Asset Pricing Notes. Chapter 5: Present Value Relations

Roman Sigalov

December 7, 2021

1 Return Predictability

Empirical Evidence The evidence for return predictability ca n be summarized as follows

- On horizons of days, weeks and even months individual stock returns are mostly negatively autocorrelated that results from compensation for providing liquidity during selling and buying pressures
- On horizons of days, weeks and even months stock indices are positively autocorrelated due to large cross-autocorrelation between stocks.
- At yearly horizons both individual stocks and indices show positive autocorrelation
- These estimates, however, have quite large standard errors $\implies$ no strong statistical significance.

2 Constant Discount Rates

2.1 Dividend Based Model

Return of an asset is

$$1 + R_{t+1} = \frac{D_{t+1} + P_{t+1}}{P_t} \implies E_t[1 + R_{t+1}] = E_t\left[\frac{D_{t+1} + P_{t+1}}{P_t}\right]$$

We start by assuming that $E_t[1 + R_{t+1}] = 1 + R = const$ over time. Under ths assumption

$$1 + R = E_t\left[\frac{D_{t+1} + P_{t+1}}{P_t}\right] \implies P_t = E_t\left[\frac{D_{t+1} + P_{t+1}}{1 + R}\right]$$

we can iterate this forward to get

$$P_t = E_t \sum_{k=1}^{\infty} \left(\frac{1}{1+R}\right)^k D_{t+k} + \lim_{k\to\infty} \left(\frac{1}{1+R}\right)^k P_{t+k} \tag{1}$$

Under the assumption that the limiting term is zero, we obtain a **dividend discount model**:

$$P_t = E_t \sum_{k=1}^{\infty} \left(\frac{1}{1+R}\right)^k D_{t+k} \tag{2}$$

One can show that the portfolio that reinvest dividends back into the asset follows a martingale. Denote the number of shares owned at time as N_t. At time $t+1$ get dividend $N_t D_{t+1}$ on N_t shares. This allows to buy $N_t D_{t+1}/P_{t+1}$ more shares of the asset. Hence, the number of shares next period is

$$N_{t+1} = N_t + N_t \frac{D_{t+1}}{P_{t+1}} = N_t \left(1 + \frac{D_{t+1}}{P_{t+1}}\right)$$

Discounted value of these shares $V_t \equiv N_t P_t/(1+R)^t$ follows a martingale[1]

$$E_t[V_{t+1}] = E_t\left[\frac{N_{t+1} P_{t+1}}{(1+R)^{t+1}}\right] = V_t$$

[1] This is a general feature of so-called self-financing portoflios. Once the value of a self-financing portfolio is properly discounted (multiplied by the SDF) it follows a random walk: $M_t W_t = E_t[M_{t+1} W_{t+1}]$. See notes on stochastic discount factor process in the end of this review

2.2 Shiller's Variance Bounds

Dividend discount model in equation (2) implies that

$$\sum_{k=1}^{\infty}\left(\frac{1}{1+R}\right)^k D_{t+k} = E_t\sum_{k=1}^{\infty}\left(\frac{1}{1+R}\right)^k D_{t+k} + \varepsilon \implies \sum_{k=1}^{\infty}\left(\frac{1}{1+R}\right)^k D_{t+k} = P_t + \varepsilon$$

$$\implies Var\left(\sum_{k=1}^{\infty}\left(\frac{1}{1+R}\right)^k D_{t+k}\right) \geq Var(P_t)$$

This means that realized dividends should move more than prices. This was rejected by Shiller (1981). This result started the literature on time-varying discount rates: if dividends don't move enough to explain the volatility of asset prices, then the other part – discount rates – should move a lot.

2.3 The Gordon Growth Formula

Gordon growth formula assumes that dividends are expected to grow at rate $G : E_t D_{t+k} = (1+G)^k E_t D_{t+1}$. Under this assumption

$$P_t = \frac{E_t D_{t+1}}{R-G}$$

if we omit subscripts we get

$$R = \underbrace{\frac{D}{P}}_{\text{income part of return}} + \underbrace{G}_{\text{capital-gain part of return}}$$

2.4 Earnings Based Models

2.5 Rational Bubbles

Let's go back to the limiting term in equation (1). If this term is not equal to zero then we have multiple solutions of the form

$$P_t = \underbrace{P_t^D}_{\text{dividend-discount model price}} + \underbrace{Q_t}_{\text{bubble term}} \quad \text{where } Q_t = E_t\frac{Q_t}{1+R}$$

i.e. the bubble term is expected to grow at rate $1+R$ so that the price is not explosive. *There is a lengthy discussion of bubbles on pages 132-134 of Campbell's book*

3 Time-Varying Discount Rates

Since there is an evidence of stock return predictability we need to account for this by introducing time-varying expected returns. In the present-value relations it is useful to think of expected returns as of compensation for risk: if future expected returns increase, this means that the systematic risk of an asset increased. To understand why exactly the risk may change we need a model. We are going to consider a few in Consumption CAPM part of the course.

Without some form of approximation working with time varying expected return becomes quite cumbersome since it combines a lot of multiplicative and additive terms in one expression. One approach is to use continuous time another is to log-linearize around the average dividend-price ratio. We do the second one

3.1 Campbell-Shiller Approximation

Campbell and Shiller (1988a) derive the following approximation for log return around the average log dividend price ratio $\overline{d-p}$

$$
\begin{aligned}
r_{t+1} &= \log(1+R_{t+1}) = \log\left(\frac{P_{t+1}+D_{t+1}}{P_t}\right) \\
&= \log(P_{t+1}+D_{t+1}) - \log(P_t) \\
&= \log\left(P_{t+1}\left(1+\frac{D_{t+1}}{P_{t+1}}\right)\right) - \log(P_t) \\
&= \log(P_{t+1}) - \log(P_t) + \log(1+\exp(\log(D_{t+1}) - \log(P_{t+1}))) \\
&= p_{t+1} - p_t + \log(1+\exp(d_{t+1}-p_{t+1})) \\
&= p_{t+1} - p_t + \log(1+\exp(\overline{d-p})) + \underbrace{\frac{\exp(\overline{d-p})}{1+\exp(\overline{d-p})}}_{1-\rho}(d_{t+1}-p_{t+1}-\overline{d-p}) \\
&= p_{t+1} - p_t + k + (1-\rho)(d_{t+1}-p_{t+1}) \\
&= k + \rho p_{t+1} + (1-\rho)d_{t+1} - p_t
\end{aligned}
$$

where k absorbed all constants and ρ is a constant that is close to and lower than 1. Importantly, the accuracy of this approximation doesn't improve as the time interval shrinks as was the case with portfolio return approximation. Instead the accuracy depends on the movements of the price-dividend ratio around its mean.

Price Decomposition We can express p_t as

$$
p_t = k + (1-\rho)d_{t+1} - r_{t+1} + \rho p_{t+1}
$$

and iterate it forward imposing a no-bubble condition $\lim_{j\to\infty}\rho^j p_{t+j} = 0$ to get

$$
\begin{aligned}
p_t &= \frac{k}{1-\rho} + \sum_{j=0}^{\infty}\rho^j(1-\rho)d_{t+1+j} - \sum_{j=0}^{\infty}\rho^j r_{t+1+j} \\
&= \frac{k}{1-\rho} + p_{CF,t} + p_{DR,t}
\end{aligned}
\tag{3}
$$

where we have two components of the price: $p_{CF,t}$ – component of the price coming from cash-flows/dividends and $p_{DR,t}$ – effect of discount rates on the price. Higher price today ($p_t \uparrow$) with unchanged future dividends ($p_{CF,t} = const$) implies lower future returns as captured by $p_{DR,t}\downarrow$. The flip side is that higher returns in the future ($p_{DR,t}\uparrow$) with unchanged dividends ($p_{CF,t} = const$) implies lower price today. This is the present value logic that is often misunderstood.

As mentioned we can think about discount rates as risk-premium. Hence, if the risk of the asset increases (either the quantity of risk or the price of risk) agents will perceive asset as relatively unattractive and, hence, the price will decline today. This is captured by the present value relation in equation (3).

Divide-Price Decomposition In case that dividends follow a unit root process Campbell and Shiller suggest to subtract dividend since $d_t - p_t$ is stationary (at least at appears to be). We get

$$\begin{aligned}
d_t - p_t &= -\frac{k}{1-\rho} + \sum_{j=0}^{\infty} \rho^j \left((1-\rho)d_{t+1+j} - (1-\rho)d_t\right) + \sum_{j=0}^{\infty} \rho^j r_{t+1+j} \\
&= -\frac{k}{1-\rho} - \sum_{j=0}^{\infty} \rho^j (1-\rho)(d_{t+1+j} - d_t) + \sum_{j=0}^{\infty} \rho^j r_{t+1+j} \\
&= -\frac{k}{1-\rho} - \sum_{j=0}^{\infty} \rho^j (1-\rho) \sum_{i=0}^{j} \Delta d_{t+1+j} + \sum_{j=0}^{\infty} \rho^j r_{t+1+j} \\
&= -\frac{k}{1-\rho} - [(1-\rho)\Delta d_{t+1+0} + \rho(1-\rho)\Delta d_{t+1+0} + \rho(1-\rho)\Delta d_{t+1+1} \\
&+ \rho^2(1-\rho)\Delta d_{t+1+0} + \rho^2(1-\rho)\Delta d_{t+1+1} + \rho^2(1-\rho)\Delta d_{t+1+2} + \dots] + \sum_{j=0}^{\infty} \rho^j r_{t+1+j} \\
&= -\frac{k}{1-\rho} - \sum_{j=0}^{\infty} \rho^j \Delta d_{t+1+j} + \sum_{j=0}^{\infty} \rho^j r_{t+1+j}
\end{aligned}$$

$$d_t - p_t = -\frac{k}{1-\rho} + dp_{CF,t} + dp_{DR,t} \tag{4}$$

Aggregate US dividends were smooth throughout history. If we don't expect dividends to move a lot or assume that it is a random walk, then variation in dividend-price ratio come from time varying discount rates. This is a theoretical reason why dividend-price ratio may predict future return. According to equation (4) higher dividend to price ratio (low valuation in terms of prices) with constant dividend component $dp_{CF,t}$ implies high $dp_{DR,t}$, i.e. higher expected returns.

Return Decomposition Take difference of expectations at time $t+1$ and t to get

$$\underbrace{(\mathbb{E}_{t+1} - \mathbb{E}_t)(d_t - p_t)}_{=0} = \underbrace{-(\mathbb{E}_{t+1} - \mathbb{E}_t)\frac{k}{1-\rho}}_{=0} - (\mathbb{E}_{t+1} - \mathbb{E}_t)\sum_{j=0}^{\infty} \rho^j \Delta d_{t+1+j} + (\mathbb{E}_{t+1} - \mathbb{E}_t)\sum_{j=0}^{\infty} \rho^j r_{t+1+j}$$

$$(\mathbb{E}_{t+1} - \mathbb{E}_t)\sum_{j=0}^{\infty} \rho^j r_{t+1+j} = (\mathbb{E}_{t+1} - \mathbb{E}_t)\sum_{j=0}^{\infty} \rho^j \Delta d_{t+1+j}$$

$$\begin{aligned}
(E_{t+1} - E_t) r_{t+1} &= (E_{t+1} - E_t)\sum_{j=0}^{\infty} \rho^j \Delta d_{t+1+j} - (E_{t+1} - E_t)\sum_{j=1}^{\infty} \rho^j r_{t+1+j} \\
(E_{t+1} - E_t) r_{t+1} &= N_{CF,t} - N_{DR,t}
\end{aligned} \tag{5}$$

unexpected return at time $t+1$ can be decomposed into news about future growth rate of dividends/cash-flows and news about future discount rates. The second component is going to be important in understanding intertemporal CAPM later on.

3.2 Short and Long Run Predictability

Set-Up Suppose that expected returns are described by a persistent AR(1) process

$$E_t[r_{t+1}] = x_{t+1} = \phi x_t + \xi_{t+1}$$

and return is a sum of expected and unexpected component

$$r_{t+1} = \overline{r} + x_t + u_{t+1}$$

Campbell-Shiller decomposition tells us that unexpected return can be written in terms of news to future discount rates and future expected returns as

$$r_{t+1} - E_t r_{t+1} = N_{CF,t+1} - N_{DR,t+1}$$

In our statistical model

$$r_{t+1} - E_t r_{t+1} = \overline{r} + x_t + u_{t+1} - E_t[\overline{r} + x_t + u_{t+1}] = u_{t+1} = N_{CF,t+1} - N_{DR,t+1}$$

Therefore, we can the return as

$$r_{t+1} = \overline{r} + x_t + N_{CF,t+1} - N_{DR,t+1}$$

Dividend Price Ratio First, use dividend price ratio in equation (4) to write down part of the price dividend ratio that is drive by discount rate

$$\begin{aligned} dp_{DR,t} &= E_t \sum_{j=0}^{\infty} \rho^j r_{t+1+j} \\ &= E_t \sum_{j=0}^{\infty} \rho^j [\overline{r} + x_{t+j} + u_{t+1+j}] \\ &= \frac{\overline{r}}{1-\rho} + \frac{x_t}{1-\phi\rho} \end{aligned}$$

The variance of this component of price dividend ratio is

$$Var(dp_{DR,t}) = \frac{Var(x_t)}{(1-\rho\phi)^2}$$

Even is the expected return has a small volatility (Var(x) is small) the effect of that on the volatility of price-dividend ratio may be very large if the expected return process is very persistent (ϕ is close to 1). To

Discount Rate News In this model discount rate news can be calculated as

$$\begin{aligned} N_{DR,t} &= (E_{t+1} - E_t) \sum_{j=1}^{\infty} \rho_j r_{t+1+j} \\ &= (E_{t+1} - E_t) \sum_{j=1}^{\infty} \rho_j \left[\overline{r} + x_{t+j} + u_{t+1+j}\right] \\ &= (E_{t+1} - E_t) \left[\rho x_{t+1} + \rho^2 x_{t+2} + \rho^3 x_{t+1} + \dots\right] \\ &= (E_{t+1} - E_t) \left[\rho(\phi x_t + \xi_{t+1}) + \rho^2(\phi^2 x_t + \phi\xi_{t+1} + \xi_{t+2}) + \rho^3(\phi^3 x_t + \phi^2\xi_{t+1} + \phi\xi_{t+2} + \xi_{t+3}) + \dots\right] \\ &= (E_{t+1} - E_t) \left[\rho\xi_{t+1} + \rho^2\phi\xi_{t+1} + \rho^3(\phi^2\xi_{t+1}) + \dots\right] \\ &= \left[\rho\xi_{t+1} + \rho^2\phi\xi_{t+1} + \rho^3(\phi^2\xi_{t+1}) + \dots\right] \\ &= \frac{\rho\xi_{t+1}}{1-\rho\phi} \approx \frac{\xi_{t+1}}{1-\phi} \end{aligned}$$

Here, we similarly get that for a given variance of expected return innovation ξ the variance of discount rate news is increasing in the persistent of expected return ϕ.

Autocorrelation of returns Now we turn to the analysis of autocorrelation of returns

$$
\begin{aligned}
Cov(r_{t+1}, r_{t+1+j}) &= Cov(x_t + N_{CF,t+1} - N_{DR,t+1}, x_{t+j} + N_{CF,t+1+j} - N_{DR,t+1+j}) \\
&= Cov(x_t + N_{CF,t+1} - N_{DR,t+1}, x_{t+j} + N_{CF,t+1+j} - N_{DR,t+1+j}) \\
&= Cov\left(x_t + N_{CF,t+1} - \frac{\rho\xi_{t+1}}{1-\rho\phi}, \phi^j x_t + \phi^{j-1}\xi_{t+1} + \phi^{j-2}\xi_{t+2} + \cdots + N_{CF,t+1+j} - \frac{\rho\xi_{t+1+j}}{1-\rho\phi}\right) \\
&= \phi^j \sigma_x^2 + Cov(N_{CF,t+1}, \phi^{j-1}\xi_{t+1}) - Cov\left(\frac{\rho\xi_{t+1}}{1-\rho\phi}, \phi^{j-1}\xi_{t+1}\right) \\
&= \phi^j \frac{\sigma_\xi^2}{1-\phi^2} + \phi^{j-1} Cov(N_{CF,t+1}, \xi_{t+1}) - \frac{\rho\phi^{j-1}}{1-\rho\phi}\sigma_\xi^2 \\
&= \phi^{j-1}\left[Cov(N_{CF,t+1}, \xi_{t+1}) + \sigma_\xi^2\left(\frac{\phi}{1-\phi^2} - \frac{\rho}{1-\rho\phi}\right)\right]
\end{aligned}
$$

In principle, all the autocovariances can be zero (or close to zero) if all the terms in the square brackets cancel each other. Hence, this will imply that there is no predictability of future returns from past returns even thought there exist a state variable x_t that predicts returns. Therefore, prices can be weak form efficient but not semi-string form efficient.

3.3 Evidence on Time Varying Expected Returns

We have the following facts about the history of US returns

- D/P ratio is around 2% today compared to the long-term average of 4%. From the Gordon growth formula $D/P = R - G$ this means 2.4% decrease in the gap between required return and growth of dividends
- D/P ratio has little power in predicting dividends and predicts returns. This means that most of the variation in D/P comes from time-varying discount rates rather than time varying dividends
- In the same spirit, Price-to-Smoothed-Earnings ratio doesn't predict future earnings but predicts future 1-year returns. One should bear in mind that in these regressions however, there are only 10 independent (not-overlapping) samples of 10 year length
- Evidence based on VAR analysis of return finds that for **broad stock indexes** the standard deviation of discount rate news is about twice the standard deviation of cash-flow news. This is related to the inability of valuation ratio to predict dividends or earnings growth.
- In contrast, on **individual stock level** there is very little time series predictability of stock characteristics on future returns. This suggests that most of the variation of individual stocks is attributed to cash-flow news.

4 Predictive Return Regressions

4.1 Stambaugh Bias and Reponses

Kendall Bias Kendall (1954) showed that there is a downward bias in estimating AR(1) coefficient ϕ in regression

$$x_{t+1} = (1-\phi)\overline{x} + \phi x_t + \xi_{t+1}$$

that comes from the fact that the mean of the process is estimated at the same time as the persistence coefficient

Stambaugh Bias Stambaugh (1999) shows that when estimating a predictive return regression

$$r_{t+1} = \alpha + \beta x_t + u_{t+1}$$

Kendall bias generates a bias in β. In particular,

$$E[\hat{\beta} - \beta] = \frac{cov(\xi, u)}{var(\xi)} E[\hat{\phi} - \phi]$$

When we use Dividend-Price ratio as a predictive variable, then innovations in returns and in the predictive variable are negatively correlated $\implies cov(\xi, u) < 0$ meaning that

$$E[\hat{\beta} - \beta] = \underbrace{\frac{cov(\xi, u)}{var(\xi)}}_{<0} \underbrace{E[\hat{\phi} - \phi]}_{<0} > 0$$

Which undermines predictive regressions.

Responses to Stambaugh (1999) There are several arguments that restore the validity of predictive regressions

1. We can construct a poor-man's correction for Stambaug bias

$$\hat{\beta}^{adj} = \hat{\beta} - \frac{cov(\xi, u)}{var(\xi)} E[\hat{\phi} - 1]$$

   using the worst case $\phi = 1$. Under this adjustment can't reject predictability

2. Cochrane (2008) uses the Campbell-Shiller approximation of dividend-price ratio

$$\begin{aligned} r_{t+1} &\approx k + \rho p_{t+1} + (1 - \rho) d_{t+1} - p_t \\ &= k - \rho(d_{t+1} - p_{t+1}) + d_{t+1} - d_t + d_t - p_t \\ &= k - \rho(d_{t+1} - p_{t+1}) + \Delta d_{t+1} + (d_t - p_t) \end{aligned}$$

   Then if we regress $r_{t+1}, d_{t+1} - p_{t+1}$ and Δd_{t+1} on $d_t - p_t$ to get coefficients β, ϕ and β_d, respectively they will be linked as

$$\beta = 1 + \beta_d - \rho\phi$$

   Suppose that $\rho = 0.96$ and $\phi \leq 1$. If $\beta = 0$, then $\beta_d = \beta - 1 + \rho\phi < 0$. However, the fact that β_d is close to zero implicitly means that $\beta > 0$