Financial Data Analysis

Dr. Markus Haas

LMU München

Summer term 2010

April 21, 2010
Financial Data Analysis

- Markus Haas, haas@stat.uni-muenchen.de
  Office hour: Tuesday 16-18 and by appointment

- Prerequisites: “Einführung in die Empirische Wirtschaftsforschung”, “Ökonometrie 1” or ”Applied Econometrics”
Course Outline

- Introduction: Basic properties of financial return series
- Review of linear time series methods
- Parametric volatility modeling
  (i) GARCH
  (ii) Stochastic volatility models
  (iii) Regime–switching models
- Modeling the dependence structure of returns
  (i) Multivariate GARCH processes
  (ii) Multivariate regime–switching and copulas
- Value–at–Risk: Regulatory framework, quantile estimation, backtesting
Textbooks


Returns

• Let $P_t$ be the asset price at time $t$.

• There is a dividend payment $D_t$ at the end of period $t$.

• Then the (single-period) discrete return is

$$R_t = \frac{P_t + D_t - P_{t-1}}{P_{t-1}}.$$  \hfill (1)

• Dividends are often excluded from return calculations.

• Often (1) is multiplied by 100 to be interpretable in terms of percentage returns.

• The continuously compounded or log return is (ignoring dividends for simplicity)

$$r_t = \log P_t - \log P_{t-1} = \log(1 + R_t).$$  \hfill (2)
• This name derives from the fact that the interest rate $i_n$ equivalent to $R_t$, when interest is paid $n$ times in the period, solves

$$
\left(1 + \frac{i_n}{n}\right)^n = 1 + R_t. \quad (3)
$$

• Continuous compounding is approached as $n \to \infty$, and then

$$
e^{i\infty} = 1 + R_t \Rightarrow i_{\infty} = \log(1 + R_t) = r_t. \quad (4)
$$

• Empirical analysis is often based on log returns. These have the advantage that they can be additively aggregated over time.

• That is, if $r_{t,t+\tau}$ denotes the (multi–period) return from time $t$ to time $t + \tau$, we have

$$
r_{t,t+\tau} = \log \left(\frac{P_{t+\tau}}{P_t}\right) = \sum_{i=1}^{\tau} r_{t+i}. \quad (5)
$$
• This is not the case for the discrete return, where

\[ R_{t, t+\tau} = \prod_{i=1}^{\tau} (1 + R_{t+i}) - 1. \]  

(6)

• On the other hand, if we consider a portfolio of \( N \) assets with weights \( a_i \), and returns \( R_{it}, i = 1, \ldots, n \), then the portfolio return is

\[ R_{p,t} = \sum_{i=1}^{N} a_i R_{it}, \]  

(7)

whereas

\[ r_{p,t} = \log(1 + R_{p,t}) \neq \sum_{i=1}^{N} a_i r_{it}, \]  

(8)

i.e., the linear combination of continuously compounded asset returns is not the continuously compounded portfolio return.
• For small $x$,¹

$$\log(1 + x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \cdots \approx x, \quad (9)$$

so that $r_t$ may serve as a reasonable approximation to the discrete return.

<table>
<thead>
<tr>
<th>$100 \times R_t$</th>
<th>-30.0</th>
<th>-20.0</th>
<th>-15.0</th>
<th>-10.0</th>
<th>-5.0</th>
<th>0</th>
<th>5.0</th>
<th>10.0</th>
<th>15.0</th>
<th>20.0</th>
<th>30.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>$100 \times r_t$</td>
<td>-35.7</td>
<td>-22.3</td>
<td>-16.3</td>
<td>-10.5</td>
<td>-5.1</td>
<td>0</td>
<td>4.9</td>
<td>9.5</td>
<td>14.0</td>
<td>18.2</td>
<td>26.2</td>
</tr>
</tbody>
</table>

$R_t$ and $r_t = \log(1 + R_t)$ are the discrete and continuously compounded returns, respectively.

• The approximation

$$r_{p,t} \approx \sum_{i=1}^{N} a_i r_{it} \quad (10)$$

is also frequently used.

¹Note that the expansion (not the approximation) is only valid for $x \in (-1, 1]$. 

Basic Statistical Properties of Returns: Return Distribution

• The traditional assumption that has long dominated empirical finance was that log–returns over longer time intervals are approximately normally distributed.

• For example, daily returns are the sum of a large number of intraday returns.

• Appealing to the central limit theorem, Osborne (1959) argued in a classical article that

  “under fairly general conditions [...] we can expect that the distribution function of \([r_t]\) will be normal.”

S&P 500 index level (daily), January 1990 to March 2010

[Graph showing the S&P 500 index level from January 1990 to March 2010. The chart displays the index level on the Y-axis ranging from 0 to 700 and the years on the X-axis from 1990 to 2010.]
S&P 500 index returns (daily), January 1990 to March 2010
DAX 30 index returns (daily), January 1990 to October 2009
Density of the S&P 500 returns

- "empirical (kernel)"
- "normal"
Log-Density of the S&P 500 returns

- empirical (kernel)
- normal
Density of the DAX returns

- empirical (kernel)
- normal
Log-Density of the DAX returns

- Black line: empirical (kernel)
- Red dashed line: normal
Basic Statistical Properties of Returns: Return Distribution

- Financial Returns at higher frequencies (higher than a month at least) are not normally distributed.

- In particular, they have much more probability mass in the center and the tails than a normal distribution with the same variance.

- This implies, among other things, that the probability of large losses is much higher than under the Gaussian assumption.

- At lower frequencies, however, the central limit theorem appears to operate, and the return distribution begins to closer resemble a Gaussian shape.
• A further simple device for detecting departures from normality (or any other hypothesized distribution) are QQ plots.

• This is a scatter plot of the empirical quantiles (vertical axis) against the theoretical quantiles (horizontal axis) of a given distribution (e.g., the normal distribution).

• Excess kurtosis means that the probability of large negative or positive values is greater than under the corresponding normal density function. So the lower quantiles are smaller than the normal quantiles, and the upper quantiles are greater.

• Consequently, fat tails show up in QQ plots as deviations below an ideal straight line at the lower quantiles, and above the straight line at the upper quantiles.
QQ plot for the S&P 500 returns

Normal quantiles

Return Quantiles

Normal quantiles
Kurtosis

- A distribution with higher peaks and fatter tails (and, consequently, less mass in the shoulders) than the normal is called “leptokurtic”.

- The standardized fourth moment is often calculated to measure the degree of leptokurtosis, i.e.,

  \[ \kappa = \text{kurtosis}(r) = \frac{E((r - \mu)^4)}{\sigma^4}, \]  

  where \( \mu \) and \( \sigma^2 \) are the mean and variance of \( r \), respectively, and the sample analogue is

  \[ \hat{\kappa} = \frac{T^{-1} \sum_{t=1}^{T} (r_t - \bar{r})^4}{\hat{\sigma}^4}, \]  

  where \( \bar{r} \) is the sample mean.

- For the normal distribution, \( \kappa = 3 \), and a distribution with \( \kappa > 3 \) is then classified as leptokurtic.

- The intuition is that the contribution of the rare and large returns in the tails is larger for the fourth moment than for the second (variance).
Kurtosis

• Although this is the typical pattern of financial return data, it should be mentioned that a high value of kurtosis does not always (in the sense of a mathematical relationship) imply fatter tails and higher peaks than the normal (or any other distribution).

• In fact, it can be shown that, if there are two densities $f$ and $g$, each being symmetrical with mean zero and common variance, and if there are numbers $a$ and $b$ such that

$$g(x) < f(x) \quad \text{for } a < |x| < b,$$

whereas

$$g(x) > f(x) \quad \text{for } |x| < a \text{ and } |x| > b,$$

then the kurtosis measure $\kappa$ (standardized fourth moment) is greater for $g$ than for $f$.\(^3\)

• The converse need not necessarily hold, however.

• The moment–based measures are still useful, however, and routinely calculated in the literature.

• However, a nonparametric density estimate or QQ plot will in any case be more informative.
**Skewness**

- Sometimes we also observe deviations from symmetry, although these tend to be less pronounced and more difficult to detect.

- The moment–based skewness measure is

\[ s = \text{skewness}(r) = \frac{\mathbb{E}(r - \mu)^3}{\sigma^3}, \]  

with sample counterpart

\[ \hat{s} = \frac{T^{-1} \sum_{t=1}^{T} (r_t - \mu)^3}{\hat{\sigma}^3}. \]

- For the normal, which is symmetric, \( s = 0 \).
Jarque–Bera test for normality

- Measures $\hat{\kappa}$ and $\hat{s}$ can be used to construct a test for normality.

- Under normality, $\hat{s}^{asy} \sim \text{Normal}(0, 6/T)$, and $\hat{\kappa}^{asy} \sim \text{Normal}(3, 24/T)$, so

$$T\hat{s}^2/6 \sim \chi^2(1), \quad T(\hat{\kappa} - 3)^2/24 \sim \chi^2(1),$$

and both are asymptotically independent, so

$$JB = T\hat{s}^2/6 + T(\hat{\kappa} - 3)^2/24 \sim \chi^2(2),$$

a $\chi^2$ distribution with two degrees of freedom.

- Note that we cannot easily use $\hat{s}$ as a basis for a test of symmetry. Although symmetric distributions always have $s = 0$, the asymptotic standard error $\sqrt{6/T}$ is valid only under normality, and it is much larger for fat–tailed symmetric distributions.
Alternative Distributions for Returns

• Mandelbrot (1963),⁴ in a famous study of cotton price changes, was one of the first to point out the fat–tailed nature of return distributions.

• Mandelbrot suggested (nonnormal) $\alpha$–stable (or stable Paretian) distributions as a generic model for asset returns, which may be viewed as a generalization of Osborne’s Gaussian model.

---

Alternative Distributions for Returns: Discrete Normal Mixtures

- A $k$–component (discrete) normal mixture distribution is described by the density

$$f(x) = \sum_{j=1}^{k} \lambda_j \phi(x; \mu_j, \sigma_j^2), \quad \phi(x; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} \exp \left\{ -\frac{(x - \mu)^2}{2\sigma^2} \right\},$$  

(17)

$\lambda_j > 0$, $j = 1, \ldots, k$, are the mixing weights, satisfying $\sum_j \lambda_j = 1$, and the $\mu_j$s and $\sigma_j^2$s are the component means and variances respectively.

- Flexible with respect to skewness and kurtosis.

- A possible interpretation of the normal mixture is that returns are normally distributed, but that return expectation and variance depend on the market regime, e.g., bull vs. bear markets.
• For the S&P 500 and the DAX, we find

<table>
<thead>
<tr>
<th></th>
<th>$\hat{\lambda}_1$</th>
<th>$\hat{\mu}_1$</th>
<th>$\hat{\sigma}^2_1$</th>
<th>$\hat{\lambda}_2$</th>
<th>$\hat{\mu}_2$</th>
<th>$\hat{\sigma}^2_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>S&amp;P 500</td>
<td>0.801</td>
<td>0.071</td>
<td>0.518</td>
<td>0.199</td>
<td>-0.125</td>
<td>4.771</td>
</tr>
<tr>
<td>DAX 30</td>
<td>0.822</td>
<td>0.098</td>
<td>1.025</td>
<td>0.178</td>
<td>-0.285</td>
<td>7.580</td>
</tr>
</tbody>
</table>

• The basic mixture specification (17) can be generalized in various directions to provide a more satisfactory return model. For example, the mixing weights can be made time–varying, so that the regimes are persistent or depend on exogenous or lagged endogenous variables.
Normal mixture QQ plot for the S&P 500 returns
QQ plot for the S&P 500 returns

Normal quantiles

Return quantiles

![QQ plot](image-url)
Alternative Distributions for Returns: Student’s $t$

- The standard Student’s $t$ distribution with mean $\mu$, scale $\sigma$ and $\nu$ degrees of freedom has density

$$f(x) = \frac{\Gamma \left(\frac{\nu+1}{2}\right)}{\Gamma(\nu/2)\sqrt{\pi\nu\sigma}} \left\{1 + \frac{(x - \mu)^2}{\nu\sigma^2}\right\}^{-(\nu+1)/2}, \quad \nu > 0. \quad (18)$$

- It can also be viewed as a (continuous) mixture of normals.

- The smaller $\nu$, the fatter the tails, and normality is approached as $\nu \to \infty$.

- Generalizations that allow for skewness exist.
Alternative Distributions for Returns: Generalized Exponential Distribution (GED)

- This has density

\[
f(x) = \frac{2^{-(1/p+1)}p}{\Gamma(1/p)s} \exp \left\{ -\frac{1}{2} \left| \frac{x - \mu}{\sigma} \right|^p \right\}, \quad p > 0, \tag{19}\]

where \( p \) measures the thickness of the tails.

- For \( p = 2 \), this nests the normal, and for \( p = 1 \) we get the Laplace (double exponential) distribution.

- Asymmetric extensions have been proposed.
Concentrating on the Tails: Extreme Value Theory

- Often (e.g., when calculating risk measures such as Value–at–Risk) we are not interested in the entire distribution of returns but only in the probability of extreme events.

- We can then use extreme value theory to fully concentrate on the tail behavior, without the need to model the central part of the distribution.

- It is often found that the tails of return distribution are well described by a power law, i.e., for large $x$, with $F$ being the distribution function (cdf),

$$1 - F(x) = P(|r| > x) \approx cx^{-\alpha},$$  \hspace{1cm} (20)

where $\alpha$ is the tail exponent. In a log–log plot of the empirical complementary cdf $(1 - F(x))$ against $|r|$ (assuming symmetry), the observations should approximately plot along a straight line (which can be drawn using linear regression, $\log(1-F(x)) = \log c - \alpha \log x$, the slope parameter is the tail exponent; but note that more efficient estimators exist for $\alpha$).

- Several distributions are characterized by power law tails. For example, Student’s $t$ has power tails with tail index $\nu$. 
Here the (regression–based) estimated tail index is 3.17, which is rather typical for stock returns.
FTSE 100 data

linear fit
(slope = −2.983)

Gaussian fit
CAC 40 data linear fit (slope = −3.105)
Gaussian fit
DAX 30

- data
- linear fit (slope = −2.987)
- Gaussian fit
• Consider the sample autocorrelation function at lag \( \tau \),

\[
\hat{\rho}(\tau) = \frac{\sum_{t=1}^{T-\tau} (r_t - \bar{r})(r_{t+\tau} - \bar{r})}{\sum_{t=1}^{T} (r_t - \bar{r})^2}, \quad \tau > 0,
\]

(21)

where

\[
\bar{r} = \frac{1}{T} \sum_{t=1}^{T} r_t,
\]

(22)

and \( T \) is the sample size.
autocorrelations of absolute (demeaned) DAX 30 returns

autocorrelations of squared (demeaned) DAX 30 returns
Temporal Properties of Returns

- Return series are characterized by *volatility clustering*, that is, “large [price] changes tend to be followed by large changes—of either sign—and small changes tend to be followed by small changes”.

- Thus variance (and thus risk) appears to be persistent and predictable (in contrast to the direction of price changes).

- Several approaches for capturing time-varying volatility have been developed, such as (G)ARCH, stochastic volatility, and regime-switching models.

---

Dependence Structure of Returns

• In basic portfolio theory, we are interested in the first two moments of the (portfolio) return distribution, i.e., mean and variance.

• In this framework, correlations between assets are of predominant interest, because the strength of the correlations determines the degree of risk (variance) reduction that can be achieved by efficient portfolio diversification.

• Simple correlation estimates may be misleading, however, due to asymmetric dependence structures.

• This refers to the observation that, for example, stock returns are more dependent in bear markets (market downturns) than in bull markets.

• Therefore, diversification might fail when the benefits from diversification are most urgently needed.
Dependence Structure of Returns

• A popular tool to describe this asymmetric dependence structure are the exceedance correlations of Longin and Solnik (2001).\(^6\)

• For a given threshold \(\theta\), the exceedance correlation between (demeaned) returns \(r_1\) and \(r_2\) is given by

\[
\rho(\theta) = \begin{cases} 
\text{Corr}(x, y|x > \theta, y > \theta) & \text{for } \theta \geq 0 \\
\text{Corr}(x, y|x < \theta, y < \theta) & \text{for } \theta \leq 0 
\end{cases}
\] (23)

• Let us consider monthly returns of MSCI stock market indices for the US and Germany from January 1970 to June 2008.

---

exceedance threshold $\theta = -5$
exceedance threshold $\theta = 5$
- Models capable of producing asymmetric dependence structures are multivariate regime-switching models as well as (general) copulas.