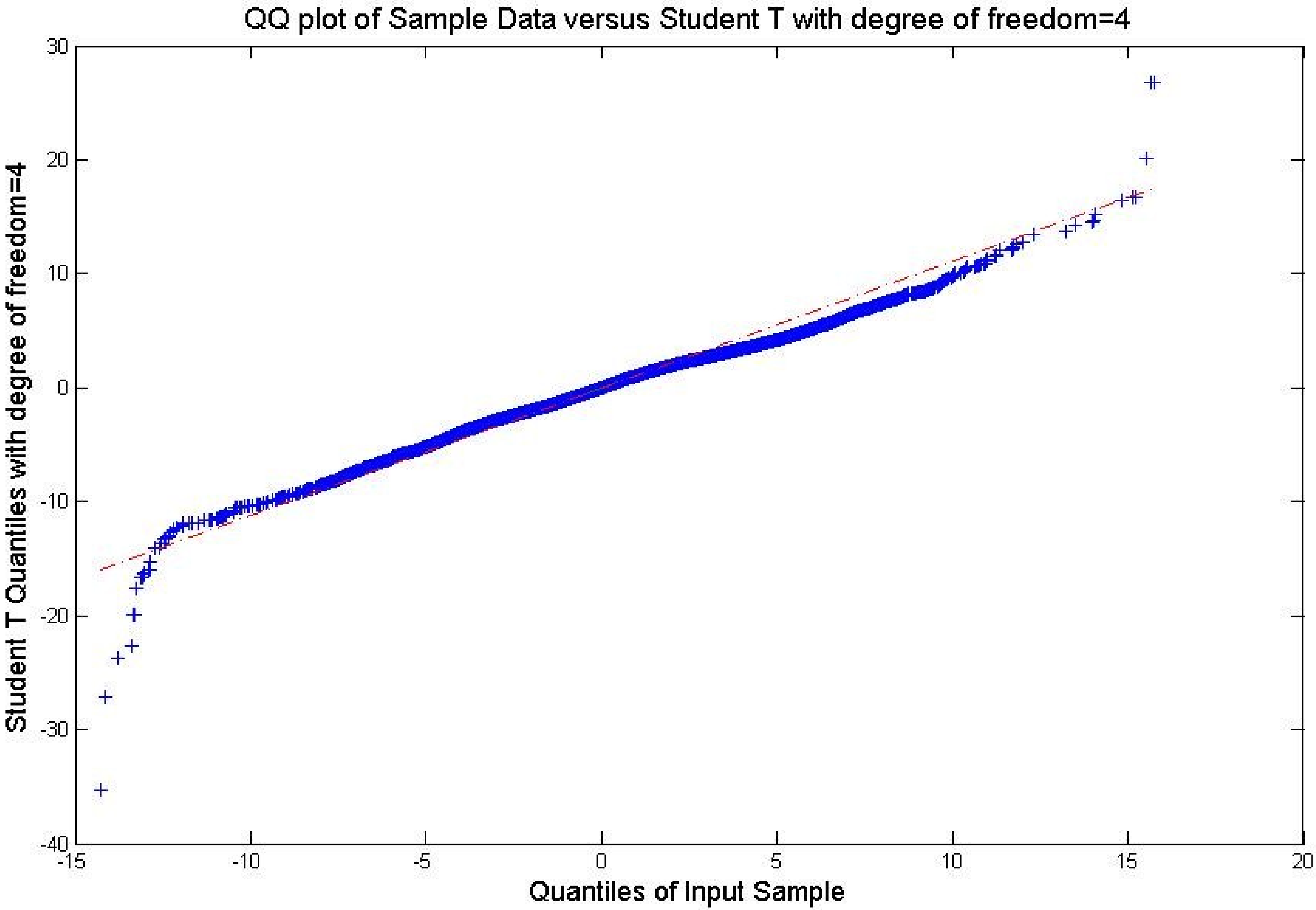
Student-t with 8 df



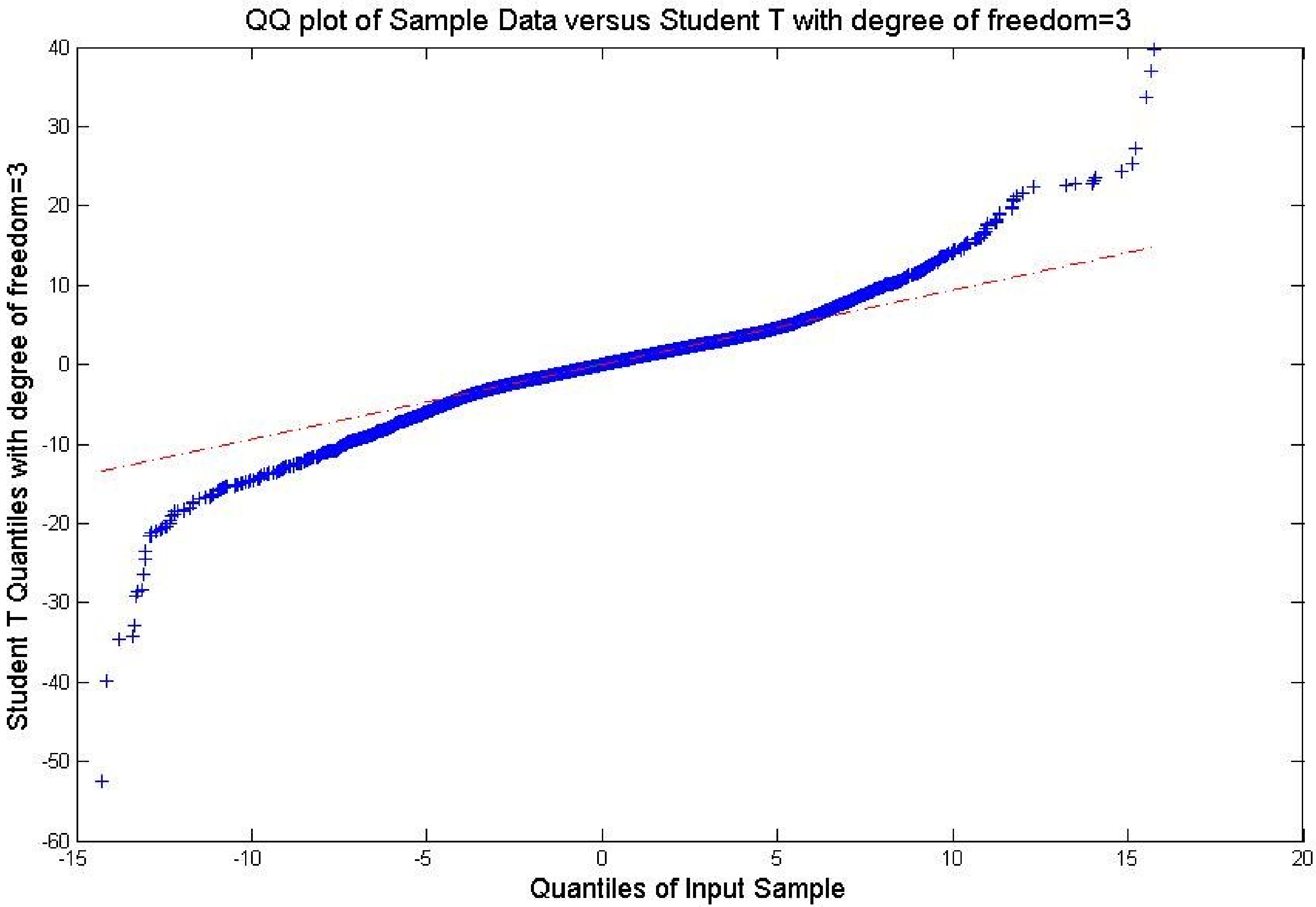
Student-t with 6 df



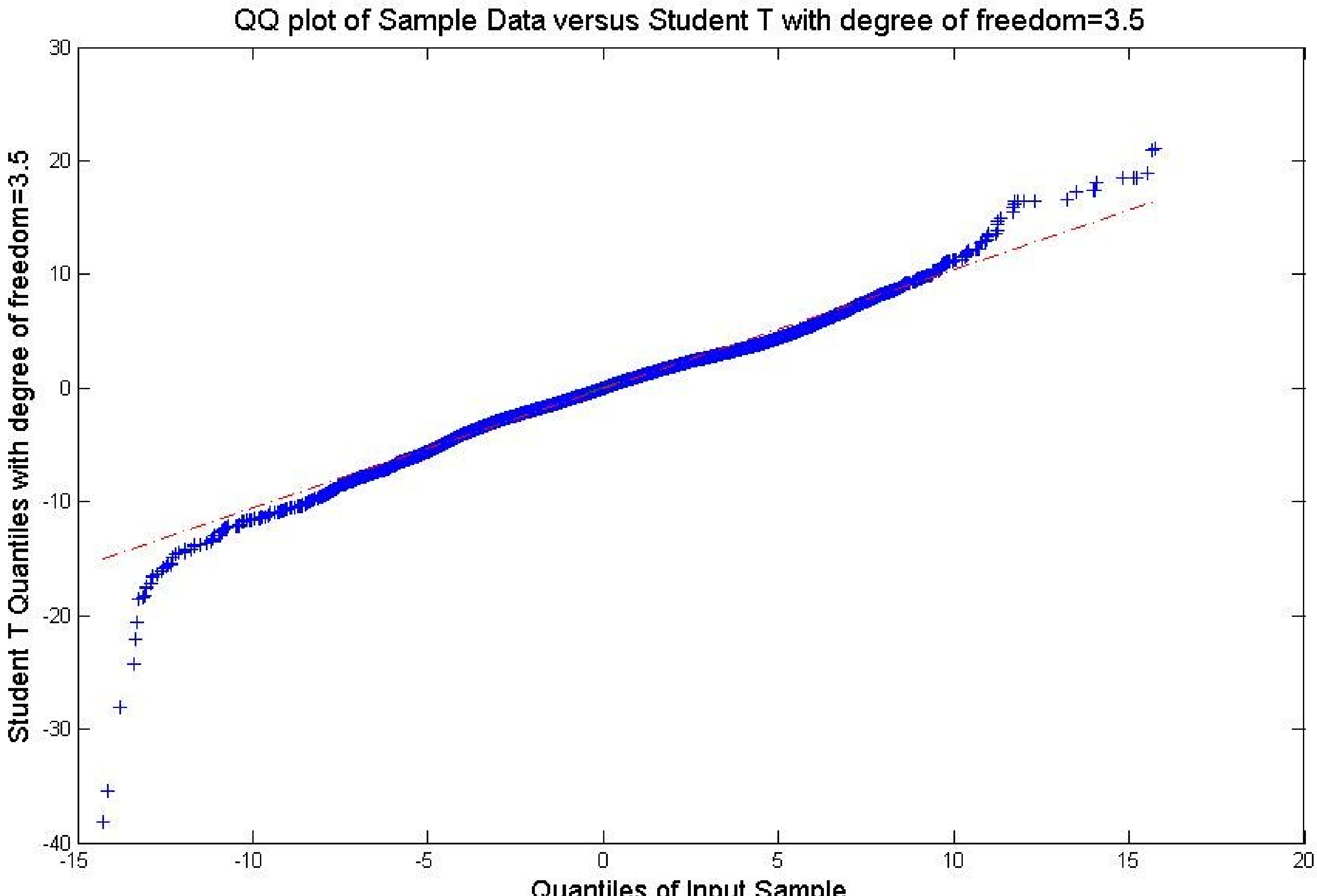
Student-t with 4 df



Student-t with 3 df



Student-t with 3.5 df



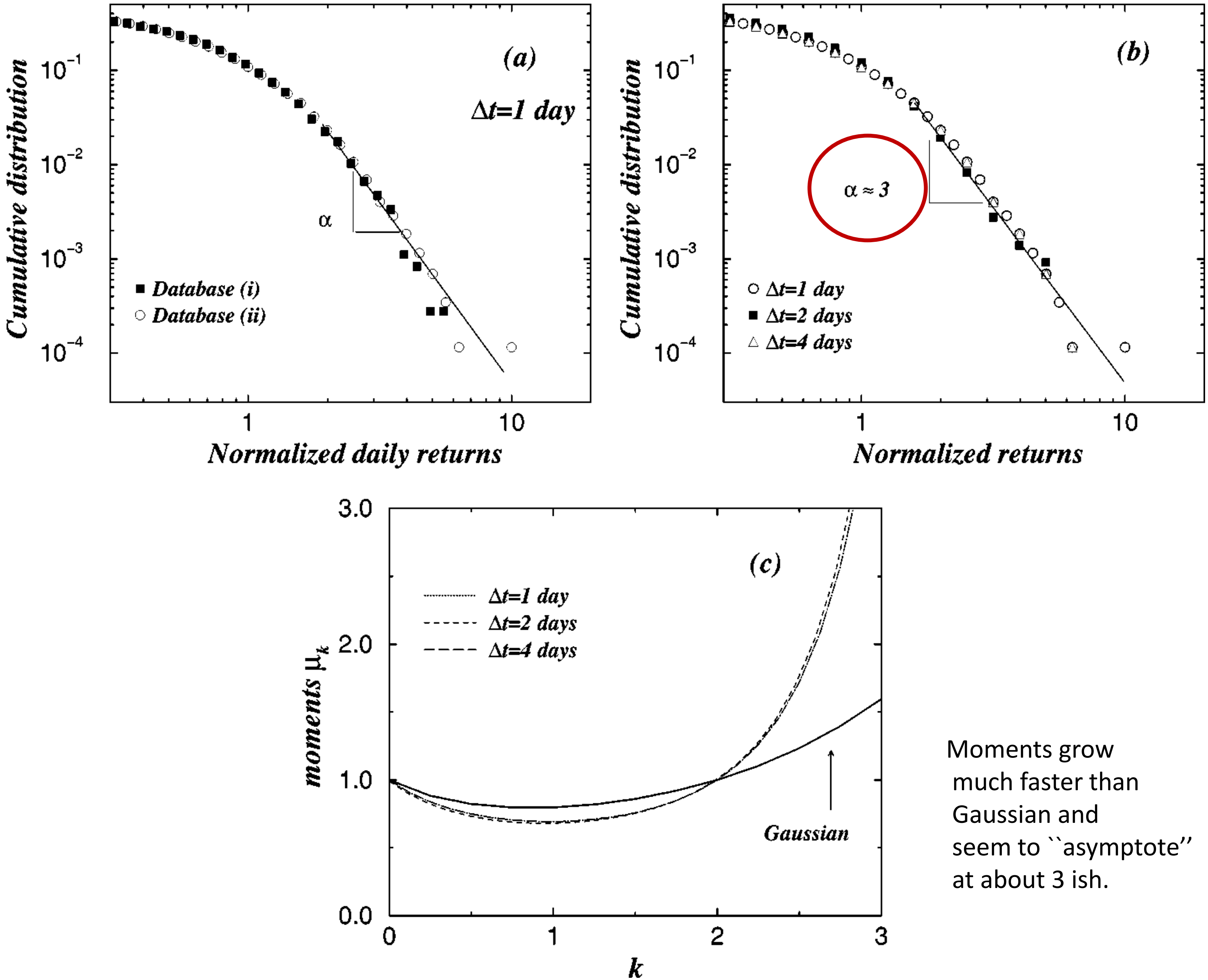
Consistent with classical result from Econophysics

Phys. Rev. E 60, 5305–5316
Scaling of the distribution of fluctuations of financial market indices

Parameswaran Gopikrishnan, Vasiliki Plerou, Luís A. Nunes Amaral, Martin Meyer, and H. Eugene Stanley

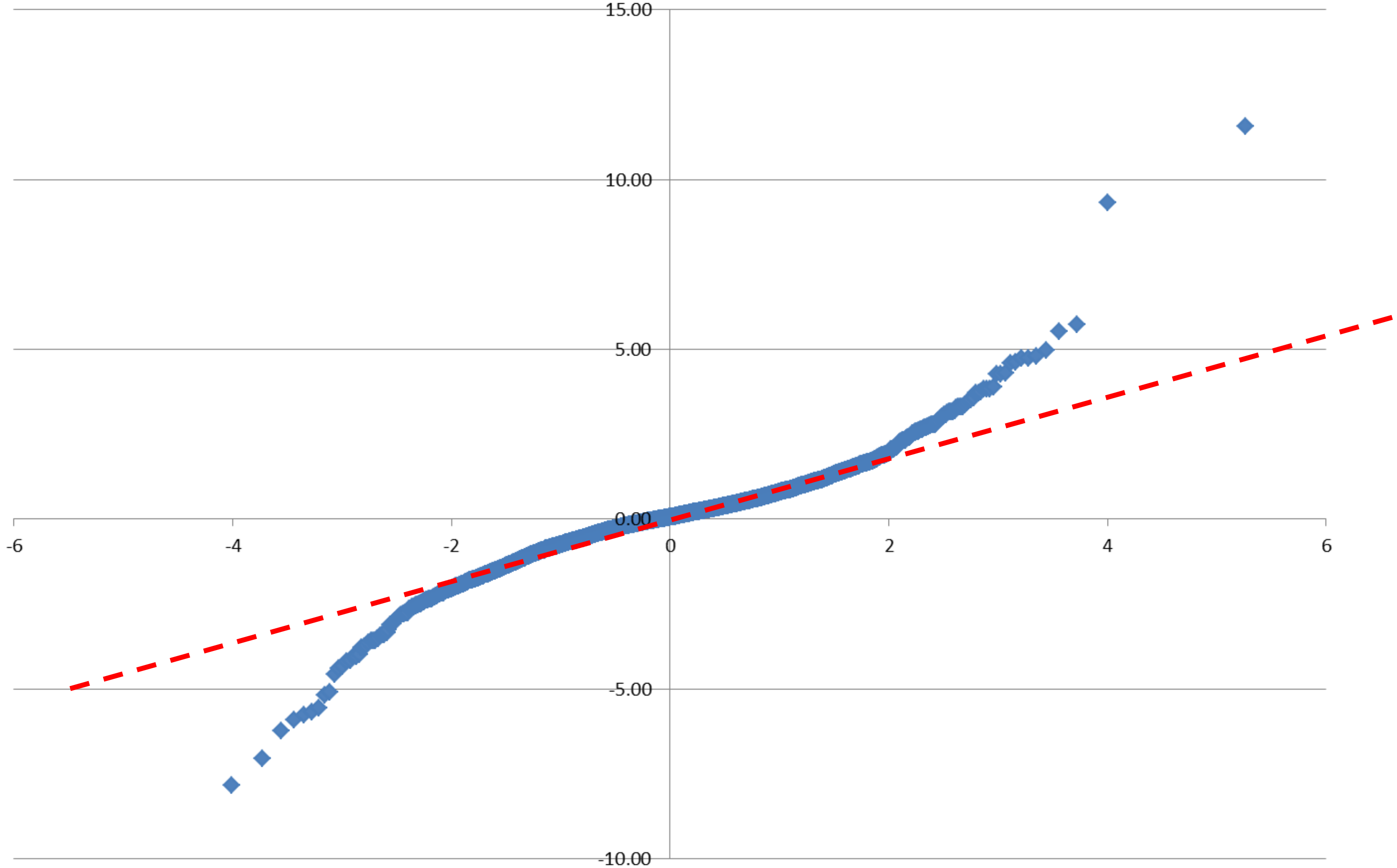
This paper studies equity returns over different time-horizons (1 minute- 1 day) and finds scaling behavior in the probability of large moves (up or down).

The tail exponent is k = 3 or 4 for the physicists, consistently with the data shown here.

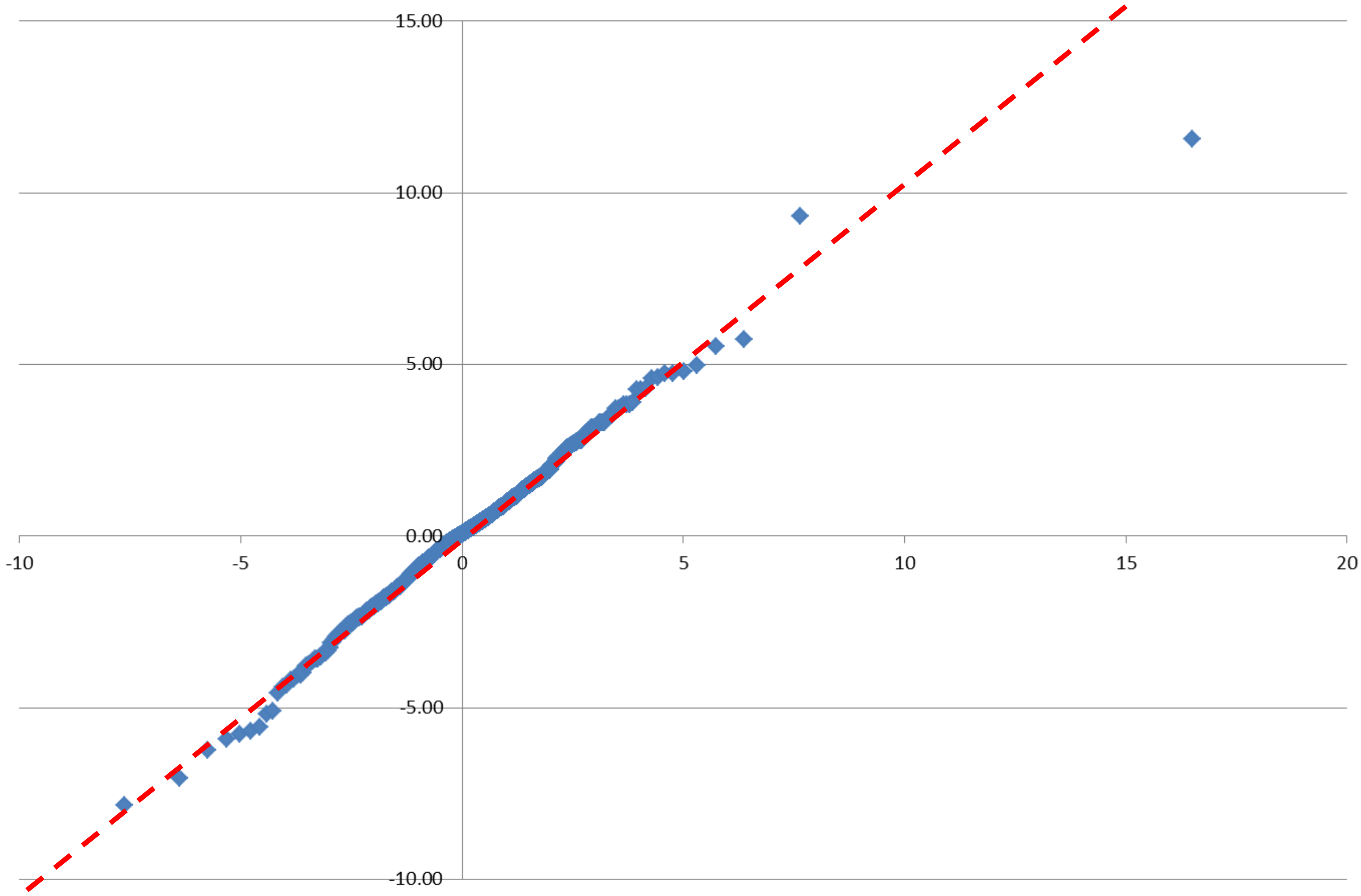


Moments grow much faster than Gaussian and seem to ``asymptote'' at about 3 ish.

S&P 500 Jan 1993-Jan 2012
DF=20, X=Student, Y=Data



S&P 500, Jan 1993 – Jan 2003 (daily returns) X= Student(DF=4), Y=Data

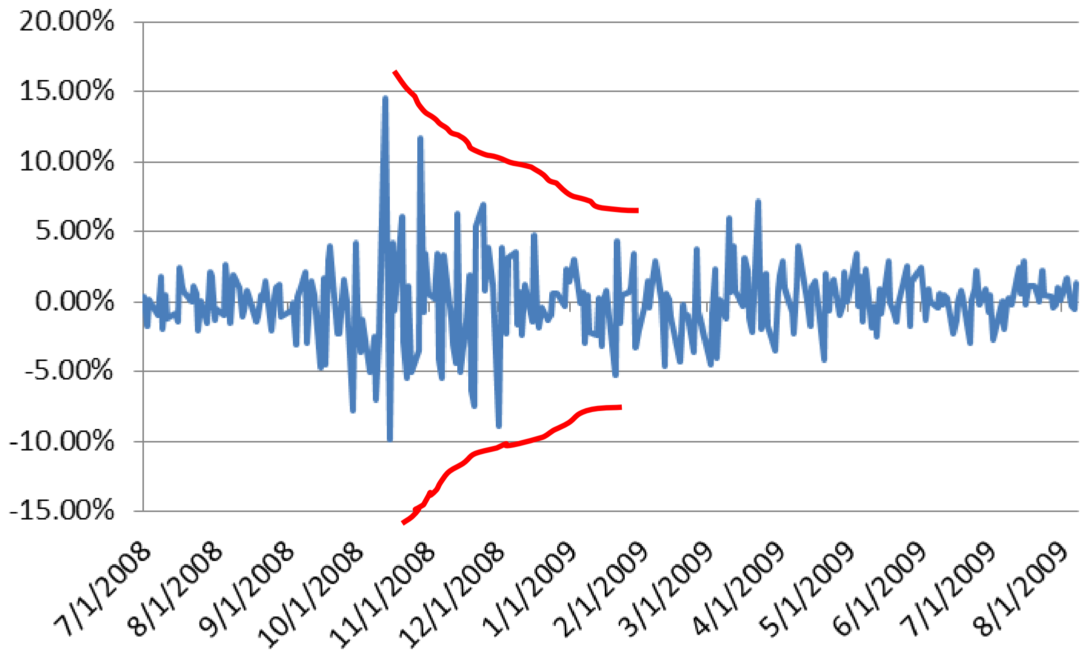


Statistical techniques for heavy tail estimation

- Hill's estimator gives statistical criteria for goodness of fit based on confidence intervals
- Rama Cont's paper has a good account of Extreme Value Theory (EVT) which can be helpful
- The main issue is to fit the tail exponents of power-laws
- Remember that there are very extreme outliers that could sometimes not fit the power-law and these are handled typically by stress-testing

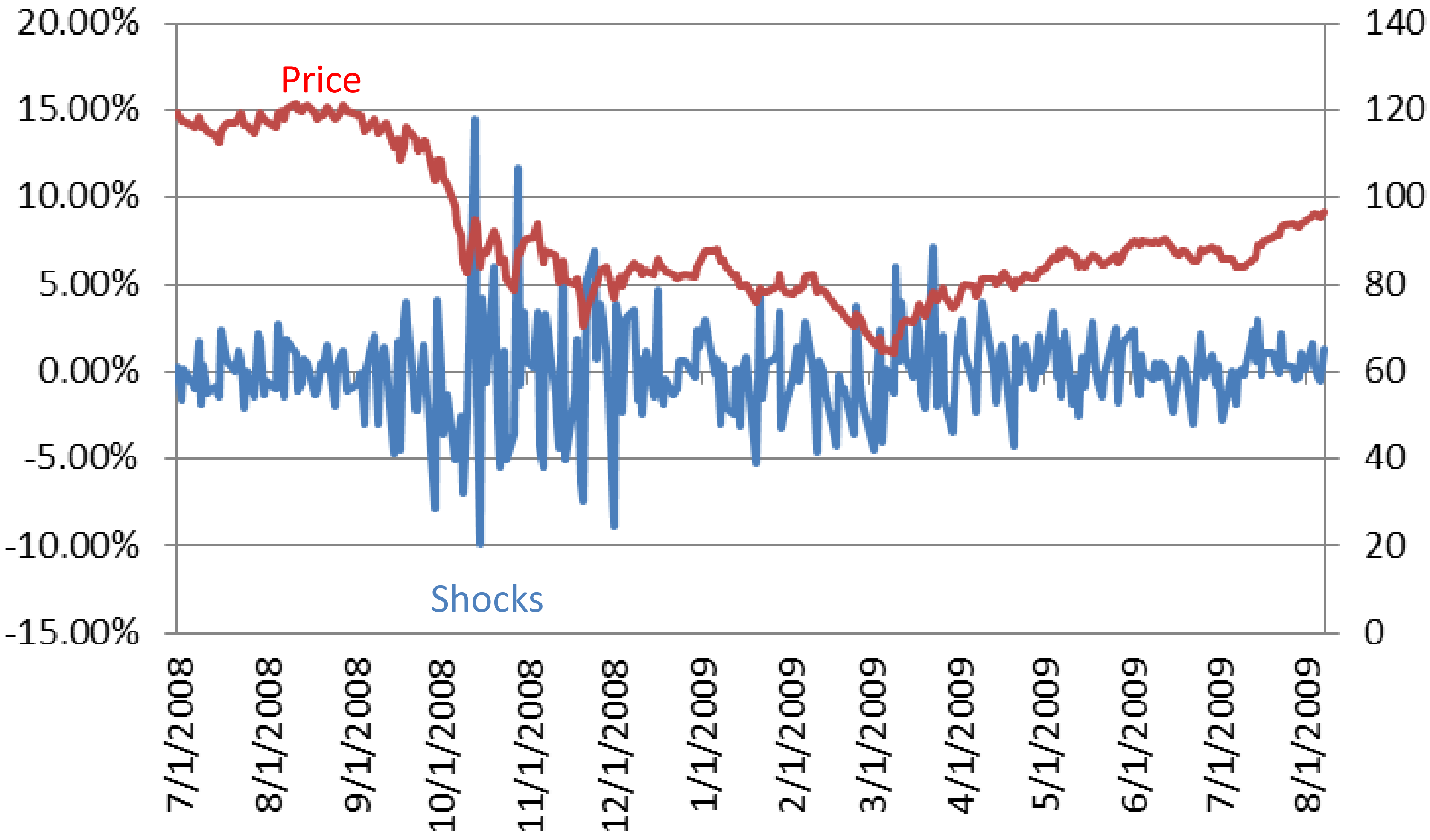
The idea of persistence

- Persistence= autocorrelation of the absolute value of returns
- After a shock to the market, we observe large shocks for a certain period of time, the amplitude of which decays slowly

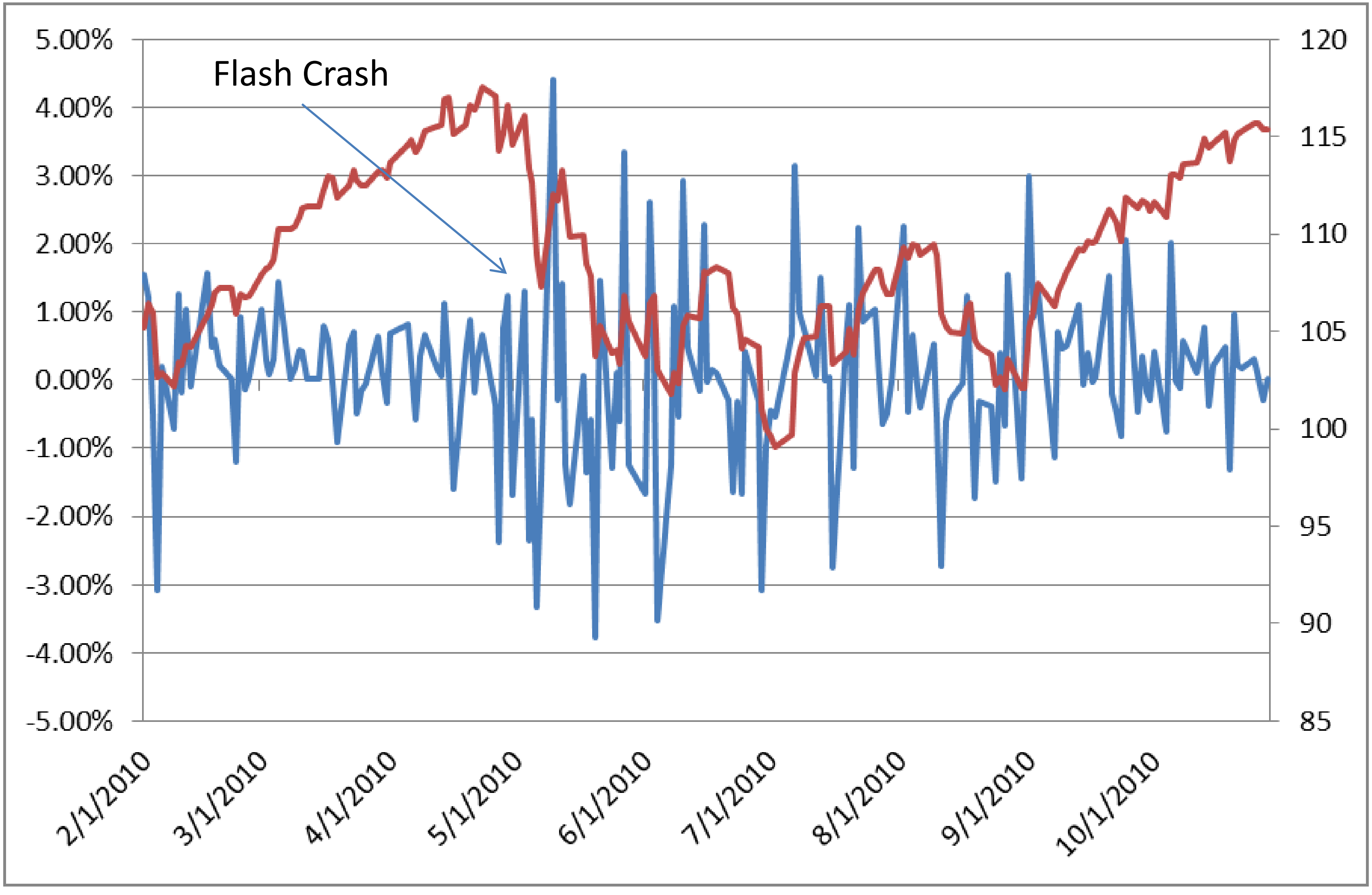


SPY-
post Lehman
shocks

Price level and market shocks post Lehman



``Flash Crash'' and onset of the Greek debt crisis



Risk Management of option portfolios: Modeling the Volatility Risk

1. Compute the historical volatility of a constant maturity series by interpolation over fixed maturities.
   (Typically, for equities: 30d , 60 d, 90 d, 180 d, etc)

2. Express the implied volatilities in terms of moneyness or deltas.
   Deltas is better because this takes into account the volatility of the underlying asset as well.

3. Study the variations of the implied volatility curve for each maturity using PCA & extreme-value theory (Student T)

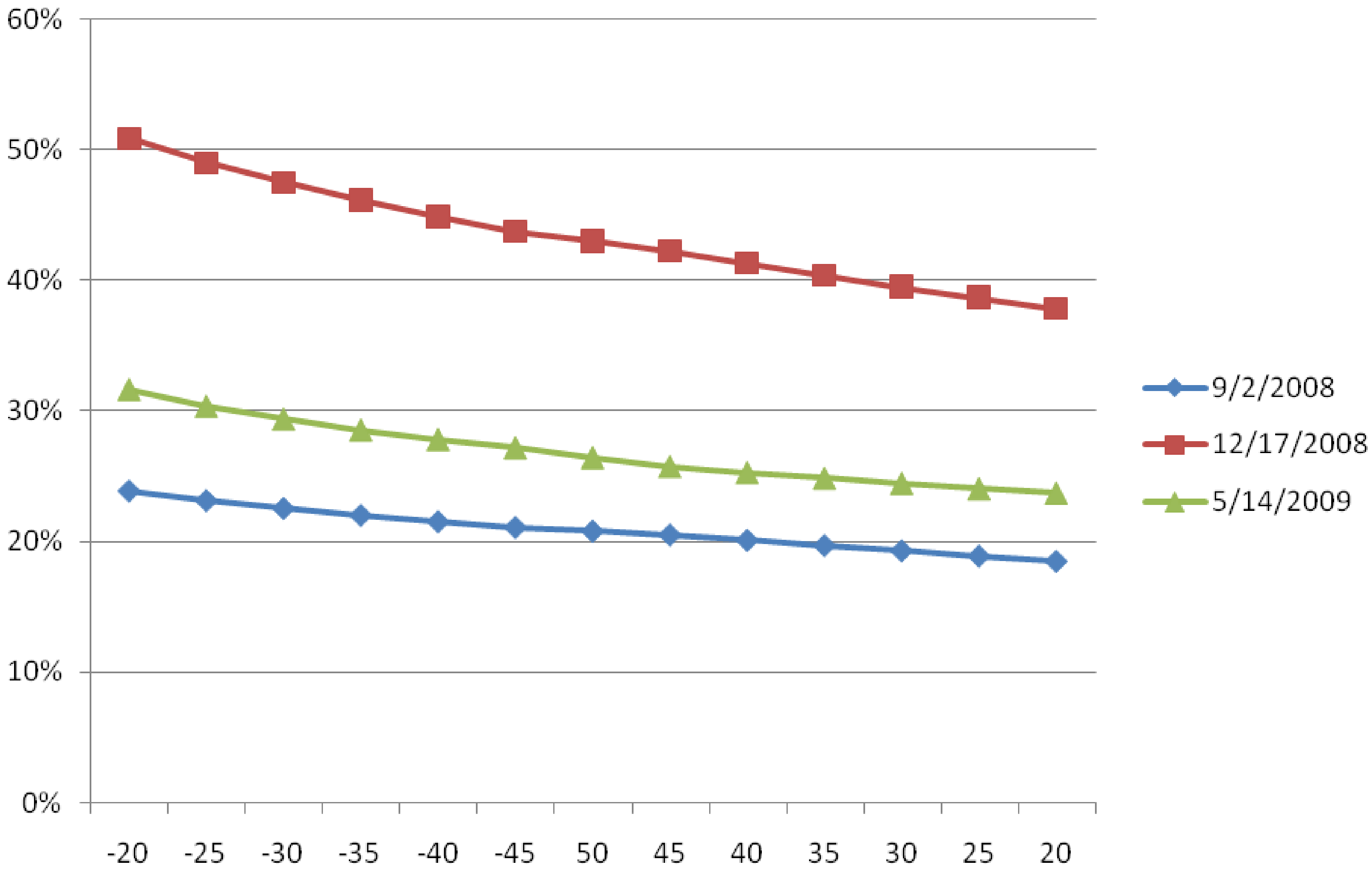
4. Deduce a model for the variation of implied volatilities for portfolio risk analysis

The Data (example with DIA)

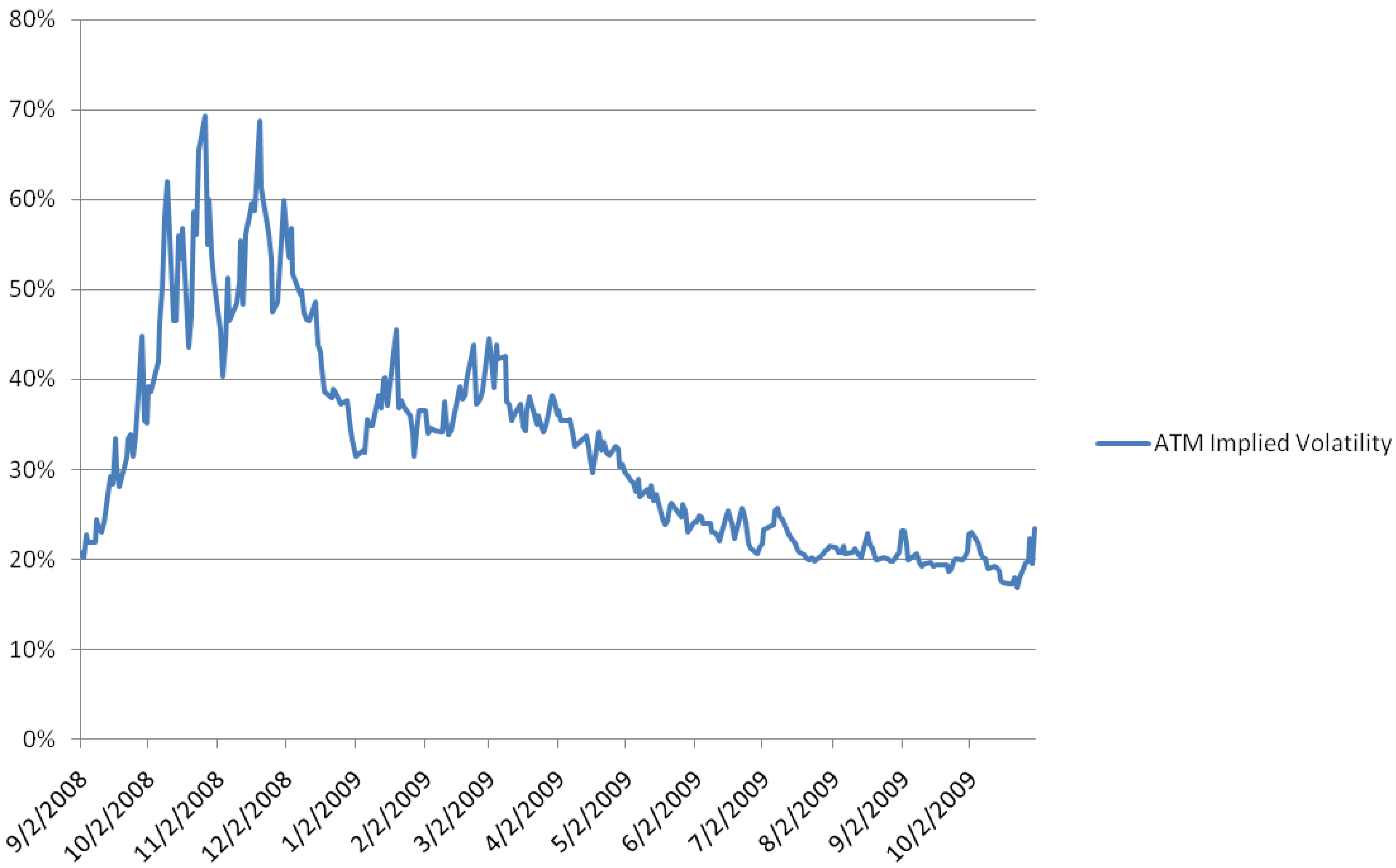
	OTM Puts										OTM Calls		
date\delta	**-20**	**-25**	**-30**	**-35**	**-40**	**-45**	**50**	**45**	**40**	**35**	**30**	**25**	**20**
9/2/2008	23.9%	23.2%	22.6%	22.0%	21.5%	21.1%	20.8%	20.5%	20.1%	19.7%	19.3%	18.9%	18.5%
9/3/2008	23.1%	22.4%	21.9%	21.3%	20.9%	20.4%	20.2%	20.1%	19.7%	19.3%	18.9%	18.5%	18.1%
9/4/2008	26.2%	25.6%	25.0%	24.6%	24.2%	23.8%	22.7%	21.6%	21.3%	21.0%	20.7%	20.4%	20.0%
9/5/2008	25.0%	24.3%	23.7%	23.2%	22.8%	22.3%	21.9%	21.5%	21.1%	20.7%	20.4%	20.0%	19.6%
9/8/2008	24.9%	24.2%	23.6%	23.0%	22.5%	22.0%	21.9%	21.7%	21.3%	20.8%	20.4%	19.9%	19.5%

We consider data from 9/2/2008 until 10/30/2009, organized by Deltas (13 strikes per day)

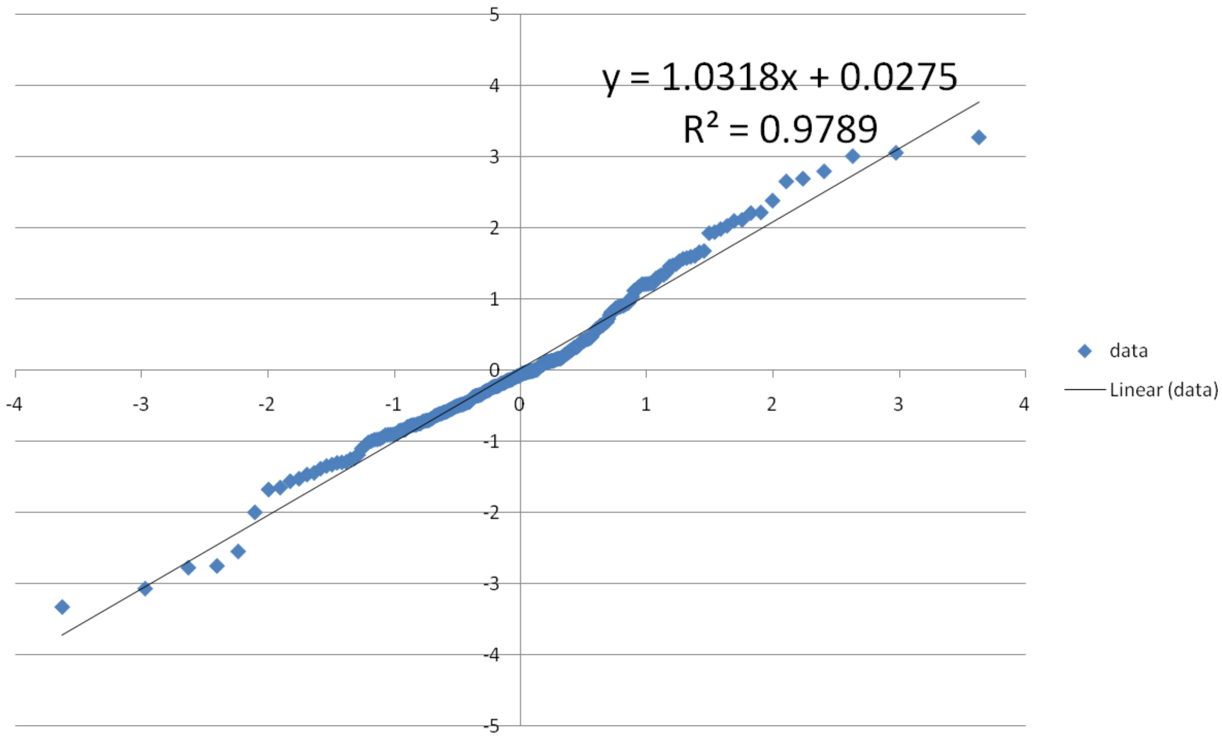
DIA 30 day Implied Vol Curves



DIA ATM Volatility
Sep 2, 2008 – Oct 30 2009



Extreme-value analysis: ATM vol



QQ-plot vs. Student T with DF=4

prob	student	data
0.0034	-3.633	-3.333
0.0068	-2.976	-3.074
0.0102	-2.633	-2.779
0.01361	-2.406	-2.755
0.01701	-2.238	-2.55
0.02041	-2.106	-1.999
0.02381	-1.997	-1.678
0.02721	-1.905	-1.651
0.03061	-1.825	-1.561
0.03401	-1.755	-1.526
0.03741	-1.693	-1.468
0.04082	-1.637	-1.444
0.04422	-1.585	-1.385
0.04762	-1.538	-1.347
0.05102	-1.495	-1.328

Left tail vs right tail using DF=4

Extreme down moves

prob	student	data
0.0034	-3.633	-3.333
0.0068	-2.976	-3.074
0.0102	-2.633	-2.779
0.01361	-2.406	-2.755
0.01701	-2.238	-2.55
0.02041	-2.106	-1.999
0.02381	-1.997	-1.678
0.02721	-1.905	-1.651
0.03061	-1.825	-1.561
0.03401	-1.755	-1.526
0.03741	-1.693	-1.468
0.04082	-1.637	-1.444
0.04422	-1.585	-1.385
0.04762	-1.538	-1.347
0.05102	-1.495	-1.328

Extreme up moves moves

prob	student	data
0.9558	1.5853	1.99021
0.9592	1.6366	2.0349
0.9626	1.6929	2.10579
0.966	1.7554	2.11977
0.9694	1.8255	2.21635
0.9728	1.9051	2.22458
0.9762	1.9971	2.39156
0.9796	2.1058	2.66136
0.983	2.2381	2.70045
0.9864	2.406	2.8036
0.9898	2.6331	3.01731
0.9932	2.9757	3.06495
0.9966	3.6328	3.28219

Modeling the Volatility of Factors

$$R_{n+1} = \sigma_n \epsilon_{n+1}$$

$${\sigma_{n+1}}^2 = {\sigma_n}^2 + \alpha {R_{n+1}}^2 - \beta {\sigma_n}^2$$

GARCH 1-1 model
(Engle & Granger)

This model has ``persistence'' built in, in the sense that the change
In volatility is affected by the contemporaneous squared-return, but
with memory loss. This can be ``estimated from data'' or can be used
as a paradigm for modeling volatility.

$${\sigma_n}^2 = \alpha \sum_{j=0}^{\infty} (1-\beta)^j {R_{n-j}}^2$$

Exponentially-Weighted Moving Average Volatility Estimator

In equilibrium,

$$E(R^2) = \mathrm{E}(\sigma^2)$$

So $\alpha = \beta$.

$$\boxed{{\sigma_n}^2 = \beta \sum_{j=0}^{\infty} (1-\beta)^j {R_{n-j}}^2}$$

Weights add up to 1. In finite estimation cases, people often use

$$\boxed{{\sigma_n}^2 = \frac{\beta}{1-(1-\beta)^{N+1}} \sum_{j=0}^{N} (1-\beta)^j {R_{n-j}}^2}$$

Estimating the distribution of factor returns (equities)

- Fat-tail distributions (eg. Student T), determined by fitting tails
- EWMA for estimating volatility on a window

In many situations, the tail estimate is taken to be **fixed** (e.g. df=3 or 3.5) following the robust physicst’s approach with pooled data.

Thus, the only parameter that needs to be estimated is the EWMA volatility for each factor.

Interpretation of the parameter β

- This parameter dictates the ``memory decay'' of the estimator or the effective number of periods that intervene in the estimation.

If t is such that $(1-\beta)^t$=1/2

then the weights count by half every t days. This is the ``half-life'' of the estimator:

$$t = \frac{\ln 0.5}{\ln(1-\beta)}$$

$$\beta = 1 - e^{-\frac{\ln(2)}{t}}$$

Example: 60 day dependence corresponds to beta=0.0114, 1-beta=0.9885

Specifications involve window (N) and half-life (1-beta).

Distribution of returns over N days

$$R_{t,t+N} \sim \sigma_{\mathrm{EWMA}} \sqrt{T} \xi \cdot \frac{\nu - 2}{\nu}, \qquad \xi = \mathrm{T\text{-}student}(\nu)$$

For other risk-factors, e.g. implied volatility, we need to estimate autocorrelation of vol returns in order the determine the scaling coefficient for N-day returns. The reason is that volatility returns may have positive correlation (as suggested by GARCH).

Modeling Dependence

- Use stocks as risk-factors and a ``clean covariance matrix’’
- Use eigenvector of the correlation matrix as risk factors

Factor models

$$R = \sum_{j=1}^{N_f} \beta_j F_j + \varepsilon$$

$F_j, \ \ j = 1, \ldots, N_f,$	Explanatory factors
$\beta_j, \ \ j = 1, \ldots, N_f,$	Factor <u>loadings</u>
$\sum_{j=1}^{N_f} \beta_j F_j$	<u>Explained</u>, or <u>systematic</u> portion
ε	<u>Residual,</u> or <u>idiosyncratic</u> portion

CAPM: a `minimalist' approach

A single explanatory factor: the ``market'', or ``market portfolio''

$$R = \beta F + \varepsilon, \quad Cov(R, \varepsilon) = 0$$

F = usually taken to be the returns of a broad-market index (e.g., S&P 500)
<u>Normative statement</u>:

$$< \varepsilon >= 0 \text{ or } < R >= \beta < F >$$

Argument: if the market is ``efficient'', or in ``equilibrium'', investors cannot make money (systematically) by picking individual stocks and shorting the index or vice-versa (assuming uncorrelated residuals). (Lintner, Sharpe. 1964)

Counter-arguments: (i) the market is not ``efficient'', (ii) residuals may be correlated (additional factors are needed).

Multi-factor models (APT)

$$R = \sum_{j=1}^{N_f} \beta_j F_j + \varepsilon, \quad Corr(F_j, \varepsilon) = 0$$

Several factors representing sub-indices in different sectors, size, financial statement variables, etc.

<u>Normative statement </u>(APT):

$$< \varepsilon >= 0 \text{ or } < R >= \sum_{j=1}^{N_f} \beta_j < F_j >$$

Argument: Generalization of CAPM, based again on no-arbitrage. (Ross, 1976)

Counter-arguments: (i) How do we actually define the factors? (ii) Is the number of factors fixed? Known? (iii) The structure of the stock market and risk-premia vary strongly (think pre & post WWW) (iv) The issue of correlation of residuals is intimately related to the number of factors.

Factor decomposition in practice

-- Putting aside normative theories (how prices should behave, which we don't know and never will), factor analysis can be quite useful in practice.

-- In risk-management: to measure exposure of a portfolio to a particular industry or market feature (size, volatility)

-- Dimension-reduction technique to study a system (the market) with a large number of degrees of freedom

-- Makes Portfolio Theory viable in practice (Markowitz to Sharpe to Ross!)

-- Useful to analyze stock investments in a relative fashion (buy ABC, sell XYZ to eliminate exposure to an industry sector, for example).

-- New investment techniques arise from factor analysis. The technique is called *defactoring* (Pole, 2007, Avellaneda and Lee, 2008) and is used in market-neutral investment strategies known as stat-arb.

Principal Components Analysis of Correlation Data

Consider a time window *t=0,1,2,…,T,* (days) a universe of *N* stocks. The returns data is represented by a *T* by N matrix $\left(R_{it}\right)$

$$\sigma_i^2 = \frac{1}{T-1}\sum_{t=1}^{T}\left(R_{it} - \overline{R_i}\right)^2, \quad \overline{R_i} = \frac{1}{T}\sum_{t=1}^{T} R_{it}$$

$$Y_{it} = \frac{R_{it}}{\sigma_i}$$

$$\Gamma_{ij} = \frac{1}{T-1}\sum_{t=1}^{T} Y_{it} Y_{jt}$$

Clearly, $Rank(\Gamma) \leq \min\left(N, T\right)$

Regularized correlation matrix

$$C_{ij} = \frac{1}{T-1}\sum_{t=1}^{T}\left(R_{it}-\overline{R_i}\right)\left(R_{jt}-\overline{R_j}\right)+\gamma\delta_{ij}, \quad \gamma = 10^{-9}$$

$$\Gamma_{ij}^{reg} = \frac{C_{ij}}{\sqrt{C_{ii}C_{jj}}}$$

This matrix is a correlation matrix and is positive definite. It is equivalent for all practical purposes to the original one but is numerically stable for inversion and eigenvector analysis (e.g. with Matlab).

Note: this is especially useful when *T<<N.*

Eigenvalues, Eigenvectors and Eigenportfolios

$$\lambda_1 > \lambda_2 \geq \ldots. \geq \lambda_N > 0$$

eigenvalues

$$\mathbf{V}^{(j)} = \left(V_1^{(j)}, V_2^{(j)}, \ldots, V_N^{(j)}\right), \quad j = 1,2,\ldots, N.$$

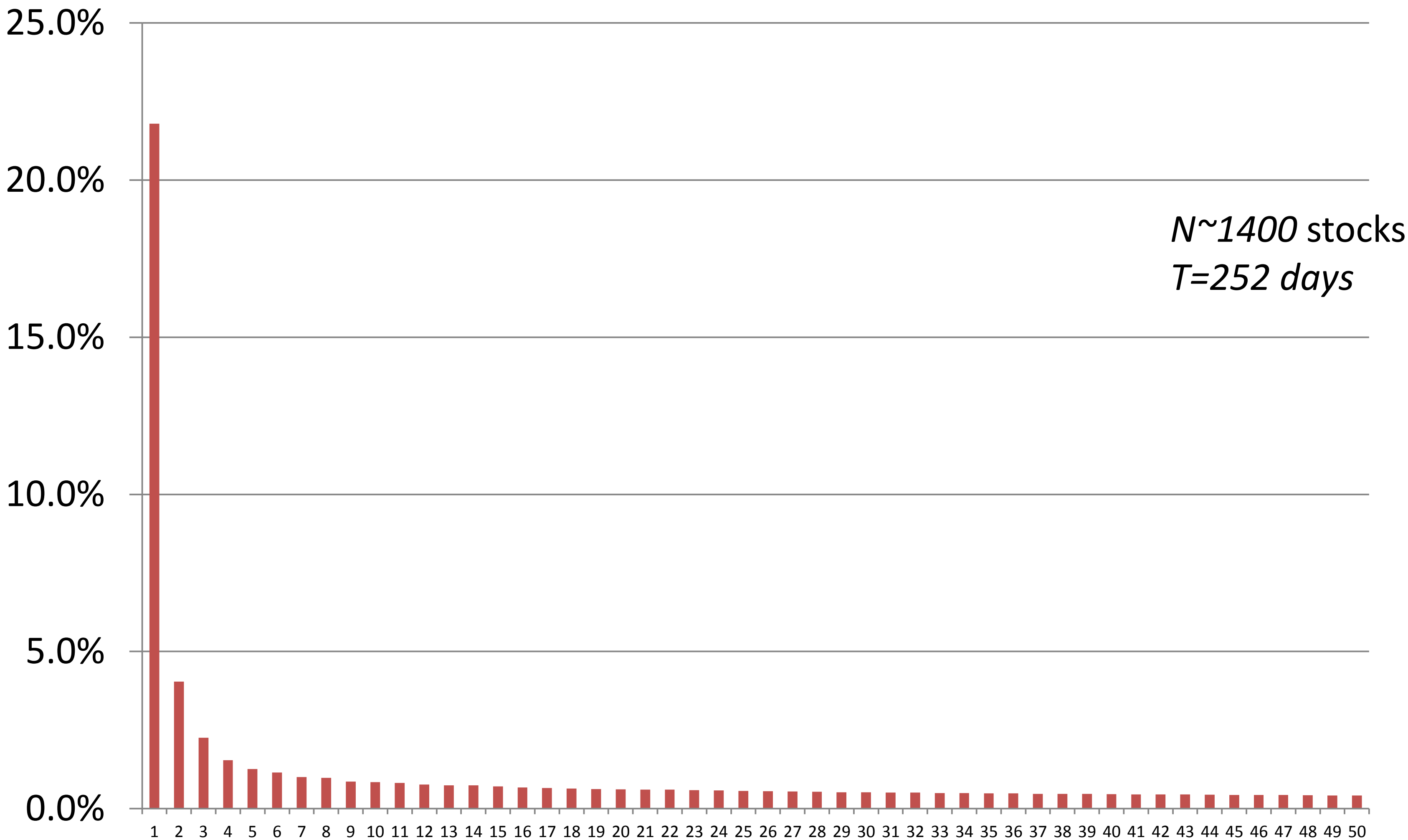
eigenvectors

$$F_{jt} = \sum_{i=1}^{N} V_i^{(j)} Y_{it} = \sum_{i=1}^{N} \left(\frac{V_i^{(j)}}{\sigma_i} \right) R_{it}$$

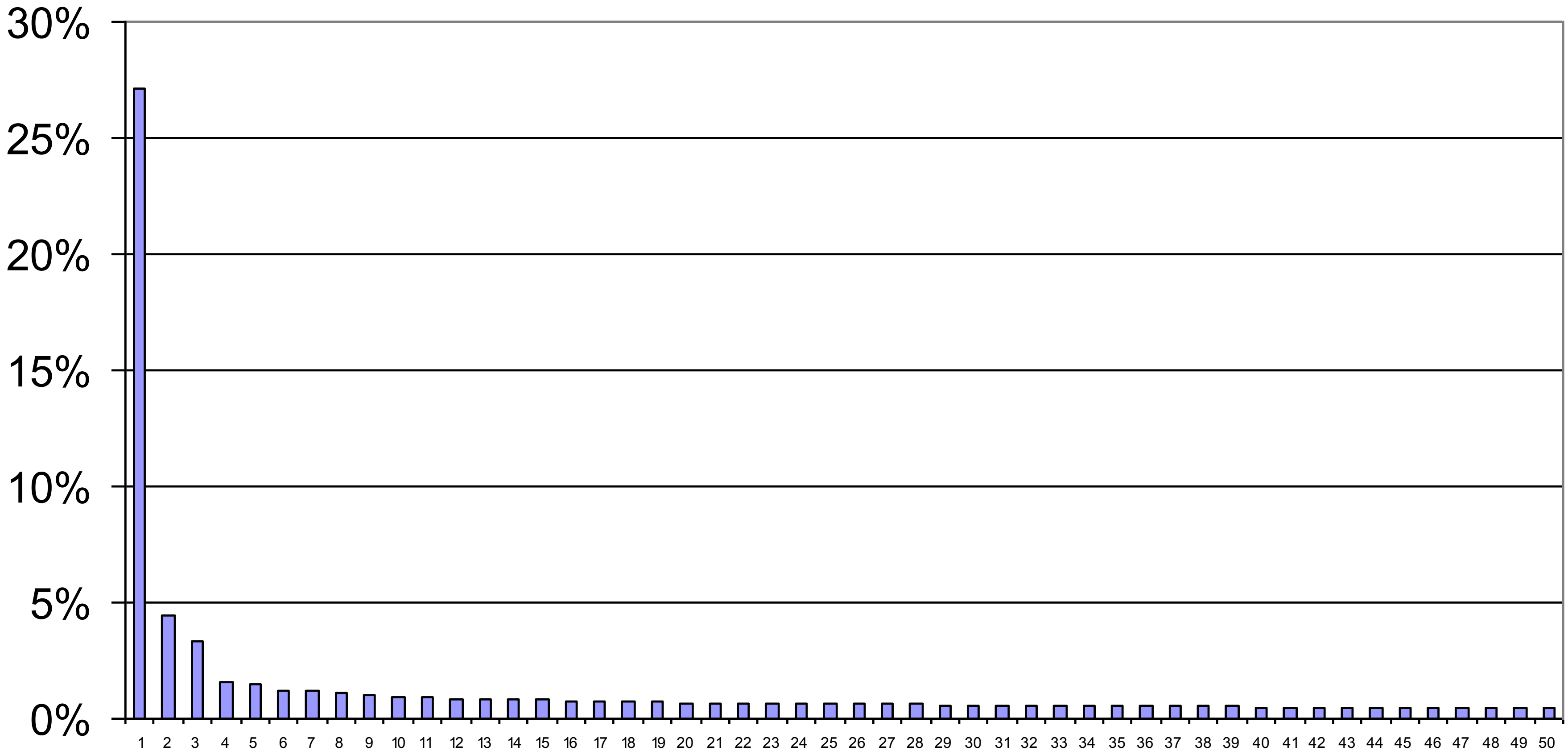
returns of "eigenportfolios"

We use the coefficients of the eigenvectors and the volatilities of the stocks to build ``portfolio weights''. These random variables span the same linear space as the original returns.

50 largest eigenvalues using the 1400 US stocks with cap >1BB cap (Jan 2007)



Top 50 eigenvalues for S&P 500 index components, May 1 2007,T=252



Model Selection Problem: How many EV are significant?

Need to estimate the significant eigenportfolios which can be used as factors.

Assuming that the correlation matrix is invertible (regularize if necessary)

$$< R_i R_j > = C_{ij} = \sum_{k=1}^{N} \lambda_k V_i^{(k)} V_j^{(k)}$$

$$F_k \equiv \sum_{i=1}^{N} \frac{V_i^{(k)}}{\sigma_i} R_i , \qquad \tilde{F}_k \equiv \frac{1}{\sqrt{\lambda_k}} \sum_{i=1}^{N} \frac{V_i^{(k)}}{\sigma_i} R_i$$

$$< F_k^2 > = \lambda_k , \qquad < \tilde{F}_k^2 > = 1, \qquad < \tilde{F}_k \, \tilde{F}_{k'} > = \delta_{kk'}$$

$$R_i = \sum_k \beta_{ik} F_k \quad \Rightarrow \quad \beta_{ik} = \sigma_i \sqrt{\lambda_k} \, V_i^{(k)}$$

Karhunen-Loeve Decomposition

$\mathbf{R}$ = vector of random variables with finite second moment, <.,>=correlation

$$\mathbf{C} = \langle \mathbf{R} \otimes \mathbf{R} \rangle = \langle \mathbf{R}\mathbf{R}' \rangle$$

Covariance matrix

$$\Omega = \mathbf{C}^{1/2}$$

Symmetric square root of C

$$\mathbf{F} = \Omega^{-1}\mathbf{R}, \qquad \mathbf{R} = \Omega\mathbf{F}$$

F has uncorrelated components

$$\mathrm{B} = \Omega = \mathbf{C}^{1/2}$$

Loadings= components of the square-root of C

Since the eigenvectors vanish or are very small in a real system, the modeling consists in defining a small number of factors and attribute the rest to ``noise''

Bai and Ng 2002, *Econometrica*

$$I(m) = \min_{\beta} \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} \left(R_{it} - \sum_{k=1}^{m} \beta_{ik} F_{kt} \right)^2$$

$$m^* = \arg\min_{m} \left(I(m) + m \cdot g(N,T) \right)$$

$$\lim_{N,T \to \infty} g(N,T) = 0, \quad \lim_{N,T \to \infty} \min(N,T) g(N,T) = \infty$$

Under reasonable assumptions on the underlying model, Bai and Ng prove that under PCA estimation, m^* converges in probability to the true number of factors as $N, T \to \infty$

Implementation of Bai & Ng on SP500 Data

g	m*	Lambda_m*	Explained Variance	Tail	Objective Fun	Convexity
1	117	0.20%	87.88%	12.12%	0.355	-
2	59	0.39%	71.44%	28.56%	0.522	-0.085085
3	29	0.59%	57.11%	42.89%	0.603	-0.041266
4	16	0.76%	48.51%	51.49%	0.643	-0.018110
5	10	0.96%	43.52%	56.48%	0.665	-0.007000
6	7	1.18%	40.43%	59.57%	0.680	-0.003096
7	6	1.22%	39.25%	60.75%	0.691	-0.004872
8	4	1.56%	36.56%	63.44%	0.698	0.001069
9	4	1.56%	36.56%	63.44%	0.706	0.000000
10	4	1.56%	36.56%	63.44%	0.714	0.000000
11	4	1.56%	36.56%	63.44%	0.722	0.000000
12	4	1.56%	36.56%	63.44%	0.730	0.000000
13	4	1.56%	36.56%	63.44%	0.738	-

If we choose the cutoff m* as the one for which the sensitivity to g is zero, then m*~5 seems appropriate.
This would lead to the conclusion that the S&P 500 corresponds to a 5-factor model.
The number is small in relation to industry sectors and to the amount of variance explained by industry factors.