1

Asset Returns

The primary goal of investing in a financial market is to make profits without taking excessive risks. Most common investments involve purchasing financial assets such as stocks, bonds or bank deposits, and holding them for certain periods. Positive revenue is generated if the price of a holding asset at the end of holding period is higher than that at the time of purchase (for the time being we ignore transaction charges). Obviously the size of the revenue depends on three factors: (i) the initial capital (i.e. the number of assets purchased), (ii) the length of holding period, and (iii) the changes of the asset price over the holding period. A successful investment pursues the maximum revenue with a given initial capital, which may be measured explicitly in terms of the so-called *return*. A return is a percentage defined as the change of price expressed as a fraction of the initial price. It turns out that asset returns exhibit more attractive statistical properties than asset prices themselves. Therefore it also makes more statistical sense to analyze return data rather than price series.

1.1 Returns

Let P_t denote the price of an asset at time t. First we introduce various definitions for the returns for the asset.

1.1.1 One-period simple returns and gross returns

Holding an asset from time $t-1$ to t, the value of the asset changes from P_{t-1} to P_t. Assuming that no dividends paid are over the period. Then the *one-period simple return* is defined as

$$R_t = (P_t - P_{t-1})/P_{t-1}. \tag{1.1}$$

It is the profit rate of holding the asset from time $t-1$ to t. Often we write $R_t = 100R_t\%$, as $100R_t$ is the percentage of the gain with respect to the initial capital P_{t-1}. This is particularly useful when the time unit is small (such as a day or an hour); in such cases R_t typically takes very small values. The returns for less risky assets such as bonds can be even smaller in a short period and are often quoted in *basis points*, which is $10,000R_t$.

The *one-period gross return* is defined as $P_t/P_{t-1} = R_t + 1$. It is the ratio of the new market value at the end of the holding period over the initial market value.

1.1.2 Multiperiod returns

The holding period for an investment may be more than one time unit. For any integer $k \geqslant 1$, the returns for over k periods may be defined in a similar manner. For example, the *k-period simple return* from time $t-k$ to t is

$$R_t(k) = (P_t - P_{t-k})/P_{t-k},$$

and the *k-period gross return* is $P_t/P_{t-k} = R_t(k) + 1$. It is easy to see that the multiperiod returns may be expressed in terms of one-period returns as follows:

$$\frac{P_t}{P_{t-k}} = \frac{P_t}{P_{t-1}} \frac{P_{t-1}}{P_{t-2}} \cdots \frac{P_{t-k+1}}{P_{t-k}}, \tag{1.2}$$

$$R_t(k) = \frac{P_t}{P_{t-k}} - 1 = (R_t + 1)(R_{t-1} + 1) \cdots (R_{t-k+1} + 1) - 1. \tag{1.3}$$

If all one-period returns $R_t, \ldots, R_{t-k+1}$ are small, (1.3) implies an approximation

$$R_t(k) \approx R_t + R_{t-1} + \cdots + R_{t-k+1}. \tag{1.4}$$

This is a useful approximation when the time unit is small (such as a day, an hour or a minute).

1.1.3 Log returns and continuously compounding

In addition to the simple return R_t, the commonly used *one-period log return* is defined as

$$r_t = \log P_t - \log P_{t-1} = \log(P_t/P_{t-1}) = \log(1 + R_t). \tag{1.5}$$

Note that a log return is the logarithm (with the natural base) of a gross return and $\log P_t$ is called the log price. One immediate convenience in using

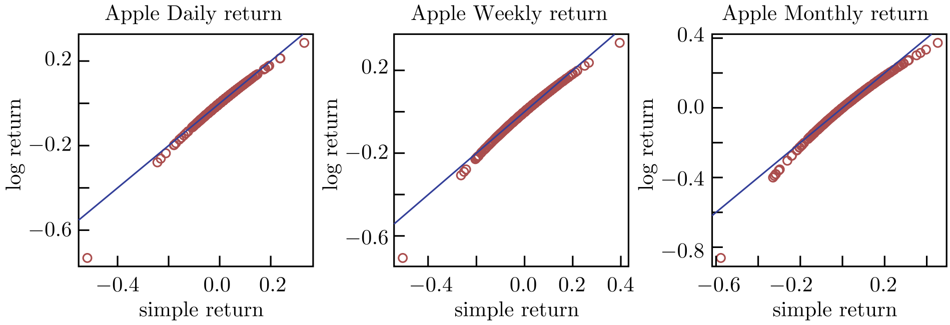

Figure 1.1 Plots of log returns against simple returns of the Apple Inc share prices in January 1985 to February 2011. The blue straight lines mark the positions where the two returns are identical.

log returns is that the additivity in multiperiod log returns, i.e. the *k-period log return* $r_t(k) \equiv \log(P_t/P_{t-k})$ is the sum of the k one-period log returns:

$$r_t(k) = r_t + r_{t-1} + \cdots + r_{t-k+1}. \tag{1.6}$$

An investment at time $t-k$ with initial capital A yields at time t the capital

$$A\exp\{r_t(k)\} = A\exp(r_t + r_{t-1} + \cdots + r_{t-k+1}) = Ae^{k\bar{r}},$$

where $\bar{r} = (r_t + r_{t-1} + \cdots + r_{t-k+1})/k$ is the average one-period log returns. In this book *returns* refer to log returns unless specified otherwise.

Note that the identity (1.6) is in contrast with the approximation (1.4) which is only valid when the time unit is small. Indeed when the values are small, the two returns are approximately the same:

$$r_t = \log(1 + R_t) \approx R_t.$$

However, $r_t < R_t$. Figure 1.1 plots the log returns against the simple returns for the Apple Inc share prices in the period of January 1985 to February 2011. The returns are calculated based on the daily close prices for the three holding periods: a day, a week and a month. The figure shows that the two definitions result almost the same daily returns, especially for those with the values between -0.2 and 0.2. However when the holding period increases to a week or a month, the discrepancy between the two definitions is more apparent with a simple return always greater than the corresponding log return.

The log return r_t is also called *continuously compounded return* due to its close link with the concept of compound rates or interest rates. For a bank deposit account, the quoted interest rate often refers to as ‘simple interest’. For example, an interest rate of 5% payable every six months will be quoted as a simple interest of 10% per annum in the market. However if an account with the initial capital \$1 is held for 12 months and interest rate remains unchanged, it follows from (1.2) that the gross return for the two periods is

$$1 \times (1 + 0.05)^2 = 1.1025,$$

i.e. the annual simple return is $1.1025 - 1 = 10.25\%$, which is called the *compound return* and is greater than the quoted annual rate of 10%. This is due to the earning from 'interest-on-interest' in the second six-month period.

Now suppose that the quoted simple interest rate per annum is r and is unchanged, and the earnings are paid more frequently, say, m times per annum (at the rate r/m each time of course). For example, the account holder is paid every quarter when $m = 4$, every month when $m = 12$, and every day when $m = 365$. Suppose m continues to increase, and the earnings are paid continuously eventually. Then the gross return at the end of one year is

$$\lim_{m \to \infty} (1 + r/m)^m = e^r.$$

More generally, if the initial capital is C, invested in a bond that compounds continuously the interest at annual rate r, then the value of the investment at time t is

$$C \exp(rt).$$

Hence the log return per annum is r, which is the logarithm of the gross return. This indicates that the simple annual interest rate r quoted in the market is in fact the annual log return if the interest is compounded continuously. Note that if the interest is only paid once at the end of the year, the simple return will be r, and the log return will be $\log(1 + r)$ which is always smaller than r.

In summary, a simple annual interest rate quoted in the market has two interpretations: it is the simple annual return if the interest is only paid once at the end of the year, and it is the annual log return if the interest is compounded continuously.

1.1.4 Adjustment for dividends

Many assets, for example some bluechip stocks, pay dividends to their shareholders from time to time. A dividend is typically allocated as a fixed amount of cash per share. Therefore adjustments must then be made in computing returns to account for the contribution towards the earnings from dividends. Let D_t denote the dividend payment between time $t - 1$ and t. Then the returns are now defined as follows:

$$R_t = (P_t + D_t)/P_{t-1} - 1, \quad r_t = \log(P_t + D_t) - \log P_{t-1},$$

$$R_t(k) = \big(P_t + D_t + \cdots + D_{t-k+1}\big)/P_{t-k} - 1,$$

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

The above definitions are based on the assumption that all dividends are cashed out and are not re-invested in the asset.

1.1.5 Bond yields and prices

Bonds are quoted in annualized yields. A so-called zero-coupon bond is a bond bought at a price lower than its face value (also called par value or principal), with the face value repaid at the time of maturity. It does not make periodic interest payments (i.e. coupons), hence the term 'zero-coupon'. Now we consider a zero-coupon bond with the face value \$1. If the current yield is r_t and the remaining duration is D units of time, with continuous compounding, its current price B_t should satisfy the condition

$$B_t \exp(Dr_t) = \$1,$$

i.e. the price is $B_t = \exp(-Dr_t)$ dollars. Thus, the annualized log-return of the bond is

$$\log(B_{t+1}/B_t) = D(r_t - r_{t+1}). \tag{1.7}$$

Here, we ignore the fact that B_{t+1} has one unit of time shorter maturity than B_t.

Suppose that we have two baskets of high-yield bonds and investment-grade bonds (i.e. the bonds with relatively low risk of default) with an average duration of 4.4 years each. Their yields spread (i.e. the difference) over the Treasury bond with similar maturity are quoted and plotted in Figure 1.2. The daily returns of bonds can then be deduced from (1.7), which is the change of yields multiplied by the duration. The daily changes of treasury bonds are typically much smaller. Hence, the changes of yield spreads can directly be used as proxies of the changes of yields. As expected, the high-yield bonds have higher yields than the investment grade bonds, but have higher volatility too (about 3 times). The yield spreads widened significantly in a period after the financial crisis following Lehman Brothers filing bankrupt protection on September 15, 2008, reflecting higher default risks in corporate bonds.

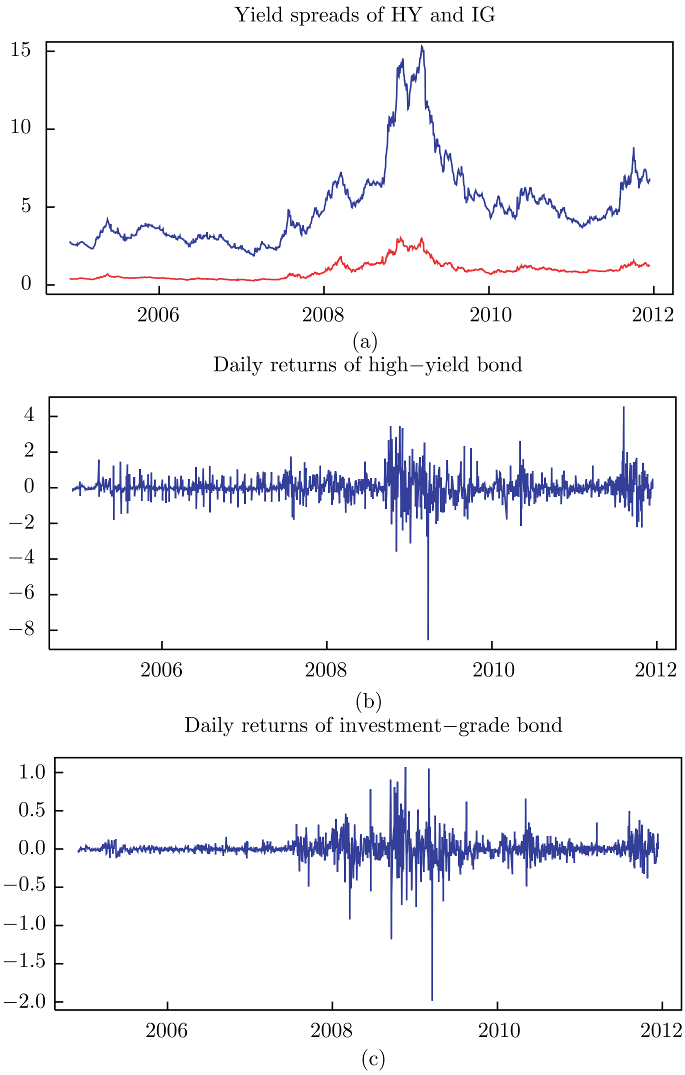

Figure 1.2 Time series of the yield spreads (the top panel) of high-yield bonds (blue curve) and investment-grade bonds (red curve), and their associated daily returns (the 2nd and 3rd panels) in November 29, 2004 to December 10, 2014.

1.1.6 Excess returns

In many applications, it is convenient to use an *excess return*, which is defined in the form $r_t - r_t^\star$, where $r_t^\star$ is a reference rate. The commonly used reference rates are, for example, bank interest rates, *LIBOR* rates (London Interbank Offered Rate: the average interest rate that leading banks in London charge when lending to other banks), log returns of a riskless asset (e.g., yields of short-term government bonds such as the 3-month US treasury bills) or market portfolio (e.g. the S&P 500 index or CRSP value-weighted index, which is the value-weighted index of all stocks traded in three major stock exchanges, created by the Center for Research in Security Prices of University of Chicago).

For bonds, *yield spread* is an excess yield defined as the difference between the yield of a bond and the yield of a reference bond such as a US treasury bill with a similar maturity.

1.2 Behavior of financial return data

In order to build useful statistical models for financial returns, we collect some empirical evidence first. To this end, we look into the daily closing indices of the S&P 500 and the daily closing share prices (in US dollars) of the Apple Inc in the period of January 1985 to February 2011. The data were adjusted for all splits and dividends, and were downloaded from *Yahoo!Finance*.

The S&P 500 is a value-weighted index of the prices of the 500 large-cap common stocks actively traded in the United States. Its present form has been published since 1957, but its history dates back to 1923 when it was a value-weighted index based on 90 stocks. It is regarded as a bellwether for the American economy. Many mutual funds, exchange-traded funds, pension funds etc are designed to track the performance of S&P 500. The first panel in Figure 1.3 is the time series plot for the daily closing indices of S&P 500. It shows clearly that there was a slow and steady increase momentum in 1985–1987. The index then reached an all-time high on March 24, 2000 during the dotcom bubble, and consequently lost about 50% of its value during the stock market downturn in 2002. It peaked again on October 9, 2007 before suffering during the credit crisis stemming from subprime mortgage lending in 2008–2010. The other three panels in Figure 1.3 show the daily, the weekly and the monthly log returns of the index. Although the profiles of the three plots are similar, the monthly return curve is a 'smoothed' version of, for example, the daily return curve which exhibits higher volatile fluctuations. In particular, the high volatilities during the 2008-2010 are more vividly depicted in the daily return plot. In contrast to the prices, the returns oscillate around a constant level close to 0. Furthermore, high oscillations tend to cluster together, reflecting more volatile market periods. Those features on return data are also apparent in the Apple stock displayed in Figure 1.4. The share prices of the Apple Inc are also non-stationary in time in the sense that the price movements over different time periods are different. For example in 1985–1998, the prices almost stayed at a low level. Then it experienced a steady increase until September 29, 2000 when the Apple's value sliced in half due to the earning warning in the last quarter of the year. The more recent surge of the price increase was largely due to Apple's success in the mobile consumer electronics market via its products the iPod, iPhone and iPad, in spite of its fluctuations during the subprime mortgage credit crisis.

We plot the normalized histograms of the daily, the weekly and the monthly log returns of the S&P 500 index in Figure 1.5. For each histogram, we also superimpose the normal density function with the same mean and variance. Also plotted in Figure 1.5 are the quantile–quantile plots for the three returns. (See an introduction to Q–Q plots in Section 1.5.) It is clear that the returns within the given holding periods are not normally distributed. Especially the tails of the return distributions are heavier than those of the normal distribution, which is highlighted explicitly in the Q–Q plots: the left tail (red circles) is below (negatively larger) the blue line, and the right tail (red circles) is above (larger) the blue line. We have also noticed that when the holding period increases from a day, a week to a month, the tails of the distributions become lighter. In particular the upper tail of the distribution for the monthly returns is about equally heavy as that of a normal distribution (red circles and blue line are about the same). All the distributions are skewed to the left due to a few large negative returns. The histograms also show that the distribution for the monthly returns is closer to a normal distribution than those for the weekly returns and the daily returns. The similar patterns are also observed in the Apple return data; see Figure 1.6.

Figures 1.7 and 1.8 plot the sample autocorrelation function (ACF) $\widehat{\rho}_k$ against the time lag k for the log returns, the squared log returns and the absolute log returns. Given a return series $r_1, \ldots, r_T$, the sample autocorrelation function is defined as $\widehat{\rho}_k = \widehat{\gamma}_k / \widehat{\gamma}_0$, where

$$\widehat{\gamma}_k = \frac{1}{T} \sum_{t=1}^{T-k} (r_t - \bar{r})(r_{t+k} - \bar{r}), \quad \bar{r} = \frac{1}{T} \sum_{t=1}^{T} r_t. \tag{1.8}$$

$\widehat{\gamma}_k$ is the sample autocovariance at lag k. It is (about) the same as the sample correlation coefficient of the paired observations $\{(r_t, r_{t+k})\}_{t=1}^{T-k}$ (the difference is in the definition of the sample mean in the calculation of the sample covariance). The sample autocorrelation functions for the squared and the absolute returns are defined in the same manner but with r_t replaced by, respectively, r_t^2 and $|r_t|$. For each ACF plot in Figures 1.7 and 1.8,

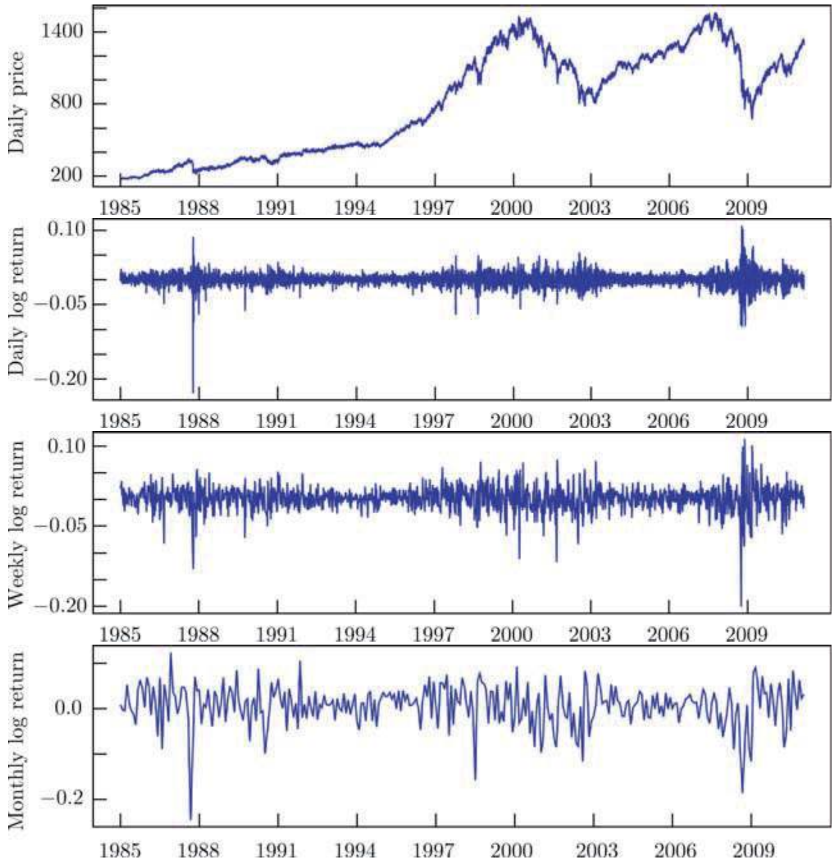

Figure 1.3 Time series plots of the daily indices, the daily log returns, the weekly log returns, and the monthly log returns of S&P 500 index in January 1985 to February 2011.

the two dashed horizontal lines, which are $\pm 1.96/\sqrt{T}$, are the bounds for the 95% confidence interval for ρ_k if the true value is $\rho_k = 0$. Hence ρ_k would be viewed as not significantly different from 0 if its estimator $\widehat{\rho}_k$ is between those two lines. It is clear from Figures 1.7 and 1.8 that all the daily, weekly and monthly returns for both S&P 500 and the Apple stock exhibit no significant autocorrelation, supporting the hypothesis that the returns of a financial asset are uncorrelated across time. However there are some small but significant autocorrelations in the squared returns and more in the absolute returns.

Furthermore the autocorrelations are more pronounced and more persistent in the daily data than in weekly and monthly data. Since the correlation coefficient is a measure of linear dependence, the above empirical evidence

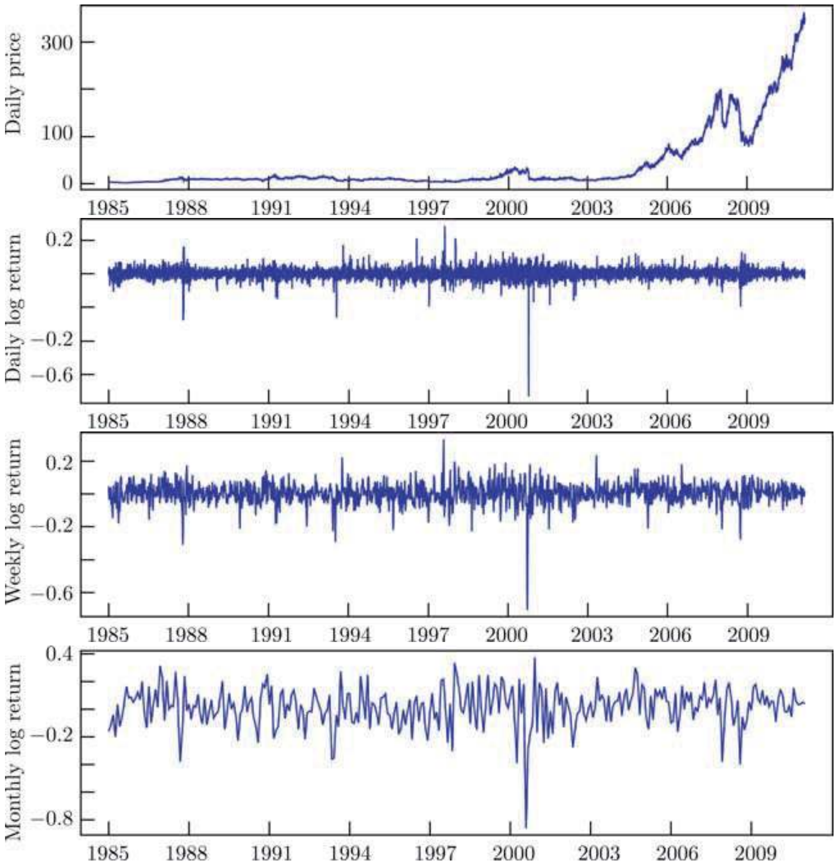

Figure 1.4 Time series plots of the daily prices, the daily log returns, the weekly log returns, and the monthly log returns of the Apple stock in January 1985 to February 2011.

indicates that the returns of a financial asset are linearly independent with each other, although there exist nonlinear dependencies among the returns at different lags. Especially the daily absolute returns exhibit significant and persistent autocorrelations – a characteristic of so-called long memory processes.

1.2.1 Stylized features of financial returns

The above findings from the two real data sets are in line with the so-called stylized features in financial returns series, which are observed across different kinds of assets including stocks, portfolios, bonds and currencies. See, e.g. Rydberg (2000). We summarize these features below.