

Empirical Properties of Asset Returns

Econ 643: Financial Economics II

Nikolay Gospodinov
Department of Economics, Concordia University

January 5, 2011

- Time series and distributional properties of asset returns

- Time series and distributional properties of asset returns
 - conditional mean

- Time series and distributional properties of asset returns
 - conditional mean
 - conditional variance: measuring and modeling volatility of returns

- Time series and distributional properties of asset returns
 - conditional mean
 - conditional variance: measuring and modeling volatility of returns
 - modeling non-normality of the unconditional distribution of standardized asset returns

- Time series and distributional properties of asset returns
 - conditional mean
 - conditional variance: measuring and modeling volatility of returns
 - modeling non-normality of the unconditional distribution of standardized asset returns
- Knowledge of the conditional moments and distribution of the underlying asset returns is of key importance for

- Time series and distributional properties of asset returns
 - conditional mean
 - conditional variance: measuring and modeling volatility of returns
 - modeling non-normality of the unconditional distribution of standardized asset returns
- Knowledge of the conditional moments and distribution of the underlying asset returns is of key importance for
 - derivative pricing (pricing of bonds and options, futures and forward contracts, and other more exotic derivative products)

- Time series and distributional properties of asset returns
 - conditional mean
 - conditional variance: measuring and modeling volatility of returns
 - modeling non-normality of the unconditional distribution of standardized asset returns
- Knowledge of the conditional moments and distribution of the underlying asset returns is of key importance for
 - derivative pricing (pricing of bonds and options, futures and forward contracts, and other more exotic derivative products)
 - risk management

- Time series and distributional properties of asset returns
 - conditional mean
 - conditional variance: measuring and modeling volatility of returns
 - modeling non-normality of the unconditional distribution of standardized asset returns
- Knowledge of the conditional moments and distribution of the underlying asset returns is of key importance for
 - derivative pricing (pricing of bonds and options, futures and forward contracts, and other more exotic derivative products)
 - risk management
 - forecasting

- Time series and distributional properties of asset returns
 - conditional mean
 - conditional variance: measuring and modeling volatility of returns
 - modeling non-normality of the unconditional distribution of standardized asset returns
- Knowledge of the conditional moments and distribution of the underlying asset returns is of key importance for
 - derivative pricing (pricing of bonds and options, futures and forward contracts, and other more exotic derivative products)
 - risk management
 - forecasting
 - approaching critically some assumptions about the underlying asset returns that make the pricing of contingent claims analytically more tractable

Returns

- Let S_t denote the price of an asset (stock, exchange rate, commodity) at time t .

Returns

- Let S_t denote the price of an asset (stock, exchange rate, commodity) at time t .
- The return from holding the asset from period $t - 1$ to period t is

$$r_t = \frac{S_t - S_{t-1}}{S_{t-1}} = \frac{S_t}{S_{t-1}} - 1$$

or $S_t = S_{t-1}(1 + r_t)$.

- Let S_t denote the price of an asset (stock, exchange rate, commodity) at time t .
- The return from holding the asset from period $t - 1$ to period t is

$$r_t = \frac{S_t - S_{t-1}}{S_{t-1}} = \frac{S_t}{S_{t-1}} - 1$$

or $S_t = S_{t-1}(1 + r_t)$.

- The k -period return can be defined similarly as

$$\begin{aligned} r_t[k] &= \frac{S_t}{S_{t-k}} - 1 = \frac{S_t}{S_{t-1}} \frac{S_{t-1}}{S_{t-2}} \cdots \frac{S_{t-k+1}}{S_{t-k}} - 1 \\ &= (1 + r_t)(1 + r_{t-1}) \cdots (1 + r_{t-k+1}) - 1 \\ &= \prod_{j=0}^{k-1} (1 + r_{t-j}) - 1. \end{aligned}$$

Continuously compounded returns

- Suppose now that instead of holding the asset for the whole period, you reinvest the return m times between time $t - 1$ and t .

Continuously compounded returns

- Suppose now that instead of holding the asset for the whole period, you reinvest the return m times between time $t - 1$ and t .
- In this case, $S_t = S_{t-1}(1 + r_t/m)^m$ and, as $m \rightarrow \infty$, $S_t = S_{t-1}e^{r_t}$ or

$$r_t = \ln \left(\frac{S_t}{S_{t-1}} \right) = \ln(S_t) - \ln(S_{t-1}). \quad (1)$$

We refer to (1) as the continuously compounded or log return.

Continuously compounded returns

- Suppose now that instead of holding the asset for the whole period, you reinvest the return m times between time $t - 1$ and t .
- In this case, $S_t = S_{t-1}(1 + r_t/m)^m$ and, as $m \rightarrow \infty$, $S_t = S_{t-1}e^{r_t}$ or

$$r_t = \ln\left(\frac{S_t}{S_{t-1}}\right) = \ln(S_t) - \ln(S_{t-1}). \quad (1)$$

We refer to (1) as the continuously compounded or log return.

- The k -period log return is given by

$$\begin{aligned} r_t[k] &= \ln\left(\frac{S_t}{S_{t-k}}\right) = \ln\left(\frac{S_t}{S_{t-1}} \frac{S_{t-1}}{S_{t-2}} \cdots \frac{S_{t-k+1}}{S_{t-k}}\right) \\ &= \ln\left(\frac{S_t}{S_{t-1}}\right) + \dots + \ln\left(\frac{S_{t-k+1}}{S_{t-k}}\right) = \sum_{j=0}^{k-1} r_{t-j}. \end{aligned}$$

If the asset pays a dividend D_t , $r_t = \ln(S_t + D_t) - \ln(S_{t-1})$.

Continuously compounded returns

- Suppose now that instead of holding the asset for the whole period, you reinvest the return m times between time $t - 1$ and t .
- In this case, $S_t = S_{t-1}(1 + r_t/m)^m$ and, as $m \rightarrow \infty$, $S_t = S_{t-1}e^{r_t}$ or

$$r_t = \ln\left(\frac{S_t}{S_{t-1}}\right) = \ln(S_t) - \ln(S_{t-1}). \quad (1)$$

We refer to (1) as the continuously compounded or log return.

- The k -period log return is given by

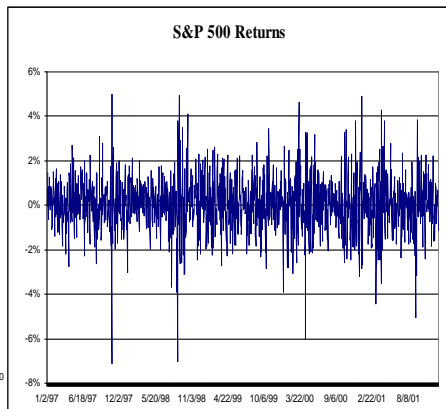
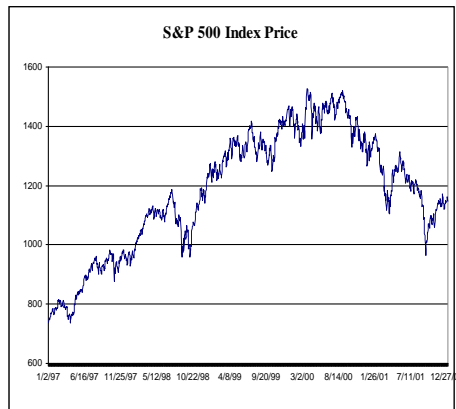
$$\begin{aligned} r_t[k] &= \ln\left(\frac{S_t}{S_{t-k}}\right) = \ln\left(\frac{S_t}{S_{t-1}} \frac{S_{t-1}}{S_{t-2}} \cdots \frac{S_{t-k+1}}{S_{t-k}}\right) \\ &= \ln\left(\frac{S_t}{S_{t-1}}\right) + \dots + \ln\left(\frac{S_{t-k+1}}{S_{t-k}}\right) = \sum_{j=0}^{k-1} r_{t-j}. \end{aligned}$$

If the asset pays a dividend D_t , $r_t = \ln(S_t + D_t) - \ln(S_{t-1})$.

- The excess return of the asset over the risk-free rate is computed as $er_r = r_t - r_{t0}$, where r_{t0} denotes the return on the risk-free asset.

Properties of asset returns

- While asset prices are characterized by high persistence and possible nonstationarity, asset returns exhibit very little or no autocorrelation.
 - Since at high frequency (intradaily, daily) $E(r_t) \approx 0$, $E_{t-1}(r_t) = 0$ in applied work.

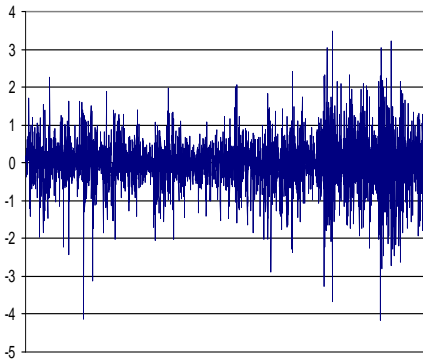


Properties of asset returns

British Pound / Deutsche Mark Daily Exchange Rate

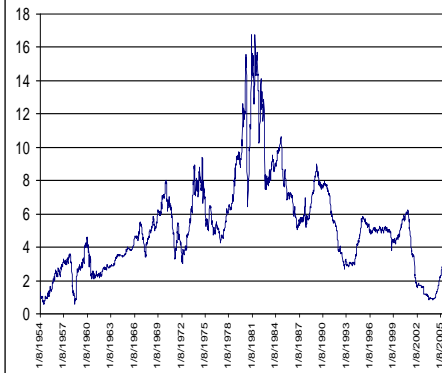


BP/DM Exchange Rate Daily Returns

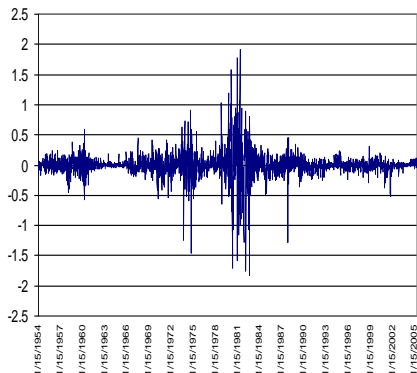


Properties of asset returns

3-Month US Treasury Bill at Weekly Frequency

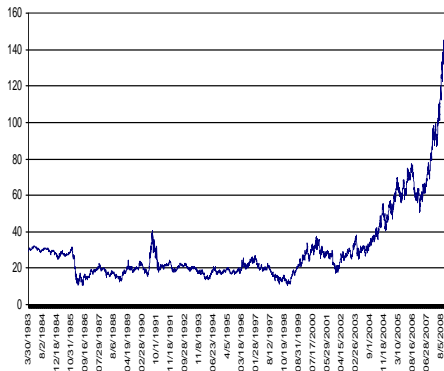


3-Month US Treasury Bill Weekly Returns

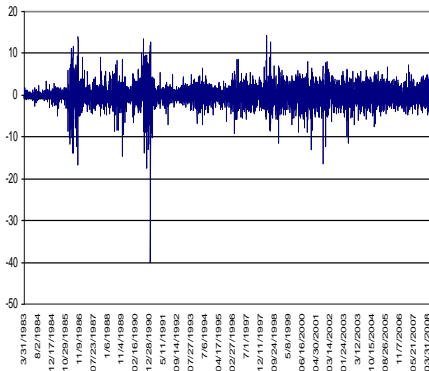


Properties of asset returns

Daily Oil Price

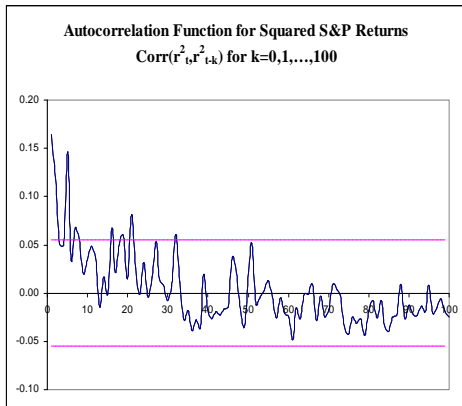
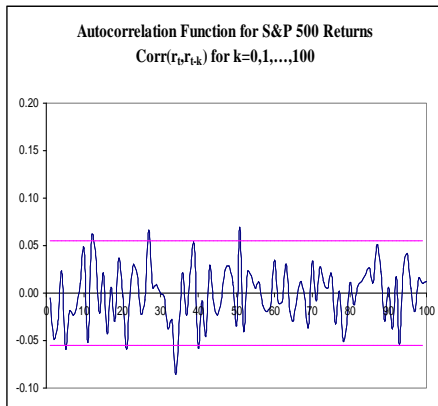


Daily Oil Returns

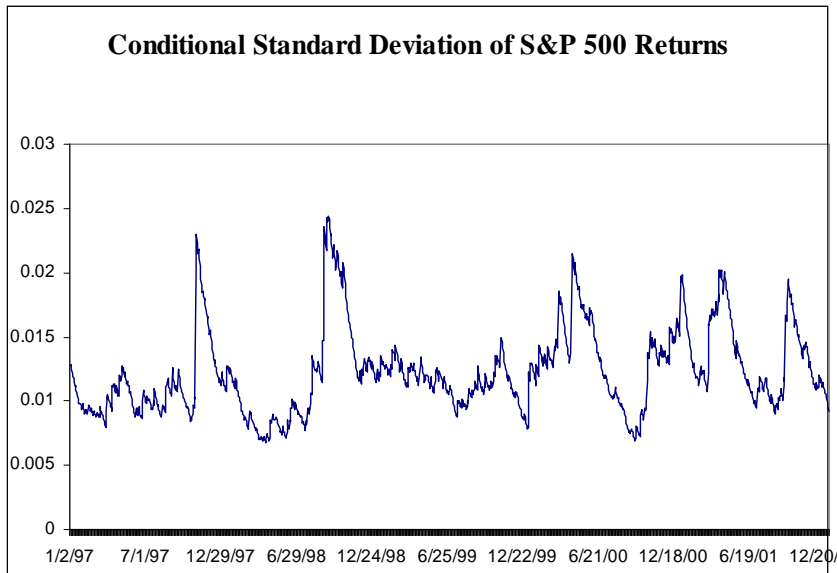


Properties of asset returns

- The squared asset returns, which are often used as a proxy of the unobserved volatility, exhibit positive autocorrelation.
 - This indicates the need to model the conditional variance of the asset returns $Var_{t-1}(r_t)$ as a time-varying process.

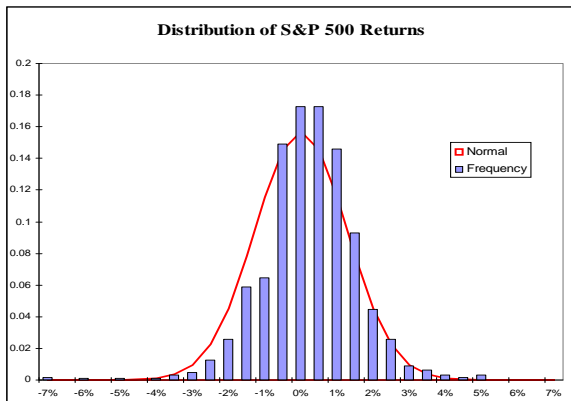


Properties of asset returns



Properties of asset returns

- At high frequency (intradaily, daily), the unconditional distribution of raw and standardized returns has fatter tails than normal distribution. The distribution of stock returns is also negatively skewed.
 - As lower frequency (monthly, quarterly etc.), the unconditional distribution of returns approaches the normal distribution.



Properties of asset returns

- The above empirical properties suggest that the dynamics of the asset returns can be modeled as

$$r_t = \sigma_t \varepsilon_t,$$

where $\sigma_t^2 = \text{Var}_{t-1}(r_t)$ and ε_t is an *iid* error term with mean 0 and variance 1 drawn from a distribution with fatter tails than the standard normal distribution.

Properties of asset returns

- The above empirical properties suggest that the dynamics of the asset returns can be modeled as

$$r_t = \sigma_t \varepsilon_t,$$

where $\sigma_t^2 = \text{Var}_{t-1}(r_t)$ and ε_t is an *iid* error term with mean 0 and variance 1 drawn from a distribution with fatter tails than the standard normal distribution.

- The stock returns are typically negatively correlated with the volatility.

Properties of asset returns

- The above empirical properties suggest that the dynamics of the asset returns can be modeled as

$$r_t = \sigma_t \varepsilon_t,$$

where $\sigma_t^2 = \text{Var}_{t-1}(r_t)$ and ε_t is an *iid* error term with mean 0 and variance 1 drawn from a distribution with fatter tails than the standard normal distribution.

- The stock returns are typically negatively correlated with the volatility.
 - This is referred to in the literature as leverage effect since a drop in the stock price, other things being equal, increases the debt-equity ratio (leverage) of the firm.

Properties of asset returns

- The above empirical properties suggest that the dynamics of the asset returns can be modeled as

$$r_t = \sigma_t \varepsilon_t,$$

where $\sigma_t^2 = \text{Var}_{t-1}(r_t)$ and ε_t is an *iid* error term with mean 0 and variance 1 drawn from a distribution with fatter tails than the standard normal distribution.

- The stock returns are typically negatively correlated with the volatility.
 - This is referred to in the literature as leverage effect since a drop in the stock price, other things being equal, increases the debt-equity ratio (leverage) of the firm.
- The covariances and the correlations between assets appear to be time varying. This has important implications for managing portfolio risk.

- Modeling and forecasting volatility is important for option pricing and risk management.

Conditional variance

- Modeling and forecasting volatility is important for option pricing and risk management.
- Unfortunately, the volatility is not directly observable from the returns.

Conditional variance

- Modeling and forecasting volatility is important for option pricing and risk management.
- Unfortunately, the volatility is not directly observable from the returns.
 - Since there is only one observation of returns at time t , it is difficult to construct a measure of volatility for this particular period.

- Modeling and forecasting volatility is important for option pricing and risk management.
- Unfortunately, the volatility is not directly observable from the returns.
 - Since there is only one observation of returns at time t , it is difficult to construct a measure of volatility for this particular period.
 - One natural but very noisy estimator of the conditional variance σ_t^2 are the squared returns r_t^2 for period t .

- Modeling and forecasting volatility is important for option pricing and risk management.
- Unfortunately, the volatility is not directly observable from the returns.
 - Since there is only one observation of returns at time t , it is difficult to construct a measure of volatility for this particular period.
 - One natural but very noisy estimator of the conditional variance σ_t^2 are the squared returns r_t^2 for period t .
- There are several alternative approaches for estimating the unobserved volatility process.

Rolling average

- The volatility process can be estimated as a simple average of the most recent m observations of the squared returns as

$$\sigma_t^2 = \frac{1}{m} \sum_{s=1}^m r_{t-s}^2$$

for $t = 1, 2, \dots$, and m fixed.

Rolling average

- The volatility process can be estimated as a simple average of the most recent m observations of the squared returns as

$$\sigma_t^2 = \frac{1}{m} \sum_{s=1}^m r_{t-s}^2$$

for $t = 1, 2, \dots$, and m fixed.

- The above rolling average assigns equal weight to all r_{t-s}^2 .

Rolling average

- The volatility process can be estimated as a simple average of the most recent m observations of the squared returns as

$$\sigma_t^2 = \frac{1}{m} \sum_{s=1}^m r_{t-s}^2$$

for $t = 1, 2, \dots$, and m fixed.

- The above rolling average assigns equal weight to all r_{t-s}^2 .
- It is appropriate to give larger weights to more recent observations as in

$$\sigma_t^2 = \sum_{s=1}^m \alpha_s r_{t-s}^2,$$

where $\alpha_l < \alpha_k$ if $l > k$ and $\sum_{s=1}^m \alpha_s = 1$. This is the exponentially weighted moving average model.

Exponentially weighted moving average (EWMA) model

- Let $\alpha_{s+1} = \lambda\alpha_s$, where $0 < \lambda < 1$. Then,

$$\sigma_t^2 = \lambda\sigma_{t-1}^2 + (1 - \lambda)r_{t-1}^2.$$

Exponentially weighted moving average (EWMA) model

- Let $\alpha_{s+1} = \lambda\alpha_s$, where $0 < \lambda < 1$. Then,

$$\sigma_t^2 = \lambda\sigma_{t-1}^2 + (1 - \lambda)r_{t-1}^2.$$

- By recursive substitution

$$\sigma_t^2 = (1 - \lambda) \sum_{s=1}^m \lambda^{s-1} r_{t-s}^2 + \lambda^m \sigma_0^2 \approx \sum_{s=1}^m \alpha_s r_{t-s}^2,$$

where $\alpha_s = (1 - \lambda)\lambda^{s-1}$ (exponential smoothing) and $\lambda^m \sigma_0^2$ is small for large m and therefore is ignored.

Exponentially weighted moving average (EWMA) model

- Let $\alpha_{s+1} = \lambda\alpha_s$, where $0 < \lambda < 1$. Then,

$$\sigma_t^2 = \lambda\sigma_{t-1}^2 + (1 - \lambda)r_{t-1}^2.$$

- By recursive substitution

$$\sigma_t^2 = (1 - \lambda) \sum_{s=1}^m \lambda^{s-1} r_{t-s}^2 + \lambda^m \sigma_0^2 \approx \sum_{s=1}^m \alpha_s r_{t-s}^2,$$

where $\alpha_s = (1 - \lambda)\lambda^{s-1}$ (exponential smoothing) and $\lambda^m \sigma_0^2$ is small for large m and therefore is ignored.

- If λ is small, the weight given to r_{t-1}^2 is large and the volatility estimates are highly volatile since r_{t-1}^2 is volatile.

Exponentially weighted moving average (EWMA) model

- Let $\alpha_{s+1} = \lambda\alpha_s$, where $0 < \lambda < 1$. Then,

$$\sigma_t^2 = \lambda\sigma_{t-1}^2 + (1 - \lambda)r_{t-1}^2.$$

- By recursive substitution

$$\sigma_t^2 = (1 - \lambda) \sum_{s=1}^m \lambda^{s-1} r_{t-s}^2 + \lambda^m \sigma_0^2 \approx \sum_{s=1}^m \alpha_s r_{t-s}^2,$$

where $\alpha_s = (1 - \lambda)\lambda^{s-1}$ (exponential smoothing) and $\lambda^m \sigma_0^2$ is small for large m and therefore is ignored.

- If λ is small, the weight given to r_{t-1}^2 is large and the volatility estimates are highly volatile since r_{t-1}^2 is volatile.
- If λ is close to one, the volatility estimates respond relatively slowly to new information provided by r_{t-1}^2 .

Exponentially weighted moving average (EWMA) model

- Let $\alpha_{s+1} = \lambda\alpha_s$, where $0 < \lambda < 1$. Then,

$$\sigma_t^2 = \lambda\sigma_{t-1}^2 + (1 - \lambda)r_{t-1}^2.$$

- By recursive substitution

$$\sigma_t^2 = (1 - \lambda) \sum_{s=1}^m \lambda^{s-1} r_{t-s}^2 + \lambda^m \sigma_0^2 \approx \sum_{s=1}^m \alpha_s r_{t-s}^2,$$

where $\alpha_s = (1 - \lambda)\lambda^{s-1}$ (exponential smoothing) and $\lambda^m \sigma_0^2$ is small for large m and therefore is ignored.

- If λ is small, the weight given to r_{t-1}^2 is large and the volatility estimates are highly volatile since r_{t-1}^2 is volatile.
- If λ is close to one, the volatility estimates respond relatively slowly to new information provided by r_{t-1}^2 .
- **EWMA with $\lambda = 0.94$ is typically used for updating daily volatilities.**

GARCH models

- One of the most popular models for modeling the time-varying volatility of asset returns is the generalized autoregressive conditionally heteroskedastic model of order one, GARCH(1,1).

GARCH models

- One of the most popular models for modeling the time-varying volatility of asset returns is the generalized autoregressive conditionally heteroskedastic model of order one, GARCH(1,1).
- The GARCH(1,1) model parameterizes $\sigma_t^2 = \text{Var}_{t-1}(r_t)$ as

$$\sigma_t^2 = \omega + \gamma r_{t-1}^2 + \beta \sigma_{t-1}^2, \quad \omega > 0, \quad \beta, \gamma \geq 0. \quad (2)$$

- One of the most popular models for modeling the time-varying volatility of asset returns is the generalized autoregressive conditionally heteroskedastic model of order one, GARCH(1,1).
- The GARCH(1,1) model parameterizes $\sigma_t^2 = \text{Var}_{t-1}(r_t)$ as

$$\sigma_t^2 = \omega + \gamma r_{t-1}^2 + \beta \sigma_{t-1}^2, \quad \omega > 0, \quad \beta, \gamma \geq 0. \quad (2)$$

- If $\beta + \gamma < 1$, the GARCH(1,1) process is weakly stationary and its unconditional variance is given by

$$E(\sigma_t^2) = \sigma^2 = \omega + \beta \sigma^2 + \gamma \sigma^2 = \frac{\omega}{1 - (\beta + \gamma)}.$$

- One of the most popular models for modeling the time-varying volatility of asset returns is the generalized autoregressive conditionally heteroskedastic model of order one, GARCH(1,1).
- The GARCH(1,1) model parameterizes $\sigma_t^2 = \text{Var}_{t-1}(r_t)$ as

$$\sigma_t^2 = \omega + \gamma r_{t-1}^2 + \beta \sigma_{t-1}^2, \quad \omega > 0, \quad \beta, \gamma \geq 0. \quad (2)$$

- If $\beta + \gamma < 1$, the GARCH(1,1) process is weakly stationary and its unconditional variance is given by

$$E(\sigma_t^2) = \sigma^2 = \omega + \beta \sigma^2 + \gamma \sigma^2 = \frac{\omega}{1 - (\beta + \gamma)}.$$

- Furthermore, the marginal distribution of r_t can have fat tails even if ε_t is standard normal. From $r_t = \varepsilon_t \sigma_t$, we have

$$E(r_t^4) = E(\varepsilon_t^4)E(\sigma_t^4) \geq E(\varepsilon_t^4) \left[E(\sigma_t^2) \right]^2 = 3\sigma^4$$

since $E[(\sigma_t^2)^2] \geq [E(\sigma_t^2)]^2$ by Jensen's inequality and $E(\varepsilon_t^4) = 3$ by the normality assumption. Then, the kurtosis of r_t is $\kappa = \frac{E(r_t^4)}{\sigma^4} \geq 3$.

- Adding r_t^2 to both sides of (2) and moving σ_t^2 to the right-hand side shows that GARCH(1,1) can be written as

$$\begin{aligned}r_t^2 &= \omega + (\beta + \gamma)r_{t-1}^2 + \eta_t - \beta\eta_{t-1} \\ &= (1 - \beta - \gamma)\frac{\omega}{1 - \beta - \gamma} + (\beta + \gamma)r_{t-1}^2 + \eta_t - \beta\eta_{t-1} \\ &= \sigma^2 + (\beta + \gamma)(r_{t-1}^2 - \sigma^2) + \eta_t - \beta\eta_{t-1}\end{aligned}$$

where $\eta_t = r_t^2 - \sigma_t^2 = \sigma_t^2(\varepsilon_t^2 - 1)$ and $E_{t-1}(\eta_t) = 0$.

- Adding r_t^2 to both sides of (2) and moving σ_t^2 to the right-hand side shows that GARCH(1,1) can be written as

$$\begin{aligned}r_t^2 &= \omega + (\beta + \gamma)r_{t-1}^2 + \eta_t - \beta\eta_{t-1} \\ &= (1 - \beta - \gamma)\frac{\omega}{1 - \beta - \gamma} + (\beta + \gamma)r_{t-1}^2 + \eta_t - \beta\eta_{t-1} \\ &= \sigma^2 + (\beta + \gamma)(r_{t-1}^2 - \sigma^2) + \eta_t - \beta\eta_{t-1}\end{aligned}$$

where $\eta_t = r_t^2 - \sigma_t^2 = \sigma_t^2(\varepsilon_t^2 - 1)$ and $E_{t-1}(\eta_t) = 0$.

- If $0 \leq \beta + \gamma < 1$ and r_{t-1}^2 is larger (smaller) than σ^2 , then r_t^2 is expected to be larger (smaller) than σ^2 (volatility clustering).

- Adding r_t^2 to both sides of (2) and moving σ_t^2 to the right-hand side shows that GARCH(1,1) can be written as

$$\begin{aligned}r_t^2 &= \omega + (\beta + \gamma)r_{t-1}^2 + \eta_t - \beta\eta_{t-1} \\ &= (1 - \beta - \gamma)\frac{\omega}{1 - \beta - \gamma} + (\beta + \gamma)r_{t-1}^2 + \eta_t - \beta\eta_{t-1} \\ &= \sigma^2 + (\beta + \gamma)(r_{t-1}^2 - \sigma^2) + \eta_t - \beta\eta_{t-1}\end{aligned}$$

where $\eta_t = r_t^2 - \sigma_t^2 = \sigma_t^2(\varepsilon_t^2 - 1)$ and $E_{t-1}(\eta_t) = 0$.

- If $0 \leq \beta + \gamma < 1$ and r_{t-1}^2 is larger (smaller) than σ^2 , then r_t^2 is expected to be larger (smaller) than σ^2 (volatility clustering).
- In applications, we typically find that $\hat{\beta} + \hat{\gamma} \approx 1$. When $\beta + \gamma = 1$, we have the integrated GARCH (IGARCH).

- Adding r_t^2 to both sides of (2) and moving σ_t^2 to the right-hand side shows that GARCH(1,1) can be written as

$$\begin{aligned}r_t^2 &= \omega + (\beta + \gamma)r_{t-1}^2 + \eta_t - \beta\eta_{t-1} \\ &= (1 - \beta - \gamma)\frac{\omega}{1 - \beta - \gamma} + (\beta + \gamma)r_{t-1}^2 + \eta_t - \beta\eta_{t-1} \\ &= \sigma^2 + (\beta + \gamma)(r_{t-1}^2 - \sigma^2) + \eta_t - \beta\eta_{t-1}\end{aligned}$$

where $\eta_t = r_t^2 - \sigma_t^2 = \sigma_t^2(\varepsilon_t^2 - 1)$ and $E_{t-1}(\eta_t) = 0$.

- If $0 \leq \beta + \gamma < 1$ and r_{t-1}^2 is larger (smaller) than σ^2 , then r_t^2 is expected to be larger (smaller) than σ^2 (volatility clustering).
- In applications, we typically find that $\hat{\beta} + \hat{\gamma} \approx 1$. When $\beta + \gamma = 1$, we have the integrated GARCH (IGARCH).
 - It is easy to see that the EWMA model is an IGARCH model with $\omega = 0$, $\beta = \lambda$ and $\gamma = 1 - \lambda$ ($\beta + \gamma = \lambda + 1 - \lambda = 1$).

- Adding r_t^2 to both sides of (2) and moving σ_t^2 to the right-hand side shows that GARCH(1,1) can be written as

$$\begin{aligned}r_t^2 &= \omega + (\beta + \gamma)r_{t-1}^2 + \eta_t - \beta\eta_{t-1} \\ &= (1 - \beta - \gamma)\frac{\omega}{1 - \beta - \gamma} + (\beta + \gamma)r_{t-1}^2 + \eta_t - \beta\eta_{t-1} \\ &= \sigma^2 + (\beta + \gamma)(r_{t-1}^2 - \sigma^2) + \eta_t - \beta\eta_{t-1}\end{aligned}$$

where $\eta_t = r_t^2 - \sigma_t^2 = \sigma_t^2(\varepsilon_t^2 - 1)$ and $E_{t-1}(\eta_t) = 0$.

- If $0 \leq \beta + \gamma < 1$ and r_{t-1}^2 is larger (smaller) than σ^2 , then r_t^2 is expected to be larger (smaller) than σ^2 (volatility clustering).
- In applications, we typically find that $\hat{\beta} + \hat{\gamma} \approx 1$. When $\beta + \gamma = 1$, we have the integrated GARCH (IGARCH).
 - It is easy to see that the EWMA model is an IGARCH model with $\omega = 0$, $\beta = \lambda$ and $\gamma = 1 - \lambda$ ($\beta + \gamma = \lambda + 1 - \lambda = 1$).
 - **But IGARCH is incompatible with intuition since the shocks to volatility should not persist infinitely long.**

GARCH models: Forecasting

- Forecasts of volatility are important for derivative pricing and risk management.

GARCH models: Forecasting

- Forecasts of volatility are important for derivative pricing and risk management.
 - For example, to price an option expiring in n periods, we need the average future volatility $\frac{1}{n} \sum_{k=0}^{n-1} E(\sigma_{t+k}^2)$.

GARCH models: Forecasting

- Forecasts of volatility are important for derivative pricing and risk management.
 - For example, to price an option expiring in n periods, we need the average future volatility $\frac{1}{n} \sum_{k=0}^{n-1} E(\sigma_{t+k}^2)$.
- Using that the unconditional variance is $\sigma^2 = \omega / (1 - \beta - \gamma)$, we can rewrite the GARCH(1,1) model as

$$\sigma_t^2 - \sigma^2 = \beta(\sigma_{t-1}^2 - \sigma^2) + \gamma(r_{t-1}^2 - \sigma^2)$$

GARCH models: Forecasting

- Forecasts of volatility are important for derivative pricing and risk management.
 - For example, to price an option expiring in n periods, we need the average future volatility $\frac{1}{n} \sum_{k=0}^{n-1} E(\sigma_{t+k}^2)$.
- Using that the unconditional variance is $\sigma^2 = \omega / (1 - \beta - \gamma)$, we can rewrite the GARCH(1,1) model as

$$\sigma_t^2 - \sigma^2 = \beta(\sigma_{t-1}^2 - \sigma^2) + \gamma(r_{t-1}^2 - \sigma^2)$$

- At time $t + k$,

$$\sigma_{t+k}^2 - \sigma^2 = \beta(\sigma_{t+k-1}^2 - \sigma^2) + \gamma(r_{t+k-1}^2 - \sigma^2)$$

or

$$E(\sigma_{t+k}^2) = \sigma^2 + (\beta + \gamma)^k (\sigma_t^2 - \sigma^2)$$

by repeated substitution and using that $E(r_{t+k-1}^2) = \sigma_{t+k-1}^2$.

GARCH models: Forecasting

- Forecasts of volatility are important for derivative pricing and risk management.
 - For example, to price an option expiring in n periods, we need the average future volatility $\frac{1}{n} \sum_{k=0}^{n-1} E(\sigma_{t+k}^2)$.
- Using that the unconditional variance is $\sigma^2 = \omega / (1 - \beta - \gamma)$, we can rewrite the GARCH(1,1) model as

$$\sigma_t^2 - \sigma^2 = \beta(\sigma_{t-1}^2 - \sigma^2) + \gamma(r_{t-1}^2 - \sigma^2)$$

- At time $t + k$,

$$\sigma_{t+k}^2 - \sigma^2 = \beta(\sigma_{t+k-1}^2 - \sigma^2) + \gamma(r_{t+k-1}^2 - \sigma^2)$$

or

$$E(\sigma_{t+k}^2) = \sigma^2 + (\beta + \gamma)^k (\sigma_t^2 - \sigma^2)$$

by repeated substitution and using that $E(r_{t+k-1}^2) = \sigma_{t+k-1}^2$.

- Note that if $\beta + \gamma = 1$ (IGARCH or EWMA model), $E(\sigma_{t+k}^2) = \sigma_t^2$.

GARCH models: Forecasting

- Forecasts of volatility are important for derivative pricing and risk management.
 - For example, to price an option expiring in n periods, we need the average future volatility $\frac{1}{n} \sum_{k=0}^{n-1} E(\sigma_{t+k}^2)$.
- Using that the unconditional variance is $\sigma^2 = \omega / (1 - \beta - \gamma)$, we can rewrite the GARCH(1,1) model as

$$\sigma_t^2 - \sigma^2 = \beta(\sigma_{t-1}^2 - \sigma^2) + \gamma(r_{t-1}^2 - \sigma^2)$$

- At time $t + k$,

$$\sigma_{t+k}^2 - \sigma^2 = \beta(\sigma_{t+k-1}^2 - \sigma^2) + \gamma(r_{t+k-1}^2 - \sigma^2)$$

or

$$E(\sigma_{t+k}^2) = \sigma^2 + (\beta + \gamma)^k (\sigma_t^2 - \sigma^2)$$

by repeated substitution and using that $E(r_{t+k-1}^2) = \sigma_{t+k-1}^2$.

- Note that if $\beta + \gamma = 1$ (IGARCH or EWMA model), $E(\sigma_{t+k}^2) = \sigma_t^2$.
- If $\beta + \gamma < 1$, the variance forecast exhibits a reversion towards the unconditional variance σ^2 .

GARCH models: Estimation

- Since $r_t = \sigma_t \varepsilon_t$, the log likelihood for the t^{th} observation is given by

$$l_t(\theta) = \ln \left[f_\varepsilon \left(\frac{r_t}{\sigma_t} \right) \right] - \frac{1}{2} \ln(\sigma_t^2),$$

where $f_\varepsilon(\varepsilon_t)$ is the density of ε_t and θ are the parameters of the conditional variance equation.

GARCH models: Estimation

- Since $r_t = \sigma_t \varepsilon_t$, the log likelihood for the t^{th} observation is given by

$$l_t(\theta) = \ln \left[f_\varepsilon \left(\frac{r_t}{\sigma_t} \right) \right] - \frac{1}{2} \ln(\sigma_t^2),$$

where $f_\varepsilon(\varepsilon_t)$ is the density of ε_t and θ are the parameters of the conditional variance equation.

- If we further assume that $\varepsilon_t \sim N(0, 1)$, then $f_\varepsilon(\varepsilon_t) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{\varepsilon_t^2}{2}\right)$ and the Gaussian quasi-likelihood is

$$l_t(\theta) = -\frac{1}{2} \ln(2\pi) - \frac{1}{2} \ln(\sigma_t^2) - \frac{r_t^2}{2\sigma_t^2}.$$

GARCH models: Estimation

- Since $r_t = \sigma_t \varepsilon_t$, the log likelihood for the t^{th} observation is given by

$$l_t(\theta) = \ln \left[f_\varepsilon \left(\frac{r_t}{\sigma_t} \right) \right] - \frac{1}{2} \ln(\sigma_t^2),$$

where $f_\varepsilon(\varepsilon_t)$ is the density of ε_t and θ are the parameters of the conditional variance equation.

- If we further assume that $\varepsilon_t \sim N(0, 1)$, then $f_\varepsilon(\varepsilon_t) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{\varepsilon_t^2}{2})$ and the Gaussian quasi-likelihood is

$$l_t(\theta) = -\frac{1}{2} \ln(2\pi) - \frac{1}{2} \ln(\sigma_t^2) - \frac{r_t^2}{2\sigma_t^2}.$$

- The quasi-maximum likelihood estimator (QMLE) is the value of θ that maximizes $\sum_{t=1}^T l_t(\theta)$.

GARCH models: Estimation

- Since $r_t = \sigma_t \varepsilon_t$, the log likelihood for the t^{th} observation is given by

$$l_t(\theta) = \ln \left[f_\varepsilon \left(\frac{r_t}{\sigma_t} \right) \right] - \frac{1}{2} \ln(\sigma_t^2),$$

where $f_\varepsilon(\varepsilon_t)$ is the density of ε_t and θ are the parameters of the conditional variance equation.

- If we further assume that $\varepsilon_t \sim N(0, 1)$, then $f_\varepsilon(\varepsilon_t) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{\varepsilon_t^2}{2})$ and the Gaussian quasi-likelihood is

$$l_t(\theta) = -\frac{1}{2} \ln(2\pi) - \frac{1}{2} \ln(\sigma_t^2) - \frac{r_t^2}{2\sigma_t^2}.$$

- The quasi-maximum likelihood estimator (QMLE) is the value of θ that maximizes $\sum_{t=1}^T l_t(\theta)$.
 - For GARCH(1,1), $\omega > 0, \beta \geq 0, \gamma \geq 0$ and $\beta + \gamma < 1$.

GARCH models: Estimation

- Since $r_t = \sigma_t \varepsilon_t$, the log likelihood for the t^{th} observation is given by

$$l_t(\theta) = \ln \left[f_\varepsilon \left(\frac{r_t}{\sigma_t} \right) \right] - \frac{1}{2} \ln(\sigma_t^2),$$

where $f_\varepsilon(\varepsilon_t)$ is the density of ε_t and θ are the parameters of the conditional variance equation.

- If we further assume that $\varepsilon_t \sim N(0, 1)$, then $f_\varepsilon(\varepsilon_t) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{\varepsilon_t^2}{2})$ and the Gaussian quasi-likelihood is

$$l_t(\theta) = -\frac{1}{2} \ln(2\pi) - \frac{1}{2} \ln(\sigma_t^2) - \frac{r_t^2}{2\sigma_t^2}.$$

- The quasi-maximum likelihood estimator (QMLE) is the value of θ that maximizes $\sum_{t=1}^T l_t(\theta)$.
 - For GARCH(1,1), $\omega > 0, \beta \geq 0, \gamma \geq 0$ and $\beta + \gamma < 1$.
 - **Constrained optimization to estimate the parameters of the model.**

GARCH models: Estimation

- Since $r_t = \sigma_t \varepsilon_t$, the log likelihood for the t^{th} observation is given by

$$l_t(\theta) = \ln \left[f_\varepsilon \left(\frac{r_t}{\sigma_t} \right) \right] - \frac{1}{2} \ln(\sigma_t^2),$$

where $f_\varepsilon(\varepsilon_t)$ is the density of ε_t and θ are the parameters of the conditional variance equation.

- If we further assume that $\varepsilon_t \sim N(0, 1)$, then $f_\varepsilon(\varepsilon_t) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{\varepsilon_t^2}{2})$ and the Gaussian quasi-likelihood is

$$l_t(\theta) = -\frac{1}{2} \ln(2\pi) - \frac{1}{2} \ln(\sigma_t^2) - \frac{r_t^2}{2\sigma_t^2}.$$

- The quasi-maximum likelihood estimator (QMLE) is the value of θ that maximizes $\sum_{t=1}^T l_t(\theta)$.
 - For GARCH(1,1), $\omega > 0, \beta \geq 0, \gamma \geq 0$ and $\beta + \gamma < 1$.
 - Constrained optimization to estimate the parameters of the model.
 - **Note that the QML estimator is consistent even if the true distribution of ε_t is non-normal.**

GARCH models: Extensions

- We argued above that one of the stylized facts of stock returns is the existence of leverage effect.

GARCH models: Extensions

- We argued above that one of the stylized facts of stock returns is the existence of leverage effect.
- One easy way to incorporate the leverage effect into the GARCH model is by specifying the conditional variance as

$$\begin{aligned}\sigma_t^2 &= \omega + \gamma(r_{t-1} - \delta\sigma_{t-1})^2 + \beta\sigma_{t-1}^2 \\ &= \omega + \gamma\sigma_{t-1}^2(\varepsilon_{t-1} - \delta)^2 + \beta\sigma_{t-1}^2\end{aligned}$$

for $\delta > 0$.

GARCH models: Extensions

- We argued above that one of the stylized facts of stock returns is the existence of leverage effect.
- One easy way to incorporate the leverage effect into the GARCH model is by specifying the conditional variance as

$$\begin{aligned}\sigma_t^2 &= \omega + \gamma(r_{t-1} - \delta\sigma_{t-1})^2 + \beta\sigma_{t-1}^2 \\ &= \omega + \gamma\sigma_{t-1}^2(\varepsilon_{t-1} - \delta)^2 + \beta\sigma_{t-1}^2\end{aligned}$$

for $\delta > 0$.

- This specification allows for different impacts on volatility of positive and negative shocks (news) with negative shocks ($\varepsilon_{t-1} < 0$) having a greater impact on the conditional variance than positive shocks ($\varepsilon_{t-1} > 0$).

GARCH models: Extensions

- We argued above that one of the stylized facts of stock returns is the existence of leverage effect.
- One easy way to incorporate the leverage effect into the GARCH model is by specifying the conditional variance as

$$\begin{aligned}\sigma_t^2 &= \omega + \gamma(r_{t-1} - \delta\sigma_{t-1})^2 + \beta\sigma_{t-1}^2 \\ &= \omega + \gamma\sigma_{t-1}^2(\varepsilon_{t-1} - \delta)^2 + \beta\sigma_{t-1}^2\end{aligned}$$

for $\delta > 0$.

- This specification allows for different impacts on volatility of positive and negative shocks (news) with negative shocks ($\varepsilon_{t-1} < 0$) having a greater impact on the conditional variance than positive shocks ($\varepsilon_{t-1} > 0$).
- Further generalizations of the GARCH model include fractionally integrated (long-memory) GARCH, exponential GARCH, GARCH in mean etc.

Stochastic volatility (SV) model

- In GARCH models, the conditional volatility of the observed series is driven by the same shocks as its conditional mean.

Stochastic volatility (SV) model

- In GARCH models, the conditional volatility of the observed series is driven by the same shocks as its conditional mean.
 - Furthermore, conditional on the information set at time $t - 1$, the current volatility is deterministic.

Stochastic volatility (SV) model

- In GARCH models, the conditional volatility of the observed series is driven by the same shocks as its conditional mean.
 - Furthermore, conditional on the information set at time $t - 1$, the current volatility is deterministic.
- In SV models, σ_t^2 is subject to an additional shock

$$r_t = \varepsilon_t \sigma_t$$

$$\ln(\sigma_t^2) = \gamma_0 + \gamma_1 \ln(\sigma_{t-1}^2) + \gamma_2 z_t,$$

where ε_t and z_t are uncorrelated $N(0, 1)$ shocks.

Stochastic volatility (SV) model

- In GARCH models, the conditional volatility of the observed series is driven by the same shocks as its conditional mean.
 - Furthermore, conditional on the information set at time $t - 1$, the current volatility is deterministic.
- In SV models, σ_t^2 is subject to an additional shock

$$r_t = \varepsilon_t \sigma_t$$

$$\ln(\sigma_t^2) = \gamma_0 + \gamma_1 \ln(\sigma_{t-1}^2) + \gamma_2 z_t,$$

where ε_t and z_t are uncorrelated $N(0, 1)$ shocks.

- The random term z_t represents shocks to the intensity of the flow of new information as measured by σ_t^2 whereas ε_t represents the contents (large/small, positive/negative) of the news.

Stochastic volatility (SV) model

- In GARCH models, the conditional volatility of the observed series is driven by the same shocks as its conditional mean.
 - Furthermore, conditional on the information set at time $t - 1$, the current volatility is deterministic.
- In SV models, σ_t^2 is subject to an additional shock

$$r_t = \varepsilon_t \sigma_t$$

$$\ln(\sigma_t^2) = \gamma_0 + \gamma_1 \ln(\sigma_{t-1}^2) + \gamma_2 z_t,$$

where ε_t and z_t are uncorrelated $N(0, 1)$ shocks.

- The random term z_t represents shocks to the intensity of the flow of new information as measured by σ_t^2 whereas ε_t represents the contents (large/small, positive/negative) of the news.
- Also, z_t adds variability to σ_t^2 and in turn induces more kurtosis than the GARCH can produce.

Stochastic volatility model: Pros and cons

- In the GARCH class, the kurtosis is related to persistence via β and γ .

Stochastic volatility model: Pros and cons

- In the GARCH class, the kurtosis is related to persistence via β and γ .
- In the SV models, the persistence and kurtosis are modeled separately.

Stochastic volatility model: Pros and cons

- In the GARCH class, the kurtosis is related to persistence via β and γ .
- In the SV models, the persistence and kurtosis are modeled separately.
 - This reduces the link between the unconditional distribution (that gives the fat tails) and the conditional distribution (that gives rise to the clustering effect).

Stochastic volatility model: Pros and cons

- In the GARCH class, the kurtosis is related to persistence via β and γ .
- In the SV models, the persistence and kurtosis are modeled separately.
 - This reduces the link between the unconditional distribution (that gives the fat tails) and the conditional distribution (that gives rise to the clustering effect).
- The SV model has the further advantage of being an exact discretization of a continuous time model that will be considered later in this course.

Stochastic volatility model: Pros and cons

- In the GARCH class, the kurtosis is related to persistence via β and γ .
- In the SV models, the persistence and kurtosis are modeled separately.
 - This reduces the link between the unconditional distribution (that gives the fat tails) and the conditional distribution (that gives rise to the clustering effect).
- The SV model has the further advantage of being an exact discretization of a continuous time model that will be considered later in this course.
- **Drawback: the estimation of SV models is quite involved.**

Stochastic volatility model: Pros and cons

- In the GARCH class, the kurtosis is related to persistence via β and γ .
- In the SV models, the persistence and kurtosis are modeled separately.
 - This reduces the link between the unconditional distribution (that gives the fat tails) and the conditional distribution (that gives rise to the clustering effect).
- The SV model has the further advantage of being an exact discretization of a continuous time model that will be considered later in this course.
- Drawback: the estimation of SV models is quite involved.
- It turns out that there are other, more direct and model-free, methods for constructing the unobserved volatility process.

- Suppose we are interested in monthly (daily) returns but we also have daily (intradaily, 5-minute) data. Define

$$r_{(m),t+s/m} = \ln(S_{t+s/m}) - \ln(S_{t+(s-1)/m}).$$

where m is the number of daily (intradaily) observations per month (day), $1/m$ is the return horizon and $t = \frac{1}{m}, \frac{2}{m}, \dots$ For example,

Realized volatility

- Suppose we are interested in monthly (daily) returns but we also have daily (intradaily, 5-minute) data. Define

$$r_{(m),t+s/m} = \ln(S_{t+s/m}) - \ln(S_{t+(s-1)/m}).$$

where m is the number of daily (intradaily) observations per month (day), $1/m$ is the return horizon and $t = \frac{1}{m}, \frac{2}{m}, \dots$. For example,

- $m = 22$ for daily data (number of trading days per month).

Realized volatility

- Suppose we are interested in monthly (daily) returns but we also have daily (intradaily, 5-minute) data. Define

$$r_{(m),t+s/m} = \ln(S_{t+s/m}) - \ln(S_{t+(s-1)/m}).$$

where m is the number of daily (intradaily) observations per month (day), $1/m$ is the return horizon and $t = \frac{1}{m}, \frac{2}{m}, \dots$ For example,

- $m = 22$ for daily data (number of trading days per month).
- $m = 78$ for 5-minute returns in 6.5-hour (stock) markets.

Realized volatility

- Suppose we are interested in monthly (daily) returns but we also have daily (intradaily, 5-minute) data. Define

$$r_{(m),t+s/m} = \ln(S_{t+s/m}) - \ln(S_{t+(s-1)/m}).$$

where m is the number of daily (intradaily) observations per month (day), $1/m$ is the return horizon and $t = \frac{1}{m}, \frac{2}{m}, \dots$ For example,

- $m = 22$ for daily data (number of trading days per month).
- $m = 78$ for 5-minute returns in 6.5-hour (stock) markets.
- $m = 288$ for 5-minute returns for 24-hour (foreign exchange) markets.

- Suppose we are interested in monthly (daily) returns but we also have daily (intradaily, 5-minute) data. Define

$$r_{(m),t+s/m} = \ln(S_{t+s/m}) - \ln(S_{t+(s-1)/m}).$$

where m is the number of daily (intradaily) observations per month (day), $1/m$ is the return horizon and $t = \frac{1}{m}, \frac{2}{m}, \dots$. For example,

- $m = 22$ for daily data (number of trading days per month).
 - $m = 78$ for 5-minute returns in 6.5-hour (stock) markets.
 - $m = 288$ for 5-minute returns for 24-hour (foreign exchange) markets.
- We construct the volatility of the asset return at time $t + 1$ as

$$RV_{(m),t+1} \equiv \sum_{s=1}^m r_{(m),t+s/m}^2 \quad (3)$$

called realized (or integrated) volatility.

Realized volatility

- Suppose we are interested in monthly (daily) returns but we also have daily (intradaily, 5-minute) data. Define

$$r_{(m),t+s/m} = \ln(S_{t+s/m}) - \ln(S_{t+(s-1)/m}).$$

where m is the number of daily (intradaily) observations per month (day), $1/m$ is the return horizon and $t = \frac{1}{m}, \frac{2}{m}, \dots$. For example,

- $m = 22$ for daily data (number of trading days per month).
 - $m = 78$ for 5-minute returns in 6.5-hour (stock) markets.
 - $m = 288$ for 5-minute returns for 24-hour (foreign exchange) markets.
- We construct the volatility of the asset return at time $t + 1$ as

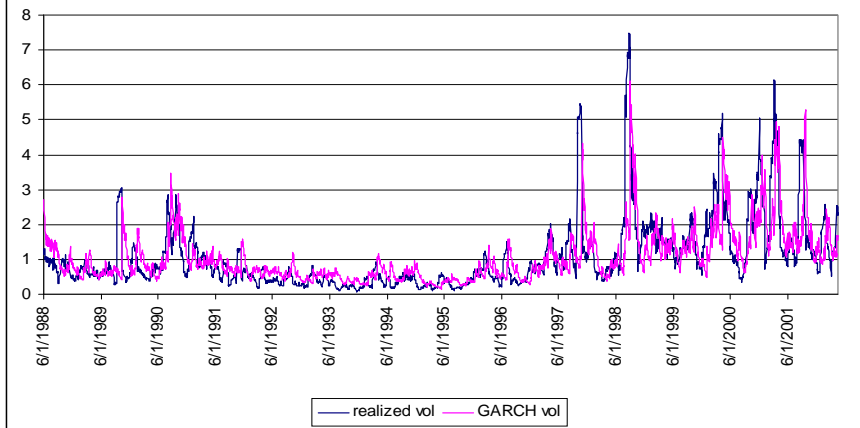
$$RV_{(m),t+1} \equiv \sum_{s=1}^m r_{(m),t+s/m}^2 \quad (3)$$

called realized (or integrated) volatility.

- As $m \rightarrow \infty$, $\sum_{s=1}^m r_{(m),t+s/m}^2 \rightarrow \int_0^1 \sigma_{t+\tau}^2 d\tau$ and the realized volatility converges to the quadratic variation of the process.

Realized and GARCH volatility

Realized and GARCH volatility of S&P 100 returns



Ranged-based volatility

- The realized volatility possesses nice properties when it is based on highly liquid securities but it is a noisy estimate of the true variance when it is constructed from assets that are traded infrequently and with large bid-ask spreads.

- The realized volatility possesses nice properties when it is based on highly liquid securities but it is a noisy estimate of the true variance when it is constructed from assets that are traded infrequently and with large bid-ask spreads.
- An alternative estimator that is more robust to market microstructure noise (but less efficient than the realized volatility in the absence of trading imperfections) is the range-based volatility computed as

$$\sigma_t^2 = c[\ln(S_t^{\max}) - \ln(S_t^{\min})]^2,$$

where $c = \frac{1}{4\ln(2)} \approx 0.361$ is a constant and $\ln(S_t^{\max}) - \ln(S_t^{\min})$ is the range between the logarithms of the highest and lowest prices observed during period t .

Realized bi-power variation

- It is common now in modern finance to allow for occasional, discontinuous jumps in the dynamics of the price process which add more mass to the tails of the empirical distribution of returns.

Realized bi-power variation

- It is common now in modern finance to allow for occasional, discontinuous jumps in the dynamics of the price process which add more mass to the tails of the empirical distribution of returns.
- Let κ denotes the size of these discrete jumps in the price process.

Realized bi-power variation

- It is common now in modern finance to allow for occasional, discontinuous jumps in the dynamics of the price process which add more mass to the tails of the empirical distribution of returns.
- Let κ denotes the size of these discrete jumps in the price process.
- Then, it can be shown that realized volatility

$$RV_{(m),t+1} \rightarrow \int_0^1 \sigma_{t+\tau}^2 d\tau + \sum_{0 < \tau \leq 1} \kappa^2(\tau) \text{ as } m \rightarrow \infty.$$

Realized bi-power variation

- It is common now in modern finance to allow for occasional, discontinuous jumps in the dynamics of the price process which add more mass to the tails of the empirical distribution of returns.
- Let κ denotes the size of these discrete jumps in the price process.
- Then, it can be shown that realized volatility

$$RV_{(m),t+1} \rightarrow \int_0^1 \sigma_{t+\tau}^2 d\tau + \sum_{0 < \tau \leq 1} \kappa^2(\tau) \text{ as } m \rightarrow \infty.$$

- *Realized bi-power variation is defined as*

$$\begin{aligned} RBP_{(m),t+1} &= \frac{\pi}{2} \sum_{s=2}^m \left| r_{(m),t+(s-1)/m} \right| \left| r_{(m),t+s/m} \right| \\ &\rightarrow \int_0^1 \sigma_{t+\tau}^2 d\tau \text{ as } m \rightarrow \infty. \end{aligned}$$

Realized bi-power variation

- It is common now in modern finance to allow for occasional, discontinuous jumps in the dynamics of the price process which add more mass to the tails of the empirical distribution of returns.
- Let κ denotes the size of these discrete jumps in the price process.
- Then, it can be shown that realized volatility

$$RV_{(m),t+1} \rightarrow \int_0^1 \sigma_{t+\tau}^2 d\tau + \sum_{0 < \tau \leq 1} \kappa^2(\tau) \text{ as } m \rightarrow \infty.$$

- *Realized bi-power variation* is defined as

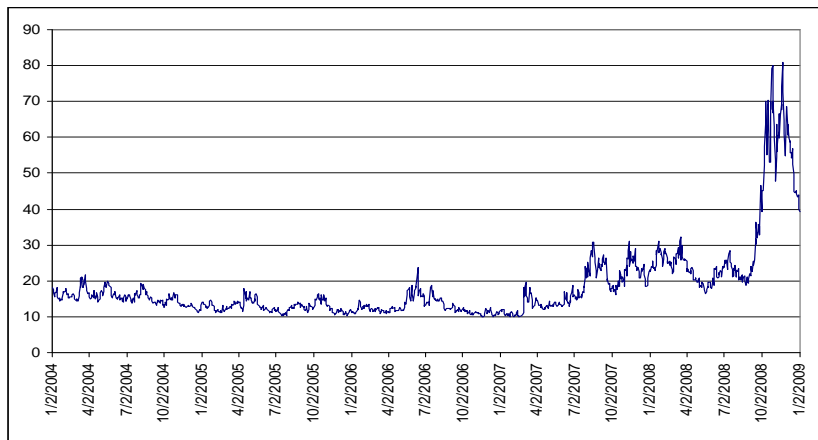
$$\begin{aligned} RBP_{(m),t+1} &= \frac{\pi}{2} \sum_{s=2}^m \left| r_{(m),t+(s-1)/m} \right| \left| r_{(m),t+s/m} \right| \\ &\rightarrow \int_0^1 \sigma_{t+\tau}^2 d\tau \text{ as } m \rightarrow \infty. \end{aligned}$$

- As a result, one could isolate the contribution of jump components as

$$RV_{(m),t+1} - RBP_{(m),t+1} \rightarrow \sum_{0 < \tau \leq 1} \kappa^2(\tau).$$

Implied volatility from option prices

- The volatility index VIX (known as market's "fear gauge") is constructed from S&P 500 (SPX) options and measures the market's expectations of volatility over the next 30-day period. More on that later in the course.



Conditional covariance

- Let $\sigma_{ij,t}$ denote the conditional covariance between assets i and j in period t .

Conditional covariance

- Let $\sigma_{ij,t}$ denote the conditional covariance between assets i and j in period t .
- Similarly to the conditional variance case, measures of the conditional covariance can be constructed as

Conditional covariance

- Let $\sigma_{ij,t}$ denote the conditional covariance between assets i and j in period t .
- Similarly to the conditional variance case, measures of the conditional covariance can be constructed as
 - historical rolling average

$$\sigma_{ij,t} = \frac{1}{m} \sum_{s=1}^m r_{i,t-s} r_{j,t-s}$$

Conditional covariance

- Let $\sigma_{ij,t}$ denote the conditional covariance between assets i and j in period t .
- Similarly to the conditional variance case, measures of the conditional covariance can be constructed as
 - historical rolling average

$$\sigma_{ij,t} = \frac{1}{m} \sum_{s=1}^m r_{i,t-s} r_{j,t-s}$$

- exponential smoother (or an integrated GARCH(1,1) model)

$$\sigma_{ij,t} = \lambda \sigma_{ij,t-1} + (1 - \lambda) r_{i,t-1} r_{j,t-1}$$

Conditional covariance

- Let $\sigma_{ij,t}$ denote the conditional covariance between assets i and j in period t .
- Similarly to the conditional variance case, measures of the conditional covariance can be constructed as
 - historical rolling average

$$\sigma_{ij,t} = \frac{1}{m} \sum_{s=1}^m r_{i,t-s} r_{j,t-s}$$

- exponential smoother (or an integrated GARCH(1,1) model)

$$\sigma_{ij,t} = \lambda \sigma_{ij,t-1} + (1 - \lambda) r_{i,t-1} r_{j,t-1}$$

- weakly stationary GARCH(1,1)

$$\sigma_{ij,t} = \omega_{ij} + \gamma r_{i,t-1} r_{j,t-1} + \beta \sigma_{ij,t-1}$$

with $\gamma + \beta < 1$.

Conditional covariance

- Let $\sigma_{ij,t}$ denote the conditional covariance between assets i and j in period t .
- Similarly to the conditional variance case, measures of the conditional covariance can be constructed as
 - historical rolling average

$$\sigma_{ij,t} = \frac{1}{m} \sum_{s=1}^m r_{i,t-s} r_{j,t-s}$$

- exponential smoother (or an integrated GARCH(1,1) model)

$$\sigma_{ij,t} = \lambda \sigma_{ij,t-1} + (1 - \lambda) r_{i,t-1} r_{j,t-1}$$

- weakly stationary GARCH(1,1)

$$\sigma_{ij,t} = \omega_{ij} + \gamma r_{i,t-1} r_{j,t-1} + \beta \sigma_{ij,t-1}$$

with $\gamma + \beta < 1$.

- Note that we restrict the persistence parameters λ , γ and β to be the same across assets in order to ensure a positive definite covariance matrix.

Conditional correlation

- Define $\rho_{ij,t} = \sigma_{ij,t} / (\sigma_{i,t}\sigma_{j,t})$ to be the conditional correlation between assets i and j in period t , where $\sigma_{i,t}$ and $\sigma_{j,t}$ denote their corresponding conditional standard deviations.

Conditional correlation

- Define $\rho_{ij,t} = \sigma_{ij,t} / (\sigma_{i,t}\sigma_{j,t})$ to be the conditional correlation between assets i and j in period t , where $\sigma_{i,t}$ and $\sigma_{j,t}$ denote their corresponding conditional standard deviations.
- Substituting the relevant expressions for $\sigma_{ij,t}$, $\sigma_{i,t}$ and $\sigma_{j,t}$ from above, we obtain

Conditional correlation

- Define $\rho_{ij,t} = \sigma_{ij,t} / (\sigma_{i,t}\sigma_{j,t})$ to be the conditional correlation between assets i and j in period t , where $\sigma_{i,t}$ and $\sigma_{j,t}$ denote their corresponding conditional standard deviations.
- Substituting the relevant expressions for $\sigma_{ij,t}$, $\sigma_{i,t}$ and $\sigma_{j,t}$ from above, we obtain
 - the rolling correlation estimator

$$\rho_{ij,t} = \frac{\sum_{s=1}^m r_{i,t-s} r_{j,t-s}}{\sqrt{\left(\sum_{s=1}^m r_{i,t-s}^2\right) \left(\sum_{s=1}^m r_{j,t-s}^2\right)}}$$

Conditional correlation

- Define $\rho_{ij,t} = \sigma_{ij,t} / (\sigma_{i,t}\sigma_{j,t})$ to be the conditional correlation between assets i and j in period t , where $\sigma_{i,t}$ and $\sigma_{j,t}$ denote their corresponding conditional standard deviations.
- Substituting the relevant expressions for $\sigma_{ij,t}$, $\sigma_{i,t}$ and $\sigma_{j,t}$ from above, we obtain
 - the rolling correlation estimator

$$\rho_{ij,t} = \frac{\sum_{s=1}^m r_{i,t-s} r_{j,t-s}}{\sqrt{\left(\sum_{s=1}^m r_{i,t-s}^2\right) \left(\sum_{s=1}^m r_{j,t-s}^2\right)}}$$

- the exponential smoothing correlation estimator

$$\begin{aligned}\rho_{ij,t} &= \frac{\lambda \sigma_{ij,t-1} + (1-\lambda) r_{i,t-1} r_{j,t-1}}{\sqrt{\left[\lambda \sigma_{i,t-1}^2 + (1-\lambda) r_{i,t-1}^2\right] \left[\lambda \sigma_{j,t-1}^2 + (1-\lambda) r_{j,t-1}^2\right]}} \\ &= \frac{\sum_{s=1}^m \lambda^{s-1} r_{i,t-s} r_{j,t-s}}{\sqrt{\left(\sum_{s=1}^m \lambda^{s-1} r_{i,t-s}^2\right) \left(\sum_{s=1}^m \lambda^{s-1} r_{j,t-s}^2\right)}},\end{aligned}$$

Conditional correlation

- Note that modeling the conditional correlation of raw returns $r_{i,t}$ and $r_{j,t}$ is equivalent to modeling the conditional covariance of the standardized returns $\varepsilon_{i,t} = r_{i,t}/\sigma_{i,t}$ and $\varepsilon_{j,t} = r_{j,t}/\sigma_{j,t}$ since

$$E_{t-1}(\varepsilon_{i,t}\varepsilon_{j,t}) = E_{t-1}\left(\frac{r_{i,t}}{\sigma_{i,t}}\frac{r_{j,t}}{\sigma_{j,t}}\right) = \frac{E_{t-1}(r_{i,t}r_{j,t})}{\sigma_{i,t}\sigma_{j,t}} = \frac{\sigma_{ij,t}}{\sigma_{i,t}\sigma_{j,t}} = \rho_{ij,t}.$$

Conditional correlation

- Note that modeling the conditional correlation of raw returns $r_{i,t}$ and $r_{j,t}$ is equivalent to modeling the conditional covariance of the standardized returns $\varepsilon_{i,t} = r_{i,t}/\sigma_{i,t}$ and $\varepsilon_{j,t} = r_{j,t}/\sigma_{j,t}$ since

$$E_{t-1}(\varepsilon_{i,t}\varepsilon_{j,t}) = E_{t-1}\left(\frac{r_{i,t}}{\sigma_{i,t}}\frac{r_{j,t}}{\sigma_{j,t}}\right) = \frac{E_{t-1}(r_{i,t}r_{j,t})}{\sigma_{i,t}\sigma_{j,t}} = \frac{\sigma_{ij,t}}{\sigma_{i,t}\sigma_{j,t}} = \rho_{ij,t}.$$

- Then, the conditional correlation estimator is $\rho_{ij,t} = \frac{q_{ij,t}}{\sqrt{q_{ii,t}q_{jj,t}}}$ whose dynamics is driven by the auxiliary variable $q_{ij,t}$ specified either as an exponential smoother (integrated GARCH(1,1))

$$q_{ij,t} = \lambda q_{ij,t-1} + (1 - \lambda)\varepsilon_{i,t-1}\varepsilon_{j,t-1} \quad (4)$$

or a weakly stationary GARCH(1,1)

$$q_{ij,t} = E(\varepsilon_{i,t-1}\varepsilon_{j,t-1}) + \gamma[\varepsilon_{i,t-1}\varepsilon_{j,t-1} - E(\varepsilon_{i,t-1}\varepsilon_{j,t-1})] + \beta[q_{ij,t-1} - E(\varepsilon_{i,t-1}\varepsilon_{j,t-1})] \quad (5)$$

with $\gamma + \beta < 1$.

Conditional correlation: Estimation

- The unobserved conditional correlation is typically estimated using a simple two-step procedure.

Conditional correlation: Estimation

- The unobserved conditional correlation is typically estimated using a simple two-step procedure.
 - First, the individual conditional variances are estimated separately for each asset using one of the methods described above and are then used to construct the standardized returns.

Conditional correlation: Estimation

- The unobserved conditional correlation is typically estimated using a simple two-step procedure.
 - First, the individual conditional variances are estimated separately for each asset using one of the methods described above and are then used to construct the standardized returns.
 - In the second step, the standardized returns are treated as actual observations and the parameters of the conditional correlation model are estimated by maximum likelihood.

Conditional correlation: Estimation

- The unobserved conditional correlation is typically estimated using a simple two-step procedure.
 - First, the individual conditional variances are estimated separately for each asset using one of the methods described above and are then used to construct the standardized returns.
 - In the second step, the standardized returns are treated as actual observations and the parameters of the conditional correlation model are estimated by maximum likelihood.
 - For example, in the case of two assets whose standardized returns are jointly normally distributed, the unknown parameters are estimated by maximizing the log-likelihood function

$$L = -\frac{1}{2} \sum_{t=1}^T \left[\ln(1 - \rho_{12,t}^2) + \frac{\varepsilon_{1,t}^2 + \varepsilon_{2,t}^2 - 2\rho_{12,t}\varepsilon_{1,t}\varepsilon_{2,t}}{1 - \rho_{12,t}^2} \right],$$

where $\rho_{12,t} = \frac{q_{12,t}}{\sqrt{q_{11,t}q_{22,t}}}$ and the dynamics of $q_{12,t}$, $q_{11,t}$ and $q_{22,t}$

is modeled either by (4) or (5) with $q_{12,0} = \frac{1}{T} \sum_{t=1}^T \varepsilon_{1,t}\varepsilon_{2,t}$, $q_{11,0} = 1$ and $q_{22,0} = 1$.

Realized covariance and correlation

- Let $r_{i,t+s/m}^{(m)} = \ln(S_{i,t+s/m}) - \ln(S_{i,t+(s-1)/m})$ be the m^{th} return for day t of asset i , where m is the number of intradaily observations for day t .

Realized covariance and correlation

- Let $r_{i,t+s/m}^{(m)} = \ln(S_{i,t+s/m}) - \ln(S_{i,t+(s-1)/m})$ be the m^{th} return for day t of asset i , where m is the number of intradaily observations for day t .
- Then, the realized covariance and correlation between assets i and j are given by

$$\sigma_{ij,t+1} = \sum_{s=1}^m r_{i,t+s/m}^{(m)} r_{j,t+s/m}^{(m)}$$

$$\rho_{ij,t+1} = \frac{\sum_{s=1}^m r_{i,t+s/m}^{(m)} r_{j,t+s/m}^{(m)}}{\sqrt{\left[\sum_{s=1}^m \left(r_{i,t+s/m}^{(m)} \right)^2 \right] \left[\sum_{s=1}^m \left(r_{j,t+s/m}^{(m)} \right)^2 \right]}}$$

respectively.

Realized covariance and correlation

- Let $r_{i,t+s/m}^{(m)} = \ln(S_{i,t+s/m}) - \ln(S_{i,t+(s-1)/m})$ be the m^{th} return for day t of asset i , where m is the number of intradaily observations for day t .
- Then, the realized covariance and correlation between assets i and j are given by

$$\sigma_{ij,t+1} = \sum_{s=1}^m r_{i,t+s/m}^{(m)} r_{j,t+s/m}^{(m)}$$

$$\rho_{ij,t+1} = \frac{\sum_{s=1}^m r_{i,t+s/m}^{(m)} r_{j,t+s/m}^{(m)}}{\sqrt{\left[\sum_{s=1}^m \left(r_{i,t+s/m}^{(m)} \right)^2 \right] \left[\sum_{s=1}^m \left(r_{j,t+s/m}^{(m)} \right)^2 \right]}}$$

respectively.

- The extension of range-based volatility to range-based covariance and correlation is less trivial.

Non-normality of returns

- As we argued above, the unconditional distribution of the (high-frequency) asset returns has fatter tails than the normal distribution even after allowing for conditional heteroskedasticity.

Non-normality of returns

- As we argued above, the unconditional distribution of the (high-frequency) asset returns has fatter tails than the normal distribution even after allowing for conditional heteroskedasticity.
- One popular choice of a fat-tailed distribution in economics and finance is the t -distribution.

Non-normality of returns

- As we argued above, the unconditional distribution of the (high-frequency) asset returns has fatter tails than the normal distribution even after allowing for conditional heteroskedasticity.
- One popular choice of a fat-tailed distribution in economics and finance is the t -distribution.
 - The shape of the t -distribution (in particular the thickness of the tails) is determined by the degrees of freedom parameter ν .

Non-normality of returns

- As we argued above, the unconditional distribution of the (high-frequency) asset returns has fatter tails than the normal distribution even after allowing for conditional heteroskedasticity.
- One popular choice of a fat-tailed distribution in economics and finance is the t -distribution.
 - The shape of the t -distribution (in particular the thickness of the tails) is determined by the degrees of freedom parameter ν .
 - For $\nu > 4$, the excess kurtosis of a standardized t -distributed random variable is $\frac{6}{\nu-4} > 0$, i.e. it exceeds the kurtosis of the standard normal distribution.

Non-normality of returns

- As we argued above, the unconditional distribution of the (high-frequency) asset returns has fatter tails than the normal distribution even after allowing for conditional heteroskedasticity.
- One popular choice of a fat-tailed distribution in economics and finance is the t -distribution.
 - The shape of the t -distribution (in particular the thickness of the tails) is determined by the degrees of freedom parameter ν .
 - For $\nu > 4$, the excess kurtosis of a standardized t -distributed random variable is $\frac{6}{\nu-4} > 0$, i.e. it exceeds the kurtosis of the standard normal distribution.
 - As ν approaches infinity, the t -distribution approaches the normal distribution.

Non-normality of returns

- As we argued above, the unconditional distribution of the (high-frequency) asset returns has fatter tails than the normal distribution even after allowing for conditional heteroskedasticity.
- One popular choice of a fat-tailed distribution in economics and finance is the t -distribution.
 - The shape of the t -distribution (in particular the thickness of the tails) is determined by the degrees of freedom parameter ν .
 - For $\nu > 4$, the excess kurtosis of a standardized t -distributed random variable is $\frac{6}{\nu-4} > 0$, i.e. it exceeds the kurtosis of the standard normal distribution.
 - As ν approaches infinity, the t -distribution approaches the normal distribution.
 - A drawback of the t -distribution is that the variance a t -distributed random variable $\frac{\nu}{\nu-2}$ is not defined uniformly (not defined for $\nu \leq 2$).

Non-normality of returns

- A more flexible distribution that is used to model the non-normality of asset returns is the generalized error distribution (GED) with density

$$f(\varepsilon) = \frac{\eta}{2\Gamma(\frac{1}{\eta})} \exp\left(-\frac{1}{2} |\varepsilon|^\eta\right),$$

where $\Gamma(\cdot)$ is the gamma function.

Non-normality of returns

- A more flexible distribution that is used to model the non-normality of asset returns is the generalized error distribution (GED) with density

$$f(\varepsilon) = \frac{\eta}{2\Gamma(\frac{1}{\eta})} \exp\left(-\frac{1}{2} |\varepsilon|^\eta\right),$$

where $\Gamma(\cdot)$ is the gamma function.

- If $\eta = 2$, $\Gamma\left(\frac{1}{2}\right) = \sqrt{2\pi}$ and $f(\varepsilon)$ is the standard normal density.

Non-normality of returns

- A more flexible distribution that is used to model the non-normality of asset returns is the generalized error distribution (GED) with density

$$f(\varepsilon) = \frac{\eta}{2\Gamma(\frac{1}{\eta})} \exp\left(-\frac{1}{2} |\varepsilon|^\eta\right),$$

where $\Gamma(\cdot)$ is the gamma function.

- If $\eta = 2$, $\Gamma\left(\frac{1}{2}\right) = \sqrt{2\pi}$ and $f(\varepsilon)$ is the standard normal density.
- If $\eta = 1$, we have the double exponential distribution.

Non-normality of returns

- A more flexible distribution that is used to model the non-normality of asset returns is the generalized error distribution (GED) with density

$$f(\varepsilon) = \frac{\eta}{2\Gamma(\frac{1}{\eta})} \exp\left(-\frac{1}{2} |\varepsilon|^\eta\right),$$

where $\Gamma(\cdot)$ is the gamma function.

- If $\eta = 2$, $\Gamma\left(\frac{1}{2}\right) = \sqrt{2\pi}$ and $f(\varepsilon)$ is the standard normal density.
- If $\eta = 1$, we have the double exponential distribution.
- When $1 < \eta < 2$, the GED has fatter tails than the normal and when $\eta > 2$, the GED has thinner tails than the normal.

Non-normality of returns

- A more flexible distribution that is used to model the non-normality of asset returns is the generalized error distribution (GED) with density

$$f(\varepsilon) = \frac{\eta}{2\Gamma(\frac{1}{\eta})} \exp\left(-\frac{1}{2} |\varepsilon|^\eta\right),$$

where $\Gamma(\cdot)$ is the gamma function.

- If $\eta = 2$, $\Gamma\left(\frac{1}{2}\right) = \sqrt{2\pi}$ and $f(\varepsilon)$ is the standard normal density.
- If $\eta = 1$, we have the double exponential distribution.
- When $1 < \eta < 2$, the GED has fatter tails than the normal and when $\eta > 2$, the GED has thinner tails than the normal.
- The excess kurtosis of returns can also be captured by a mixture of normals $(1 - \delta)N(\mu_1, \sigma_1^2) + \delta N(\mu_2, \sigma_2^2)$, where $0 \leq \delta \leq 1$.

Non-normality of returns

- A more flexible distribution that is used to model the non-normality of asset returns is the generalized error distribution (GED) with density

$$f(\varepsilon) = \frac{\eta}{2\Gamma(\frac{1}{\eta})} \exp\left(-\frac{1}{2}|\varepsilon|^\eta\right),$$

where $\Gamma(\cdot)$ is the gamma function.

- If $\eta = 2$, $\Gamma\left(\frac{1}{2}\right) = \sqrt{2\pi}$ and $f(\varepsilon)$ is the standard normal density.
- If $\eta = 1$, we have the double exponential distribution.
- When $1 < \eta < 2$, the GED has fatter tails than the normal and when $\eta > 2$, the GED has thinner tails than the normal.
- The excess kurtosis of returns can also be captured by a mixture of normals $(1 - \delta)N(\mu_1, \sigma_1^2) + \delta N(\mu_2, \sigma_2^2)$, where $0 \leq \delta \leq 1$.
 - Typically, δ is small ($\delta = 0.05$ or 0.1) and σ_2^2 is larger than σ_1^2 .

Non-normality of returns

- A more flexible distribution that is used to model the non-normality of asset returns is the generalized error distribution (GED) with density

$$f(\varepsilon) = \frac{\eta}{2\Gamma(\frac{1}{\eta})} \exp\left(-\frac{1}{2}|\varepsilon|^\eta\right),$$

where $\Gamma(\cdot)$ is the gamma function.

- If $\eta = 2$, $\Gamma\left(\frac{1}{2}\right) = \sqrt{2\pi}$ and $f(\varepsilon)$ is the standard normal density.
- If $\eta = 1$, we have the double exponential distribution.
- When $1 < \eta < 2$, the GED has fatter tails than the normal and when $\eta > 2$, the GED has thinner tails than the normal.
- The excess kurtosis of returns can also be captured by a mixture of normals $(1 - \delta)N(\mu_1, \sigma_1^2) + \delta N(\mu_2, \sigma_2^2)$, where $0 \leq \delta \leq 1$.
 - Typically, δ is small ($\delta = 0.05$ or 0.1) and σ_2^2 is larger than σ_1^2 .
 - A large value of σ_2^2 allows for thicker tails of the mixture distribution.

Non-normality of returns

- A more flexible distribution that is used to model the non-normality of asset returns is the generalized error distribution (GED) with density

$$f(\varepsilon) = \frac{\eta}{2\Gamma(\frac{1}{\eta})} \exp\left(-\frac{1}{2} |\varepsilon|^\eta\right),$$

where $\Gamma(\cdot)$ is the gamma function.

- If $\eta = 2$, $\Gamma\left(\frac{1}{2}\right) = \sqrt{2\pi}$ and $f(\varepsilon)$ is the standard normal density.
- If $\eta = 1$, we have the double exponential distribution.
- When $1 < \eta < 2$, the GED has fatter tails than the normal and when $\eta > 2$, the GED has thinner tails than the normal.
- The excess kurtosis of returns can also be captured by a mixture of normals $(1 - \delta)N(\mu_1, \sigma_1^2) + \delta N(\mu_2, \sigma_2^2)$, where $0 \leq \delta \leq 1$.
 - Typically, δ is small ($\delta = 0.05$ or 0.1) and σ_2^2 is larger than σ_1^2 .
 - A large value of σ_2^2 allows for thicker tails of the mixture distribution.
 - **The mixture of normals maintains the tractability of the normal distribution, ensures that the higher-order moments are finite and allows for possible skewness.**

Non-normality of returns

- Finally, departures from normality such as negative skewness and excess kurtosis can be modeled explicitly using the Gram-Charlier expansion of the density of asset returns.

Non-normality of returns

- Finally, departures from normality such as negative skewness and excess kurtosis can be modeled explicitly using the Gram-Charlier expansion of the density of asset returns.
 - The Gram-Charlier expansion of the density of ε has the form

$$f(\varepsilon) = \phi(\varepsilon) - \frac{k_3}{3!} \frac{\partial^3 \phi(\varepsilon)}{\partial \varepsilon^3} + \frac{k_4}{4!} \frac{\partial^4 \phi(\varepsilon)}{\partial \varepsilon^4} - \frac{k_5}{5!} \frac{\partial^5 \phi(\varepsilon)}{\partial \varepsilon^5} + \dots,$$

where $\phi(\varepsilon)$ denotes the standard normal density and k_j is the j^{th} cumulant of the distribution of ε . In particular, $k_3 = E(\varepsilon)^3$ and $k_4 = E(\varepsilon)^4 - 3$ are the skewness and the excess kurtosis of the standardized returns ε , respectively.

Non-normality of returns

- Finally, departures from normality such as negative skewness and excess kurtosis can be modeled explicitly using the Gram-Charlier expansion of the density of asset returns.
 - The Gram-Charlier expansion of the density of ε has the form

$$f(\varepsilon) = \phi(\varepsilon) - \frac{k_3}{3!} \frac{\partial^3 \phi(\varepsilon)}{\partial \varepsilon^3} + \frac{k_4}{4!} \frac{\partial^4 \phi(\varepsilon)}{\partial \varepsilon^4} - \frac{k_5}{5!} \frac{\partial^5 \phi(\varepsilon)}{\partial \varepsilon^5} + \dots,$$

where $\phi(\varepsilon)$ denotes the standard normal density and k_j is the j^{th} cumulant of the distribution of ε . In particular, $k_3 = E(\varepsilon)^3$ and $k_4 = E(\varepsilon)^4 - 3$ are the skewness and the excess kurtosis of the standardized returns ε , respectively.

- **Dropping the terms beyond k_4 gives an approximation to an arbitrary density with non-zero skewness and kurtosis.**

Non-normality of returns

- Finally, departures from normality such as negative skewness and excess kurtosis can be modeled explicitly using the Gram-Charlier expansion of the density of asset returns.
 - The Gram-Charlier expansion of the density of ε has the form

$$f(\varepsilon) = \phi(\varepsilon) - \frac{k_3}{3!} \frac{\partial^3 \phi(\varepsilon)}{\partial \varepsilon^3} + \frac{k_4}{4!} \frac{\partial^4 \phi(\varepsilon)}{\partial \varepsilon^4} - \frac{k_5}{5!} \frac{\partial^5 \phi(\varepsilon)}{\partial \varepsilon^5} + \dots,$$

where $\phi(\varepsilon)$ denotes the standard normal density and k_j is the j^{th} cumulant of the distribution of ε . In particular, $k_3 = E(\varepsilon)^3$ and $k_4 = E(\varepsilon)^4 - 3$ are the skewness and the excess kurtosis of the standardized returns ε , respectively.

- Dropping the terms beyond k_4 gives an approximation to an arbitrary density with non-zero skewness and kurtosis.
- We will use this and the related Cornish-Fisher approximations for derivative pricing and value at risk purposes.

Size of financial markets and derivative products

Counter weight

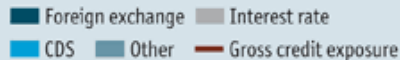
Size of financial markets*, \$trn



Sources: Bank for International Settlements, World Federation of Exchanges
 *June figures †Notional value ‡Notional principal §Market value

Exposed

Global OTC derivatives, gross market value outstanding, by type, \$trn



Source: Bank for International Settlements