Return predictability*

Jonas Nygaard Eriksen**

Empirical Asset Pricing, Spring 2022

*The note outlines the derivation of key relations in return predictability. The note is prepared for use only in the Master's course "Empirical Asset Pricing". Please do not cite, circulate, or use for purposes other than this course.

**CREATES, Department of Economics and Business Economics, Aarhus University, Fuglesangs Allé 4, DK-8210 Aarhus V, Denmark. E-mail: jeriksen@econ.au.dk.

Contents

1. The Campbell-Shiller log-linear relation

This section outlines the steps necessary to arrive at the Campbell and Shiller (1988) log-linear relation for the return on a stock. We depart from the definition of a gross stock return

$$R_{t+1} = 1 + r_{t+1} = \frac{P_{t+1} + D_{t+1}}{P_t} \tag{1}$$

and then take the natural logarithm on both the left-hand and the right-hand side to obtain

$$\begin{aligned}
r_{t+1} &= \ln\left(R_{t+1}\right) &(2)\\
&= \ln\left(P_{t+1} + D_{t+1}\right) - \ln\left(P_t\right) &(3)\\
&= \ln\left(P_{t+1}\left[1 + \frac{D_{t+1}}{P_{t+1}}\right]\right) - \ln\left(P_t\right) &(4)\\
&= \ln\left(P_{t+1}\right) + \ln\left(1 + \frac{D_{t+1}}{P_{t+1}}\right) - \ln\left(P_t\right). &(5)
\end{aligned}$$

We can write the second term, the dividend-price ratio part, as

$$\begin{aligned}
\ln\left(1 + \frac{D_{t+1}}{P_{t+1}}\right) &= \ln\left(1 + \exp\left\{\ln\left(D_{t+1}\right) - \ln\left(P_{t+1}\right)\right\}\right) &(6)\\
&= \ln\left(1 + \exp\left\{d_{t+1} - p_{t+1}\right\}\right), &(7)
\end{aligned}$$

where $d_{t+1} = \ln\left(D_{t+1}\right)$ and $p_{t+1} = \ln\left(P_{t+1}\right)$ denote log dividends and log prices, respectively. This term is clearly nonlinear, and thus constitutes the nonlinear part of log returns. To arrive at a log-linear approximation, we can make use of a first-order Taylor expansion of $\ln\left(1 + \exp\left\{d_{t+1} - p_{t+1}\right\}\right)$ around the average log dividend-price ratio $\overline{d-p}$.[1] In our case, this amounts to

$$\begin{aligned}
\ln\left(1 + \frac{D_{t+1}}{P_{t+1}}\right) &= \ln\left(1 + \exp\left\{d_{t+1} - p_{t+1}\right\}\right)\\
&\approx \ln\left(1 + \exp\left\{\overline{d-p}\right\}\right)\\
&\quad + \underbrace{\frac{1}{1 + \exp\left\{\overline{d-p}\right\}}}_{\equiv 0<\phi<1} \underbrace{\exp\left\{\overline{d-p}\right\}}_{\equiv \frac{1-\phi}{\phi}} \left(\left[d_{t+1} - p_{t+1}\right] - \overline{d-p}\right) &(8)\\
&\approx \ln\left(1 + \exp\left\{\overline{d-p}\right\}\right) + (1-\phi)\left(\left[d_{t+1} - p_{t+1}\right] - \overline{d-p}\right), &(9)
\end{aligned}$$

[1] A first-order Taylor expansion states that a nonlinear function $f(x)$ can be made approximately linear, i.e. $f(x) \approx f(\overline{x}) + \frac{\partial f(x)}{\partial x}\Big|_{x=\overline{x}}(x - \overline{x})$.

where $\phi = 1/(1 + D/P)$ if the dividend-price ratio is constant. This then implies that we can write log returns as

$$r_{t+1} \approx p_{t+1} + (1-\phi)(d_{t+1} - p_{t+1}) + \kappa - p_t \tag{10}$$
$$= \kappa + \phi p_{+1} + (1-\phi) d_{t+1} - p_t, \tag{11}$$

where

$$\kappa = \ln\left(1 + \exp\left\{\overline{d-p}\right\}\right) - (1-\phi)\left(\overline{d-p}\right) \tag{12}$$
$$= \ln\left(\frac{1}{\phi}\right) - (1-\phi)\ln\left(\frac{1-\phi}{\phi}\right) \tag{13}$$
$$= -\ln(\phi) - (1-\phi)\left(\frac{1}{\phi} - 1\right). \tag{14}$$

The expression in (11) is a log-linear expression for log returns. We can re-write the expression into an log-linear approximation for the log stock price

$$p_t = \kappa + \phi p_{+1} + (1-\phi) d_{t+1} - r_{t+1}. \tag{15}$$

We can solve (15) forward to obtain

$$p_t = \kappa + \phi \left\{ \underbrace{\kappa + \phi p_{+2} + (1-\phi) d_{t+2} - r_{t+2}}_{p_{t+1}} \right\} + (1-\phi) d_{t+1} - r_{t+1} \tag{16}$$
$$= \kappa + \phi\kappa + \phi^2 p_{t+2} + \phi(1-\phi) d_{t+2} - \phi r_{t+2} + (1-\phi) d_{t+1} - r_{t+1}. \tag{17}$$

Continuing in this fashion recursively, taking conditional expectations, and imposing a so-called transversality condition $\lim_{j\to\infty} \phi^j p_{t+j} = 0$ that rules out speculative bubbles, we obtain an approximate pricing formula

$$p_t = \frac{\kappa}{1-\phi} + \mathbb{E}_t\left[\sum_{j=0}^{\infty} \phi^j \left[(1-\phi) d_{t+1+j} - r_{t+1+j}\right]\right]. \tag{18}$$

We can similarly obtain an approximate expression for the log dividend-price ratio. We start from the definition of the log price and re-arrange

$$p_t = \kappa + \phi p_{+1} + (1-\phi) d_{t+1} - r_{t+1} \tag{19}$$
$$= \kappa + \phi p_{+1} + d_{t+1} - \phi d_{t+1} - r_{t+1} \tag{20}$$
$$= \kappa\phi(p_{t+1} - d_{t+1}) + d_{t+1} - r_{t+1} \tag{21}$$

Next, subtract d_t from both side to obtain an expression for the price-dividend ratio

$$p_t - d_t = \kappa + \phi\left(p_{t+1} - d_{t+1}\right) + \Delta d_{t+1} - r_{t+1}, \tag{22}$$

where $\Delta d_{t+1} = d_{t+1} - d_t$ denotes the change in log dividends. Last, change signs to obtain an expression for the dividend-price ratio

$$d_t - p_t = -\kappa + \phi\left(d_{t+1} - p_{t+1}\right) - \Delta d_{t+1} + r_{t+1}. \tag{23}$$

We can iterate (23) forward as above, take conditional expectations, and apply a transversality condition to obtain

$$d_t - p_t = \frac{-\kappa}{1-\phi} + \mathbb{E}_t\left[\sum_{j=0}^{\infty} \phi^j\left[-\Delta d_{t+1+j} + r_{t+1+j}\right]\right]. \tag{24}$$

This equation is central to the predictability debate because it implies that the log dividend-price must predict either future expected dividends, future expected returns, or a combination of the two. It also tells us that the log dividend-price ratio should predict returns with a positive coefficient.

2. The Stambaugh finite sample bias

Consider the following system in which the predictor x_t follows an AR(1) process (Nelson and Kim, 1993, Kothari and Shanken, 1997, Stambaugh, 1999)

$$r_{t+1} = \alpha + \beta x_t + \varepsilon_{t+1}, \quad \varepsilon_{t+1} \sim iid\left(0,\sigma_\varepsilon^2\right) \tag{25}$$

$$x_{t+1} = \lambda + \rho x_t + \nu_{t+1}, \quad \nu_{t+1} \sim iid\left(0,\sigma_\nu^2\right), \tag{26}$$

where ε_{t+1} and ν_{t+1} are white noise errors that are contemporaneously correlated with covariance $\sigma_{\varepsilon\nu}$. To facilitate the derivation, consider the matrix version of the system at time $t = 1$ specified as

$$\boldsymbol{r} = \boldsymbol{X}\boldsymbol{b} + \boldsymbol{\varepsilon} \tag{27}$$

$$\boldsymbol{x} = \boldsymbol{X}\boldsymbol{\theta} + \boldsymbol{\nu} \tag{28}$$

where

$$\boldsymbol{r} = \begin{bmatrix} r_2 \\ r_3 \\ \vdots \\ r_T \end{bmatrix}, \quad \boldsymbol{X} = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_{T-1} \end{bmatrix}, \quad \boldsymbol{x} = \begin{bmatrix} x_2 \\ x_3 \\ \vdots \\ x_T \end{bmatrix} \quad \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_2 \\ \varepsilon_3 \\ \vdots \\ \varepsilon_T \end{bmatrix}, \quad \boldsymbol{\nu} = \begin{bmatrix} \nu_2 \\ \nu_3 \\ \vdots \\ \nu_T \end{bmatrix}, \quad \boldsymbol{b} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}, \quad \boldsymbol{\theta} = \begin{bmatrix} \lambda \\ \rho \end{bmatrix}.$$

The OLS estimates of the parameter vectors b and θ are then given by

$$\widehat{\boldsymbol{b}} = \left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\boldsymbol{r} = \boldsymbol{b} + \left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}X'\boldsymbol{\varepsilon} \tag{29}$$

$$\widehat{\boldsymbol{\theta}} = \left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\boldsymbol{x} = \boldsymbol{\theta} + \left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\boldsymbol{\nu}. \tag{30}$$

If $\boldsymbol{\varepsilon}$ and $\boldsymbol{\nu}$ are correlated so that $\gamma \neq 0$ in the regression

$$\boldsymbol{\varepsilon} = \gamma\boldsymbol{\nu} + \boldsymbol{\eta}, \tag{31}$$

then we can insert (31) into (29) and obtain the result that

$$\begin{aligned}
\widehat{\boldsymbol{b}} - \boldsymbol{b} &= \left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\boldsymbol{\varepsilon} & (32)\\
&= \left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\left(\gamma\boldsymbol{\nu} + \boldsymbol{\eta}\right) & (33)\\
&= \gamma\left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\boldsymbol{\nu} + \left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\boldsymbol{\eta} & (34)\\
&= \gamma\left(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}\right) + \left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\boldsymbol{\eta}, & (35)
\end{aligned}$$

where the last equality follows from (30). Last, take expectations to arrive at the result

$$\mathbb{E}\left[\widehat{\boldsymbol{b}} - \boldsymbol{b}\right] = \gamma\mathbb{E}\left[\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}\right]. \tag{36}$$

3. Cochrane's defence using unpredictable dividends

We start from the Campbell and Shiller (1988) loglinear approximation in (11) and rewrite it in the following way

$$\begin{aligned}
r_{t+1} &\approx \kappa + \phi p_{+1} + (1-\phi)\,d_{t+1} - p_t & (37)\\
&\approx \kappa + \phi p_{+1} + d_{t+1} - \phi d_{t+1} - p_t & (38)\\
&\approx \kappa + \phi\left(p_{t+1} - d_{t+1}\right) + d_{t+1} - p_t. & (39)
\end{aligned}$$

We can then use this expression to find an approximation for the log stock price

$$p_t \approx \kappa + \phi\left(p_{t+1} - d_{t+1}\right) - r_{t+1} + d_{t+1}. \tag{40}$$

Next, we can obtain an expression for the log price-dividend ratio by subtracting d_t from both sides of (11), i.e.,

$$p_t - d_t \approx \kappa + \phi\left(p_{t+1} - d_{t+1}\right) - r_{t+1} + \Delta d_{t+1}, \tag{41}$$

where $\Delta d_{t+1} = d_{t+1} - d_t$. Finally, change signs to obtain the log dividend-price ratio

$$d_t - p_t \approx -\kappa + \phi\left(d_{t+1} - p_{t+1}\right) + r_{t+1} - \Delta d_{t+1}. \tag{42}$$

Cochrane (2008) defends return predictability by directing attention to the inability of the log dividend-price to predict dividend growth. Consider the following system

$$
\begin{aligned}
\Delta d_{t+1} &= \alpha_d + \beta_d \left(d_t - p_t\right) + \varepsilon_{d,t+1} && (43)\\
r_{t+1} &= \alpha_r + \beta_r \left(d_t - p_t\right) + \varepsilon_{r,t+1} && (44)\\
d_{t+1} - p_{t+1} &= \lambda + \rho \left(d_t - p_t\right) + \varepsilon_{dp,t+1} && (45)
\end{aligned}
$$

and insert these relations in the approximation for the log dividend-price ratio in (42) to obtain

$$
\begin{aligned}
d_t - p_t &\approx -\kappa + \phi \left(\lambda + \rho \left(d_t - p_t\right) + \varepsilon_{dp,t+1}\right) \\
&\quad + \left(\alpha_r + \beta_r \left(d_t - p_t\right) + \varepsilon_{r,t+1}\right) - \left(\alpha_d + \beta_d \left(d_t - p_t\right) + \varepsilon_{d,t+1}\right) && (46)\\
&\approx \left(\phi\rho + \beta_r - \beta_d\right) \left(d_t - p_t\right) + \left(\phi\varepsilon_{dp,t+1} + \varepsilon_{r,t+1} - \varepsilon_{d,t+1}\right) \\
&\quad + \left(-\kappa + \phi\lambda + \alpha_r - \alpha_d\right). && (47)
\end{aligned}
$$

As noted by Cochrane (2008), the Campbell and Shiller (1988) loglinearization then implies the following restrictions on the parameters of the system

$$
\begin{aligned}
1 &= \phi\rho + \beta_r - \beta_d && (48)\\
0 &= \phi\varepsilon_{dp,t+1} + \varepsilon_{r,t+1} - \varepsilon_{d,t+1} && (49)\\
0 &= -\kappa + \phi\lambda + \alpha_r - \alpha_d && (50)
\end{aligned}
$$

4. The optimal portfolio weights for utility analysis

Suppose that the investor has mean-variance preferences of the form

$$
\max_{\omega_t} \ \mathbb{E}\left[U\left(r_{p,t+1}\right)\right] = \mathbb{E}_t\left[r_{p,t+1}\right] - \frac{1}{2}\gamma \mathrm{Var}\left[r_{p,t+1}\right], \tag{51}
$$

where $r_{p,t+1} = r_{f,t+1} + \omega_t \left(r_{t+1} - r_{f,t+1}\right)$ and γ is the Pratt-Arrow coefficient of relative risk aversion. We can rewrite the optimization problem as

$$
\max_{\omega_t} \ \mathbb{E}\left[U\left(r_{p,t+1}\right)\right] = r_{f,t+1} + \omega_t \mathbb{E}_t\left[r_{t+1} - r_{f,t+1}\right] - \frac{1}{2}\gamma\omega_t^2 \mathrm{Var}\left[r_{t+1} - r_{f,t+1}\right]. \tag{52}
$$

The first-order condition (FOC) for this problem then becomes

$$
\frac{\partial \mathbb{E}\left[U\left(r_{p,t+1}\right)\right]}{\partial \omega_t} = \mathbb{E}_t\left[r_{t+1} - r_{f,t+1}\right] - \gamma\omega \mathrm{Var}\left[r_{t+1} - r_{f,t+1}\right] = 0, \tag{53}
$$

which implies that the optimal portfolio weights are given by the expression

$$
\omega_t = \left(\frac{1}{\gamma}\right) \frac{\mathbb{E}_t\left[r_{t+1} - r_{f,t+1}\right]}{\mathrm{Var}\left[r_{t+1} - r_{f,t+1}\right]} \tag{54}
$$

References

Campbell, J.Y., Shiller, R.J., 1988. The dividend-price ratio and expectations of future dividends and discount factors. Review of Financial Studies 1, 195–228.

Cochrane, J.H., 2008. The dog that did not bark: A defense of return predictability. Review of Financial Studies 21, 1533–1575.

Kothari, S.P., Shanken, J., 1997. Book-to-market, dividend yield, and expected market returns: A time-series analysis. Journal of Financial Economics 44, 169–203.

Nelson, C.R., Kim, M.J., 1993. Predictable stock returns: The role of small sample bias. Journal of Finance 48, 641–661.

Stambaugh, R.F., 1999. Predictive regressions. Journal of Financial Economics 54, 375–421.