ARCH and GARCH models

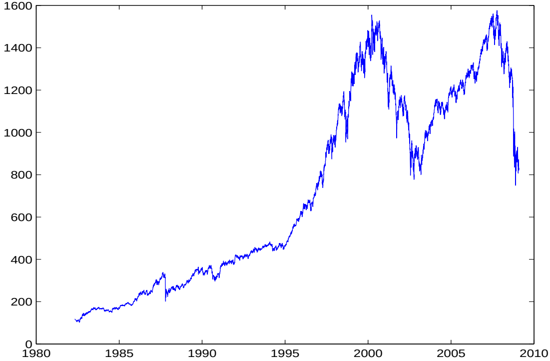
Fulvio Corsi

SNS Pisa

5 Dic 2011

Asset prices

S&P 500 index from 1982 to 2009



asset prices are typically integrated of order one processes $I(1)$

Asset returns

the standard solution is to take the first difference of prices. Two different type of returns:

1. simple net return

$$R_t = \frac{P_t - P_{t-\Delta}}{P_{t-\Delta}}$$

2. log return or continuously compounded returns

$$r_t = \log P_t - \log P_{t-\Delta} = p_t - p_{t-\Delta}$$

over short horizon Δ, r_t is typically small ($|r_t| << 10\%$) so that $R_t \approx r_t$ being

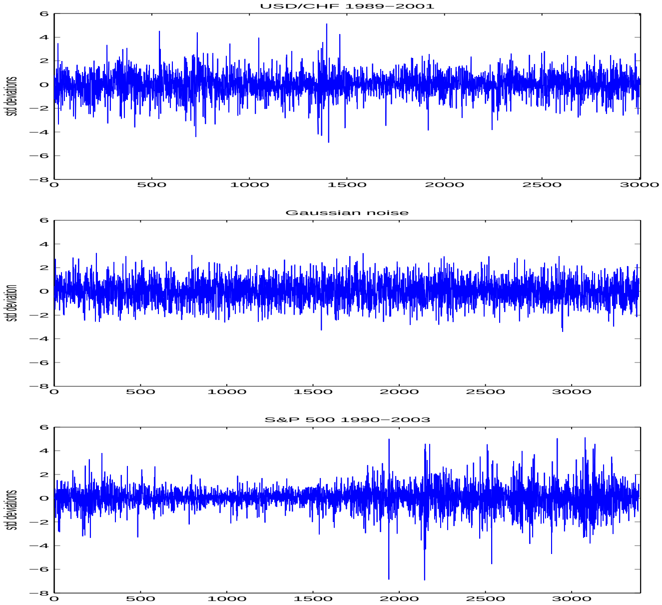
$$1 + R_t = exp(r_t) = 1 + r_t + \frac{1}{2} r_t^2 + \dots$$

The main advantage of the log return is that a k-period return $r_t(k)$ is simply:

$$r_t(k) = p_t - p_{t-k\Delta} = r_{t-(k-1)\Delta} + \dots + r_{t-\Delta} + r_t$$

hence multi-period log returns are simply the sum of single-period log returns.

Asset returns dynamics



Time changing volatility

Dynamics in the volatility of asset returns has paramount consequences in important finance applications:

- asset allocation
- risk management
- derivative pricing

What makes volatility change over time? Still unclear.

- "event-driven volatility": different information arrival rate, consistent with EMH
- "error-driven volatility": due to over- and underreaction of the market to incoming information
- "price-driven volatility": endogenously generated by trading activities of heterogeneous agents $\Rightarrow$ strong positive correlation between volatility and market presence

Different volatility notions

Different types of volatility approaches:

- Parametric: volatility measure is model-dependent

  - Discrete-time: ARCH/GARCH models

  - Continuous-time: Stochastic Volatility models

- Non-Parametric: volatility measure is model-independent (or model-free)

  - Realized Volatility (exploiting the information in High Frequency data)

Different notions of volatility:

- ex-ante conditional volatility
- spot/instantaneous volatility
- ex-post integrated volatility

Basic Structure and Properties of ARMA model

- standard time series models have:

$$Y_t = \mathbb{E}[Y_t|\Omega_{t-1}] + \epsilon_t$$
$$\mathbb{E}[Y_t|\Omega_{t-1}] = f\left(\Omega_{t-1};\theta\right)$$
$$Var\left[Y_t|\Omega_{t-1}\right] = \mathbb{E}\left[\epsilon_t^2|\Omega_{t-1}\right] = \sigma^2$$

  hence,

  - Conditional mean: varies with Ω_{t-1}
  - Conditional variance: constant (unfortunately)

  - k-step-ahead mean forecasts: generally depends on Ω_{t-1}
  - k-step-ahead variance of the forecasts: depends only on k, not on Ω_{t-1} (again unfortunately)

  - Unconditional mean: constant
  - Unconditional variance: constant

AutoRegressive Conditional Heteroskedasticity (ARCH) model

- Engle (1982, Econometrica) intruduced the ARCH models:

$$Y_t = \mathbb{E}[Y_t|\Omega_{t-1}] + \epsilon_t$$
$$\mathbb{E}[Y_t|\Omega_{t-1}] = f(\Omega_{t-1};\theta)$$
$$Var[Y_t|\Omega_{t-1}] = \mathbb{E}\left[\epsilon_t^2|\Omega_{t-1}\right] = \sigma(\Omega_{t-1};\theta) \equiv \sigma_t^2$$

hence,

- Conditional mean: varies with Ω_{t-1}
- Conditional variance: varies with Ω_{t-1}

- k-step-ahead mean forecasts: generally depends on Ω_{t-1}
- k-step-ahead variance of the forecasts: generally depends on Ω_{t-1}

- Unconditional mean: constant
- Unconditional variance: constant

ARCH(q)

- How to parameterize $\mathbb{E}\left[\epsilon_t^2|\Omega_{t-1}\right] = \sigma\left(\Omega_{t-1};\theta\right) \equiv \sigma_t^2$?

- ARCH(q) postulated that the conditional variance is a linear function of the past q squared innovations

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 = \omega + \alpha(L)\epsilon_{t-1}^2$$

- Defining $v_t = \epsilon_t^2 - \sigma_t^2$, the ARCH(q) model can be written as

$$\epsilon_t^2 = \omega + \alpha(L)\epsilon_{t-1}^2 + v_t$$

  Since $\mathbb{E}_{t-1}(v_t) = 0$, the model corresponds directly to an AR(q) model for the squared innovations, ϵ_t^2.

- The process is covariance stationary if and only if the sum of the positive AR parameters is less than 1 i.e. $\sum_{i=1}^{q} \alpha_i < 1$. Then, the unconditional variance of ϵ_t is

$$Var(\epsilon_t) = \sigma^2 = \omega/(1 - \alpha_1 - \alpha_2 - ... - \alpha_q).$$

ARCH and fat tails

- Note that the unconditional distribution of ϵ_t has Fat Tail.

  In fact, the unconditional kurtosis of ϵ_t is

$$\frac{\mathbb{E}(\epsilon_t^4)}{\mathbb{E}(\epsilon_t^2)^2}$$

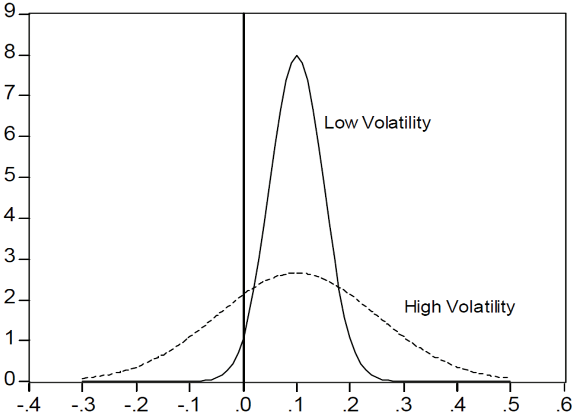
  where the numerator is

$$\begin{aligned}
\mathbb{E}\left[\epsilon_t^4\right] &= \mathbb{E}\left[\mathbb{E}(\epsilon_t^4|\Omega_{t-1})\right] \\
&= 3\mathbb{E}\left[\sigma_t^4\right] \\
&= 3[Var(\sigma_t^2) + \mathbb{E}(\sigma_t^2)^2] \\
&= 3[\underbrace{Var(\sigma_t^2)}_{>0} + \mathbb{E}(\epsilon_t^2)^2] \\
&> 3\mathbb{E}(\epsilon_t^2)^2.
\end{aligned}$$

  Hence,

$$Kurtosis(\epsilon_t) = \frac{\mathbb{E}(\epsilon_t^4)}{\mathbb{E}(\epsilon_t^2)^2} > 3$$

ARCH and fat tails: intuition



$r_t \sim N(\mu, \sigma_t)$ is a mixture of Normals with different $\sigma_t \Rightarrow$ fatter tails

AR(1)-ARCH(1)

- Example: the AR(1)-ARCH(1) model

$$
\begin{aligned}
Y_t &= \phi Y_{t-1} + \epsilon_t \\
\sigma_t^2 &= \omega + \alpha \epsilon_{t-1}^2 \\
\epsilon_t &\sim N(0, \sigma_t^2)
\end{aligned}
$$

  - Conditional mean: $\mathbb{E}(Y_t | \Omega_{t-1}) = \phi Y_{t-1}$
  - Conditional variance: $\mathbb{E}([Y_t - \mathbb{E}(Y_t|\Omega_{t-1})]^2 | \Omega_{t-1}) = \omega + \alpha \epsilon_{t-1}^2$
  - Unconditional mean: $\mathbb{E}(Y_t) = 0$
  - Unconditional variance: $\mathbb{E}(Y_t - \mathbb{E}(Y_t))^2 = \frac{1}{(1-\phi^2)} \frac{\omega}{(1-\alpha)}$

Other characteristics of ARCH(1)

for the ARCH(1) model

$$
\begin{aligned}
Y_t &= \mu + \epsilon_t \\
\sigma_t^2 &= \omega + \alpha \epsilon_{t-1}^2 \\
\epsilon_t &\sim N(0, \sigma_t^2)
\end{aligned}
$$

we also have:

- Excess kurtosis:

$$
Kurtosis(\epsilon_t) = \frac{\mathbb{E}(\epsilon_t^4)}{\mathbb{E}(\epsilon_t^2)^2} = 3\,\frac{1-\alpha^2}{1-3\alpha^2}
$$

  kurtosis is equal to 3 iff $\alpha = 0$

- Stationarity condition for finite variance of ϵ^2 (or for finite kurtosis of ϵ)

$$
\alpha < \frac{1}{\sqrt{3}} \approx 0.577
$$

- Autocorrelation of σ_t: $Corr(\sigma_t, \sigma_{t-h}) = \alpha^h \Rightarrow$ difficult to replicate empirical persistence of σ_t.
  $\Rightarrow$ ARCH(q) models but ...

GARCH(p,q)

- ARCH(q) problem: empirical volatility very persistent $\Rightarrow$ Large q i.e. too many α's
- Bollerslev (1986, J. of Econometrics) proposed the Generalized ARCH model.

  The GARCH(p,q) is defined as

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \, \epsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \, \sigma_{t-j}^2 = \omega + \ \alpha(L)\epsilon_{t-1}^2 + \beta(L)\sigma_{t-1}^2$$

- As before, defining $v_t = \epsilon_t^2 - \sigma_t^2$, the GARCH(p,q) can be also rewritten as

$$\epsilon_t^2 = \omega + [\alpha(L) + \beta(L)] \, \epsilon_{t-1}^2 - \beta(L) v_{t-1} + v_t$$

  which defines an ARMA[max(p, q),p] model for ϵ_t^2.

GARCH(1,1)

- By far the most commonly used is the GARCH(1,1):

$$\begin{aligned} \epsilon_t &= \sigma_t z_t & \text{with} \quad & z_t \sim i.i.d.N(0,1), \\ \sigma_t^2 &= \omega + \alpha\, \epsilon_{t-1}^2 + \beta_j\, \sigma_{t-1}^2 & \text{with} \quad & \omega > 0, \alpha > 0, \beta > 0 \end{aligned}$$

- Stationarity conditions:

  being

$$\sigma_t^2 = \omega + (\alpha z_t^2 + \beta)\sigma_{t-1}^2$$

  and hence,

$$\mathbb{E}[\sigma_t^2] = \omega + (\alpha + \beta)\mathbb{E}[\sigma_{t-1}^2]$$

  then the process is covariance stationary iff

$$\alpha + \beta < 1$$

- Unconditional variance:

$$Var(\epsilon) = \frac{\omega}{1 - (\alpha + \beta)}$$

GARCH(1,1)

- By recursive substitution, the GARCH(1,1) may be written as the following ARCH(∞):

$$\sigma_t^2 = \omega(1-\beta) + \alpha \sum_{i=1}^{\infty} \beta^{i-1} \epsilon_{t-i}^2$$

which reduces to an exponentially weighted moving average filter for $\omega = 0$ and $\alpha + \beta = 1$ (sometimes referred to as Integrated GARCH or IGARCH(1,1)).

- Moreover, GARCH(1,1) implies an ARMA(1,1) representation in the ϵ_t^2

$$\epsilon_t^2 = \omega + (\alpha+\beta)\epsilon_{t-1}^2 - \beta v_{t-1} + v_t$$

- From the ARMA(1,1) representation we can guess that

$$\rho_h \equiv Corr(\sigma_t, \sigma_{t-h}) \approx (\alpha+\beta)^h.$$

The precise calculations give:

$$\rho_1 = \frac{\alpha(1-\beta^2-\alpha\beta)}{1-\beta^2-2\alpha\beta}, \qquad \rho_h = (\alpha+\beta)\rho_{h-1} \quad \text{for} \quad h > 1$$

- Forecasting. Denoting the unconditional variance $\sigma^2 \equiv \omega(1-\alpha-\beta)^{-1}$ we have:

$$\hat{\sigma}_{t+h|t}^2 = \sigma^2 + (\alpha+\beta)^{h-1}(\sigma_{t+1}^2 - \sigma^2)$$

showing that the forecasts of the conditional variance revert to the long-run unconditional variance at an exponential rate dictated by $\alpha + \beta$

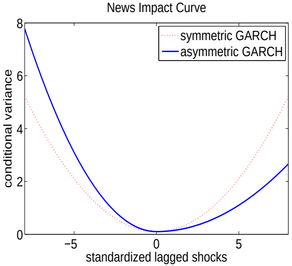
Asymmetric GARCH

- In standard GARCH model:

$$\sigma_t^2 = \omega + \alpha r_{t-1}^2 + \beta \sigma_{t-1}^2$$

σ_t^2 responds symmetrically to past returns.

The so called **“news impact curve”** is a parabola



- Empirically negative r_{t-1} impact more than positive ones $\rightarrow$ asymmetric news impact curve
- GJR or T-GARCH

$$\sigma_t^2 = \omega + \alpha r_{t-1}^2 + \gamma r_{t-1}^2 D_{t-1} + \beta \sigma_{t-1}^2 \quad \text{with} \quad D_t = \begin{cases} 1 & \text{if} \quad r_t < 0 \\ 0 & \text{otherwise} \end{cases}$$

  - Positive returns (good news): α
  - Negative returns (bad news): $\alpha + \gamma$
  - Empirically $\gamma > 0 \rightarrow$ “Leverage effect”
- Exponential GARCH (EGARCH)

$$ln(\sigma_t^2) = \omega + \alpha \left| \frac{r_{t-1}}{\sigma_{t-1}} \right| + \gamma \frac{r_{t-1}}{\sigma_{t-1}} + \beta ln(\sigma_{t-1}^2)$$

GARCH in mean

- In the GARCH-M (Garch-in-Mean) model Engle, Lilien and Robins (1987) introduce the (positive) dependence of returns on conditional variance, the so called "risk-return tradeoff".

- The specification of the model is:

$$\begin{aligned} r_t &= \mu + \gamma\sigma_t^2 + \sigma_t z_t \\ \sigma_t^2 &= \omega + \alpha r_{t-1}^2 + \beta\sigma_{t-1}^2 \end{aligned}$$

- Given the inherent noise of financial returns r_t, the estimates of γ are often very difficult, typically long time series are required to find significant results.

Engle test for heteroscedasticity or ARCH test

The ARCH test of Engle assesses the null hypothesis that a series of residuals ϵ_t exhibits no conditional heteroscedasticity (ARCH effects),

The test is performed by running the following regression

$$\epsilon_t^2 = c + a_1 \epsilon_{t-1}^2 + a_2 \epsilon_{t-2}^2 + \ldots + a_L \epsilon_{t-L}^2$$

then computes the Lagrange multiplier statistic $T \times R^2$, where T is the sample size and R^2 is the coefficient of determination of the regression.

Under the null, we have that

$$T \times R^2 \rightarrow \chi_L^2$$

i.e. the asymptotic distribution of the test statistic is chi-square with L degrees of freedom.

Estimation

- A GARCH process with gaussian innovation:

$$r_t|\Omega_{t-1} \sim N(\mu_t(\theta), \sigma_t^2(\theta))$$

- has conditional densities:

$$f(r_t|\Omega_{t-1};\theta) = \frac{1}{\sqrt{2\pi}}\sigma_t^{-1}(\theta)\exp\left(-\frac{1}{2}\frac{(r_t-\mu_t(\theta))}{\sigma_t^2(\theta)}\right)$$

- using the **“prediction–error”** decomposition of the likelihood

$$L(r_T, r_{T-1}, \dots, r_1;\theta) = f(r_T|\Omega_{T-1};\theta) \times f(r_{T-1}|\Omega_{T-2};\theta) \times \dots \times f(r_1|\Omega_0;\theta)$$

the log-likelihood becomes:

$$\log L(r_T, r_{T-1}, \dots, r_1;\theta) = -\frac{T}{2}\log(2\pi) - \sum_{t=1}^{T}\log\sigma_t(\theta) - \frac{1}{2}\sum_{t=1}^{T}\frac{(r_t-\mu_t(\theta))}{\sigma_t^2(\theta)}$$

- Non–linear function in $\theta \Rightarrow$ Numerical optimization techniques.
- When innovations not Normal $\rightarrow$ PMLE standard errors (“sandwich form”)

Stochastic Volatility models (idea)

- In ARCH-GARCH models, the variance at time t, σ_t^2 is completely determined by the information at time $t-1$, i.e. σ_t^2 it is conditionally deterministic or $\mathcal{F}_{t-1}$ measurable
- Another possibility is to have σ_t^2 being also a (positive) stochastic process i.e. variance is also affected by an idiosyncratic noise term $\Rightarrow$ Stochastic Volatility models
- Example: the Heston model

$$\begin{aligned} dP(t) &= \mu P(t)dt + \sqrt{h(t)}P(t)dW^P(t) \\ h(t) &= k(\theta - h(t)) + \nu\sqrt{h(t)}dW^h(t) \qquad \text{CIR process} \end{aligned}$$

where $dW^P(t)$ and $dW^h(t)$ are two (possibly correlated) Brownian processes.