Financial Time Series and Their Characteristics

Mei-Yuan Chen
Department of Finance
National Chung Hsing University

February 26, 2013
Financial Time Series and Their Characteristics

The future is neither completely knowable nor totally obscure; it is full of uncertainty. In our daily life we make forecasts from time to time, either implicitly or explicitly, and rely on these forecasts to make our decisions. We usually believe that the better the forecasts, the better will be the decisions. There are numerous forecasting methods for different problems. We are primarily interested in the methods that can be justified scientifically. The behavior of a variable usually consists of a systematic component and an idiosyncratic component. The systematic part is characterized by a suitably constructed model from which forecasts can be obtained.
Time Value of Money

Consider an amount $V$ invested for $n$ years at a simple interest rate of $r$ per annum (where $r$ is expressed as a decimal). If compounding takes place only at the end of the year, the future value after $n$ years is:

$$F_n = V \times (1 + r)^n.$$ 

If interest is paid $m$ times per year then the future value after $n$ years is:

$$F_{n}^{m} = V \times \left(1 + \frac{r}{m}\right)^{n \times m}.$$
As \( m \), the frequency of compounding, increases the rate becomes \textit{continuously compounded} and it can be shown that the future value becomes:

\[
F_n^c = \lim_{m \to \infty} V \times \left(1 + \frac{r}{m}\right)^{n \times m} = V \exp(r \times n).
\]

On the other hand,

\[
V = F_n^c \exp(-r \times n)
\]

is referred to as the \textit{present value} of an asset that is worth \( F_n^c \) dollars \( n \) years from now.
## Continuous Compounding Return

<table>
<thead>
<tr>
<th>Type</th>
<th>Number of payments</th>
<th>Interest rate per period</th>
<th>Net value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Annual</td>
<td>1</td>
<td>0.1</td>
<td>$1.10000</td>
</tr>
<tr>
<td>Semiannual</td>
<td>2</td>
<td>0.05</td>
<td>$1.10250</td>
</tr>
<tr>
<td>Quarterly</td>
<td>4</td>
<td>0.025</td>
<td>$1.10381</td>
</tr>
<tr>
<td>Monthly</td>
<td>12</td>
<td>0.0083</td>
<td>$1.10471</td>
</tr>
<tr>
<td>Weekly</td>
<td>52</td>
<td>0.1/52</td>
<td>$1.10506</td>
</tr>
<tr>
<td>Daily</td>
<td>365</td>
<td>0.1/365</td>
<td>$1.10516</td>
</tr>
<tr>
<td>Continuous</td>
<td>∞</td>
<td></td>
<td>$1.10517</td>
</tr>
</tbody>
</table>

In general, if the bank pays interest $m$ times a year, then the interest rate of each payment is $0.1/m$ and the net value of the deposit becomes 

$1(1 + 0.1/m)^m$ one year later. As $m \to \infty$, 

$(1 + 0.1/m)^m = \exp(0.1) = 1.10517$ which is referred to as the continuous compounding.
Asset Returns

Instead of prices, asset returns are the objects of interest in financial studies. Two main reasons are raised by Campbell, Lo, and MacKinlay (1997). First, for general investors, return of an asset is a complete and scale-free summary of the investment opportunity. Second, asset prices are commonly observed empirically to be nonstationary which makes the statistical analysis difficult. There are several definitions of asset returns.
One-period Simple Returns

Let $P_t$ be the price of an asset at time $t$. For the time being, no dividend being paid for the asset is assumed.

1. Simple Gross Return:

   $$1 + R_{t,1} = \frac{P_t}{P_{t-1}}.$$

2. Simple Net Returns (Simple Return):

   $$R_{t,1} = \frac{P_t}{P_{t-1}} - 1 = \frac{P_t - P_{t-1}}{P_t}.$$
Multi-period Simple Returns

Holding the asset for one period from date $t - k$ to $t$, the $k$-period simple gross return (or called a compound return) is

$$1 + R_{t,k} = \frac{P_t}{P_{t-k}} = \frac{P_t}{P_{t-1}} \times \frac{P_{t-1}}{P_{t-2}} \times \cdots \times \frac{P_{t-k+1}}{P_{t-k}} = (1 + R_{t,1})(1 + R_{t-1,1}) \cdots (1 + R_{t-k,1})$$

$$= \prod_{j=0}^{k-1} (1 + R_{t-j,1}),$$

and the $k$-period simple net return is

$$R_{t,k} = \frac{P_t - P_{t-k}}{P_{t-k}}.$$
The natural logarithm of the simple gross return of an asset is called the **continuously compounded return** or **log return**:

\[ r_{t,1} = \ln(1 + R_{t,1}) = \ln \frac{P_t}{P_{t-1}} = \ln P_t - \ln P_{t-1} = p_t - p_{t-1}. \]

As to the multi-period returns, we have

\[ r_{t,k} = \ln(1 + R_{t,k}) = \ln[(1 + R_{t,1})(1 + R_{t-1,1}) \cdots (1 + R_{t-k,1})] = \ln(1 + R_{t,1}) + \ln(1 + R_{t-1,1}) + \cdots + \ln(1 + R_{t-k,1}) = r_{t,1} + r_{t-1,1} + \cdots + r_{t-k,1}. \]
As the simple return is defined as

\[ R_{t,1} = \frac{P_t}{P_{t-1}} - 1 = \frac{P_t - P_{t-1}}{P_{t-1}}. \]

Then, the simple return of small time change \( dt \) is

\[ R_{t,dt} = \frac{P_t}{P_{t-dt}} - 1 = \frac{P_t - P_{t-dt}}{P_{t-dt}} \]

\[ \approx \frac{1}{P_{t-dt}} \frac{dP_{t-dt}}{dt} = \frac{d \ln(P_{t-dt})}{dt} \]

\[ = \ln(P_t) - \ln(P_{t-dt}) = r(t, dt) \]

Therefore, the simple return \( (R_{t,dt}) \) will equal to the continuous compound return \( (r_{t,dt}) \) whenever the change of time unit \( dt \) is close to zero.
The simple net return of a portfolio consisting of $N$ assets is a weighted average of the simple net returns of the assets involved, $R_{p,t} = \sum_{i=1}^{N} w_i R_{it}$. The weights are usually determined by the percentages of trading values (value weighted portfolio) and/or trading volumes (volume weighted portfolio) of the assets among total assets.
Dividend Payment

Suppose an asset pays dividend \((D_t)\), periodically. The simple net return and continuously compound return at time \(t\) are

\[
R_t = \frac{P_t + D_t}{P_{t-1}}, \quad r_t = \ln(P_t + D_t) - \ln(P_{t-1}).
\]
Excess Return

The excess return of an asset at time $t$ is defined as the difference between its return and the return on some reference asset.
Homework 1

1. Download daily close prices of a stock in Taiwan stock market from TEJ.

2. Calculate the daily gross return and compound return of the stock you select.

3. Test the equality of the distributions of daily gross return and compound return.

4. Calculate the weekly gross return and compound return of the stock you select.

5. Test the equality of the distributions of weekly gross return and compound return.
Stylized Facts for Financial Returns

1. The distribution of returns is not normal, but it has the following empirical properties:
   1.1 Stationarity.
   1.2 It is approximately symmetric.
   1.3 It has fat tails.
   1.4 It has a high peak.

2. There is almost no correlation between returns for different days.

3. The correlation between magnitudes of returns on nearby days are positive and statistically significant.

Distributional Properties of Returns

Consider a collection of $N$ assets held for $T$ periods. For each asset $i$, let $r_{it}$ be the log return at time $t$. The log returns under study are $\{r_{it}; i = 1, \ldots, N; t = 1, \ldots, T\}$. The most general model for the log returns is their joint distribution function:

$$F_r(r_{11}, \ldots, r_{N1}; r_{12}, \ldots, r_{N2}; \ldots; r_{1T}, \ldots, r_{NT}|\mathbf{Y}, \theta),$$

where $\mathbf{Y}$ is a state vector consisting of variables that summarize the environment in which asset returns are determined and $\theta$ is a vector of parameters that uniquely determine the distributional function $F_r(\cdot)$. 

M.-Y. Chen

time-ser
Some financial theories such as CAPM focus on the joint distribution of $N$ returns at a single time index $t$, i.e., \( \{r_{1t}, r_{2t}, \ldots, r_{Nt}\} \). Other theories emphasize the dynamic structure of individual asset returns, i.e., \( \{r_{i1}, \ldots, r_{iT}\} \). In the univariate time series analysis, our main concern is the joint distribution of \( \{r_{it}\}_{t=1}^T \) for asset $i$. The joint distribution of \( \{r_{it}\}_{t=1}^T \) can be partitioned as

\[
F(r_{i1}, \ldots, r_{iT}; \theta) = F(r_{i1}; \theta)F(r_{i2}|r_{i1}; \theta)F(r_{i3}|r_{i1}, r_{i2}; \theta) \cdots F(r_{iT}|r_{i1}, \ldots, r_{i,T-1}; \theta) \\
= F(r_{i1}; \theta) \prod_{t=2}^{T} F(r_{it}|r_{i,t-1}, \ldots, r_{i1}; \theta).
\]
This partition highlights the temporal dependence of the log return $r_{it}$. The main issue then is the specification of the conditional distribution $F(r_{it}|r_{i,t-1}, \ldots, r_{i1})$ – in particular, how the conditional distribution evolves over time. The partition can also be represented in density functions:

$$f(r_{i1}, \ldots, r_{iT}; \theta) = f(r_{i1}; \theta) \prod_{t=2}^{T} f(r_{it}|r_{i,t-1}, \ldots, r_{i1}; \theta).$$
Assumptions on the Distribution of Returns

Several statistical distributions have been proposed in the literature for the marginal distributions of asset returns, including normal, lognormal, stable, and scale-mixture of normal distributions.
Normal Distribution:

\{ R_{it} | t = 1, \ldots, T \} have been assumed as be independent and identically distributed as normal with fixed mean and variance. Drawbacks of this assumption are as follows. First, \( R_{it} \) has lower bound -1 however there is no bound for realizations of a normal distribution. Second, the multi-period simple return \( R_{it}[k] \) will not be normally distributed even \( R_{it} \) is normally distributed. Third, the normality assumption is not supported by most empirical empirical asset returns.
Jarque-Bera Test for Normal Distribution:

For a random variable $X$, a random sample $\{x_1, x_2, \ldots, x_n\}$ is obtained from $X$ randomly. The Jarque-Bera test is to test the null $H_0 : X$ is normally distributed and its statistic is

$$JB = \frac{1}{6} \left( S_k + \frac{(K - 3)^2}{4} \right)$$

which converges to $\chi^2(2)$ asymptotically, where

$$S_k = \frac{1}{n-1} \sum_{i=1}^{n} \frac{(x_i - \bar{x}_n)^3}{s_x^3}$$

$$K = \frac{1}{n-1} \sum_{i=1}^{n} \frac{(x_i - \bar{x}_n)^4}{s_x^4}.$$
Empirical Density Estimation

This empirical density estimate has been calculated using nonparametric kernel density estimation:

\[ \hat{f}(z) = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{h} K \left( \frac{z - z_i}{h} \right), \]

where \( K(\cdot) \) is a kernel function and \( h = h(n) \to 0 \) as \( n \to \infty \) is called bandwidth. In practice, \( h = cn^{-0.2} \) for some positive \( c \) dependent on the features of data. For details, see the book by Fan and Gijbels (1996).
Homework 1

1. Download daily close prices of a stock in Taiwan stock market from TEJ.
2. Calculate the daily simple returns and compound returns of the stock you select.
3. Test the equality of the distributions of daily simple return and compound return.
4. Calculate the weekly simple returns and compound returns of the stock you select.
5. Test the equality of the distributions of weekly simple return and compound return.
6. Test the normality for the daily simple returns and compound returns.
Lognormal Distribution:

The log returns $r_t$ of an asset is commonly assumed to be i.i.d. normally distributed with mean $\mu$ and variance $\sigma^2$. The simple returns are then i.i.d. lognormal distributed with

$$E(R_t) = \exp \left( \mu + \frac{\sigma^2}{2} \right) - 1, \quad \text{var}(R_t) = \exp(2\mu + \sigma^2)[\exp(\sigma^2) - 1].$$

Alternatively, let $m_1$ and $m_2$ be the mean and variance of the simple return $R_t$, which is lognormally distributed. Then the mean and variance of the corresponding log return $r_t$ are

$$E(r_t) = \ln \left[ \frac{m_1 + 1}{\sqrt{1 + \frac{m_2}{(1+m_1)^2}}} \right], \quad \text{var}(r_t) = \ln \left[ 1 + \frac{m_2}{(1 + m_1)^2} \right].$$
Stable Distribution:

$r_t$ is stable iff its characteristic function $h$ can be expressed as $h = e^g$, where $g$ has one of the following forms: For $0 < \alpha < 1$ or $1 < \alpha \leq 2$,

$$g(u) = iu\delta - d|u|^{\alpha} \left(1 + i\beta \frac{u}{|u|} \tan\left(\frac{\pi}{2} \alpha\right)\right), \quad (1)$$

and for $\alpha = 1$,

$$g(u) = iu\delta - d|u| \left(1 + i\beta \frac{u}{|u|} \frac{2}{\pi} \ln |u|\right), \quad (2)$$

where $\delta \in R$, $d \geq 0$, $|\beta| \leq 1$, and take $u/|u| = 0$ when $u = 0$. Usually, equations (1) and (2) are called the characteristic function of the family of stable Paretian distribution.
The parameter $\delta$ is a location parameter, $d$ a scale parameter, $\beta$ is a measure of of skewness, and $\alpha$ is the characteristic exponent. The characteristic exponent, $\alpha$, determines the total probability in the extreme tails. The smaller the value of $\alpha$, the thicker the tails of the distribution (Fama, 1963). The general form of the symmetric stable characteristic function located at zero, i.e., $\delta = 0$, $\beta = 0$, is

$$h(u) = \exp[-d|u|^\alpha], \quad d \geq 0, \quad 0 < \alpha \leq 2.$$ 

When $\alpha = 2$, $r_t$ is normal $(0, 2d)$; when $\alpha = 1$, $r_t$ has the Cauchy density with parameter $d$. 

M.-Y. Chen
If \( r_t \) is stable (not necessary symmetric) and \( 0 < \alpha \leq 1 \), then \( h \) is not differentiable at \( u = 0 \), so \( E(|r_t|) = \infty \). In the symmetric case, \( E(r_t) \) does not exist. If \( 1 < \alpha < 2 \), \( h \) can be differentiated once but not twice at \( u = 0 \), so that \( E(r_t^2) = \infty \). This is to be expected, for if \( r_t \) has finite mean and variance, the fact that \( r_t \) can be obtained as a limit of a sequence of normalized sums implies that \( r_t \) must be normal. It can be shown that if \( r_t \) is stable, \( r_t \) has a finite \( r \)th moment for all \( r \in (0, \alpha) \).
It is known that a normal random variable is a stable random variable with $\alpha = 2$, while a Cauchy is a stable random variable with $\alpha = 1$. Ash (1972, pp 345–346) pointed out that the normalized sum of i.i.d. Cauchy (special stable distribution with $\alpha = 1$) random variables has a limit which is also a Cauchy distribution. Moreover, the normalized sum of stable random variables has the same stable distribution as its limit. That means, the normalized sums of stable random variables will not follow the central limit theorem so that the functional central limit theorem breaks down. Due to its heavy dependence on the results of functional central limit theorem, the conventional large sample tests will be problematic in models with stable distributed errors.
Scale Mixture of Normal Distributions:

An example of finite mixture of normal distributions is

\[ r_t \sim (1 - \alpha)N(\mu, \sigma_1^2) + \alpha N(\mu, \sigma_2^2), \]

where \( 0 \leq \alpha \leq 1 \), \( \sigma_1^2 \) is small and \( \sigma_2^2 \)
Likelihood Function of Returns

Suppose the conditional distribution \( f(r_t | r_{t-1}, \ldots, r_1; \theta) \) (the subscript \( i \) is omitted) is independently normal with mean \( \mu_t \) and variance \( \sigma_t^2 \), then \( \theta \) consists of the parameters in \( \mu_t \) and \( \sigma_t^2 \) and the likelihood function of the data is

\[
f(r_1, \ldots, r_T; \theta) = f(r_1; \theta) \prod_{t=2}^{T} \frac{1}{\sqrt{2\pi\sigma_t}} \exp \left[ -\frac{(r_t - \mu_t)^2}{2\sigma_t^2} \right],
\]

and the log likelihood function is

\[
\ln f(r_1, \ldots, r_T; \theta) = \ln f(r_1; \theta) - \frac{1}{2} \sum_{t=2}^{T} \left[ \ln(2\pi) + \ln(\sigma_t^2) + \frac{(r_t - \mu_t)^2}{\sigma_t^2} \right].
\]
Empirical Properties of Returns

(1) Daily returns of the market indexes and individual stocks tend to have higher excess kurtoses than monthly returns. For monthly series, the returns of market indexes have higher excess kurtoses than individual stocks.

(2) The mean of a daily return series is close to zero, whereas that of a monthly return series is slightly higher.

(3) Monthly returns have higher standard deviations than daily returns.

(4) Among the daily returns, market indexes have smaller standard deviations than individual stocks.
(5) The skewness is not a serious problem for both daily and monthly returns.

(6) The descriptive statistics show that the difference between simple and log returns is not substantial.

(7) The empirical density function has a higher peak around its mean, but fatter tails than that of the corresponding normal distribution. In other words, the empirical density function is taller, skinner, but with a wider support than the corresponding normal density.
Growth vs Value Stock

A growth stock is a stock of a company that generates substantial and sustainable positive cash flow and whose revenues and earnings are expected to increase at a faster rate than the average company within the same industry. A growth company typically has some sort of competitive advantage (a new product, a breakthrough patent, overseas expansion) that allows it to fend off competitors. Growth stocks usually pay smaller dividends, as the company typically reinvests retained earnings in capital projects.
Common Characteristics Growth Stocks

1. Strong growth rate: both historic and projected forward. Historically, you want to see smaller companies with a 10%+ growth rate for the past five years and larger companies with 5% to 7%. You might want these same rates and more for projected five-year growth rates. Big companies will not grow as fast (normally) as small companies, so you need to make some accommodation.

2. Strong Return on Equity. How does the company’s return on equity (ROE) compare with the industry and its five-year average?

3. What about earnings per share (EPS)? Especially look at pre-tax profit margins. Is the company translating sales
Value Stocks

Value stocks are not cheap stocks, although one of the places you can look for candidates is on the list of stocks that have hit 52-week lows. Investors like to think of value stocks as bargains. The market has under valued the stock for a variety of reasons and the investor hopes to get in before the market corrects the price.
Some Characteristics of Value Stock

1. The price earnings ratio (P/E) should be in the bottom 10% of all companies.

2. A price to earning growth ration (PEG) should be less than 1, which indicates the company is undervalued.

3. There should be at least as much equity as debt.

4. Current assets at twice current liabilities.

5. Share price at tangible book value or less.
Value Investing

Value investing is an investment paradigm that derives from the ideas on investment that Ben Graham and David Dodd began teaching at Columbia Business School in 1928 and subsequently developed in their 1934 text Security Analysis. Although value investing has taken many forms since its inception, it generally involves buying securities whose shares appear underpriced by some form of fundamental analysis. As examples, such securities may be stock in public companies that trade at discounts to book value or tangible book value, have high dividend yields, have low price-to-earning multiples or have low price-to-book ratios.
High-profile proponents of value investing, including Berkshire Hathaway chairman Warren Buffett, have argued that the essence of value investing is buying stocks at less than their intrinsic value. The discount of the market price to the intrinsic value is what Benjamin Graham called the “margin of safety”. The intrinsic value is the discounted value of all future distributions. However, the future distributions and the appropriate discount rate can only be assumptions. For the last 25 years, Warren Buffett has taken the value investing concept even further with a focus on “finding an outstanding company at a sensible price” rather than generic companies at a bargain price.
Components of a Time Series

A time series, $y_t$, can be represented as

$$y_t = \{\text{systematic component}\} + \{\text{idiosyncratic component}\}$$

$$= \{\text{[deterministic part] + stochastic part}\} + \{\text{idiosyncratic component}\}$$

$$= \{\text{[business cycle + time trend] + seasonality}\} + \{\text{idiosyncratic component}\}$$

$$= \{[f(t)] + g(s_t) + ARMA(p, q)\} + \sqrt{h_t}e_t.$$

The aim of conventional time series analysis is to explore the functional forms of $f(t)$, $g(s_t)$, $h_t$ and the orders of $p$ and $q$. Tools of discovering $f(t)$, $g(s_t)$, $h_t$ and $p$ and $q$ include the regression analysis, smoothing techniques, and the method of Box-Jenkins.
For examples, $f(t)$ can be formulated as a linear ($\alpha_0 + \alpha_1 t$) or quadratic ($\alpha_0 + \alpha_1 t + \alpha_2 t^2$) function and $g(s_t)$ be modeled as $\gamma_1 s_{t1} + \gamma_2 s_{t2} + \gamma_3 s_{t3}$ for quarterly time series, where $s_{ti}, i = 1, 2, 3$ are the seasonal dummy variables. As to the Box-Jenkins method, we will have complete discussion lately.
Time Series Smoothing

Smoothing techniques are ways of discovering time trend pattern of a time series. In other words, Smoothing techniques remove the jagged path of a time series.
Smoothing via Moving Averages

A time series usually exhibits a rather jagged time path so that its underlying regularities may be difficult to identify. To get a clearer picture of a time series, it is important to smooth its time path. A simple way of smoothing is to compute moving averages of the original series. Let $y_t, t = 1, \ldots, T$, be time series observations. The simple moving average with $m$ periods is

$$y_t^* = \frac{y_{t-m+1} + \cdots + y_{t-1} + y_t}{m}, \quad t = m, \ldots, T.$$
In technical analysis, $y_t^*$ is usually taken as the 1-step ahead forecasting of $y_{t+1}$ at time $t$. The 1-step ahead forecast error is defined as $e_t = y_{t+1} - y_t^*$. Observe that

$$
y_t^* = \frac{y_{t-m+1} + \cdots + y_{t-1} + y_t}{m}
$$

$$
y_{t-1}^* = \frac{y_{t-m} + \cdots + y_{t-2} + y_{t-1}}{m}
$$

we have

$$
y_t^* - y_{t-1}^* = \frac{y_t - y_{t-m}}{m}
$$

$$
y_t^* = y_{t-1}^* + \frac{y_t - y_{t-m}}{m}.
$$

This updating scheme makes the forecasting process much easier.
Notes:

1. Moving average is an easy and efficient way to understand and forecast the time path.
2. The drawback of using moving average is its inability to capture the peaks and troughs of the time series.
3. Under-prediction is obtained for data moving up persistently and over-prediction is for data moving down persistently.
4. The moving average is fail to deal with nonstationary time series.
5. Seasonality is eliminated by the moving average method.
6. Equal weight are given to all the data.
Simple Exponential Smoothing

A different approach of smoothing a time series is the so-called **exponential smoothing**. We first discuss **simple exponential smoothing** which assigns a weight to the current observation \( y_t \) and exponentially decaying weights to previous observations as:

\[
y_t^* = \alpha y_t + \alpha (1 - \alpha) y_{t-1} + \alpha (1 - \alpha)^2 y_{t-2} + \alpha (1 - \alpha)^3 y_{t-3} + \cdots
\]

\[
= \alpha \sum_{j=0}^{\infty} (1 - \alpha)^j y_{t-j},
\]

where \( 0 < \alpha < 1 \) is a **smoothing constant** to be determined by practitioners.
As

\[ \alpha \sum_{j=0}^{\infty} (1 - \alpha)^j = 1, \]

\( y^*_t \) is a weighted average (linear combination) of current and past \( y_t \) and can be interpreted as an “estimate” of the current level of \( y_t \). It is also easy to verify that \( y^*_t \) can be computed via the following simple algorithm:

\[ y^*_t = \alpha y_t + (1 - \alpha) y^*_{t-1}, \]

so that \( y^*_t \) is a linear combination of \( y_t \) and previously smoothed \( y^*_{t-1} \). This algorithm typically starts with \( y^*_1 = y_1 \). We do not discuss other choices of initial value because their effect on forecasts eventually vanish when \( t \) becomes large.
A drawback of simple exponential smoothing is that it yields constant forecasts for all future values. To see this, the forecast of $y_{t+2}$ at $t + 1$ is

$$y_{t+2|t+1} = y^*_t = \alpha y_{t+1} + (1 - \alpha)y^*_t.$$

To make a 2-step ahead forecast, we may replace $y_{t+1}$ by its forecast $y^*_t$ and obtain

$$y_{t+2|t} = y_{t+2|t+1|t} = \left[\alpha y_{t+1} + (1 - \alpha)y^*_t\right]|t = \alpha y_{t+1|t} + (1 - \alpha)y^*_t = \alpha y^*_t + (1 - \alpha)y^*_t = y^*_t.$$

Following the same line we find that the $h$-step ahead forecasts are $y_{t+h|t} = y_t$, $h = 1, 2, \ldots$. **Homework!**
The error-correction form of the simple exponential smoothing algorithm is:

\[ y_t^* = \alpha(y_{t-1}^* + e_t) + (1 - \alpha)y_{t-1}^* = y_{t-1}^* + \alpha e_t. \]

This expression shows that positive (negative) forecast errors result in upward (downward) adjustments.
Another difficult problem associated with the simple exponential smoothing algorithm is the choice of smoothing constant $\alpha$. An analyst may choose a smoothing constant subjectively based on his/her experience with similar time series. When the behavior of a time series is rather erratic so that an observation may contain a large irregular component, one would tend to adopt a smaller smoothing constant which gives less weight to the most recent observation but more weight to the previously smoothed estimate. For a smoother time series, a larger smoothing constant is then needed to give more weight to the most recent observation.
This method relies on visual inspection of the time series; the exact weight to be assigned is determined quite arbitrarily. An objective way to determine a smoothing constant is the method of grid search. By selecting a grid of values for smoothing constant, we can compute sequences of smoothed series $y_t^*(\alpha)$ and their one-step forecast errors $e_t(\alpha)$. The “optimal” smoothing constant is the $\alpha$ for which the sum of squared one-step forecast errors, $\sum_{t=3}^{T} e_t(\alpha)^2$, is the smallest. Clearly, the effectiveness of this method depends on the choice of the grid.
grtel Demonstration
Holt’s Linear Trend Algorithm

Writing \( y_t = y_{t-1} + (y_t - y_{t-1}) \), a better estimate of \( y_t \) may then be obtained by combining estimates of the level and change in level (local trend) of the underlying series. This motivates Holt’s linear trend algorithm:

\[
\begin{align*}
    y_t^* &= \alpha y_t + (1 - \alpha)(y_{t-1}^* + \tau_{t-1}), \\
    \tau_t &= \beta (y_t - y_{t-1}^*) + (1 - \beta) \tau_{t-1},
\end{align*}
\]

where both \( \alpha \) and \( \beta \) are smoothing constants between zero and one. This algorithm typically starts with \( y_2^* = y_2 \) and \( \tau_2 = y_2 - y_1 \).
The algorithm can be expressed explicitly as

\[
\begin{align*}
y_2^* &= y_2, \\
\tau_2 &= y_2 - y_1, \\
y_3^* &= \alpha y_3 + (1 - \alpha)(y_2^* + \tau_2), \\
\tau_3 &= \beta(y_3 - y_2^*) + (1 - \beta)\tau_2, \\
\vdots &= \vdots \\
y_T^* &= \alpha y_T + (1 - \alpha)(y_{T-1}^* + \tau_{T-1}), \\
\tau_T &= \beta(y_T - y_{T-1}^*) + (1 - \beta)\tau_{T-1}.
\end{align*}
\]
The forecast of $y_{t+1}$, $y_{t+1|t}$, is based on the current estimates of level and change in level, i.e., $y_{t+1|t} = y_t^* + \tau_t$. Hence,

$$y_{t+2|t+1} = y_{t+1}^* + \tau_{t+1} = \alpha y_{t+1} + (1 - \alpha)(y_t^* + \tau_t) + \tau_{t+1}.$$
The 2-step ahead forecast is derived as:

\[
y_{t+2|t} = y_{t+2|t+1|t} \\
= \{y_{t+1}^* + \tau_{t+1} \}|t\} \\
= \{\alpha y_{t+1} + (1 - \alpha)(y_t^* + \tau_t) + \tau_{t+1}|t\} \\
= \alpha y_{t+1|t} + (1 - \alpha)(y_{t|t}^* + \tau_{t|t}) + \tau_{t+1|t} \\
= \alpha y_{t+1|t} + (1 - \alpha)(y_t^* + \tau_t) + \tau_t \\
= \alpha(y_t^* + \tau_t) + (1 - \alpha)(y_t^* + \tau_t) + \tau_t \\
= y_t^* + 2\tau_t.
\]

Similarly, the \( h \)-step ahead forecasts can be written as

\[
y_{t+h|t} = y_t^* + h\tau_t, \quad h = 1, 2, \ldots \quad \text{Homework!!!}
\]
Let $e_t = y_t - y_{t-1}^* - \tau_{t-1}$ be the one-step forecast error. The error-correction form of Holt’s algorithm becomes

$$
\begin{align*}
    y_t^* &= \alpha(y_{t-1}^* + \tau_{t-1} + e_t) + (1 - \alpha)(y_{t-1}^* + \tau_{t-1}) \\
         &= y_{t-1}^* + \tau_{t-1} + \alpha e_t,
\end{align*}
$$

$$
\begin{align*}
    \tau_t &= \beta(y_t^* - y_{t-1}^*) + (1 - \beta)\tau_{t-1} = \tau_{t-1} + \alpha\beta e_t.
\end{align*}
$$

Note that previous forecast errors affect both the estimates of level and local trend and that the adjustment of $\tau_t$ depends on $\alpha$ and $\beta$ simultaneously.
To choose appropriate smoothing constants, we may still employ a grid search of pairs of values \((\alpha, \beta)\) to find the one minimizing the sum of squared one-step forecast errors. This method now must search for the best combination of two smoothing constants, and hence is computationally more demanding than for the simple exponential smoothing algorithm.
The Holt-Winter Algorithm

To allow for seasonality, we consider an extension of Holt’s algorithm, which are known as the Holt-Winters algorithm. Let \( \varphi \) denote the seasonal factor and \( s \) its number of periods per year. Given additive seasonality, the Holt-Winters algorithm is

\[
    y_t^* = \alpha(y_t - \varphi_{t-s}) + (1 - \alpha)(y_{t-1}^* + \tau_{t-1}),
\]

\[
    \tau_t = \beta(y_t^* - y_{t-1}^*) + (1 - \beta)\tau_{t-1},
\]

\[
    \varphi_t = \gamma(y_t - y_t^*) + (1 - \gamma)\varphi_{t-s},
\]

where \( \alpha, \beta, \) and \( \gamma \) are smoothing constants between zero and one.
It should be clear that, owing to additive seasonality, the forecasts of $y_{t+h}$ are

$$y_{t+h|t} = \begin{cases} 
    y^*_t + h\tau_t + \varphi_{t+h-s}, & h = 1, \ldots, s, \\
    y^*_t + h\tau_t + \varphi_{t+h-2s}, & h = s + 1, \ldots, 2s, \\
    y^*_t + h\tau_t + \varphi_{t+h-3s}, & h = 2s + 1, \ldots, 3s, \\
    \vdots & \vdots
\end{cases}$$

Note that each seasonal factor repeats every $s$ periods.
Let $e_t = y_t - y^*_{t-1} - \tau_{t-1} - \varphi_{t-s}$ be the one-step forecast error. The error-correction form of the Holt-Winter algorithm becomes:

$$
y_t^* = \alpha(y^*_{t-1} + \tau_{t-1} + e_t) + (1 - \alpha)(y^*_{t-1} + \tau_{t-1}) = y^*_{t-1} + \tau_{t-1} + \alpha e_t,
$$

$$
\tau_t = \beta(y^*_t - y^*_{t-1}) + (1 - \beta)\tau_{t-1} = \tau_{t-1} + \alpha \beta e_t,
$$

$$
\varphi_t = \gamma(y_t - y^*_t) + (1 - \gamma)\varphi_{t-s}, = \varphi_{t-s} + \gamma(1 - \alpha)e_t.
$$

Observe that the first two equations are the same as those of the Holt’s algorithm, and the adjustment of $\varphi_t$ also depends on $\alpha$. 
Similarly, given multiplicative seasonality, the Holt-Winters algorithm is:

\[
\begin{align*}
y_t^* &= \alpha(y_t/\varphi_{t-s}) + (1 - \alpha)(y_{t-1}^* + \tau_{t-1}), \\
\tau_t &= \beta(y_t^* - y_{t-1}^*) + (1 - \beta)\tau_{t-1}, \\
\varphi_t &= \gamma(y_t/y_t^*) + (1 - \gamma)\varphi_{t-s}.
\end{align*}
\]

The initial values \(y_s^*\) and \(\tau_s\) are the same as those for additive seasonality, and the initial values for seasonal factor are \(\varphi_i = y_i/y_s^*, i = 1, \ldots, s\).
The $h$-step ahead forecasts are:

$$y_{t+h|t} = \begin{cases} 
(y_t^* + h\tau_t)\varphi_{t+h-s}, & h = 1, \ldots, s, \\
(y_t^* + h\tau_t)\varphi_{t+h-2s}, & h = s + 1, \ldots, 2s, \\
(y_t^* + h\tau_t)\varphi_{t+h-3s}, & h = 2s + 1, \ldots, 3s, \\
\vdots & \vdots 
\end{cases}$$
Let $e_t = y_t - (y_{t-1}^* - \tau_{t-1})\varphi_{t-s}$. The error-correction form is:

$$
\begin{align*}
{y}_t^* &= {y}_{t-1}^* + \tau_{t-1} + \alpha(e_t/\varphi_{t-s}), \\
\tau_t &= \tau_{t-1} + \alpha\beta(e_t/\varphi_{t-s}), \\
\varphi_t &= \varphi_{t-s} + \gamma(1 - \alpha)(e_t/y_t^*).
\end{align*}
$$

Although a grid search of appropriate smoothing constants is still plausible, it involves triples of values $(\alpha, \beta, \gamma)$ and is much more difficult to implement.
Damped-trend Exponential Smoothing Algorithms

Given an estimated local trend \( \tau_t \), the predicted local trends may evolve as \( c\tau_t \) at time \( t + 1 \), \( c^2\tau_t \) at \( t + 2 \), and so on, where \( 0 < c \leq 1 \) is the damping factor. The larger the damping factor, the slower the predicted trend diminishes. This leads to the damped trend algorithm:

\[
\begin{align*}
y^*_t &= \alpha y_t + (1 - \alpha)(y^*_{t-1} + c\tau_{t-1}), \\
\tau_t &= \beta(y^*_t - y^*_{t-1}) + (1 - \beta)c\tau_{t-1},
\end{align*}
\]

and the \( h \)-step forecasts are:

\[
y_{t+h|t} = y^*_t + \left( \sum_{j=1}^{h} c^j \right) \tau_t, \quad h = 1, 2, \ldots
\]
Let $e_t = y_t - y_{t-1} - c\tau_{t-1}$. The error-correction form of this algorithm is:

\[
\begin{align*}
y_t^* &= y_{t-1}^* + c\tau_{t-1} + \alpha e_t, \\
\tau_t &= c\tau_{t-1} + \alpha \beta e_t.
\end{align*}
\]

Clearly, for $c = 1$, this algorithm simply reduces to Holt’s algorithm.
Exponential-trend Exponential Smoothing Algorithms

Let $\tau_t$ denote growth rate, rather than local trend. The exponential trend algorithm is:

$$y_t^* = \alpha y_t + (1 - \alpha)y_{t-1}^* \tau_{t-1},$$
$$\tau_t = \beta(y_t^*/y_{t-1}^*) + (1 - \beta)\tau_{t-1},$$

and the $h$-step forecasts are:

$$y_{t+h|t} = y_t^* \tau_t^h, \quad h = 1, 2, \ldots$$

If there is a growth, i.e., $\tau_t > 1$, the predicted future values will increase exponentially with a constant growth rate.
Let $e_t = y_t - y_{t-1}^* \tau_{t-1}$. The error-correction form of this algorithm is:

$$y_t^* = y_{t-1}^* \tau_{t-1} + \alpha e_t,$$

$$\tau_t = \tau_{t-1} + \alpha \beta \left( \frac{e_t}{y_{t-1}^*} \right).$$
Application to BCC

1) Two year sine-wave fluctuation. Lead of business expectations over business situation: 6 months.
Application to BCC

資料來源：Eurostat
圖: BCC and Business Cycle Plot
以所選取的經濟指標的長期趨勢作為比較基準，觀察經濟指標之現況是高於、等於或低於長期趨勢。當觀察點越逼近縱軸兩端，表示經濟指標現況偏離長期趨勢的幅度越大，其中，落於上方的點，表示經濟指標現況正向偏離長期趨勢，落於下方的點，表示經濟指標現況負向偏離長期趨勢。
Horizontal Axis

以所選取指標之上期表現為比較基準，觀察當期經濟指標較上期擴張、持平抑或萎縮的幅度。當觀察點越往橫軸兩端靠近，表示當期經濟指標與前期之間的差異越大，其中，落於右方的點，表示當期經濟指標較上期擴張，落於左方的點，表示當期經濟指標較前期萎縮。
原則上需採用經過季節調整後的時間序列資料，但價格類數據是一例外，不需經過季節調整即可直接使用。所有統計資料須為絕對指數 (absolute index) 或絕對數字 (absolute number) 的形式，亦即所使用的數據不以成長率的樣貌呈現。
消除每組時間序列資料的趨勢(de-trend)。資料經過濾波器(filter)處理抽離出循環因子。本研究參照Eurostat的概念，以Hodrick-Prescott filter去除時間趨勢。此外，為求資料呈現相對平穩，我們亦針對個別數據經過適當的平滑化處理。將處理後之數據繪製於縱軸為偏離長期趨勢程度、橫軸為較前期增減幅度的圖上，即得景氣循環時鐘圖。