ECMT3150: The Econometrics of Financial Markets

2a. Conditional Heteroskedastic Models

Simon Kwok
University of Sydney

Semester 1, 2021

Outline

Volatility

Q: What is Volatility?

- Volatility is a measure of the degree of asset price/market fluctuation.
- Volatility is high during more turbulent periods/market crashes, and low during quiet or booming periods.

Q: Why do should we care about Volatility?

- It is important for pricing options and other financial derivatives.
- VIX volatility index is being traded on CBOE in the form of futures.

Some Stylized Facts

daily

Q: What are some stylized facts about stock returns?

- Stock return has mean that is close to zero.
- Stock return tends to be serially uncorrelated.[1]

Q: What are some stylized facts about volatility?

- Volatility is not directly observable.
- Volatility level of stock price typically goes up fast and then decays slowly over time.
- Volatility clustering: volatility level is autocorrelated, i.e., high volatility periods tend to cluster together (and similarly for low volatility periods).

ARCH effect

- Leverage effect: volatility level tends to be inversely associated with price changes.

[1] This is not true for high frequency data, however.

Model Structure

P_t: asset price at time t.
r_t: log-return of the asset at time t, given by

$$r_t = \log(P_t) - \log(P_{t-1}).$$

Let $\mathcal{F}_{t-1}$ be the information set at time $t-1$. The conditional mean and variance of r_t are:

$$\begin{aligned} \mu_t &= E(r_t|\mathcal{F}_{t-1}), && (1) \\ \sigma_t^2 &= Var(r_t|\mathcal{F}_{t-1}) = E[(r_t - \mu_t)^2|\mathcal{F}_{t-1}]. && (2) \end{aligned}$$

We need a parametric model to describe the evolution of μ_t and σ_t^2 over time.
Let $a_t = r_t - \mu_t$. The dynamic regression model is:

$$r_t = \mu_t + a_t.$$

The error a_t may exhibit *conditional* heteroskedasticity, so that $\sigma_t^2 := Var\left(a_t|\mathcal{F}_{t-1}\right)$ is not a constant over t.

Model Structure

- We may want to explain r_t with its lagged terms $r_{t-1}, r_{t-2}, \ldots, r_{t-p}$:

$(t)-1$ μ_t

$t-1$ t

$$\begin{aligned}
\mu_t &= E(r_t|\mathcal{F}_{t-1}) \\
&= \phi_0 + \phi_1 r_{t-1} + \cdots + \phi_p r_{t-p}.
\end{aligned}$$

- Sometimes we may want to include additional (weakly) exogenous regressors $\mathbf{x}_t$ as well:

$$\begin{aligned}
\mu_t &= E(r_t|\mathcal{F}_{t-1}, \mathbf{x}_t) \\
&= \phi_0 + \cancel{\phi_1 r_{t-1}} + \cdots + \cancel{\phi_p r_{t-p}} + \gamma' \mathbf{x}_t.
\end{aligned}$$

eg r_{t-1}^2

- A popular class of models for σ_t^2 is the *(generalized) autoregressive conditional heteroskedasticity* model, or ARCH (GARCH) model.

Testing for ARCH Effect

Let $\hat{a}_t = r_t - \hat{\mu}_t$ be the residuals from the dynamic regression model.

data cond. mean.

Q: How to check for conditional heteroskedasticity (ARCH effect)?

proxy for volatility.

A: Two approaches, both require choosing a maximum lag order m for testing.

- Ljung-Box test $Q(m)$ on $\{\hat{a}_t^2\}$ (McLeod-Li (1983)).
- Lagrange multiplier (LM) test (Engle (1982)). Carry out an F test in the linear regression ($t = m+1, \ldots, T$)

$$\hat{a}_t^2 = \alpha_0 + \alpha_1 \hat{a}_{t-1}^2 + \cdots + \alpha_m \hat{a}_{t-m}^2 + e_t. \tag{3}$$

We want to test $H_0 : \alpha_i = 0$ for all $i = 1, \ldots, m$ vs $H_a : \alpha_i \neq 0$ for some $i = 1, \ldots, m$.
Define the restricted and unrestricted sums of squared residuals: $SSR_r = \sum_{t=1}^{T} (\hat{a}_t^2 - \overline{\hat{a}^2})^2$ and $SSR_{ur} = \sum_{t=1}^{T} \hat{e}_t^2$, where $\overline{\hat{a}^2} = \frac{1}{T} \sum_{t=1}^{T} \hat{a}_t^2$, and $\hat{e}$ is the residuals from regression (3).
The F test statistic is

$$F = \frac{(SSR_r - SSR_{ur})/m}{SSR_{ur}/(T - 2m - 1)},$$

ARCH Model

$a_t \sim ARCH(m)$ if   log-return $\mu_t = 0$

model error $\Rightarrow$ $a_t = \sigma_t \varepsilon_t,$

where $\varepsilon_t \sim iid(0,1)$ (i.e., *iid* with mean 0 and variance 1), and

$$\boxed{\sigma_t^2 = \alpha_0 + \sum_{i=1}^{m} \alpha_i a_{t-i}^2,} \tag{4}$$

proxy for vola. at $t-i$.

$\sigma_t^2 = Var(a_t | \mathcal{F}_{t-1})$ is interpreted as the conditional variance at time t given the history.

$\alpha_i > 0$ due to vola. clustering.

If a_t is weakly stationary, the *unconditional* (long-run) variance $\sigma^2 := Var(a_t)$ exists. We may take unconditional expectations on both sides of (4) and solve for σ^2:

$$\sigma^2 = \frac{\alpha_0}{1 - \sum_{i=1}^{m} \alpha_i}.$$

$\neq$

The necessary condition for the existence of σ^2 is $\sum_{i=1}^{m} \alpha_i < 1$.

GARCH Model

$\mu_t = 0$ (hence pure GARCH model).

$a_t \sim GARCH(m, s)$ if

$$a_t = \sigma_t \varepsilon_t, \tag{5}$$

where $\boxed{\varepsilon_t \sim iid(0, 1)}$, and

$$\boxed{\sigma_t^2 = \alpha_0 + \sum_{i=1}^{m} \alpha_i a_{t-i}^2 + \sum_{i=1}^{s} \beta_i \sigma_{t-i}^2.} \tag{6}$$

$\sigma^2_{t+1} = \alpha_0 + \sum \alpha_i a^2_{t+1-i} + \sum \beta_i \sigma^2_{t+1-i}$

If a_t is weakly stationary, the *unconditional* (long-run) variance exists and is equal to: $\alpha_1 a_t^2 + \alpha_2 a_{t-1}^2 + \dots$

$$\sigma^2 = \frac{\alpha_0}{1 - \sum_{i=1}^{m} \alpha_i - \sum_{i=1}^{s} \beta_i}. \tag{7}$$

The necessary condition for the existence of σ^2 is

$$\sum_{i=1}^{m} \alpha_i + \sum_{i=1}^{s} \beta_i < 1. \tag{8}$$

ARMA Representation of Squared Errors GARCH(m,s).

error of squared log-return.

Let $\eta_t := a_t^2 - \sigma_t^2$. Recall that $\sigma_t^2 = Var(a_t|\mathcal{F}_{t-1}) = E(a_t^2|\mathcal{F}_{t-1})$, so we see that $\eta_t = a_t^2 - E(a_t^2|\mathcal{F}_{t-1})$.

Ex: Show that $\eta_t \sim mds$, i.e., $E(\eta_t|\mathcal{F}_s) = 0$ for all $t > s$.

We may then rewrite (6) as follows:

$$\begin{aligned}
\overbrace{a_t^2 - \eta_t}^{\sigma_t^2} &= \alpha_0 + \sum_{i=1}^{m} \alpha_i a_{t-i}^2 + \sum_{i=1}^{s} \beta_i \overbrace{(a_{t-i}^2 - \eta_{t-i})}^{\sigma_{t-i}^2} \\
a_t^2 &= \alpha_0 + \underbrace{\sum_{i=1}^{\max(m,s)} (\alpha_i + \beta_i) a_{t-i}^2}_{AR} + \underbrace{\eta_t - \sum_{i=1}^{s} \beta_i \eta_{t-i}}_{MA.}. \qquad (9)
\end{aligned}$$

In (9), we adopt the convention that $\alpha_i = 0$ for $i > m$, and $\beta_i = 0$ for $i > s$.

Interpretation: The squared innovation $a_t^2 \sim ARMA(\max(m,s), s)$ with mds errors η_t.

GARCH Volatility Forecasting

known by time t (due to GARCH) assumption.

At the forecast origin t, the 1-step ahead forecast of σ^2_{t+1} is

$$
\begin{aligned}
\hat{\sigma}^2_t(1) &= E(\sigma^2_{t+1}|\mathcal{F}_t) \quad \text{(see p.20 of slides 1)} \\
&= \sigma^2_{t+1} \quad (\text{since } \sigma^2_{t+1} \text{ is measurable to } \mathcal{F}_t) \\
&= \alpha_0 + \sum_{i=1}^{m} \alpha_i a^2_{t+1-i} + \sum_{i=1}^{s} \beta_i \sigma^2_{t+1-i}.
\end{aligned}
$$

Optimal forecast of σ^2_{t+1}

$\leftarrow$ from (6).

Q: What is the ℓ-step ahead forecast of $\sigma^2_{t+\ell}$?

$\hat{\sigma}^2_t(1) = E(\sigma^2_{t+1} | \mathcal{F}_t)$. $\leftarrow$ min. MSE.

GARCH Volatility Forecasting

A: First, using (5), rewrite the recursive equation (6) as

$$\sigma^2_{t+\ell} = \alpha_0 + \underbrace{\sum_{i=1}^{m} \alpha_i \sigma^2_{t+\ell-i}\varepsilon^2_{t+\ell-i}}_{a^2_{t+\ell-i}} + \sum_{i=1}^{s} \beta_i \sigma^2_{t+\ell-i} + \sum \alpha_i \sigma^2_{t+\ell-i} - \sum \alpha_i \sigma^2_{t+\ell-i}$$

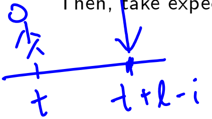
$\sigma^2_{t+\ell-i}$ is known with $\mathcal{F}_{t+\ell-i}$

$$= \alpha_0 + \sum_{i=1}^{\max(m,s)} (\alpha_i + \beta_i)\sigma^2_{t+\ell-i} + \sum_{i=1}^{m} \alpha_i \sigma^2_{t+\ell-i}(\varepsilon^2_{t+\ell-i} - 1).$$

ARMA rep.

error.

Then, take expectations conditional on $\mathcal{F}_t$, and note that



$$E(\sigma^2_{t+\ell-i}|\mathcal{F}_t) = \begin{cases} \hat{\sigma}^2_t(\ell - i) & \text{if } \ell - i > 0, \\ \sigma^2_{t+\ell-i} & \text{otherwise;} \end{cases}$$

constant as at time t.

and

Tower's rule

$$E(\sigma^2_{t+\ell-i}(\varepsilon^2_{t+\ell-i} - 1)|\mathcal{F}_t) = \begin{cases} E[\sigma^2_{t+\ell-i}E(\varepsilon^2_{t+\ell-i} - 1|\mathcal{F}_{t+\ell-i-1})|\mathcal{F}_t] & \text{if } \ell - i > 0, \\ \sigma^2_{t+\ell-i}(\varepsilon^2_{t+\ell-i} - 1) & \text{otherwise,} \end{cases}$$

$\sigma^2_{t+\ell-i}$ is known at time $t+\ell-i$

by iid of ε_t

$$= \begin{cases} E[\sigma^2_{t+\ell-i}E(\varepsilon^2_{t+\ell-i} - 1)|\mathcal{F}_t] = 0 & \text{if } \ell - i > 0, \\ a^2_{t+\ell-i} - \sigma^2_{t+\ell-i} & \text{otherwise.} \end{cases}$$

$E(\varepsilon^2_{t+\ell-i}) = Var(\varepsilon_{t+\ell-i}) = 1$

GARCH Volatility Forecasting

The ℓ-step ahead forecast of $\sigma^2_{t+\ell}$ is thus given by $E(\sigma^2_{t+\ell-i}(\varepsilon^2_{t+\ell-i}-1)|\mathcal{F}_t)$

$E(\sigma^2_{t+\ell}|\mathcal{F}_t)$ $\qquad$ $E(\sigma^2_{t+\ell-i}|\mathcal{F}_t)$

$$\hat{\sigma}^2_t(\ell) = \alpha_0 + \sum_{i=1}^{\max(m,s)} (\alpha_i + \beta_i)\hat{\sigma}^2_t(\ell - i) + \sum_{i=\ell}^{m} \alpha_i (a^2_{t+\ell-i} - \sigma^2_{t+\ell-i}).$$

We adopt the convention that $\hat{\sigma}^2_t(\ell - i) = \sigma^2_{t+\ell-i}$ if $\ell - i \leq 0$.
The last summation is gone if $\ell > m$.

E.g.: For the $GARCH(1,1)$ model (i.e., $m = s = 1$), the ℓ-step ahead forecast of $\sigma^2_{t+\ell}$ is:

$$\underbrace{\hat{\sigma}^2_t(\ell)}_{c} = \begin{cases} \alpha_0 + \alpha_1 a^2_t + \beta_1 \sigma^2_t & \text{if } \ell = 1, \\ \alpha_0 + (\alpha_1 + \beta_1)\underbrace{\hat{\sigma}^2_t(\ell - 1)}_{c} & \text{if } \ell > 1. \end{cases}$$

$\Rightarrow c = \dfrac{\alpha_0}{1-(\alpha_1+\beta_1)}$

Q: What happens to $\hat{\sigma}^2_t(\ell)$ as $\ell \to \infty$?
A: If $\sum_{i=1}^{\max(m,s)} (\alpha_i + \beta_i) < 1$, then $\hat{\sigma}^2_t(\ell)$ converges to the long-run variance $\alpha_0 \Big/ \left[1 - \sum_{i=1}^{\max(m,s)} (\alpha_i + \beta_i)\right]$.

$m, s = 1.$

Properties of GARCH(1,1) Model

The GARCH(1,1) model is given by $E(r_t | \mathcal{F}_{t-1})$.

$$
\begin{aligned}
r_t &= \mu_t + a_t, \\
a_t &= \sigma_t \varepsilon_t, \\
\sigma_t^2 &= \alpha_0 + \alpha_1 a_{t-1}^2 + \beta_1 \sigma_{t-1}^2,
\end{aligned}
$$

with $\alpha_1, \beta_1 > 0$ and $\alpha_1 + \beta_1 < 1$, ε_t is iid noise with $E(\varepsilon_t) = 0$, and $Var(\varepsilon_t) = E(\varepsilon_t^2) = 1$. eg. $\varepsilon_t \sim N(0,1) \Rightarrow K_\varepsilon = 0$

$\varepsilon_t \sim t_4 \Rightarrow K_\varepsilon > 0$

By (1) and (2), μ_t and σ_t^2 are measurable with respect to the information set $\mathcal{F}_{t-1}$.

We assume for now that r_t is covariance stationary (so that its first two unconditional moments are constant over time).

Conditional Mean and Variance of GARCH(1,1)

Q: What are the conditional mean and variance of r_t?

$$
\begin{aligned}
E(r_t|\mathcal{F}_{t-1}) &= E(\mu_t + a_t|\mathcal{F}_{t-1}) \\
&= \mu_t + E(a_t|\mathcal{F}_{t-1}) \\
&= \mu_t + \sigma_t E(\varepsilon_t|\mathcal{F}_{t-1}) \\
&= \mu_t + \sigma_t E(\varepsilon_t) \\
&= \mu_t.
\end{aligned}
$$

$$
\begin{aligned}
Var(r_t|\mathcal{F}_{t-1}) &= E[(r_t - \mu_t)^2|\mathcal{F}_{t-1}] \\
&= E(a_t^2|\mathcal{F}_{t-1}) \\
&= E(\sigma_t^2\varepsilon_t^2|\mathcal{F}_{t-1}) \\
&= \sigma_t^2 E(\varepsilon_t^2|\mathcal{F}_{t-1}) \\
= \sigma_t^2 \quad &= \sigma_t^2 E(\varepsilon_t^2) = \sigma_t^2 \\
&= (\alpha_0 + \alpha_1 a_{t-1}^2 + \beta_1\sigma_{t-1}^2) \times 1 \\
&= \alpha_0 + \alpha_1 a_{t-1}^2 + \beta_1\sigma_{t-1}^2.
\end{aligned}
$$

Unconditional Variance of GARCH(1,1)

$E(r_t) = E(\mu_t)$

uncond. mean

Q: What is the unconditional variance of r_t?

A: By the law of total variance,

$$\begin{aligned} Var(r_t) &= E[Var(r_t|\mathcal{F}_{t-1})] + Var[E(r_t|\mathcal{F}_{t-1})] \\ &= E(\alpha_0 + \alpha_1 a_{t-1}^2 + \beta_1 \sigma_{t-1}^2) + Var(\mu_t) \\ &= \alpha_0 + \alpha_1 E(a_{t-1}^2) + \beta_1 E(\sigma_{t-1}^2) + 0 \\ &= \alpha_0 + \alpha_1 Var(r_{t-1}) + \beta_1 E(\sigma_{t-1}^2). \end{aligned}$$

By the stationarity of r_t, $Var(r_t) = Var(r_{t-1})$, i.e., the variance is constant over t. Also,

$$\begin{aligned} Var(r_t) &= E(a_t^2) \\ &= E[\sigma_t^2 E(\varepsilon_t^2|\mathcal{F}_{t-1})] \\ &= E(\sigma_t^2 E(\varepsilon_t^2)) \\ &= E(\sigma_t^2). \end{aligned}$$

By letting $V := Var(r_t) = Var(r_{t-1}) = E(\sigma_t^2)$, we have

$$V = \alpha_0 + \alpha_1 V + \beta_1 V.$$

Provided that $|\alpha_1 + \beta_1| < 1$, the solution exists, and is equal to

$$V := Var(r_t) = \frac{\alpha_0}{1 - \alpha_1 - \beta_1}.$$

Unconditional Kurtosis of GARCH(1,1)

Assume further that r_t is stationary up to the fourth moment, that $E(\varepsilon_t^4)$ exists, and $E(\varepsilon_t^4) = 3 + K_\varepsilon$ (i.e., K_ε is the excess kurtosis of ε_t).

Q: What is the unconditional excess kurtosis of r_t?

A: We start with the conditional fourth central moment of r_t

$$\begin{aligned}
& E[(r_t - \mu_t)^4 | \mathcal{F}_{t-1}] \\
= \quad & E(a_t^4 | \mathcal{F}_{t-1}) \\
= \quad & E(\sigma_t^4 \varepsilon_t^4 | \mathcal{F}_{t-1}) \\
= \quad & \sigma_t^4 E(\varepsilon_t^4) \\
= \quad & (\alpha_0 + \alpha_1 a_{t-1}^2 + \beta_1 \sigma_{t-1}^2)^2 (3 + K_\varepsilon) \\
= \quad & (\alpha_0^2 + \alpha_1^2 a_{t-1}^4 + \beta_1^2 \sigma_{t-1}^4 + 2\alpha_0\alpha_1 a_{t-1}^2 + 2\alpha_0\beta_1\sigma_{t-1}^2 + 2\alpha_1\beta_1 a_{t-1}^2 \sigma_{t-1}^2)(3 + K_\varepsilon).
\end{aligned}$$

Let us then work out the unconditional fourth moment of a_t

$$\begin{aligned}
E(a_t^4) \quad = \quad & E[(r_t - \mu_t)^4] \\
= \quad & (3 + K_\varepsilon)[\alpha_0^2 + \alpha_1^2 E(a_{t-1}^4) + \beta_1^2 E(\sigma_{t-1}^4) \\
& + 2\alpha_0\alpha_1 E(a_{t-1}^2) + 2\alpha_0\beta_1 E(\sigma_{t-1}^2) + 2\alpha_1\beta_1 E(a_{t-1}^2\sigma_{t-1}^2)]. \qquad (10)
\end{aligned}$$

Unconditional Kurtosis of GARCH(1,1)

By iterated expectations, the last expectation can be simplified as follows

$$
\begin{aligned}
E(a_{t-1}^2\sigma_{t-1}^2) &= E[E(a_{t-1}^2\sigma_{t-1}^2|\mathcal{F}_{t-2})] \\
&= E[\sigma_{t-1}^2 E(a_{t-1}^2|\mathcal{F}_{t-2})] \\
&= E[\sigma_{t-1}^2 E(\sigma_{t-1}^2\varepsilon_{t-1}^2|\mathcal{F}_{t-2})] \\
&= E[\sigma_{t-1}^4 E(\varepsilon_{t-1}^2)] = E(\sigma_{t-1}^4).
\end{aligned}
$$

By stationarity in the fourth moment, we have $E(a_t^4) = E(a_{t-1}^4)$. By iterated expectations,

$$
\begin{aligned}
E(a_t^4) &= E(\sigma_t^4\varepsilon_t^4) = E[\sigma_t^4 E(\varepsilon_t^4|\mathcal{F}_{t-1})] \\
&= E[\sigma_t^4 E(\varepsilon_t^4)] = E(\sigma_t^4)(3+K_\varepsilon),
\end{aligned}
$$

so that $E(\sigma_t^4) = E(a_t^4)/(3+K_\varepsilon)$. We can simplify (10) as follows

$$
\begin{aligned}
E(a_t^4) &= (3+K_\varepsilon)[\alpha_0^2 + \alpha_1^2 E(a_t^4)] + (\beta_1^2 + 2\alpha_1\beta_1)E(a_t^4) \\
&\quad + (3+K_\varepsilon)(2\alpha_0\alpha_1 + 2\alpha_0\beta_1)E(a_t^2).
\end{aligned}
$$

Together with the fact that $Var(r_t) = E(a_t^2) = E(\sigma_t^2) = \frac{\alpha_0}{1-\alpha_1-\beta_1}$, we may then solve for $E(a_t^4)$:

$$
[1-(3+K_\varepsilon)\alpha_1^2 - \beta_1^2 - 2\alpha_1\beta_1]E(a_t^4) = (3+K_\varepsilon)[\alpha_0^2 + 2\alpha_0(\alpha_1+\beta_1)E(a_t^2)],
$$

so that

$$
E(a_t^4) = \frac{(3+K_\varepsilon)[\alpha_0^2 + \frac{2\alpha_0^2(\alpha_1+\beta_1)}{1-\alpha_1-\beta_1}]}{1-(3+K_\varepsilon)\alpha_1^2 - \beta_1^2 - 2\alpha_1\beta_1} = \frac{\alpha_0^2(3+K_\varepsilon)(1+\alpha_1+\beta_1)}{(1-\alpha_1-\beta_1)[1-2\alpha_1^2 - K_\varepsilon\alpha_1^2 - (\alpha_1+\beta_1)^2]}.
$$

Unconditional Kurtosis of GARCH(1,1)

$\mathcal{N}($ $\quad E[(X-\mu)^4]=3.$

$X \sim N(0,1) \quad K_X = 0$

excess kur

$:= E[(X-\mu)^4]-3.$

The unconditional kurtosis of r_t is thus

$$\begin{aligned}\frac{E[(r_t-\mu_t)^4]}{\{E[(r_t-\mu_t)^2]\}^2} &= \frac{(3+K_\varepsilon)(1+\alpha_1+\beta_1)(1-\alpha_1-\beta_1)}{1-2\alpha_1^2-K_\varepsilon\alpha_1^2-(\alpha_1+\beta_1)^2} \\ &= \frac{[1-(\alpha_1+\beta_1)^2](3+K_\varepsilon)}{1-2\alpha_1^2-K_\varepsilon\alpha_1^2-(\alpha_1+\beta_1)^2}.\end{aligned}$$

By simple algebra, the unconditional excess kurtosis is given by

$$K_r := \frac{E[(r_t-\mu_t)^4]}{\{E[(r_t-\mu_t)^2]\}^2} - 3 = \frac{6\alpha_1^2 + K_\varepsilon[1-(\alpha_1+\beta_1)^2+3\alpha_1^2]}{1-2\alpha_1^2-K_\varepsilon\alpha_1^2-(\alpha_1+\beta_1)^2}. > 0$$

ARCH par

We see that K_r is zero if $\alpha_1 = 0$ and $K_\varepsilon = 0$. In other words, the ARCH parameter α_1 and the excess kurtosis of ε_t contribute to the fat-tailedness of the unconditional distribution of the log-return r_t.