Predictability of Stock Market Returns

February 26, 2020

1 Log-linearized Models of Stock and Bond Returns

1.1 Stock Returns and the dynamic dividend growth model

Consider the one-period total holding returns in the stock market, that are defined as follows:[1]

$$H_{t+1}^{s} \equiv \frac{P_{t+1}+D_{t+1}}{P_t}-1=\frac{P_{t+1}-P_t+D_{t+1}}{P_t}=\frac{\Delta P_{t+1}}{P_t}+\frac{D_{t+1}}{P_t}, \quad (1)$$

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:

$$r_{t+1}^{s} \equiv \log\left(1+H_{t+1}^{s}\right)=\log\left(\frac{P_{t+1}+D_{t+1}}{P_t}\right)=\log\left(P_{t+1}+D_{t+1}\right)-\log\left(P_t\right). \quad (2)$$

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 (1) by $\left(1+H_{t+1}^{s}\right)$ and multiplying both sides by P_t/D_t we have:

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

[1] The use of '$\equiv$' emphasizes that (1) 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:[2]

$$p_t - d_t = -r^s_{t+1} + \Delta d_{t+1} + \ln\left(1 + e^{p_{t+1}-d_{t+1}}\right) \tag{3}$$

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 (3), we obtain that the log price-dividend ratio is defined as:[3]

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

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, $\rho\left(p_{t+1} - d_{t+1}\right) - \left(p_t - d_t\right) =$

[2] $-r^s_{t+1}$ 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

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

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

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

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

Under the assumption that there can be no rational bubbles, i.e., that[4]

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

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, $r^s_{t+j} \equiv r^f + (r^s_{t+j} - r^f)$.[5] 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 $m \longrightarrow \infty$. Note that, under the null hypothesis of constancy of returns, the volatility of the price dividend ratio should be completelyt explained

[4] 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 $\rho^m (p_{t+m} - d_{t+m})$ is zero.

[5] 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.

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.

1.2 Bond Returns

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:

$$
\begin{aligned}
D^c_{t,T} &= Y^c_{t,T}\sum_{i=1}^{T-t}\frac{i}{\left(1+Y^c_{t,T}\right)^i}+\frac{(T-t)}{\left(1+Y_{t,T}\right)^{T-t}} \\
&= Y^c_{t,T}\frac{\left((T-t)\frac{1}{1+Y^c_{t,T}}-(T-t)-1\right)\frac{1}{\left(1+Y^c_{t,T}\right)^{T-t+1}}+\frac{1}{1+Y^c_{t,T}}}{\left(1-\frac{1}{1+Y^c_{t,T}}\right)^2}+\frac{(T-t)}{\left(1+Y_{t,T}\right)^{T-t}} \\
&= \frac{1-\left(1+Y^c_{t,T}\right)^{-(T-t)}}{1-\left(1+Y^c_{t,T}\right)^{-1}},
\end{aligned}
$$

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:

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

1.3 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:

$$E\left[\frac{y_{t,T} - \gamma y_{t+1,T}}{1-\gamma} \mid I_t\right] = r_t + \phi_{t,T} \tag{4}$$

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

$$y_{t,T} = y^*_{t,T} + E[\Phi_T \mid I_t] = \frac{1-\gamma}{1-\gamma^{T-t}} \sum_{j=0}^{T-t-1} \gamma^j E[r_{t+j} \mid I_t] + E[\Phi_T \mid I_t] \tag{5}$$

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:

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

2 Linearized Present Value models for Consumption

The accumulation equation for aggregate wealth may be written as:

$$W_{t+1} = (1 + R_{m,t+1})(W_t - C_t) \tag{6}$$

Define $r_{m,t+1} = \log(1 + R_{m,t+1})$, and use lowercase letters to denote log variables throughout. As LL we follow Campbell and Mankiw (1989) and assume that the consumption–aggregate wealth ratio is stationary. In this case the budget constraint may be approximated by taking a first-order Taylor expansion of equation (6) , to obtain

$$\begin{aligned} \Delta w_{t+1} &= r_{m,t+1} + k + \left(1 - \frac{1}{\rho}\right)(c_t - w_t) \\ \rho &= 1 - \exp\left(\overline{c - w}\right) \end{aligned} \tag{7}$$

where k, is a constant of normalization, not relevant for the problem at hand. By solving (7)forward, we have :

$$c_t - w_t = E_t\left[\sum_{j=1}^{\infty}\rho^j\left(r_{m,t+j} - \Delta c_{t+j}\right)\right] + \frac{\rho k}{1-\rho} \tag{8}$$

LL point out that (8) shows that the consumption–wealth ratio is a function of expected future returns to the market portfolio in a broad range of optimal consumption models, so they concentrate in finding a proxy for $c_t - w_t$ and in assessing its performance for forecasting market returns.

To illustrate how we take equation (**??**) to the data note that, following Campbell(1996), we approximate the log of total wealth as:

$$w_t = va_t + (1-v)\,h_t$$

where v is a constant of linearization, equal to the average share of asset holdings in total wealth, a_t is the log of asset holdings and h_t is the log of human capital. While we have available data for financial wealth, the measurement of h_t is not immediate. To find an empirical counterpart of this variable consider that labour income can be interpreted as a dividend on human capital (see Julliard(2004)):

$$1 + R_{h,t+1} = \frac{H_{t+1} + Y_{t+1}}{H_t}$$

Log-linearizing this relation around the steady state human capital-labor income ratio ($\frac{Y}{H} = \frac{1}{\rho_h} - 1$) we have:

$$r_{h,t+1} = (1-\rho_h)\,k_h + \rho_h\,(h_{t+1} - y_{t+1}) - (h_t - y_t) + \Delta y_{t+1}$$

By solving this relation forward and by imposing the transversality condition we have:

$$h_t = y_t + \sum_{i=1}^{\infty}\rho_h^{i-1}\left(\Delta y_{t+i} - r_{h,t+i}\right) + k_h$$

so the log of human capital to income ratio is determined by discounted sum of future labour income growth and human capital returns.

Consistently with our linearization for wealth, the total return on wealth can be approximated by:

$$r_{m,t} = v r_{a,t} + (1-v)\, r_{h,t} + k_r$$

we decompose the unobservable $r_{h,t}$ into a part correlated with $r_{a,t}$ and a part orthogonal to it:

$$r_{h,t} = \beta r_{a,t} + \epsilon_t$$

By substituting all these relationships in the optimality condition we have:

$$\begin{aligned}
c_t - v a_t - (1-v)\, y_t &= (1-\psi)\, E_t \left[\sum_{j=1}^{\infty} \rho^j \left(v + (1-v)\, \beta \right) r_{a,t+j} \right] + k + \\
&\quad \sum_{j=1}^{\infty} E_t \rho_h^{j-1} \left(\Delta y_{t+j} - \beta r_{a,t+j} \right) + \eta_t \qquad (9) \\
\eta_t &= \sum_{j=1}^{\infty} \rho_h^{j-1} \epsilon_{t+j}
\end{aligned}$$

where η_t is an unobservable stationary component.

because when $|x| < 1$,

3 Present Value Models and Forecasting Regressions for Stock market Returns

Forecasting regressions for stock market returns can be interpreted in the framework of the dynamic dividend growth model.

The dynamic dividend growth model of Campbell and Shiller(1988) uses a loglinear approximation to the definition of returns on the stock market to express the log of the price-dividend ratio $p_t - d_t$ as a linear function of the future discounted dividend growth,Δd_{t+j}, and of future returns,h^s_{t+j}:

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

This basic relation allows to classify virtually all forecasting regression of stock market returns in terms of different approaches to proxying the future expected variables included in the linearized relations.

- The classical Gordon-growth model posits constant dividend growth $E_t \Delta d_{t+j} = g$ and constant returns $E_t h^s_{t+j} = r$.In this case we have

$$p^*_t = k + d_t + \frac{g}{1-\rho g} - \frac{r}{1-\rho r}$$

- The Lander et al.(1997) model also known as the FED model can be understood by substituting out the no-arbitrage restrictions in (**??**) $E_t h^s_{t+j} = E_t \left(r_{t+j} + \phi^s_{t+j} \right)$ and then by assuming constant dividend growth and a close relation between the risk premium on long-term bonds and the risk premium on stocks in this case we have:

$$p^*_t = k + d_t + \frac{g}{1 - \rho g} - \beta R_t$$

  where R_t is th yield to maturity on long-term bonds

- Asness(2003) considers the assumption of proportionality between the stock market risk premium and the bond market risk premium as problematic and corrects the FED model with the following specification:

$$p^*_t = k + d_t + \frac{g}{1 - \rho g} - \beta_1 R_t - \beta_2 \frac{\sigma^s_t}{\sigma^B_t}$$

  where $\frac{\sigma^s_t}{\sigma^B_t}$ is the ratio between the historical volatility of stock and bonds.

- Lamont(2004) argues that the log dividend payout ratio $(d_t - e_t)$ is the most appropriate proxy for future stock market returns to consider the following model:

$$p^*_t = k + d_t + \frac{g}{1 - \rho g} + \beta_1 (d_t - e_t)$$

- Ribeiro(2005) higlight the importance of labour income in predicting future dividends and posits VECM error correction model for dividend growth and future returns with two cointegrating vectors defined as $(d_t - y_t)$ and $(d_t - p_t)$, hence the implicit equilibrium stock market price is:

$$p^*_t = k + d_t + \frac{g}{1 - \rho g} + \beta_1 (d_t - y_t)$$

The empirical investigation of the dynamic dividend growth model has established a few empirical results:

(i) dp_t is a very persistent time-series and forecasts stock market returns and excess returns over horizons of many years (Fama and French (1988), Campbell and Shiller (1988), Cochrane (2005, 2007).

(ii) dp_t does not have important long-horizon forecasting power for future discounted dividend-growth (Campbell (1991), Campbell, Lo and McKinlay (1997) and Cochrane (2001)).

(iii) the very high persistence of dp_t has led some researchers to question the evidence of its forecasting power for returns, especially at short-horizons. Careful statistical analysis that takes full account of the persistence in dp_t provides little evidence in favour of the stock-market return predictability based on this financial ratio (Nelson and Kim (1993); Stambaugh (1999); Ang and Bekaert (2007); Valkanov (2003); Goyal and Welch (2003) and Goyal and Welch (2008)). Structural breaks have also been found in the relation between dp_t and future returns (Neely and Weller (2000), Paye and Timmermann (2006) and Rapach and Wohar (2006)).

(iv) More recently, Lettau and Ludvigson (2001, 2005) have found that dividend growth and stock returns are predictable by long-run equilibrium relationships derived from a linearized version of the consumer's intertemporal budget constraint. The excess consumption with respect to its long run equilibrium value is defined by the authors alternatively as a linear combination of labour income and financial wealth, cay_t, or as a linear combination of aggregate dividend payments on human and non-human wealth, cdy. cay_t and cdy_t are much less persistent than dp_t, they are predictors of stock market returns and dividend-growth, and, when included in a predictive regression relating stock market returns to dp_t, they swamp the significance of this variable. Lettau and Ludvigson (2005) interpret this evidence in the light of the presence of a common component in dividend growth and stock market returns. This component cancels out from (**??**), cay_t and cdy_t are instead able to capture it as the linearized intertemporal consumer budget constraint delivers a relationship between excess consumption and expected dividend growth or future stock market returns that is independent from their difference.

4 The Dog that did not bark

Consider the following DGP:

$$
\begin{aligned}
r^s_{t+1} &= \alpha_r + b_r dp_t + \varepsilon_{1,t+1} \\
\Delta d_{t+1} &= \alpha_d + b_d dp_t + \varepsilon_{2,t+1} \\
dp_{t+1} &= \alpha_{dp} + \varphi dp_t + \varepsilon_{3,t+1}
\end{aligned}
$$

given that

$$
\begin{aligned}
r^s_{t+1} &= \alpha_d + b_d dp_t + \varepsilon_{2,t+1} - \rho(\alpha_{dp} + \varphi dp_t + \varepsilon_{3,t+1}) + dp_t \\
r^s_{t+1} &= (\alpha_d - \rho\alpha_{dp}) + (1 - \rho\varphi + b_d)\, dp_t + \varepsilon_{2,t+1} - \rho\varepsilon_{3,t+1}
\end{aligned}
$$

we have

$$
r^s_{t+1} = \Delta d_{t+1} - \rho dp_{t+1} + dp_t
$$

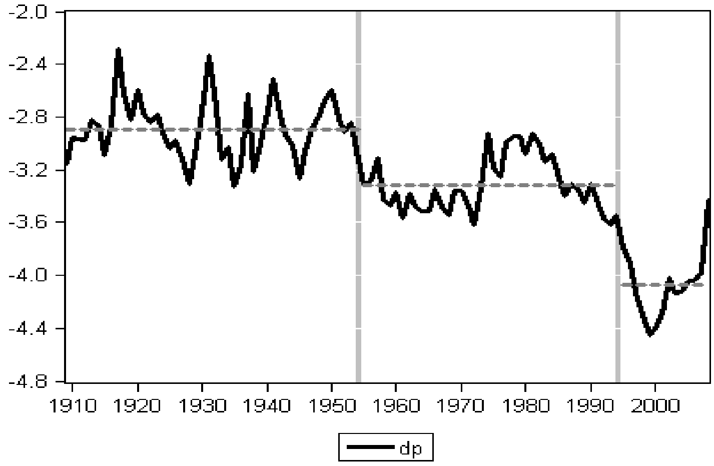
$$
\begin{aligned}
b_r &= 1 - \rho\varphi + b_d \\
\varepsilon_{1,t+1} &= \varepsilon_{2,t+1} - \rho\varepsilon_{3,t+1}
\end{aligned}
$$

In particular, as long as φ is nonexplosive, $\varphi < 1/\rho$■ 1.04, we cannot choose a null in which both dividend growth and returns are unforecastable, i.e. in which both $b_r = 0$ and $b_d = 0$. To generate a coherent null with b_r =0, we must assume an equally large b_d of the opposite sign, and then we must address the failure of this dividend growth forecastability in the data.

5 Time varying mean for dp

A recent strand of the empirical literature has related the contradictory evidence on the dynamic dividend growth model to the potential weakness of its fundamental hypothesis that the log dividend-price ratio is a stationary process (Lettau and Van Nieuwerburgh (2008), LVN henceforth). LVN use a century of US data to show evidence on the breaks in the constant mean $\overline{\text{dp}}$. We report the time series of US data on dp_t over the last century in Figure 1. As a matter of fact, the evidence from univariate test for non-stationarity and bivariate cointegration tests does not lead to the rejection of the null of the presence of a unit-root in dp. So far lineariztion of the dividend price ratio has been implemented around a constant. An interesting possibility is to allow for a time-varying mean in dp :

Time Series of Log Dividend-Price Ratio



$$\begin{aligned}
\Delta d_{t+1} &= \varepsilon_{1,t+1} && (10)\\
dp_{t+1} &= \varphi_{22} dp_t + \varphi_{23} X_{t+1} + \varepsilon_{2,t+1} && (11)\\
r^s_{t+1} &= \Delta d_{t+1} - \rho \left[dp_{t+1} - \overline{dp}_{t+1} \right] + \left[dp_t - \overline{dp}_t \right] + \varepsilon_{3,t+1} && (12)\\
\begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \\ \varepsilon_{3,t} \end{bmatrix} &\sim \left[\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \Sigma \right]
\end{aligned}$$

Where X_{t+1} is a a time-varying variable that determines the mean of dp_{t+1}.In Appendix B of their paper, Lettau&VanNieuwerburgh (2008) derive the following log-linear approximation of returns[6]:

$$\left(r^s_{t+1} - \overline{r}_{t+1} \right) = \left(\Delta d_{t+1} - \overline{\Delta d}_{t+1} \right) - \rho_t \left[dp_{t+1} - \overline{dp}_{t+1} \right] + \left[dp_t - \overline{dp}_t \right] + \Delta \overline{dp}_{t+1} \qquad (13)$$

We obtain our equation (**??**) from (13) by assuming that of the three processes for returns, dividend growth and the dividend-price ratio only the last one is persistent $\left(\overline{r}_t = \overline{\Delta d}_t = const \right)$, that the time-varying mean of the dividend-price ratio is very slowly evolving, i.e. $\Delta \overline{dp}_{t+1} \approx 0$ and that the linearization parame-

[6] $r^s_{t+1} = \ln \left(\frac{P_{t+1} + D_{t+1}}{P_t} \right)$,
$dp_t = \ln \left(\frac{D_t}{P_t} \right)$

ter is constant, $\rho_t = \rho$.[7] We introduce an error term $\varepsilon_{3,t+1}$ to capture the effect of our approximation[8].

The speed of mean reversion towards a constant mean of the dividend-price ratio is very different from that of annual real returns and annual real dividend growth. We model this feature of the data by introducing a time-varying mean for the dividend-price ratio, driven by X_{t+1}. In practice, without this step it would be very hard to reconcile the time-series properties of dp_t with those of r_t^s and Δd_{t+j}.

By solving eq. (**??**) forward we obtain:

$$\sum_{j=1}^{m} \rho^{j-1} \left(r_{t+j}^s \right) = \left[dp_t - \overline{dp}_t \right] + \sum_{j=1}^{m} \rho^{j-1} \left(\Delta d_{t+j} \right) - \rho^m \left[dp_{t+m} - \overline{dp}_{t+m} \right] + \sum_{j=1}^{m} \rho^{j-1} \left(\varepsilon_{1,t+j} + \varepsilon_{3,t+j} \right) \quad (14)$$

Eq. (14) clearly shows that deviations of the dividend/price ratio from its equilibrium value at time t have a predictive power for m-period ahead stock market returns (and/or dividend growth) that increases with the horizon, as the larger is m the smaller is the effect of future noise in the dividend-price ratio $\left[dp_{t+m} - \overline{dp}_{t+m} \right]$. However, this term cannot be ignored in the computation of the term structure of stock market risk that considers typically horizons from 1-year onwards. To bring (14) to the data, an observable counterpart of the time varying linearization value for the dividend-price must be considered. Consistently with (**??**), we assume that the relevant linearization value for computing returns from time t to time $t + m$ is the conditional expectation of the dividend-yield for time $t + m$, given the information available at time t. We then have

$$\begin{aligned} \sum_{j=1}^{m} \rho^{j-1} \left(r_{t+j}^s \right) &= dp_t - \left[\varphi_{22}^m dp_t + \sum_{j=1}^{m} \varphi_{22}^{j-1} \varphi_{23} X_{t+m+1-j} \right] + u_{t+m} \quad (15) \\ &= \left(1 - \varphi_{22}^m \right) dp_t - \sum_{j=1}^{m} \varphi_{22}^{j-1} \varphi_{23} X_{t+m+1-j} + u_{t+m} \\ u_{t+m} &= \sum_{j=1}^{m} \rho^{j-1} \left(\varepsilon_{1,t+j} + \varepsilon_{3,t+j} \right) - \rho^m \sum_{j=1}^{m} \varphi_{22}^{j-1} \varepsilon_{2,t+m+1-j} \end{aligned}$$

Note that the specification of the model requires that future values of X are used to predict return. One alternative, followed by Lettau and Van Niuewen-

[7] Rytchkov (2008) estimates a system of equation similar to ours and study how sensitive ML parameters are to variation in this parameter. He concludes that there is almost no sensitivity to the choice of ρ (see Table 1 in his paper).

[8] The validity of these assumptions will be subject to a test in the empirical section where we shall evaluate the "restricted" model generated by all our assumptions against an unrestricted one.

burgh is to allow for shift in the mean, another possibility, followed by Favero et al., is to make X function of demographic variables.

Table 3: System Estimation (1910-2008)

$$dp_{t+1} = \varphi_{20} + \varphi_{22} dp_t + \varphi_{23} MY_{t+1} + \varepsilon_{2t+1}$$

$$\text{UM:} \frac{1}{\sqrt{m}} \sum_{j=1}^{m} \left(r^s_{t+j} \right) = \delta_{0m} + \frac{\delta_{1m}}{\sqrt{m}} dp_t + \frac{\delta_{2m}}{\sqrt{m}} \left(\sum_{j=1}^{m} \varphi_{22}^{j-1} MY_{t+j} \right) + u_{t+m} \quad m = 1, ..., 10$$

$$\text{RM:} \frac{1}{\sqrt{m}} \sum_{j=1}^{m} \left(r^s_{t+j} \right) = \delta_{0m} + \frac{1}{\sqrt{m}} \left(1 - \varphi_{22}^{m}\right) dp_t - \frac{\varphi_{23}}{\sqrt{m}} \left(\sum_{j=1}^{m} \varphi_{22}^{j-1} MY_{t+j} \right) + u_{t+m}$$

		horizon m in years									
UM		1	2	3	4	5	6	7	8	9	10
δ_{1m}		0.49	0.65	0.71	0.76	0.77	0.77	0.77	0.78	0.79	0.78
$(t-stat)$		(11.93)	(17.39)	(21.15)	(24.89)	(24.24)	(28.02)	(27.75)	(26.87)	(24.96)	(18.80)
δ_{2m}		0.96	0.86	0.83	0.83	0.83	0.84	0.86	0.90	0.94	0.98
$(t-stat)$		(7.19)	(7.62)	(7.54)	(7.21)	(7.01)	(6.83)	(6.77)	(6.85)	(6.90)	(6.95)
φ_{22}	0.52										
$(t-stat)$	(10.07)										
φ_{23}	-0.95										
$(t-stat)$	(-6.28)										
RM											
φ_{22}	0.77										
$(t-stat)$	(20.94)										
φ_{23}	-0.51										
$(t-stat)$	(4.12)										
χ^2_{13}	17.28 (0.19)	χ^2_{20}	106.44 (0.00)								
σ_{DepVar}		0.198	0.201	0.191	0.190	0.188	0.181	0.180	0.178	0.174	0.171
$\sigma_{u_{t+m}}$	UM	0.213	0.191	0.170	0.154	0.141	0.132	0.124	0.119	0.114	0.117
$\sigma_{u_{t+m}}$	RM	0.191	0.181	0.165	0.154	0.143	0.135	0.127	0.122	0.117	0.114
$adjR^2$	UM	–	0.10	0.20	0.34	0.43	0.47	0.51	0.56	0.58	0.59
$adjR^2$	RM	0.07	0.19	0.25	0.34	0.42	0.44	0.49	0.54	0.55	0.55

Table 1: The restricted system is estimated using GMM with optimal Newey-West bandwidth selection to compute GMM standard errors. Ann.Unc.σ is the annualized unconditional standard deviation. Ann.$\sigma_r(m)$ is the annualized conditional standard deviation of the compounded (over m periods) returns, i.e. our measure of stock market risk. The effective sample period is 1910-2008.

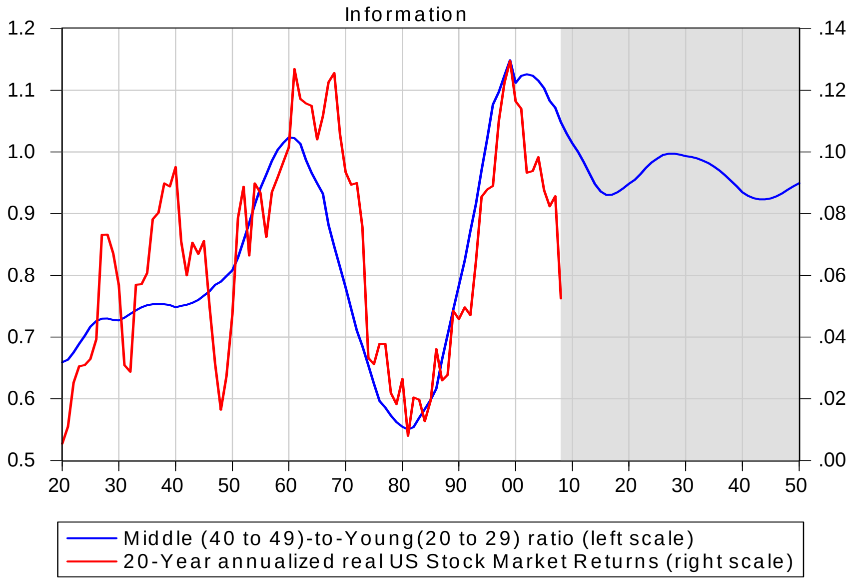


Figure 1.2: 20-year real US stock market returns and demographic trends

5.1 Spurious Regressions and the predictability of returns at different frequencies

The evidence for cointegration between price and dividends is not so clear-cut. In fact, the log of the price-dividends ratio is a very persistent time series and the possibility that it contains a unit root cannot be rule-out a priori. As a matter of fact we have used in our empirical analysis so far the UK dividend price ratio but the evidence from US data speaks less favorably in favour of a mean-reverting (log) of dividend price ratio. A widespread use empirical evidence in favour of the dynamic dividend growth model, that supports the stationarity of the log dividend yield, is the one based on multi-period predictive regressions for stock market returns. The performance of the log dividend yield as a predictor of stock market returns improves as the length of the horizon at which returns are defined increases. Table 3.4 illustrates this evidence by reporting the performance of predictive regression for stock UK 1-quarter, 1-year, 2-year and 3-year stock market returns based on the dividend yield.

TABLE 3.4. Forecasting UK Stock-Market Returns at different horizons

Dependent variable $\sum_{j=1}^{k} (h_{t+j}^{s})$, regression by OLS, 1973:1-2011:4

horizon	β_0^k	β_1^k	R^2	S.E
1-quarter	0.304 (0.093)	0.087 (0.028)	0.0631	0.0069
1-year	1.08 (0.16)	0.31 (0.05)	0.21	0.02
2-year	1.85 (0.198)	0.52 (0.06)	0.34	0.0325
3-year	2.49 (0.2)	0.70 (0.06)	0.46	0.0432

$$\sum_{j=1}^{k} (h_{t+j}^{s,uk}) = \beta_0^k - \beta_1^k (p_t - d_t) + \varepsilon_{t,t+j}$$

$k = 1, 4, 8, 12$

h_t^s are log total annualized real UK stock market returns.

The evidence that long-horizon variables seem to find significant results where "short-term" approaches have failed, has been questioned. Valkanov(2003) argues that long-horizon regressions will always produce "significant" results, whether or not there is a structural relation between the underlying variables. This result depend on the fact that a rolling summation of series integrated of order zero behaves asymptotically as a series integrated of order one and, whenever the regressor is persistent, the well-know occurrence of spurious regression between I(1) variables emerges. Having established that estimation and testing using long-horizon variables cannot be carried out using the usual regression methods, Valkanov(2003) provides a simple guide on how to conduct estimation and inference using long-horizon regressions. The author proposes propose a rescaled t-statistic, t/$\sqrt{T}$, for testing long-horizon regressions. the asymptotic distribution of this statistic, although non-normal, is easy to simulate and the results are applicable to a general class of long-horizon regressions. In deriving his correction Valkanov also illustrates that the problem related to spurious regression goes beyond the inadequacy of statistical asymptotic approximation when using overlapping variables. In fact he shows that, even after correcting for serially correlated errors , using Hansen and Hodrick (1980) or Newey-West (1987) standard errors, the small-sample distribution of the estimators and the t-statistics are very different from the asymptotic normal distribution.

To illustrate the Valkanov rescaling procedure consider the following DGP:

$$\begin{aligned} r^1_{t+1} &= \alpha + \beta lpd_t + \epsilon_{1t}, \\ (1+\phi L)\, lpd_t &= \mu + \epsilon_{2t}. \\ \phi &= 1 + \frac{c}{T} \\ \begin{pmatrix} \epsilon_{1t} \\ \epsilon_{2t} \end{pmatrix} &\sim N.I.D.\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix} \right) \end{aligned} \quad (16)$$

where the parameter c measures deviations from the unit root in a decreasing (at rate T) neighborhood of 1. The unit-root case corresponds to $c = 0$.

The long-horizon variables are

$$r^k_{t+1} = \sum_{j=1}^{k} r^1_{t+1}$$

The regression at different horizon is run by projecting Z^k_t on lpd_t.The simulation of the relevant distribution requires an estimate of the nuisance parameter c. To this end long-run restrictions implied by the dynamic dividend growth model can be used.

As shown in the introductory chapter, the model implies that one-period total return can be approximated as follows:

$$r^1_{t+1} = \rho_0 + \rho lpd_{t+1} + \Delta d_{t+1} - lpd_t \quad (17)$$

assuming that the log-dividends follows an autoregressive process:

$$lpd_{t+1} = \phi lpd_t + u_t. \quad (18)$$

by substituting from (18) into (17) we have that

$$\begin{aligned} r^1_{t+1} &= \rho_0 - \beta_1 lpd_t + \varepsilon_{t+1} \\ \varepsilon_{t+1} &= \Delta d_{t+1} + u_t. \\ \beta_1 &= (1 - \rho\phi) \end{aligned} \quad (19)$$

where ε_{t+1} is a stationary variable and therefore the $E_t r^k_{t+1} = \beta_1 lpd_t$. The k-period horizon return can then be written as follows:

$$\begin{aligned} r_{t+1}^{k} &\approx \tilde{k} - \beta_k x_t + \tilde{\epsilon}_{t+1} \\ \beta_k &= \left[(1-\rho\phi) \sum_{i=0:}^{k-1} \phi^i \right] \end{aligned} \tag{20}$$

Now, we can write

$$\beta_k = \left[(1-\rho\phi) \frac{1-\phi^k}{1-\phi} \right] \tag{21}$$

If $\rho = 1$ then:

$$\beta_k = 1 - \phi^k \tag{22}$$

Now, remember that $\phi = 1 + \frac{c}{T}$ and we can express k in terms of the total length of the available sample as,$k = \lfloor \lambda T \rfloor$, from which $T \approx \frac{k}{\lambda}$. Then:

$$\begin{aligned} &\beta_k = 1 - \left(1 + \frac{c}{T}\right)^k = 1 - \left(1 + \frac{c\lambda}{k}\right)^k \\ &\lim_{k \longrightarrow \infty} \left(1 + \frac{c\lambda}{k}\right)^k = e^{c\lambda} \\ &\lim_{k \longrightarrow \infty} \beta_k = \lim_{T \longrightarrow \infty} \beta_{\lfloor \lambda T \rfloor} = 1 - e^{c\lambda} \end{aligned} \tag{23}$$

Since we can estimate β_k consistently, we can also find a consistent estimate of c by using the transformation:

$$c^{CONSISTENT} = \frac{1}{\lambda} \log(1 - \beta_k)$$

Given the knowledge of c and λ, the model can be simulated under the null to obtain the critical values of the Valkanov t-statistics.

Note that the empirical literature on predictability also cast doubts on the validity of the cointegrating relationships between dividend and prices and different models have been proposed based on alternative cointegrating relationships (see, for example the FED model by Lander et al.(1997) or the cay model by Lettau and Ludgvison(2004)). The instability of parameter estimates in econometric models has generated alternative approaches based on stationary representations of the return dynamics (Ferreira and Santa Clara(2011)).

6 Risk, Returns and Portfolio Allocation with Cointegrated VARs

Consider the continuously compounded stock market return from time t to time $t+1$, $\mathbf{r}_{t+1}$. Define $\boldsymbol{\mu}_t$, the conditional expected log return given information up to time t, as follows:

$$\mathbf{r}_{t+1} = \boldsymbol{\mu}_t + \mathbf{u}_{t+1}$$

where $\mathbf{u}_{t+1}$ is the unexpected log return. Define the k-period cumulative return from period $t+1$ through period $t+k$, as follows:

$$\mathbf{r}_{t,t+k} = \sum_{i=1}^{k} \mathbf{r}_{t+i}$$

The term structure of risk is defined as the conditional variance of cumulative returns, given the investor's information set, scaled by the investment horizon

$$\Sigma_r(k) \equiv \frac{1}{k} Var(\mathbf{r}_{t,t+k} \mid D_t) \tag{24}$$

where $D_t \equiv \sigma\{z_k : k \leq t\}$ consists of the full histories of returns as well as predictors that investors use in forecasting returns.

6.1 Inspecting the mechanism: a bivariate case

We illustrate the econometrics of the term structure of stock market risk by considering a simple bi-variate first-order VAR for continuously compounded total stock market returns, r_t^s, and the log dividend price,dp_t:

$$\begin{aligned} (z_t - E_z) &= \Phi_1 (z_{t-1} - E_z) + \nu_t \\ \nu_t &\sim \mathcal{N}(0, \Sigma_\nu) \end{aligned}$$

where

$$\begin{aligned} z_t &= \begin{bmatrix} r_t^s \\ dp_t \end{bmatrix}, E_z = \begin{bmatrix} E_{r^s} \\ E_{d-p} \end{bmatrix} \\ \Phi_1 &= \begin{bmatrix} 0 & \varphi_{1,2} \\ 0 & \varphi_{2,2} \end{bmatrix} \\ \begin{bmatrix} v_{1,t} \\ v_{2,t} \end{bmatrix} &\sim \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{matrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{matrix} \right] \end{aligned}$$

The bivariate model for returns and the predictor features a restricted dynamics such that only the lagged predictor is significant to determine current returns $\left(\varphi_{1,1}=0\right)$ and the predictor is itself a strongly exogenous variable $\left(\varphi_{2,1}=0\right)$.

Given the VAR representation and the assumption of constant Σ_ν

$$\begin{aligned} Var_t\left[(z_{t+1}+...+z_{t+k}) \mid D_t\right] &= \Sigma_\nu + (I+\Phi_1)\Sigma_\nu(I+\Phi_1)' + \\ & (I+\Phi_1+\Phi_1^2)\Sigma_\nu(I+\Phi_1+\Phi_1^2)' + ... \\ & +(I+\Phi_1+...+\Phi_1^{k-1})\Sigma_\nu(I+\Phi_1+...+\Phi_1^{k-1})' \end{aligned} \quad (25)$$

from which we can derive:

$$\begin{aligned} \Sigma_r(k) &= \frac{1}{k}\sum_{i=0}^{k-1} D_i \Sigma D_i' \\ D_i &= I + \Phi_1 \Xi_{i-1} \quad i>0 \\ \Xi_i &= \Xi_{i-1} + \Phi_1^i \quad i>0 \\ D_0 &\equiv I, \ \ \Xi_0 \equiv I \end{aligned}$$

Note that, under the chosen specification of the matrix Φ_1 we can write the generic term $D_i\Sigma D_i'$, as follows:

$$D_i\Sigma D_i' = \left(\begin{array}{c|c} M_{11} & M_{12} \\ \hline M_{12}' & M_{22} \end{array}\right) \quad (26)$$

$$M_{11} = \Sigma_{1,1} + \Phi_{1,2}\Xi_{i-1}^{(22)}\Sigma_{1,2}' + \Sigma_{1,2}\Xi_{i-1}^{(22)\prime}\Phi_{1,2}' + \Phi_{1,2}\Xi_{i-1}^{(22)}\Sigma_{2,2}\Xi_{i-1}^{(22)\prime}\Phi_{1,2}'$$
$$M_{12}' = \Xi_i^{(22)}\Sigma_{1,2}' + \Xi_i^{(22)}\Sigma_{2,2}\Xi_{i-1}^{(22)\prime}\Phi_{1,2}'$$
$$M_{22} = \Xi_i^{(22)}\Sigma_{2,2}\Xi_i^{(22)\prime}$$

where we have used the fact that

$$\begin{aligned} \Xi_i &= \sum_{j=0}^{i}\Phi_1^i \\ &= \left(\begin{array}{c|c} 0 & \phi_{1,2}\sum_{j=0}^{i-1}\phi_{2,2}^j \\ \hline 0 & \sum_{j=0}^{i}\phi_{2,2}^j \end{array}\right) \end{aligned}$$

and

$$\begin{aligned} D_i &= I + \Phi_1\Xi_{i-1} \\ &= \left(\begin{array}{c|c} I & \phi_{1,2}\sum_{j=0}^{i-1}\phi_{2,2}^j \\ \hline 0 & \sum_{j=0}^{i}\phi_{2,2}^j \end{array}\right) \end{aligned}$$

Eq. (26) implies that, in our simple bivariate example, the term structure of stock market risk takes the form

$$\sigma_r^2(k) = \sigma_1^2 + 2\varphi_{1,2}\sigma_{1,2}\psi_1(k) + \varphi_{1,2}^2\sigma_{2,2}^2\psi_2(k) \quad (27)$$

where

$$\psi_1(k) = \frac{1}{k}\sum_{l=0}^{k-2}\sum_{i=0}^{l}\varphi_{2,2}^{i} \quad k>1$$
$$\psi_2(k) = \frac{1}{k}\sum_{l=0}^{k-2}\left(\sum_{i=0}^{l}\varphi_{2,2}^{i}\right)^2 \quad k>1$$
$$\psi_1(1) = \psi_2(1) = 0$$

The total stock market risk can be decomposed in three components: i.i.d uncertainty, σ_1^2 , mean reversion, $2\varphi_{1,2}\sigma_{1,2}\psi_1(k)$, and uncertainty about future predictors, $\varphi_{1,2}^2\sigma_{2,2}^2\psi_2(k)$. Without predictability $(\varphi_{1,2}=0)$ the entire term structure is flat at the level σ_1^2. This is the classical situation where portfolio choice is independent of the investment horizon. The possible downward slope of the term structure of risk depends on the second term, and it is therefore crucially affected by predictability and a negative correlation between the innovations in dividend price ratio and in stock market returns $(\sigma_{1,2})$, the third term is always positive and increasing with the horizon when the autoregressive coefficient in the dividend yield process is positive.

6.1.1 An illustration

$$\begin{aligned} r_{t+1}^{s} &= \alpha_r + b_r dp_t + \varepsilon_{1,t+1} \\ dp_{t+1} &= \alpha_{dp} + \varphi dp_t + \varepsilon_{2,t+1} \\ \begin{bmatrix} \varepsilon_{1,t+1} \\ \varepsilon_{2,t+1} \end{bmatrix} &\sim \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{matrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{matrix}\right] \end{aligned}$$

$$\begin{aligned} E(r_{t,t+1}^{s} \quad &| \quad D_t) = E(\alpha_r + b_r dp_t + \varepsilon_{1,t+1} \mid D_t) \\ &= \quad \alpha_r + b_r dp_t \end{aligned}$$

$$\begin{aligned} E(r_{t,t+1}^{s} \quad &| \quad D_t) = E(\alpha_r + b_r dp_t + \varepsilon_{1,t+1} + \mid D_t) \\ E(r_{t+1,t+2}^{s} \quad &| \quad D_t) = E(\alpha_r + b_r\alpha_{dp} + b_r\varphi dp_t + b_r\varepsilon_{2,t+1} + \varepsilon_{1,t+2} \mid D_t) \end{aligned}$$

$$\begin{aligned} \sigma_r^2(1) &= Var(r_{t,t+1}^{s} \mid D_t) = \sigma_1^2 \\ \sigma_r^2(2) &= \frac{1}{2}Var(r_{t,t+2}^{s} \mid D_t) \\ &= \frac{1}{2}\sigma_1^2 + \frac{1}{2}\sigma_1^2 + \frac{1}{2}b_r^2\sigma_2^2 + b_r\sigma_{12} \end{aligned}$$