GARCH-type Models

Econometrics

Esben Høg

(based on slides by J. Eduardo Vera-Valdés)

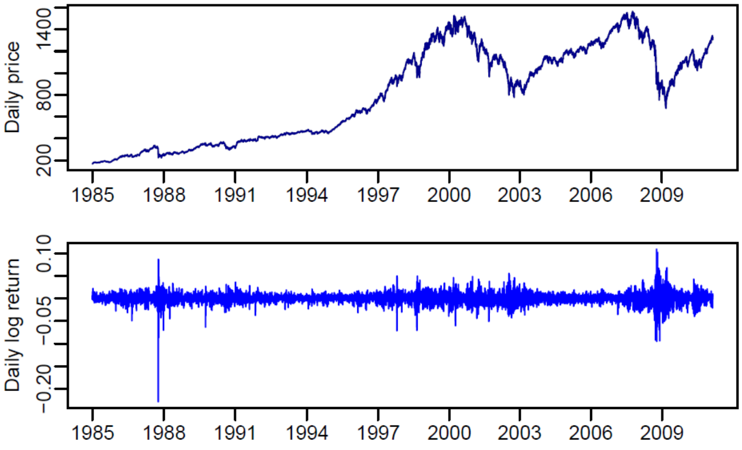
Department of Mathematical Sciences
Aalborg University
Denmark

AALBORG UNIVERSITY
DENMARK

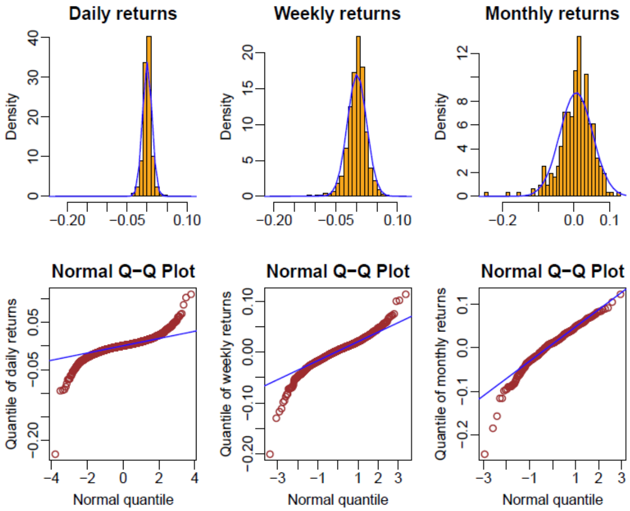
Outline

Outline

Department of Mathematical Sciences
Aalborg University
Denmark

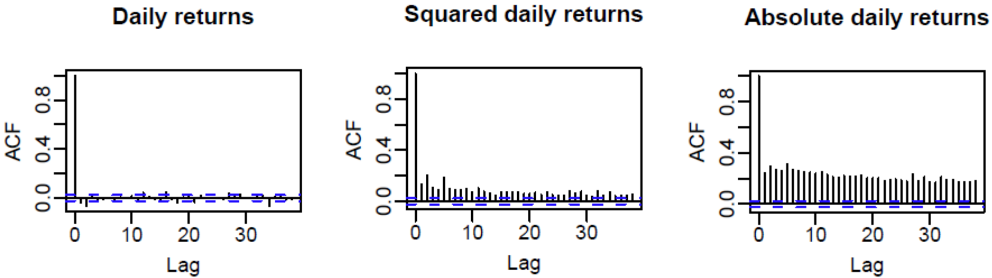


S&P 500 index from January 1985 to February 2011.



Histogram and Q-Q plots for S&P 500 from Jan. 1985 to Feb. 2011.

Department of Mathematical Sciences
Aalborg University
Denmark



S&P 500 from January 1985 to February 2011.

Introduction

With time-series data, it is not uncommon for residuals to be small in absolute value for a number of periods of time, then much larger for a while, then smaller again, and so on.

Moreover, returns tend to be skewed and have heavier tails.

Furthermore, returns seem non-autocorrelated, but squared and absolute returns show strong persistence.

These dynamics are often encountered in models for stock returns, foreign exchange rates, and other series that are determined in financial markets.

Stylized facts

Stylized Fact 1 Time series of share prices X_t and other basic financial instruments are not stationary time series and possess a local trend at the least.

Stylized Fact 2 Returns r_t have a leptokurtic distribution. The empirically estimated kurtosis is mostly greater than 3.

Stylized Fact 3 The return process is white noise since the sample autocorrelations are not significantly different from 0. Furthermore the white noise is not independent since the sample autocorrelations of squared and absolute returns are clearly greater than 0.

Outline

ARCH

The basic idea of ARCH models is that the variance of the error term at time t depends on the realized values of the squared error terms in previous time periods.

Up to today we have studied the model given by

$$Y_t = X_t \beta + u_t$$

where u_t are independent.

Engle proposed to modify the model to allow for conditional heteroskedasticity, letting

$$u_t = \sigma_t \varepsilon_t,$$

where the ε_t's are i.i.d. (0,1)

ARCH

The $ARCH(q)$ process is defined by

$$u_t = \varepsilon_t \sigma_t,$$

with

$$\sigma_t^2 = E(u_t^2 | \Omega_{t-1}) = \alpha_0 + \sum_{i=0}^{q} \alpha_i u_{t-i}^2,$$

where $\alpha_i > 0 \ \forall i$, and ε_t are white noise innovations with variance 1.

In practice, information criteria like AIC and BIC can be used to estimate q.

ARCH

By construction, u_t is not independent from u_{t-1},

Nonetheless, they are uncorrelated

$$E(u_t u_{t-1}) = E(E(u_t u_{t-1}|\Omega_{t-1})) = E(u_{t-1}\sigma_t E(\varepsilon_t|\Omega_{t-1})) = 0,$$

given that $\sigma_t \in \Omega_{t-1}$.

Moreover, the unconditional variance is given by

$$\sigma^2 = \frac{\alpha_0}{1 - \sum_{i=1}^{q} \alpha_i},$$

which imposes some restrictions on the α_i.

Department of Mathematical Sciences
Aalborg University
Denmark

ARCH

From the $ARCH(q)$, we note that

$$u_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i u_{t-i}^2 + (u_t^2 - \sigma_t^2),$$

where $E(u_t^2 - \sigma_t^2) = 0$. This suggest that we can estimate the parameters by linear regression.

Unfortunately, the $ARCH(q)$ process has not proven to be very satisfactory in applied work. Financial time series display time-varying volatility that is highly persistent.

This suggest that q must be large. But if q is large, the $ARCH(q)$ process has a lot of parameters to estimate.

Consider the $ARCH(1)$ process with innovations sampled from a normal, note that the conditional fourth moment is given by

$$E(u_t^4|\Omega_{t-1}) = 3(\alpha_0 + \alpha_1 u_{t-1}^2)^2.$$

Assuming the fourth moment exists and denoting it by m_4,

$$m_4 = \frac{3\alpha_0^2(1+\alpha_1)}{(1-\alpha_1)(1-3\alpha_1^2)}.$$

Comparing m_4 against the fourth moment of a normal, $3\sigma^2$, shows that the $ARCH$ process possess heavier tails than a normal.

Department of Mathematical Sciences
Aalborg University
Denmark

Outline

Department of Mathematical Sciences
Aalborg University
Denmark

GARCH

Bollerslev proposed the Generalized ARCH model given by

$$u_t = \varepsilon_t \sigma_t,$$

with

$$\sigma_t^2 = E(u_t^2 | \Omega_{t-1}) = \alpha_0 + \sum_{i=0}^{q} \alpha_i u_{t-i}^2 + \sum_{j=1}^{p} \delta_j \sigma_{t-j}^2,$$

where again ε_t are white noise innovations with variance 1.

The $GARCH(p, q)$ model resembles the $ARMA$ model in the same vain as $ARCH$ processes resemble AR processes.

Some GARCH properties

Unconditional variance

$$\sigma^2 = \frac{\alpha_0}{1 - \sum_{i=1}^{q} \alpha_i - \sum_{j=1}^{p} \delta_j},$$

with $\sum_{i=1}^{q} \alpha_i + \sum_{j=1}^{p} \delta_j < 1$.

Assuming the fourth moment exists, it can be shown that the kurtosis is

$$Kurt = 3 + \frac{6\alpha_1^2}{1 - \delta_1^2 - 2\alpha_1\delta_1 - 3\alpha_1^2}.$$

Note that the kurtosis equals 3 only in the case of the boundary value $\alpha_1 = 0$ where the conditional heteroscedasticity disappears and a Gaussian white noise takes place.

GARCH models are typically estimated by maximum likelihood.

Assuming that the innovations are normally distributed, the conditional density of Y_t is given by

$$f(Y_t|X_t, \Omega_{t-1}) = \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp -\frac{\varepsilon_t^2}{2\sigma_t^2},$$

which we maximise conditioning on some initial observations.

Although *GARCH* models have error terms with thicker tails than those of the normal distribution, data from financial markets often have tails even thicker than those implied by a *GARCH* model with normal innovations.

Thus, distributions with thicker tails than the standard normal can be used.

Department of Mathematical Sciences
Aalborg University
Denmark

GARCH

It has been shown in practice that the $GARCH(1,1)$ consistently beats more general processes.

An empirical regularity, almost a "stylized fact", is that α_1 is often estimated small and positive, while δ_1 is much larger, and the sum of the two between 0.9 and 1.

These parameter values imply that the time-varying volatility is highly persistent.

Department of Mathematical Sciences
Aalborg University
Denmark

Outline

IGARCH

The *IGARCH* model allows for more persistence in volatility by introducing a unit root.

For instance, the *IGARCH*$(1,1)$ is given by

$$\sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \delta_1 \sigma_{t-1}^2,$$

with $\alpha_1 + \delta_1 = 1$ so that volatility shocks have a permanent effect.

It can be shown that the *IGARCH* is strongly stationary but not weakly stationary.

IGARCH

When estimating the parameters in GARCH models, it is observed that for shorter samples, the estimated parameters α and β sum up to values significantly different from 1 while for longer samples, their sum approaches 1. The combination of these phenomenon is called the integrated GARCH (IGARCH) effect of return data.

IGARCH

Note that, by iteration, the *IGARCH* model can be written as an exponential smoothing model by

$$\sigma_t^2 = \frac{\alpha_0}{1-\delta_1} + (1-\delta_1)(u_{t-1}^2 + \delta_1 u_{t-2}^2 + \delta_1^2 u_{t-3}^2 + \cdots).$$

For instance, JP Morgan's RiskMetrics uses this model with $\alpha_0 = 0$ and $\delta_1 = 0.94$ for daily volatility.

Special asymmetric effects

Standard GARCH models assume that positive and negative error terms have a symmetric effect on volatility. In other words, good and bad news have the same effect on the volatility in this model.

In practice this assumption is frequently violated, in particular by stock returns, in that the volatility increases more after bad news than after good news. This is called *Leverage Effect*.

A drop in the value of the firm will cause a negative return on its stock, and will usually increase the leverage of the stock. That rise in the debt-equity ratio will surely mean a rise in the volatility of the stock.

Department of Mathematical Sciences
Aalborg University
Denmark

Special asymmetric effects (cont.)

A very simple but plausible explanation for the leverage effect: Negative returns imply a larger proportion of debt through a reduced market value of the firm, which leads to a higher volatility.

The risk, i.e. the volatility reacts first to larger changes of the market value, nevertheless it is empirically shown that there is a high volatility after smaller changes.

From an empirical point of view the volatility reacts asymmetrically to the sign of the shocks and therefore parameterized extensions of the standard GARCH model have been suggested: EGARCH and the threshold GARCH (TGARCH) models.

EGARCH

To deal with the stylized fact that investors react more strongly to negative news, Nelson proposed the Exponential *GARCH* given by

$$\log(\sigma_t^2) = \alpha_0 + \delta \log(\sigma_{t-1}^2) + \gamma \frac{\varepsilon_{t-1}}{\sigma_{t-1}} + \alpha_1 \frac{|\varepsilon_{t-1}|}{\sigma_{t-1}},$$

where ε_{t-1} are white noise.

The *EGARCH* model is asymmetric as long as $\gamma \neq 0$. When $\gamma < 0$, positive shocks generate less volatility than negative ones.

The logarithmic transformation guarantees that variances will never become negative.

EGARCH (cont.)

In empirical studies it has been found that EGARCH often overweighs the effects of larger shocks on volatility and thus often results in poorer fits than standard GARCH models.

TGARCH - Threshold GARCH models

The idea of the Threshold GARCH (TGARCH) model is to divide the distribution of the innovations into two disjoint intervals and then approximate a piecewise linear function for the conditional standard deviation.

For most practical purposes the TGARCH model is identical to the so-called GJR-GARCH (Glosten-Jagannathan-Runkle, 1993) model.

The TGARCH model of order $(1, 1)$ can be written as

$$\sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \delta_1 \sigma_{t-1}^2 + \phi u_{t-1}^2 I_{t-1},$$

where I_{t-1} is the indicator function: $I_{t-1} = 1_{(u_{t-1}<0)}$.

APARCH

Alternatively, we can model asymmetric effects by the Assymetric Power GARCH,

$$\sigma_t^{\nu} = \alpha_0 + \sum_{i=1}^{q} \alpha_i (|u_{t-i}| - \gamma_i u_{t-i})^{\nu} + \sum_{j=1}^{p} \delta_j \sigma_{t-j}^{\nu},$$

where $\nu > 0$ and $\gamma_i \in (-1, 1)$. ν represents a Box-Cox transformation of σ_t and γ_i the leverage effect.

Department of Mathematical Sciences
Aalborg University
Denmark

FIGARCH

Introducing Fractional Integration into the model allows for volatility to exhibit a slow hyperbolic rate of decay, or long memory.

The *FIGARCH* is given by

$$(1-\delta(L))(1-L)^d\sigma_t^2 = \alpha_0 + \alpha(L)u_t^2,$$

where $\delta(L)$ and $\alpha(L)$ are polynomials in the lag operator, and $(1-L)^d$ is the fractional difference operator given by

$$(1-L)^d = \sum_{i=0}^{\infty} \frac{\Gamma(i-d)}{\Gamma(-d)\Gamma(i+1)} L^i.$$

This model allows for short memory in returns and long memory in volatility.

GARCH-M

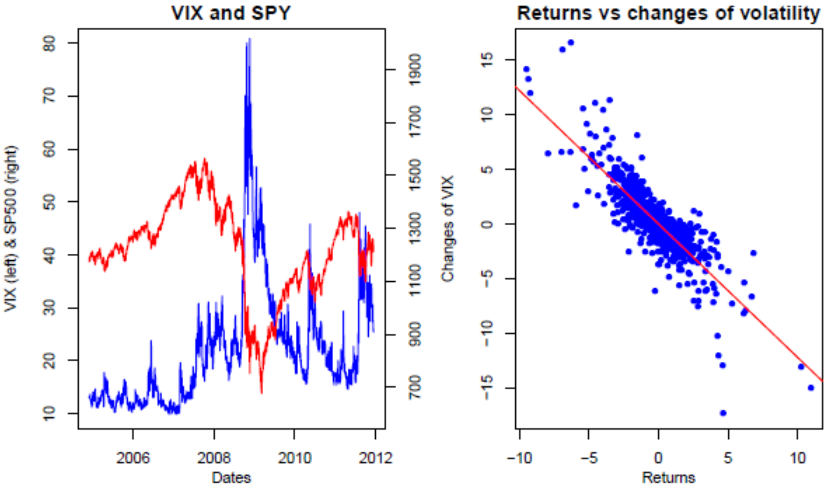
In finance theory (for example in the simple CAPM model) the relationship between risk and return plays an important role.

This is used to estimate the conditional variances in GARCH and then these estimates are used in the conditional expectations estimation.

This is the so-called GARCH in Mean (GARCH–M) model.

In other words the idea is to have the conditional variance in the mean part of the model: Expected return is a function of risk (actually an increasing function).

GARCH-M (cont.)



VIX (blue), and S&P500 (red) from Nov. 2004 to Dec 2011.

Department of Mathematical Sciences
Aalborg University
Denmark

GARCH-M

To capture the risk premium, the GARCH-in-Mean or GARCH-M directly incorporates volatility as a regression variable.

A GARCH-M model takes the form

$$
\begin{aligned}
y_t &= \mu + \lambda\sigma_t + u_t, \\
u_t &= \sigma_t \varepsilon_t, \\
\sigma_t^2 &= \alpha_0 + \alpha_1 u_{t-1}^2 + \delta_1 \sigma_{t-1}^2.
\end{aligned}
$$

The parameter λ captures the risk premium.

The presence of volatility as a statistically significant predictor of returns contributes to serial correlation.

Outline

Multivariate GARCH

Many ideas translate naturally to the multivariate setting.

Simplest generalization of univariate $GARCH(1,1)$ is given by the model

$$\begin{aligned} Y_t &= \mu + u_t, \\ u_t &= \Sigma_t^{1/2} \varepsilon_t, \\ \Sigma_t &= \alpha_0 + \Psi u_t u_t^T + \Lambda \Sigma_{t-1}, \end{aligned}$$

where u_t, ε_t are vectors, and Σ, Ψ, Λ are matrices.

The curse of dimensionality makes this model unmanageable very quickly.

Department of Mathematical Sciences
Aalborg University
Denmark

EWMA

The Exponentially Weighted Moving Average (EWMA) model reduces complexity by depending on a single parameter.

The model mimics the exponential smoothing form previously discussed for the *IGARCH* model. It is defined by

$$\Sigma_t = (1-\lambda)u_t u_t^T + \lambda \Sigma_{t-1},$$

where $\lambda \in (0,1)$.

This single parameter model is simple to estimate regardless of the dimension.

However, the dynamics can be too restrictive in practice since all components have the same discounting factor.

Outline

Introduction

ARCH

GARCH

Extensions

Multivariate

Summing Up

Department of Mathematical Sciences
Aalborg University
Denmark

Summing Up

- Some time series, particularly in Finance, show volatility clustering, heavy tails, asymmetry, and long memory.
- We can model volatility clustering by the *GARCH* model.
- *GARCH* models have heavier tails than normal distributions even if the innovations come from a normal.
- Extensions to the *GARCH* model can be used to model asymmetry and/or long memory.
- The risk premium hypothesis can be modelled by incorporating volatility directly in the regression.
- Multivariate extensions are available but suffer from the curse of dimensionality.

Department of Mathematical Sciences
Aalborg University
Denmark