Financial Econometrics I

Lecture 1

Fu Ouyang

November 22, 2018

Outline

Outline

One Period Returns

The *return* for an asset (e.g., stock, bond, etc) is a percentage defined as the change of the asset price as a fraction of its initial price. Returns exhibit more attractive statistical properties than prices. Financial econometrics typically models and analyzes return data rather than price series.

There are various definitions for the asset returns: Let P_t denote the price of an asset at time t.

- From time $t-1$ to t, assuming no dividend paid, the *one-period simple return* is defined as
$$R_t = (P_t - P_{t-1})/P_{t-1}$$
- The *one-period gross return*: $P_t/P_{t-1} = R_t + 1$.
- The *one-period log return*: $r_t = \log(P_t/P_{t-1}) = \log(R_t + 1)$. Note that when R_t is small, $r_t \approx R_t$, though theoretically $r_t \leq R_t$.

Multiperiod Returns

Consider multiperiod returns.

- the *k-period simple return* from time $t-k$ to t can be similarly defined as

$$R_t\left(k\right) = \left(P_t - P_{t-k}\right)/P_{t-k}$$

- The *k-period gross return*: $P_t/P_{t-k} = R_t(k) + 1$.
- The *k-period log return*: $r_t(k) = \log(P_t/P_{t-k}) = \log(R_t(k) + 1)$.
- $r_t(k)$ is the sum of the k one-period log returns

$$\begin{aligned}
r_t\left(k\right) &= \log(R_t\left(k\right) + 1) \\
&= \log[(R_t + 1)(R_{t-1} + 1)\cdots(R_{t-k+1} + 1)] \\
&= \log(R_t + 1) + \log(R_{t-1} + 1) + \cdots + \log(R_{t-k+1} + 1) \\
&= r_t + r_{t-1} + \cdots + r_{t-k+1}
\end{aligned}$$

In this class, returns refer to log returns unless specified otherwise.

Adjustment for Dividends

Let D_t denote the dividend payment between time $t-1$ and t. Assume that no dividends are re-invested in the asset.

$$R_t\left(k\right) = \frac{P_t + D_t + \cdots + D_{t-k+1}}{P_{t-k}} - 1$$

$$r_t\left(k\right) = \sum_{j=0}^{k-1} \log\left(\frac{P_{t-j} + D_{t-j}}{P_{t-j-1}}\right) = r_t + \cdots + r_{t-k+1}$$

Continuously Compounded Returns and Bond Yields

r_t is also called *continuously compounded return*: If a simple annual interest rate r for a bank deposit is fixed and earnings are equally paid m times per year. Then the gross return at the end of one year is

$$(1 + r/m)^m \to e^r \text{ as } m \to \infty$$

i.e., for continuously compounded interest, r = the annual log return.

Bonds are quoted in annualized yields. Consider a zero-coupon bond with the face value 1, the current yield is r_t and the remaining duration is D. Its current price B_t satisfies $B_t \exp(Dr_t) = 1$ and so the annualized log return of holding the bond (ignoring that the remaining duration for B_{t+1} is $D - 1$) is

$$\log(B_{t+1}/B_t) = D(r_t - r_{t+1})$$

Excess Returns

In many applications, it is often convenient to use an *excess return*

$$r_t - r_t^*$$

where r_t^* is a reference rate. Commonly used r_t^* include

- Bank interest rates
- LIBOR
- Log return of riskless assets
- Market portfolio

The excess yield for a bond is called *yield spread* which is the difference between the yield of a bond and the yield of a reference bond.

Outline

Behavior of Financial Return Data: Stationarity and Volatility Clustering

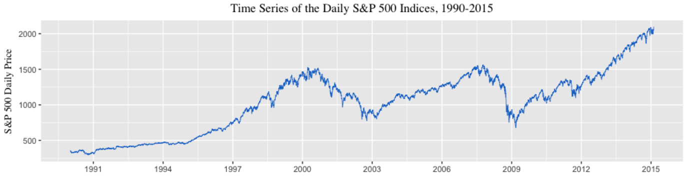
Stationarity

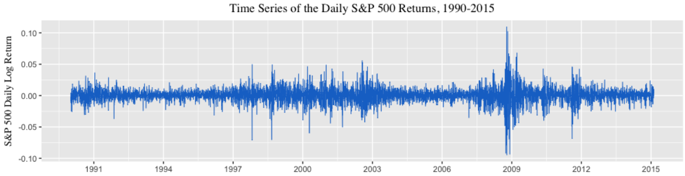
- Asset prices are often not stationary over time (e.g., financial crisis, economic growth/recession, technology improvement, etc).
- Most return sequences show certain level of stationarity (i.e., time invariant first two moments).
- The term *stationarity* will be formally defined in next lecture.

Volatility Clustering

- The volatility of asset returns is time varying.
- Large price changes (i.e., absolute returns) often occur in clusters.

Behavior of Financial Return Data: Stationarity and Volatility Clustering





Behavior of Financial Return Data: Heavy tails

Heavy tails

- The distribution of r_t often exhibits "heavier-than-normal" tails.
- A random variable X has "heavier tails" than Z if for sufficiently large $M \in \mathbb{R}_+$, $P(|X| > M) > P(|Z| > M)$.
- Example: t-distribution (with degree of freedom ν) whose PDF is

$$f_v(x) \propto (1 + x^2/\nu)^{-(\nu+1)/2}$$

  with tails of the order $|x|^{-(\nu+1)}(\asymp |x|^{-(\nu+1)})$. $N(0,1)$ has tails of order $exp(-x^2/2)$. It is not hard to see that $\lim_{x\to\infty} \exp(x^2/2)/|x|^{\nu+1} = \infty$.
- Tail behavior is closely related to the existence of $E(|r_t|^p)$. For $X \sim N(0, \sigma^2)$, $E(|X|^p) < \infty$ for all $p \in \mathbb{R}_+$, but if $X \sim t_\nu$, $E(|X|^p) < \infty$ only for $p \in (0, \nu)$. r_t is typically assumed to have at least $E(r_t^2) < \infty$.

Behavior of Financial Return Data: Asymmetry and Aggregational Gaussianity

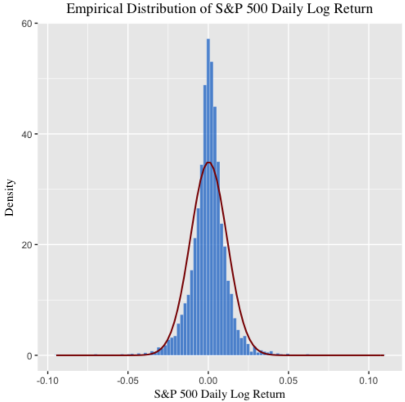
Asymmetry

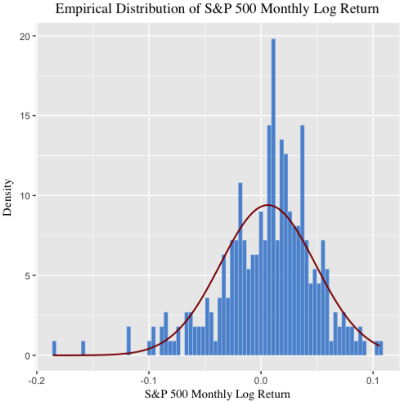
- The distribution of r_t is often negatively skewed (long left tails).
- Risk aversion: the market tends to react more strongly to negative shocks than positive ones.

Aggregational Gaussianity

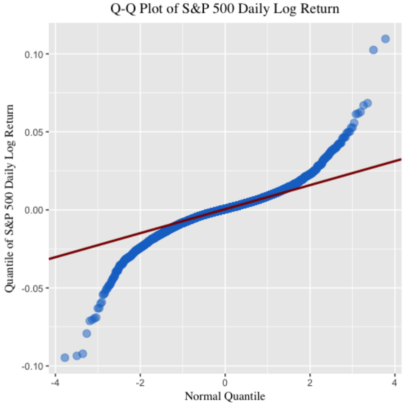
- Recall $r_t(k) = r_t + r_{t-1} + \cdots + r_{t-k+1}$.
- As k increases (long time horizon), the distribution of $r_t(k)$ tends to a normal distribution (think the central limit theorem).

Behavior of Financial Return Data: Heavy tails, Asymmetry and Aggregational Gaussianity





Behavior of Financial Return Data: Heavy tails, Asymmetry and Aggregational Gaussiananity



Testing Normality Assumption

QQ plots check normality assumption by informal graphical inspection. To formally test this hypothesis, one can perform

- Kolmogorov-Smirnov test, or
- Jarque-Bera test (H_0: $x_1, ..., x_n$ are IID from a normal distribution),

$$JB = \frac{n}{6}[S^2 + (K-3)^2/4] \xrightarrow{d} \chi_2^2$$

where S and K are *sample skewness* and *kurtosis*, respectively, i.e.,

$$S = \left(\sum_{i=1}^{n}(x_i - \bar{x})^3/n\right) \Big/ \left(\sum_{i=1}^{n}(x_i - \bar{x})^2/n\right)^{3/2}$$

$$K = \left(\sum_{i=1}^{n}(x_i - \bar{x})^4/n\right) \Big/ \left(\sum_{i=1}^{n}(x_i - \bar{x})^2/n\right)^{2}$$

Behavior of Financial Return Data: Long Range Dependence

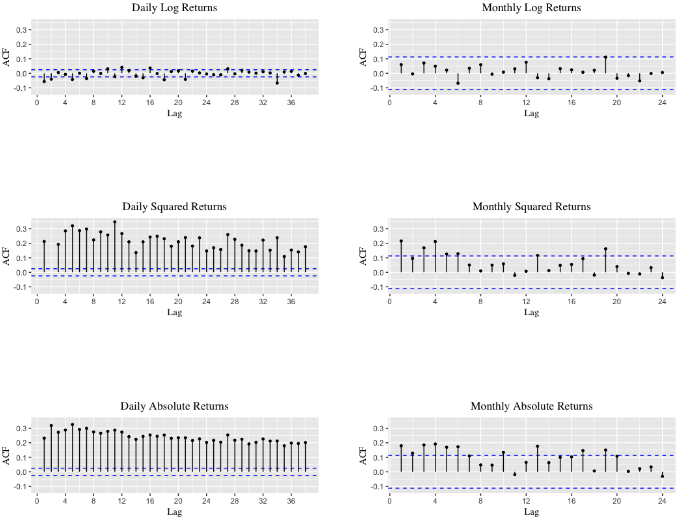
Long Range Dependence

- The returns are almost serially uncorrelated, but not independent!
- Use the sample *autocorrelation function* (ACF) $\widehat{\rho}_k$ to measure the serial correlation of a return series $\{r_t\}_{t=1}^T$. Define $\widehat{\rho}_k = \widehat{\gamma}_k / \widehat{\gamma}_0$ where

$$\widehat{\gamma}_k = \frac{1}{T} \sum_{t=1}^{T-k} (r_t - \bar{r})(r_{t+k} - \bar{r})$$

  is the sample *autocovariance* at lag k and $\bar{r} = \sum_{t=1}^T r_t / T$.
- $\{r_t^2\}_{t=1}^T$ and $\{|r_t|\}_{t=1}^T$ (*long memory*) exhibit significant (95% confidence interval for $H_0 : \rho_k = 0$) ACF $\Rightarrow$ nonlinear dependencies over time.
- ρ_k's become weaker and less persistent for longer sampling intervals.

Behavior of Financial Return Data: Long Range Dependence

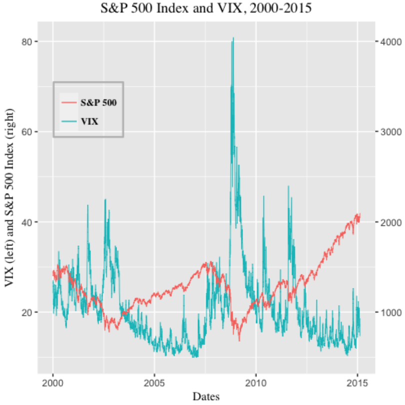


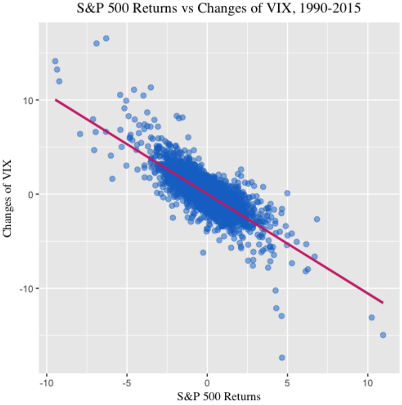
Behavior of Financial Return Data: Leverage Effect

Leverage Effect

- Assets returns are often negatively correlated with the changes of their volatilities.
- Prices down $\Rightarrow$ firms more leveraged (riskier) $\Rightarrow$ more volatile prices.
- Volatility up $\Rightarrow$ investors demand risk premiums $\Rightarrow$ prices down.
- Risk aversion: volatilities caused by price decline are typically larger than the appreciations due to declined volatilities.

Behavior of Financial Return Data: Leverage Effect





Outline

Efficient Market Hypothesis (EMH)

Strong EMH: Security prices P_t of traded assets reflect instantly all available information up to time t, public or private. Individuals do not have comparative advantages in the acquisition of information, and so there is no arbitrage opportunities.

Semi-strong EMH: Security prices merely reflect efficiently all past public information, leaving rooms for the value of private information.

Under the EMH, an asset return process may be expressed as

$$r_t = \mu_t + \epsilon_t, E(\epsilon_t) = 0, Var(\epsilon_t) = \sigma_t^2$$

- $E(\epsilon_t) = 0 \Rightarrow \mu_t = E(r_t)$, μ_t is the *rational expectation* of r_t at time $t-1$.
- ϵ_t is called an *innovation* representing the return due to unpredictable "news" that arrives between time $t-1$ and t.

Efficient Market Hypothesis

Combining the EMH and the stationarity feature discussed in last lecture, it makes sense to assume

$$\mu_t = \mu$$

For $\{\epsilon_t\}$, there are three different types of assumptions:

1. White Noise (WN) Innovations

$\{\epsilon_t\}$ are white noise, denoted as $\epsilon_t \sim WN(0, \sigma^2)$. Under WN assumption, for all $t \neq s$,

$$Cov(\epsilon_t, \epsilon_s) = 0$$

2. Martingale Difference (MD) Innovations

For any t,

$$E(\epsilon_t | r_{t-1}, r_{t-2}, ...) = E(\epsilon_t | \epsilon_{t-1}, \epsilon_{t-2}, ...) = 0$$

Efficient Market Hypothesis

3. IID Innovations

$\{\epsilon_t\}$ are independent and identically distributed (IID), denoted as $\epsilon_t \sim IID(0, \sigma^2)$.

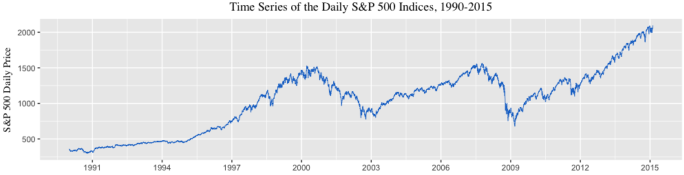
IID $\Rightarrow$ MD: $E(\epsilon_t | \epsilon_{t-1}, \epsilon_{t-2}, ...) = E(\epsilon_t) = 0$
MD $\Rightarrow$ WN: For any $t > s$, by the law of iterated expectations (LIE),

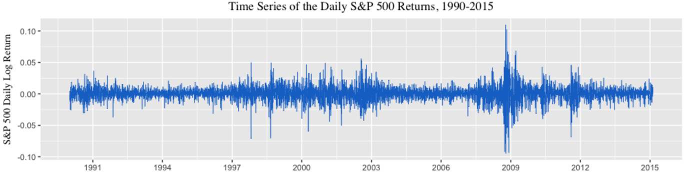
$$
\begin{aligned}
Cov(\epsilon_t, \epsilon_s) &= E(\epsilon_t \epsilon_s) \\
&= E[E(\epsilon_t \epsilon_s | \epsilon_{t-1}, \epsilon_{t-2}, ...)] \\
&= E[\epsilon_s E(\epsilon_t | \epsilon_{t-1}, \epsilon_{t-2}, ...)] \\
&= 0
\end{aligned}
$$

To sum up, IID $\Rightarrow$ MD $\Rightarrow$ WN.

Revisiting S&P 500

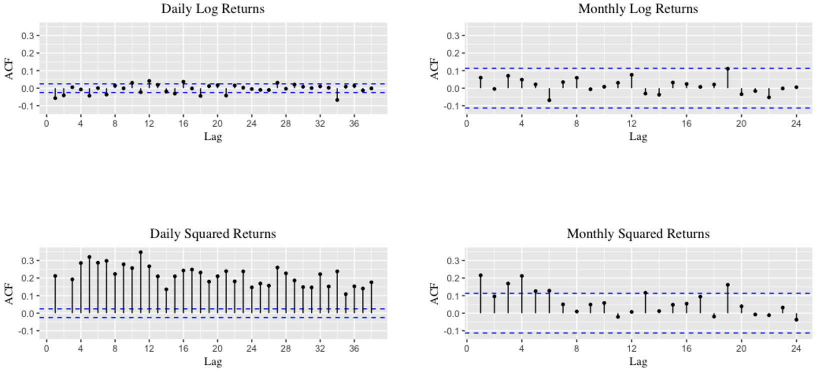
Figure 1: Time Series of P_t and r_t





Revisiting S&P 500

Figure 2: ACF of r_t and r_t^2



IID Innovations and Random Walk

Recall $r_t = \log(P_t/P_{t-1})$. Under IID assumption, $\{\log P_t\}$ form a *random walk*, i.e.,

$$\log P_t = \mu + \log P_{t-1} + \epsilon_t$$

- $\{P_t\}$ form a geometric random walk.
- If $\epsilon_t \overset{iid}{\sim} N(0, \sigma^2)$, $\{P_t\}$ form a log normal geometric random walk.
- As $\Delta t \to 0$, $\log P_t \rightsquigarrow$ Brownian motion ($P_t \rightsquigarrow$ geometric Brownian motion).

IID assumption is too strong to be true for modeling $\{r_t\}$ as it implies $Cov[f(r_t), f(r_s)] = 0$ for any function f (see Figure 2).

White Noise Innovations

The WN assumption is consistent with the stylized features of returns (see Figure 1 & 2) and is essential for the EMH.

Suppose $Corr(\epsilon_{t+1}, \epsilon_t) = \rho \neq 0$. Let $\widehat{r}_{t+1} \equiv \mu$ which is the fair predictor for r_{t+1} at time t under the EMH. Consider an alternative predictor

$$\tilde{r}_{t+1} = \mu + \rho(r_t - \mu)$$

We have

$$E(\tilde{r}_{t+1}) = E(\widehat{r}_{t+1}) = \mu$$

$$E[(\tilde{r}_{t+1} - r_{t+1})^2] = (1 - \rho^2)\sigma^2 < \sigma^2 = E[(\widehat{r}_{t+1} - r_{t+1})^2]$$

$\tilde{r}_{t+1}$ dominates $\widehat{r}_{t+1}$ as it is the same unbiased but has smaller mean squared predictive error, which violates the EMH.

Testing White Noise Hypothesis

In many applications, white nose assumption is appropriate, i.e., $\{r_t\}$ are uncorrelated (linearly independent) but may depend on each other in some nonlinear manners. To formally test H_0: $\{r_t\}$ is a white noise, one can use the Ljung-Box test

$$Q_m = T(T+2)\sum_{k=1}^{m}\frac{1}{T-k}\widehat{\rho}_k^2 \xrightarrow{d} \chi_m^2$$

or the related Box-Pierce test

$$Q_m^* = T\sum_{k=1}^{m}\widehat{\rho}_k^2 \xrightarrow{d} \chi_m^2$$

where m is a prescribed integer. In practice, we often perform the test with different values of m. We reject H_0 at the significance level α if $Q_m > \chi^2_{\alpha,m}$ ($Q_m^* > \chi^2_{\alpha,m}$) or $1 - F_{\chi_2^2}(Q_m) < \alpha$ ($1 - F_{\chi_2^2}(Q_m^*) < \alpha$).

Martingale Difference Innovations

The MD assumption is consistent with the EMH. It is not hard to show that

$$E(r_{t+1}|r_t, r_{t-1}, ...) = \arg\inf_{g \in \mathcal{G}} E[(r_{t+1} - g(r_t, r_{t-1}, ...))^2]$$

and under the MD assumption

$$E(r_{t+1}|r_t, r_{t-1}, ...) = \mu + E(\epsilon_{t+1}|r_t, r_{t-1}, ...) = \mu$$

i.e., given $r_t, r_{t-1}, ...$, the best point predictor for r_{t+1} is μ.

The MD assumption is the most appropriate mathematical form of the EMH as it is assures that asset returns cannot be predicted by any rules, but allows certain nonlinear dependence.