Chapter 2: Financial Returns

August 25, 2016

1. **Returns**

Consider an asset that does not pay any intermediate cash income (a zero-coupon bons, such as a Treasury Bill, or a share in a company that pays no dividends). Let P_t be the price of the security at time t.

1.1. *Simple and log Returns*

The linear or simple return between times t and $t-1$ is defined as[1]:

$$R_t = P_t/P_{t-1} - 1 \tag{1}$$

The log return is defined as:

$$r_t = \ln(P_t/P_{t-1})$$

Note that, while P_t means "price at time t", r_t is a shorthand for "return between time $t-1$ and t" so that the notation is not really complete and its interpretation depends on the context. When needed for clarity, we shall specify returns as indexed by the start and the end point of the interval in which they are computed as, for instance, in $r_{t-1,t}$.

The two definitions of return yield different numbers when the ratio between consecutive prices is far from 1.

Consider the Taylor formula for $\ln(x)$ for x in the neighbourhood of 1:

$$\ln(x) = \ln(1) + (x-1)/1 - (x-1)^2/2 + ...$$

[1]Note (1) defines period returns, there is usually an accrual convention applied to returns according to which they are transformed on a yearly basis.

if we truncate the series at the first order term we have:

$$\ln(x) \cong 0 + x - 1$$

so that if x is the ratio between consecutive prices, then for x close to one the two definitions give similar values. Note however that $\ln(x) \leq x - 1$. In fact $x - 1$ is equal to and tangent to $\ln(x)$ in $x = 1$ and above it anywhere else (in fact, the second derivative of $\ln(x)$ is negative). This implies that if one definition of return is used in the place of the other, the approximation errors shall be all of the same sign. This fact has important consequences when multi-period returns are computed as the difference between the two definition will become larger and larger.

1.2. *Multi-period returns and annualized returns*

What are multiperiod returns? Multiperiod returns are returns to an investment which is made with an horizon larger than one. Let us consider the case of the returns to an investment made in time t until time t+n. In this case, we define the simple multi-period return as:

$$\begin{aligned} R_{t,t+n} &= P_{t+n}/P_t - 1 \\ &= \frac{P_{t+n}}{P_{t+n-1}}\frac{P_{t+n-1}}{P_{t+n-2}}...\frac{P_{t+1}}{P_t} - 1 \\ &= \prod_{i=1}^{n}\left(1 + R_{t+i,t+i-1}\right) - 1 \end{aligned} \tag{2}$$

in the case of log returns we have instead:

$$\begin{aligned} r_{t,t+n} &= \ln\left(P_{t+n}/P_t\right) \\ &= \ln\left(\frac{P_{t+n}}{P_{t+n-1}}\frac{P_{t+n-1}}{P_{t+n-2}}...\frac{P_{t+1}}{P_t}\right) \\ &= \sum_{i=1}^{n} r_{t+i,t+i-1} \end{aligned} \tag{3}$$

Consider the case in which the length of our period in one year, given any multiperiod returns one can define its annualized value i.e. as the constant annual rate of return equivalent to the multiperiod returns of an investment in asset i over the period t,... t+n.

In the case of simple returns we have

$$\left(1+R^A_{t,t+n}\right)^n = 1+R_{t,t+n}$$
$$= \prod_{i=1}^{n}\left(1+R_{t+i,t+i-1}\right)$$
$$R^A_{t,t+n} = \left(\prod_{i=1}^{n}\left(1+R_{t+i,t+i-1}\right)\right)^{\frac{1}{n}} - 1$$

the annualized simple rate of return is the geometric mean of the annual returns over the period t, t+n.

Consider now continuosly compounded returns:

$$nr^A_{t,t+n} = r_{t,t+n}$$
$$= \sum_{i=1}^{n} r_{t+i,t+i-1}$$
$$r^A_{t,t+n} = \frac{1}{n}\sum_{i=1}^{n} r_{t+i,t+i-1}$$

The annualized log return is the arithmetic mean of annual log returns.

1.3. *Working with Returns*

Consider the value of a buy and hold portfolio invested in shares of k different companies, that pay no dividend, at time t be:

$$V_t = \sum_{i=1}^{k} n_i P_{it}$$

The simple one-period return of the portfolio shall be a linear function of the returns of each stock.

$$R_t = \frac{V_t}{V_{t-1}} - 1 = \sum_{i=1..k} \frac{n_i P_{it}}{\sum_{j=1..k} n_j P_{jt-1}} - 1$$
$$= \sum_{i=1..k} \frac{n_i P_{it-1}}{\sum_{j=1..k} n_j P_{jt-1}} \frac{P_{it}}{P_{it-1}} - 1 =$$
$$= \sum_{i=1..k} w_{it}(R_{it}+1) - 1 = (\sum_{i=1..k} w_{it}R_{it} + \sum_{i=1..k} w_{it}1) - 1 = \sum_{i=1}^{k} w_{it}R_{it}$$

Where $w_{it} = \frac{n_i P_{it-1}}{\sum_i n_i P_{it-1}}$ are non negative ”‘weights”’ summing to 1 which represent the percentage of the portfolio invested in the i-th stock at time $t-1$.

This simple result is very useful. Suppose, for instance, that you know at time $t-1$ the expected values for the returns between time $t-1$ and t. Since the expected value is a linear operator (the expected value of a sum is the sum of the expected values, moreover additive and multiplicative constants can be taken out of the expected value) and the weights w_{it} are known, hence non stochastic, at time $t-1$ we can easily compute the return for the portfolio as:

$$E(R_t) = \sum_{i=1..k} w_{it} E(R_{it})$$

Moreover if we know all the covariances between r_{it} and r_{jt} (if $i = j$ we simply have a variance) we can find the variance of the portfolio return as:

$$V(R_t) = \sum_{i=1..k} \sum_{j=1..k} w_i w_j Cov(R_{it}; R_{jt})$$

This cross-sectional additivity property does not apply to log-returns. In fact we have:

$$\begin{aligned} r_t &= \ln(\frac{V_t}{V_{t-1}}) \\ &= \ln(\frac{\sum_{i=1}^{k} n_i P_{it-1}}{\sum_{i=1}^{k} n_i P_{it-1}} \frac{P_{it}}{P_{it-1}}) = \ln(\sum_{i=1}^{k} w_{it} \exp(r_{it})) \end{aligned}$$

The log return of the portfolio is not a linear function of the log (and also the linear) returns of the components. In this case assumptions on the expected values and covariances of the components cannot be translated into assumptions on the expected value and the variance of the portfolio by simple use of basic "expected value of the sum" and "variance of the sum" formulas.

On the other hand log returns are additive when we consider the time-series of returns

$$r_{t,t+n} = \sum_{i=1}^{n} r_{t+i,t+i-1}$$

It is then easy, for instance, given the expected values and the covariances of the sub period returns, to compute the expected value and the variance of the full period return. Interestingly, additivity does not apply to simple returns.

$$R_{t,t+n} = \prod_{i=1}^{n} R_{t+i,t+i-1} - 1$$

In general the expected value of a product is difficult to evaluate and does not depend only on the expected values of the terms.

To sum up: the two definition of returns yield different values when the ratio between consecutive prices is not in the neighbourhood of the unit value. The linear definition works very well for portfolios over single periods, in the sense that expected values and variances of portfolios can be derived by expected values variances and covariances of the components, as the portfolio linear return over a time period is a linear combination of the returns of the portfolio components. For analogous reasons the log definition works very well for single securities over time.

2. Stock and Bond Returns

Computation of returns for stock and bonds must take into account the existence of intermediate cash income. In this section we show how this is performed and how linearization can help the empirical analysis of the stock and bond markets.

2.1. *Stock Returns and the dynamic dividend growth model*

Consider the one-period total holding returns in the stock market, that are defined as follows:[2]

$$H_{t+1}^{s} \equiv \frac{P_{t+1}+D_{t+1}}{P_t} - 1 = \frac{P_{t+1}-P_t+D_{t+1}}{P_t} = \frac{\Delta P_{t+1}}{P_t} + \frac{D_{t+1}}{P_t}, \tag{4}$$

where P_t is the stock price at time t, D_t is the (cash) dividend paid at time t, and the superscript s denotes "stock". The last equality decomposes a discrete holding period return as the sum of the percentage capital gain and of (a definition of) the *dividend yield*, D_{t+1}/P_t. Given that one-period returns are usually small, it is sometimes convenient to approximate them with logarithmic, continuously compounded returns, defined as:

$$r_{t+1}^{s} \equiv \log\left(1+H_{t+1}^{s}\right) = \log\left(\frac{P_{t+1}+D_{t+1}}{P_t}\right) = \log\left(P_{t+1}+D_{t+1}\right) - \log\left(P_t\right). \tag{5}$$

Interestingly, while linear returns are additive in the percentage capital gain and the dividend yield components, log returns are not as

$$\log\left(\frac{P_{t+1}+D_{t+1}}{P_t}\right) \neq \log\left(\frac{P_{t+1}}{P_t}\right) + \log\left(\frac{D_{t+1}}{P_t}\right)$$

However, it is still possible to express log returns as a linear function of the log of the price dividend and the (log) dividend growth. Dividing both sides of (4) by $\left(1+H_{t+1}^{s}\right)$ and multiplying both sides by P_t/D_t we have:

$$\frac{P_t}{D_t} = \frac{1}{\left(1+H_{t+1}^{s}\right)} \frac{D_{t+1}}{D_t}\left(1+\frac{P_{t+1}}{D_{t+1}}\right).$$

[2] The use of '$\equiv$' emphasizes that (4) provides a definition. Moreover, ΔX_{t+1} denotes the first difference of a generic variable, or $\Delta X_{t+1} \equiv X_{t+1} - X_t$.

Taking logs (denoted by lower case letters, i.e., $x_t \equiv \log X_t$ for a generic variable X_t), we have:[3]

$$p_t - d_t = -r^s_{t+1} + \Delta d_{t+1} + \ln\left(1 + e^{p_{t+1}-d_{t+1}}\right) \tag{6}$$

as $\log(D_{t+1}/D_t) = \log D_{t+1} - \log D_t = \Delta \log D_{t+1} = \Delta d_{t+1}$. Taking a first-order Taylor expansion of the last term about the point $\bar{P}/\bar{D} = e^{\bar{p}-\bar{d}}$ (where the bar denotes a sample average), the logarithm term on the right-hand side can be approximated as:

$$\begin{aligned}
\ln\left(1 + e^{p_{t+1}-d_{t+1}}\right) &\simeq \ln(1 + e^{\bar{p}-\bar{d}}) + \frac{e^{\bar{p}-\bar{d}}}{1 + e^{\bar{p}-\bar{d}}}[(p_{t+1} - d_{t+1}) - (\bar{p} - \bar{d})] \\
&= -\ln(1-\rho) - \rho \ln\left(\frac{1}{1-\rho} - 1\right) + \rho(p_{t+1} - d_{t+1}) \\
&= \kappa + \rho(p_{t+1} - d_{t+1})
\end{aligned}$$

where

$$\rho \equiv \frac{e^{\bar{p}-\bar{d}}}{1 + e^{\bar{p}-\bar{d}}} = \frac{\bar{P}/\bar{D}}{1 + (\bar{P}/\bar{D})} < 1 \qquad \kappa \equiv -\ln(1-\rho) - \rho \ln\left(\frac{1}{1-\rho} - 1\right).$$

Although $\rho \in (0,1)$ is just a factor that depends on the average price-dividend ratio, in what follows it will be used in a way that resembles a discount factor. At this point, substituting the expression for the approximated term in (6), we obtain that the log price-dividend ratio is defined as:[4]

$$p_t - d_t \simeq \kappa - r^s_{t+1} + \Delta d_{t+1} + \rho(p_{t+1} - d_{t+1}).$$

Re-arranging this expression shows that total stock market returns can be written as:

$$r^s_{t+1} = \kappa + \rho\left(p_{t+1} - d_{t+1}\right) + \Delta d_{t+1} - \left(p_t - d_t\right),$$

or a constant κ, plus the log dividend growth rate (Δd_{t+1}), plus the (discounted, at rate ρ) change in the log price-dividend ratio, $\rho\left(p_{t+1} - d_{t+1}\right) - \left(p_t - d_t\right) = \Delta(p_{t+1} - d_{t+1}) -$

[3] $-r^s_{t+1}$ follows from

$$\begin{aligned}
\log \frac{1}{\left(1 + H^s_{t+1}\right)} &= \log 1 - \log\left(1 + H^s_{t+1}\right) \\
&= -\log\left(1 + H^s_{t+1}\right) = -r^s_{t+1}
\end{aligned}$$

based on our earlier definitions and the fact that $\log 1 = 0$ for natural logs. Moreover, notice that

$$\frac{P_{t+1}}{D_{t+1}} = e^{\log(P_{t+1}/D_{t+1})} = e^{\log P_{t+1} - \log D_{t+1}} = e^{p_{t+1}-d_{t+1}}$$

[4] The approximation notation '$\simeq$' appears to emphasize that this expression is derived from an application of a Taylor expansion.

$(1-\rho)(p_{t+1}-d_{t+1})$. Moreover, by *forward* recursive substitution one obtains:

$$\begin{aligned}(p_t-d_t) &= \kappa - r^s_{t+1} + \Delta d_{t+1} + \rho(p_{t+1}-d_{t+1}) \\ &= \kappa - r^s_{t+1} + \Delta d_{t+1} + \rho\left(\kappa - r^s_{t+2} + \Delta d_{t+2} + \rho(p_{t+2}-d_{t+2})\right) \\ &= (\kappa+\rho\kappa) - (r^s_{t+1}+\rho r^s_{t+2}) + (\Delta d_{t+1}+\rho\Delta d_{t+2}) + \rho^2(p_{t+2}-d_{t+2}) \\ &= (\kappa+\rho\kappa) - (r^s_{t+1}+\rho r^s_{t+2}) + (\Delta d_{t+1}+\rho\Delta d_{t+2}) + \\ &+ \rho^2(\kappa - r^s_{t+3} + \Delta d_{t+3} + \rho(p_{t+3}-d_{t+3})) \\ &= \kappa(1+\rho+\rho^2) - (r^s_{t+1}+\rho r^s_{t+2}+\rho^2 r^s_{t+3}) + (\Delta d_{t+1}+\rho\Delta d_{t+2}+\rho^2\Delta d_{t+3}) + \rho^3(p_{t+3}-d_{t+3}) \\ &= \ldots = \kappa\sum_{j=1}^{m}\rho^{j-1} + \sum_{j=1}^{m}\rho^{j-1}(\Delta d_{t+j} - r^s_{t+j}) + \rho^m\left(p_{t+m}-d_{t+m}\right).\end{aligned}$$

Under the assumption that there can be no rational bubbles, i.e., that[5]

$$\lim_{m\longrightarrow\infty}\rho^m\left(p_{t+m}-d_{t+m}\right)=0,$$

from

$$\lim_{m\longrightarrow\infty}\sum_{j=1}^{m}\rho^{j-1}=\frac{1}{1-\rho}$$

if $\rho\in(0,1)$, we get

$$(p_t-d_t)=\frac{\kappa}{1-\rho}+\sum_{j=1}^{m}\rho^{j-1}\left(\Delta d_{t+j}-r^s_{t+j}\right).$$

This result shows that the log price-dividend ratio, (p_t-d_t), measures the value of a very long-term investment strategy (buy and hold) which—apart from a constant $\kappa/(1-\rho)$—is equal to the stream of future dividend growth discounted at the appropriate rate, which reflects the risk free rate plus risk premium required to hold risky assets, $r^s_{t+j}\equiv r^f+(r^s_{t+j}-r^f)$.[6] Therefore, for long investment horizons, econometric methods may hope to infer from the data two different types of "information": information concerning the forecasts of future (continuously compounded) dividend growth rates, i.e., Δd_{t+1}, Δd_{t+2}, ..., Δd_{t+m} as $m\longrightarrow\infty$, which are measures of the cash-flows paid out by the risky assets (e.g., how well a company will do); information concerning future discount rates, and in particular future risk premia, i.e., $(r^s_{t+1}-r^f)$, $(r^s_{t+2}-r^f)$, ..., $(r^s_{t+m}-r^f)$ as $m\longrightarrow\infty$.

[5] This assumption means that as the horizon grows without bounds, the log price-dividend ratio (hence, the underlying price-dividend ratio) may grow without bounds, but this needs to happen at a speed that is inferior to $1/\rho>1$, so that when $p_{t+m}-d_{t+m}$ is discounted at the rate ρ^m, the limit of the quantity $\rho^m\left(p_{t+m}-d_{t+m}\right)$ is zero.

[6] Here we have assumed that the risk-free interest rate is approximately constant. We shall see that, at least as a first approximation, this is an assumption that holds in practice.

Note that, under the null hypothesis of constancy of returns, the volatility of the price dividend ratio should be completelyt explained by that of the dividend process. The empirical evidence is strongly against this prediction (see the Shiller(1981) and Campbell-Shiller(1987)).

If we decompose future variables into their expected component and the unexpected one (an error term) we can write the relationship between the dividend-yield and the returns one-period ahead and over the long-horizon as follows:

$$r_{t+1}^{s} = \kappa + \rho E_t \left(p_{t+1} - d_{t+1} \right) + E_t \Delta d_{t+1} - \left(p_t - d_t \right) + \rho u_{t+1}^{pd} + u_{t+1}^{\Delta d}$$

$$\sum_{j=1}^{m} \rho^{j-1} r_{t+j}^{s} = \frac{\kappa}{1-\rho} + \sum_{j=1}^{m} \rho^{j-1} E_t \left(\Delta d_{t+j} \right) - \left(p_t - d_t \right) + \rho^m E_t \left(p_{t+m} - d_{t+m} \right) +$$

$$\rho^m u_{t+m}^{pd} + \sum_{j=1}^{m} \rho^{j-1} u_{t+j}^{\Delta d}$$

These two expressions illustrate that when the price dividends ratio is a noisy process, such noise dominates the variance of one-period returnsand the statistical relation between the price dividend ratio and one period returns is weak. However, as the horizon over which returns are defined gets longer, noise tends to be dampened and the predictability of returns given the price dividend ratio increases.

2.2. *Bond Returns: Yields-to-Maturity, Duration and Holding Period Returns*

We turn now to bonds. We distinguish between two type of bonds: those paying a coupon each given period and those that do noy pay a coupon but just reimburse the entire capital upon maturity (zero-coupon bonds).

2.2.1. **Zero-Coupon Bonds**

Define the relationship between price and yield to maturity of a zero-coupon bond as follows:

$$P_{t,T} = \frac{1}{\left(1 + Y_{t,T}\right)^{T-t}}, \tag{7}$$

where $P_{t,T}$ is the price at time t of a bond maturing at time T, and $Y_{t,T}$ is yield to maturity. Taking logs of the left and the right-hand sides of the expression for $P_{t,T}$, and defining the continuously compounded *yield,* $y_{t,T}$, as $\log(1 + Y_{t,T})$, we have the following relationship:

$$p_{t,T} = -\left(T - t\right) y_{t,T}, \tag{8}$$

which clearly illustrates that the elasticity of the yield to maturity to the price of a zero-coupon bond is the maturity of the security. Therefore the duration of the bond equals maturity as no coupons are paid. The one-period uncertain holding-period return on a bond maturing at time T, $r_{t,t+1}^{T}$, is then defined as follows:

$$r_{t,t+1}^{T} \equiv p_{t+1,T} - p_{t,T} = -(T-t-1)\, y_{t+1,T} + (T-t)\, y_{t,T} \tag{9}$$

$$= y_{t,T} - (T-t-1)\left(y_{t+1,T} - y_{t,T}\right),$$

$$= (T-t)\, y_{t,T} - (T-t-1)\, y_{t+1,T}, \tag{10}$$

which means that yields and returns differ by the a scaled measure of the change between the yield at time $t+1$, $y_{t+1,T}$, and the yield at time t, $y_{t,T}$.

2.2.2. **Coupon Bonds**

The relationship between price and yield to maturity of a constant coupon (C) bond is given by:

$$P_{t,T}^{c} = \frac{C}{\left(1+Y_{t,T}^{c}\right)} + \frac{C}{\left(1+Y_{t,T}^{c}\right)^{2}} + ... + \frac{1+C}{\left(1+Y_{t,T}\right)^{T-t}}.$$

When the bond is selling at par, the yield to maturity is equal to the coupon rate. To measure the length of time that a bondholder has invested money for we need to introduce the concept of duration:

$$D_{t,T}^{c} = \frac{\frac{C}{\left(1+Y_{t,T}^{c}\right)} + 2\frac{C}{\left(1+Y_{t,T}^{c}\right)^{2}} + ... + (T-t)\frac{1+C}{\left(1+Y_{t,T}\right)^{T-t}}}{P_{t,T}^{c}}$$

$$= \frac{C\sum_{i=1}^{T-t}\frac{i}{\left(1+Y_{t,T}^{c}\right)^{i}} + \frac{(T-t)}{\left(1+Y_{t,T}\right)^{T-t}}}{P_{t,T}^{c}}.$$

Note that when a bond is floating at par we have:

$$D_{t,T}^{c} = Y_{t,T}^{c}\sum_{i=1}^{T-t}\frac{i}{\left(1+Y_{t,T}^{c}\right)^{i}} + \frac{(T-t)}{\left(1+Y_{t,T}\right)^{T-t}}$$

$$= Y_{t,T}^{c}\frac{\left((T-t)\frac{1}{1+Y_{t,T}^{c}} - (T-t) - 1\right)\frac{1}{\left(1+Y_{t,T}^{c}\right)^{T-t+1}} + \frac{1}{1+Y_{t,T}^{c}}}{\left(1-\frac{1}{1+Y_{t,T}^{c}}\right)^{2}} + \frac{(T-t)}{\left(1+Y_{t,T}\right)^{T-t}}$$

$$= \frac{1-\left(1+Y_{t,T}^{c}\right)^{-(T-t)}}{1-\left(1+Y_{t,T}^{c}\right)^{-1}},$$

because when $|x| < 1$,

$$\sum_{k=0}^{n} kx^{k} = \frac{(nx-n-1)\,x^{n+1}+x}{(1-x)^{2}}.$$

Duration can be used to find approximate linear relationships between log-coupon yields and holding period returns. Applying the log-linearization of one-period returns to a coupon bond we have:

$$\begin{aligned} p_{c,t,T} - c &= -r^c_{t+1} + k + \rho\left(p_{c,t+1,T} - c\right) \\ r^c_{t+1} &= k + \rho p_{c,t+1,T} + (1-\rho)\, c - p_{c,t,T}. \end{aligned}$$

When the bond is selling at par, $\rho = (1+C)^{-1} = \left(1 + Y^c_{t,T}\right)^{-1}$. Solving this expression forward to maturity delivers:

$$p_{c,t,T} = \sum_{i=0}^{T-t-1} \rho^i \left(k + (1-\rho)\, c - r^c_{t+1+i}\right).$$

The log yield to maturity $y^c_{t,T}$ satisfies an expression with the same structure:

$$\begin{aligned} p_{c,t,T} &= \sum_{i=0}^{T-t-1} \rho^i \left(k + (1-\rho)\, c - y^c_{t,T}\right) = \frac{1-\rho^{T-t-1}}{1-\rho}\left(k + (1-\rho)\, c - y^c_{t,T}\right) \\ &= D^c_{t,T}\left(k + (1-\rho)\, c - y^c_{t,T}\right). \end{aligned}$$

By substituting this expression back in the equation for linearized returns we have the expression

$$r^c_{t+1} = D^c_{t,T} y^c_{t,T} - \left(D^c_{t,T} - 1\right) y^c_{t+1,T},$$

that illustrates the link between continuously compounded returns and duration.

2.3. *A simple model of the term structure*

Consider the relation between the return on a riskless one period short-term bill, r_t, and a long term bond bearing a coupon C, the one-period return on the long-term bond $H_{t,T}$ is a non-linear function of the log yield to maturity $R_{t,T}$. Shiller (1979) proposes a *linearization* which takes duration as constant and considers the following approximation in the neighborhood $y_{t,T} = y_{t+1,T} = \bar{y} = C$:

$$\begin{aligned} H_{t,T} &\simeq D_T y_{t,T} - (D_T - 1)\, y_{t+1,T} \\ D_T &= \frac{1-\gamma^{T-t-1}}{1-\gamma} = \frac{1}{1-\gamma_T} \\ \gamma_T &= \left\{1 + \bar{y}\left[1 - 1/(1+\bar{y})^{T-t-1}\right]^{-1}\right\}^{-1} \\ \lim_{T\longrightarrow\infty} \gamma_T &= \gamma = 1/(1+\bar{y}) \end{aligned}$$

solving this expression forward we generate the equivalent of the DDG model in the bond market:

$$y_{t,T} = \sum_{j=0}^{T-t-1} \gamma^j (1-\gamma) H_{t+j,T} + \gamma^{T-t} y_{T-1,T}$$

In this case, by equating one-period risk-adjusted returns, we have:

$$E\left[\frac{y_{t,T} - \gamma y_{t+1,T}}{1-\gamma} \mid I_t\right] = r_t + \phi_{t,T} \tag{11}$$

From the above expression, by recursive substitution, under the terminal condition that at maturity the price equals the principal,we obtain:

$$y_{t,T} = y^*_{t,T} + E[\Phi_T \mid I_t] = \frac{1-\gamma}{1-\gamma^{T-t}} \sum_{j=0}^{T-t-1} \gamma^j E[r_{t+j} \mid I_t] + E[\Phi_T \mid I_t] \tag{12}$$

where the constant $\Phi_{t,T}$ is the term premium over the whole life of the bond:

$$\Phi_{t,T} = \frac{1-\gamma}{1-\gamma^{T-t}} \sum_{j=0}^{T-t-1} \gamma^j \phi_{t+j,T}$$

For long-bonds, when $T-t$ is very large, we have :

$$y_{t,T} = y^*_{t,T} + E[\Phi_T \mid I_t] = (1-\gamma) \sum_{j=0}^{T-t-1} \gamma^j E[r_{t+j} \mid I_t] + E[\Phi_T \mid I_t]$$

Subtracting the risk-free rate from both sides of this equation we have:

$$\begin{aligned} S_{t,T} = y_{t,T} - r_t &= \sum_{j=1}^{T-1} \gamma^j E[\Delta r_{t+j} \mid I_t] + E[\Phi_T \mid I_t] \\ &= S^*_{t,T} + E[\Phi_T \mid I_t] \end{aligned}$$

3. **Graphical Analysis of Returns**

Graphics

- time series graphics of returns (change of prices)
- time series graphics of portfolio perfomance (level of prices)
- scatter plots
- multiple graphs
- density estimates (histograms)
-

4. **Matrix Representation of Returns**

A matrix is a double array of i rows and j columns, whose generic element can be written as a_{ij}, it is a convenient way of collecting simultaneusly information on the time-series and the cross-section of returns:

$$A = \begin{bmatrix} a_{11} & . & . & a_{1j} \\ & & & \\ & & & \\ a_{i1} & & & a_{ij} \end{bmatrix}, 0 = \begin{bmatrix} 0 & . & . & 0 \\ & & & \\ & & & \\ 0 & & & 0 \end{bmatrix}$$

$$I = \begin{bmatrix} 1 & . & . & 0 \\ & & & \\ & & & \\ 0 & & & 1 \end{bmatrix}$$

4.1. *Matrix Operations*

- Transposition $a'_{ij} = a_{ji}$
- Addition: For A and B nxm $(a+b)_{ij} = a_{ij} + b_{ij}$
- Multiplication: For A nxm and B mxp $(ab)_{ij} = \sum_{k=1}^{m} a_{ik} b_{kj}$
- Inversion for non-singular A nxn, A^{-1} satisfies $A^{-1}A = AA^{-1} = I$

5. **Modeling Returns**

Given the description of the data let us now entertain the possibility of specifying a statistical model for them. The (naive) log random walk (LRW) hypothesis on the evolution of prices states that, prices evolve approximately according to the stochastic difference equation:

$$\ln P_t = \mu\Delta + \ln P_{t-\Delta} + \epsilon_t$$

where the 'innovations' ϵ_t are assumed to be uncorrelated across time ($cov(\epsilon_t; \epsilon_{t'}) = 0 \quad \forall t \neq t'$), with constant expected value 0 and constant variance $\sigma^2\Delta$. Sometimes, a further hypothesis is added and the ϵ_t are assumed to be jointly normally distributed. In this case the assumption of non correlation becomes equivalent to the assumption of independence.

Since $\ln P_t - \ln P_{t-\Delta} = r^*_{t-\Delta;t}$ the LRW is obviously equivalent to the assumption that log returns are uncorrelated random variables with constant expected value and variance.

A linear random walk in prices was sometimes considered in the earliest times of quantitative financial research, but it does not seems a good model for prices since a sequence of negative innovations may result in negative prices. Moreover, while the hypothesis of constant variance for (log) returns may be a good first order approximation of what we observe in markets, the same hypothesis for prices is not empirically sound: in general price changes tend to have a variance which is an increasing function of the price level.

Note that :Δ is the "fraction of time" over which the return is defined. This may be expressed in any unit of time measurement: $\Delta = 1$ may mean one year, one month, one day, at the choice of the user. However, care must be taken so that μ and σ^2 are assigned consistently with the choice of the unit of measurement of Δ. In fact μ and σ^2 represent return, expected value and variance over an horizon of Length $\Delta = 1$ and they shall be completely different if 1 means, say, one year or one day (see below for a particular convention in translating the values of μ and σ^2 between different units of measurement of time.

Second: if the model is valid for a time interval of Δ consider what happens over a time span of, say, 2Δ.

By simply composing the model twice we have:

$$\ln P_t = 2\mu\Delta + \ln P_{t-2\Delta} + \epsilon_t + \epsilon_{t-\Delta} = \ln P_{t-2\Delta} + u_t$$

having set $u_t = \epsilon_t + \epsilon_{t-\Delta}$. The model appears similar to the single Δ one and in fact it is but it must be noticed that the u_t while uncorrelated (due to the hypothesis on the ϵ_t) on a time span of 2Δ shall indeed be correlated on a time span of Δ. This means, roughly, that the log random walk model can be aggregated over time if decimate the data, i.e. if we "drop" the observation in between each aggregated interval (in our example the model shall be valid if we drop every other original observation).

Consider now the case in which the time interval is of the length of 1-month. If we take prices as inclusive of dividends we can write the following model for log-returns

$$r_{t,t+1} = \mu + \sigma\epsilon_t$$
$$\epsilon_t \sim i.i.d.(0,1)$$

This simple specification has some appealing properties for the n period returns $r_{t,t+n}$:

If we assume the LRW and consider a sequence of n log returns r_t^* at times $t, t-1, t-2, ..., t-n+1$ (just for the sake of simplicity in notation we suppose each time interval Δ to be of length 1 and drop the generic Δ) we have that:

$$E(r_{t,t+n}) = E(\sum_{i=1}^{n} r_{t+i,t+i-1}) = \sum_{i=1}^{n} E(r_{t+i,t+i-1}) = n\mu$$

$$Var(r_{t,t+n}) = Var(\sum_{i=1}^{n} r_{t+i,t+i-1}) = \sum_{i=1}^{n} Var(r_{t+i,t+i-1}) = n\sigma^2$$

This obvious result, which is a direct consequence of the assumption of constant expected value and variance and of non correlation of innovations at different times, is typically applied, for annualization purposes, also when the LRW is not considered to be valid.

So, for instance, given an evaluation of σ^2 on monthly data, this evaluation is annualized multiplying it by 12

This is not a convention, but the correct procedure, if the LRW model holds. In this case, in fact, the variance over n time periods is equal to n times the variance over one time period. If the LRW model is not believed to hold, for instance, if the expected value and-or the variance of return is not constant over time or if we have correlation among the ϵ_t, this procedure becomes just as a convention.[7]

5.1. *Normal and log-normal distributions*

Under the log random walk model for prices with Gaussian residuals, log returns are normally distributed, this implies that single period gross returns are i.i.d lognormal variables, as $r_{t+1} \equiv \log\left(1 + R_{t+1}\right)$. Note that, under the lognormal model

$$r_{t,t+1} \sim n.i.d.(\mu, \sigma^2)$$
$$E\left(R_{t,t+1}\right) = \exp\left(\mu + \frac{1}{2}\sigma^2\right) - 1$$
$$Var\left(R_{t,t+1}\right) = \exp\left(2\mu + \sigma^2\right)\left(e^{\sigma^2} - 1\right)$$

In the case we have a vector of log returns that are normally distributed we have

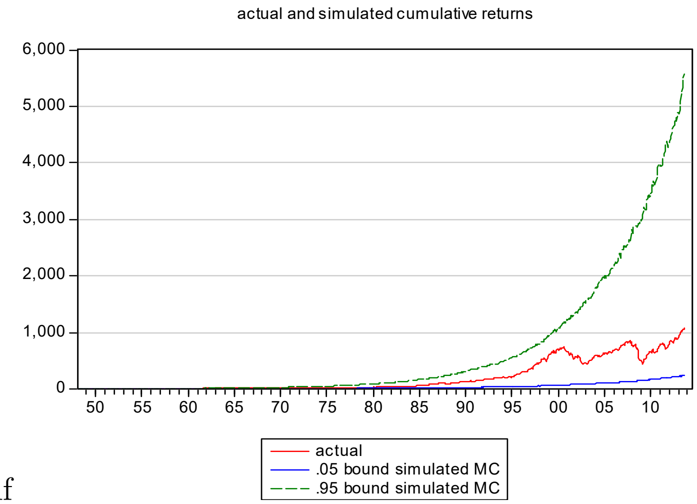
[7] Empirical computation of variances over different time intervals typically result in sequences which increase less than linearly wrt the increase of the time interval between consecutive observations. This could be interpreted as the existence of (small) on average negative correlations between returns.

$$\mathbf{r}_{t,t+1} \sim i.i.d.(\mu, \Sigma)$$
$$E\left(R_{t,t+1}^{i}\right) = \exp\left(\mu_i + \frac{1}{2}\sigma_{ii}\right) - 1$$
$$Cov\left(R_{t,t+1}^{i}, R_{t,t+1}^{j}\right) = \exp\left(\mu_i + \mu_j + \frac{1}{2}\left(\sigma_{ii} + \sigma_{jj}\right)\right)\left(e^{\sigma_{ij}} - 1\right)$$

5.2. *Assessing Models by Simulation: Monte-Carlo and Bootstrap Methods*

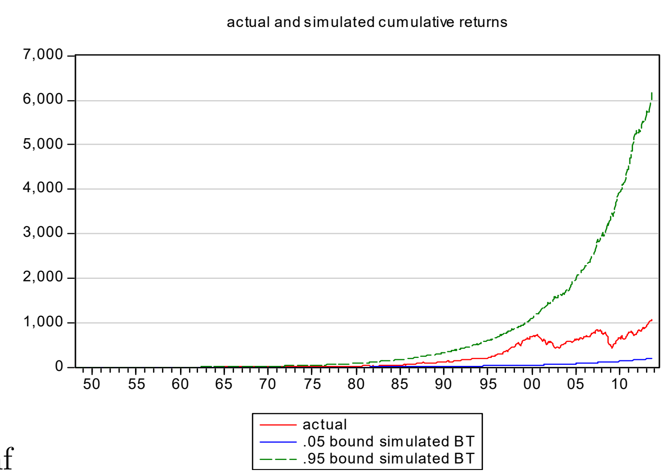
How well does the CER model in describing the data? One way to answer to this question is to use computer simulation methods to generate pseudo data from the model and compare them with actual data. We shall consider two ways of simulating pseudo-data: Monte-Carlo Simulation and Bootstrap. To use Monte Carlo Simulation to generate pseudo data from the CER model, some estimates of μ σ ar necessary. Given this estimates an assumption must be made on the distribution of ϵ_t. Then an artificial sample for ϵ_t of the length matching that of the available can be computer simulated. The simulated residuals are then mapped into simulated returns via μ, σ. This exercise can be replicated N times (and therefore a Monte-Carlo simulation generates a matrix of computer simulated returns whose dimension are defined by the sample size T and by the number of replications N). The distribution of model predicted returns can be then constructed and one can ask the question if the observed data can be considered as one draw from this distribution.

To illustrate these procedure we have generated artifical samples of stock market returns from 1948 onwards and tracked for each simulation the value of 1 dollar invested in 1948 overtime. One thousand possible histories are generated and then the 5th percentile and the 95th percentile can be plotted with the actual data. This allows to verify if the actual data are in the range of the CER, and to measure the uncertainty associated with the model.



2- Returns/OCGHMJ00.wmf

One of the possible limitation of the Monte-Carlo approach is the choice a distribution from which the residuals are to be drawn. It might be very well the case that the model goes wrong because the choice of the statistical distribution is not the correct one. Bootstrap methods overcome this problem by sampling residuals from their empirical distribution. All the steps in a boostrap simulation are the same with the Monte-Carlo simulation except that different observations for residuals are constructed by taking the deviation of returns from their sample mean putting them in urn and resampling from the urn with replacement.



2- Returns/OCGHMJ01.wmf

5.3. *Stocks for the long run*

The fact that, under the LRW, the expected value grows linearly with the length of the time period while the standard deviation (square root of the variance) grows with the square root of the number of observations, has created a lot of discussion about the existence of some time horizon beyond which it is always proper to hold a stock portfolio.

This problem, conventionally called 'time diversification', and more popularly 'stocks for the long run', has attracted some considerable attention.

We have three flavors of the "stocks for the long run" argument. The first and the second are a priori arguments depending on the log random walk hypothesis or something equivalent to it, the third is an a posteriori argument based on historical data.

The basic idea of the first version of the argument can be sketched as follows. Assume that single period (log) returns have (positive) expected value μ and variance σ^2. Moreover, assume for simplicity that the investor requires a Sharpe ratio of say S out of his-her investment. Under the above hypotheses, plus the log random walk hypothesis, the Sharpe ratio over n time periods is given by

$$S = \frac{n\mu}{\sqrt{n}\sigma} = \sqrt{n}\frac{\mu}{\sigma}$$

so that, if n is large enough, any required value can be reached. Another way of phrasing the same argument, when we add the hypothesis of normality on returns, is that, for any given probability α and any given required return C there is always an horizon for which the probability for n period return less than C is less than α.

$$\Pr\left(R^p < C\right) = \alpha.$$

$$\Pr\left(R^p < C\right) = \alpha \Longleftrightarrow \Pr\left(\frac{R^p - n\mu}{\sqrt{n}\sigma} < \frac{C - n\mu}{\sqrt{n}\sigma}\right) = \alpha$$
$$\Longleftrightarrow \Phi\left(\frac{C - \mu_p}{\sigma_p}\right) = \alpha,$$
$$C = n\mu + \Phi^{-1}\left(\alpha\right)\sqrt{n}\sigma$$

But $n\mu + \Phi^{-1}\left(\alpha\right)\sqrt{n}\sigma$, for $\sqrt{n} > \frac{1}{2}\frac{\Phi^{-1}(\alpha)}{\mu}\sigma$ is an increasing function in n so that for any α and any chosen value C, there exists a n such that from that n onward, the probability for an n period return less than C is less than α.

The investment implication could be that for a time horizon of an undetermined number n of years, the investment that has the highest expected return per unit of standard deviation is optimal even if the standard deviation is very high. This investment in very risky too risky in the "short run", but there is always a time horizon for which, the probability of any given loss is as small as you like or, that is the same, the Sharpe ratio as big as you like. Typically, such high return (and high volatility) investment are stocks, so: "stocks for the long run".

Note, however, that the value of n for which this lower bound crosses a given C level is the solution of

$$n\mu + \Phi^{-1}(\alpha)\sqrt{n}\sigma \geq C$$

In particular, for $C = 0$ the solution is

$$\sqrt{n} \geq -\frac{\Phi^{-1}(\alpha)\sigma}{\mu}$$

Consider now the case of a stock with σ/μ ratio for one year is of the order of 6. Even allowing for a large α,say 0.25, so that $\Phi^{-1}(\alpha)$ is near minus one , the required n shall be in the range of 36 which is only slightly shorter than the average working life.

As a matter of fact, based on the analysis of historical prices and risk adjusted returns, stocks have been almost always a good long run investment. However, some care must be exercised in intepreting this evidence because history is what we have observed and one could doubt on the possibility for an institution such as the stock market to survive without giving a sustainable impression of offering some opportunities.

Such arguments, if not accompanied by further argument, become somewhat empty as could be the analogue to being surprised observing that the food I most frequently eat, is also among those I like the most or, more in the extreme, that old people did not die young or, again, that wen we are in one of many queues, we spend most time in the slowest queue.

Unfortunately the arrow of time is uni-directional and experimental data for financial time-series are not available.

References

[1] Boudoukh, J.; R. Michaely; M. Richardson; and M. Roberts. On the Importance of Measuring Payout Yield: Implications for Empirical Asset Pricing. *Journal of Finance*, (2007), 877-915.

[2] Campbell J.Y. and R.J. Shiller (1987) "Cointegration and Tests of Present Value Models", Journal of Political Economy, 95, 5, 1062-1088

[3] Campbell, John Y., and Robert Shiller, 1988, Stock Prices,Earnings, and Expected Dividends, *Journal of Finance*, 43, 661-676.

[4] Campbell, John Y., and Robert Shiller, 1988, The Dividend-Price Ratio and Expectations of future Dividends and Discount Factors, *Review of Financial Studies*, 1:195-228

[5] Campbell, John Y., and Robert Shiller, 1998 Valuation Ratios and The Long-Run Stock Market Outlook, 1998, *Journal of Portfolio Management*

[6] Campbell, John Y., and Robert Shiller, 2001 Valuation Ratios and The Long-Run Stock Market Outlook, an update, Cowles Foundation DP 1295

[7] John Y. Campbell and Tuomo Vuolteenaho (2004), Inflation Illusion and Stock Prices (Cambridge: NBER Working Paper 10263)

[8] Cochrane, J. H. The Dog that Did Not Bark: A Defense of Return Predictability. *Review of Financial Studies*, 21 (2008), 4, 1533-1575.

[9] Cochrane J.(1999) New Facts in Finance

[10] Davidson, R., and E. Flachaire. The Wild bootstrap, Tamed at Last., *Journal of Econometrics*,146(2008), 1, 162-169.

[11] Fama, Eugene and Kenneth R. French, 1988, Dividend Yields and Expected Stock Returns, *Journal of Financial Economics*, 22, 3-26.

[12] Ferreira and Santa Clara (2010)

[13] Lander J., Orphanides A. and M.Douvogiannis(1997) Earning forecasts and the predictability of stock returns: evidence from trading the S&P, *Board of Governors of the Federal Reserve System*, http:www.bog.frb.fed.org

[14] Lettau, Martin, and Sydney Ludvigson, 2005, Expected Returns and Expected Dividend Growth, *Journal of Financial Economics*, 76, 583-626

[15] Lettau, Martin, and Stijn Van Nieuwerburgh, 2008, Reconciling the Return Predictability Evidence, *Review of Financial Studies*, 21, 4, 1607-1652

[16] Littermannn B.(2003), Modern Investment Management. An Equilibrium Approach, *Wiley Finance*

[17] Meznly, Lior; Tano Santos, and Pietro Veronesi, Understanding Predictability, *Journal of Political Economy, 112 (2004),1,1-47.*

[18] Franco Modigliani and Richard Cohn (1979), Inflation, Rational Valuation, and the Market, *Financial Analysts' Journal*.

[19] Newey, W. K. and K. D. West. A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix., *Econometrica*, 55 (1987), 3, 703-08.

[20] Newey, W. K., and D. K. West. Automatic Lag Selection in Covariance Matrix Estimation. *Review of Economic Studies*, 61 (1994), 631-653.

[21] Robert J. Shiller. Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends? *American Economic Review* 71 (June 1981), 421-436. 21

[22] Valkanov, R. Long-Horizon Regressions: Theoretical Results and Applications. *Journal of Financial Economics*, 68 (2003), 201–232. 33