1 Return Predictability and Volatility

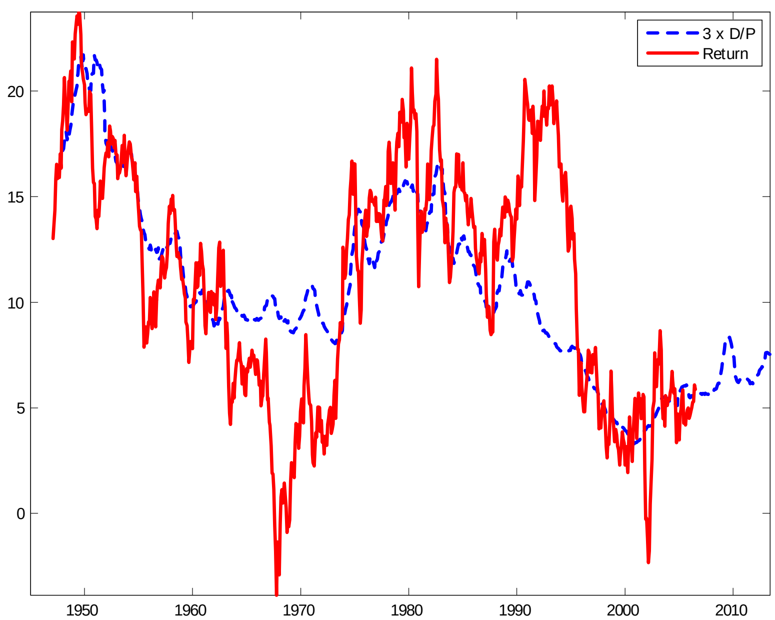
1.1 Three facts

1. Variables such as (D/P) forecast returns

   (a) "Discount rates" Table 1

Horizon k	b	t(b)	R^2	$\sigma\left[E_t(R^e)\right]$	$\frac{\sigma[E_t(R^e)]}{E(R^e)}$
1 year	3.8	(2.6)	0.09	5.46	0.76
5 years	20.6	(3.4)	0.28	29.3	0.62

- t is significant, though not dramatic.
- Return coefficient is economically large.
  i. 1% more dividend = 3.8% more return.
  ii. R^2 is big at long horizons *long-horizon returns are much more predictable.* Discount rates graph

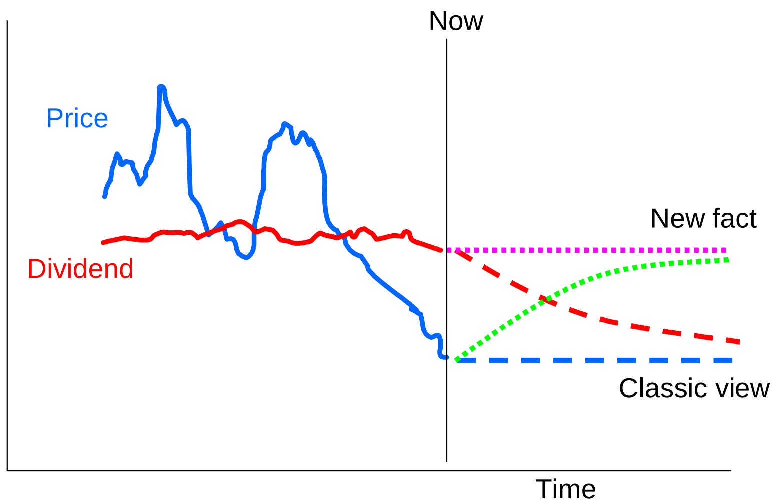


  iii. $R^2 = \sigma\left[E_t(R^e)\right]/\sigma(R^e)$ is not very interesting. $\sigma\left[E_t(R^e)\right]$ is very big compared to $E(R^e)$

2. DP does not forecast dividend growth, though it "should." The *sign* is "wrong"

Horizon k	$R^e_{t\to t+k} = a + b\frac{D_t}{P_t} + \varepsilon_{t+k}$			$\frac{D_{t+k}}{D_t} = a + b\frac{D_t}{P_t} + \varepsilon_{t+k}$		
(years)	b	t(b)	R^2	b	t(b)	R^2
1	4.0	2.7	0.08	0.07	0.06	0.0001
5	20.6	2.6	0.22	2.42	1.11	0.02

(Source: "Financial markets and the real economy") If expected returns are constant, dividends *should* be forecastable.



3. Prices seem awfully volatile, predictable or not. How can last week's 10% drop be "rational" change in expectations about dividends?

   (a) Shiller 1981

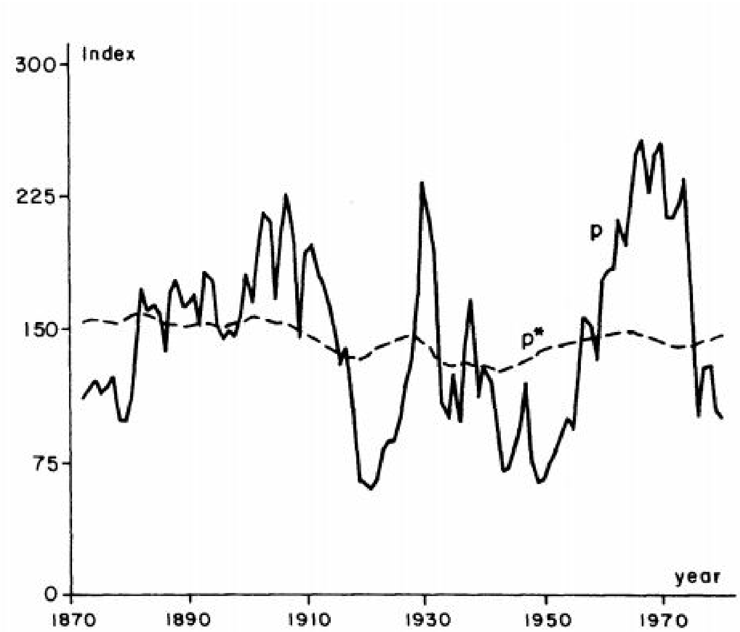


FIGURE 1

Note: Real Standard and Poor's Composite Stock Price Index (solid line p) and *ex post* rational price (dotted line p^*), 1871–1979, both detrended by dividing a long-run exponential growth factor. The variable p^* is the present value of actual subsequent real detrended dividends, subject to an assumption about the present value in 1979 of dividends thereafter. Data are from Data Set 1, Appendix.

   (b) Here is the same graph updated, from Shiller's Nobel Prize lecture

(c) Shiller equations P^* is the "ex-post rational" price

$$P_t^* = \sum_{j=1}^{\infty} \frac{1}{R^j} D_{t+j}$$

The actual price is its expectation

$$P_t = E_t\left(P_t^*\right) = E_t\left[\sum_{j=1}^{\infty} \frac{1}{R^j} D_{t+j}\right]$$

Expected values should vary less than the thing they are expecting.

$$\begin{aligned} P_t^* &= P_t + \varepsilon_t \\ \sigma^2(P_t^*) &= \sigma^2(P_t) + \sigma^2\left(\varepsilon_t\right) \\ \sigma^2(P_t^*) &> \sigma^2(P_t) \end{aligned}$$

(d) Q: A new and different test of efficiency having nothing to do with predictability?

(e) A: It is the same as prices don't predict dividends!

(f) A: It is the same as prices do predict returns!

4. Our task: tie all these ideas together.

1.2 Present value identity

1.2.1 One period version

1. Present value formula idea, for a security that lasts one period.

   (a) Take logs,

$$R_{t+1} = \frac{D_{t+1}}{P_t}$$
$$r_{t+1} = d_{t+1} - p_t$$
$$p_t - d_t = (d_{t+1} - d_t) - r_{t+1}$$
$$p_t - d_t = E_t(\Delta d_{t+1}) - E_t\left(r_{t+1}\right)$$

   (b) Prices are higher if expected returns are lower, or dividend growth is higher.

   (c) Content: none. This is the definition of returns. Content comes from models for expected return.

   (d) Quiz: If the expected return rises, the stock is more attractive, so people will push prices *up*, no?

2. Conclusions:

   (a) P-d can only vary if expected dividend growth is high, or expected returns are low. If Δd and r are coin flips (iid) then the p-d ratio is *constant.*

   (b) If traders see high $E_t \Delta d_{t+1}$, they drive up prices $p_t - d_t$. On average, we see higher Δd_{t+1} after high $p_t - d_t$; $p_t - d_t$ forecasts Δd_{t+1} *If price variation comes from news about dividend growth, then price-dividend ratios should forecast dividend growth.* (They don't) Conversely,

   (c) If traders see high $E_t r_{t+1}$, drive down p_t. On average, we see higher r_{t+1} after high $p_t - d_t$. *If price variation comes from news about changing discount rates, then price-dividend ratios should forecast returns.* (They do)

   (d) Our regressions are about how *prices* – the right hand variable – are formed!

3. Run a regression of both sides of

$$d_t - p_t = r_{t+1} - \Delta d_{t+1}$$

   on $d_t - p_t$, i.e. (notation)

$$\begin{aligned} r_{t+1} &= b_r dp_t + \varepsilon^r_{t+1} \\ \Delta d_{t+1} &= b_d dp_t + \varepsilon^d_{t+1} \end{aligned}$$

   Then

$$dp_t = \left[b_r dp_t + \varepsilon^r_{t+1}\right] - \left[b_d dp_t + \varepsilon^d_{t+1}\right]$$

   Result:

$$1 = b_r - b_d.$$
$$0 = \varepsilon^r_{t+1} - \varepsilon^d_{t+1}$$

   (a) The return and dividend growth coefficients must add up to one (if pd varies at all). If returns are not predictale, dividend growth is, and vice versa.

   (b) Tie volatility to predictability.

$$b_r = \frac{cov(r_{t+1}, dp_t)}{var(dp_t)}$$
$$var(dp_t) = cov(r_{t+1}, dp_t) - cov(\Delta d_{t+1}, dp_t)$$

   which is it? A:*all* $E\left(r\right)$

   - *Variation in price-dividend ratios corresponds entirely to changes in expected returns, not to changes in expected dividend growth.*

4. Agenda: do this for multiperiod securities.

1.2.2 Campbell-Shiller Present Value Formula

1. Linearized return identity

$$r_{t+1} \approx \rho(p_{t+1} - d_{t+1}) + \Delta d_{t+1} - (p_t - d_t)$$

small letters = logs, all variables deviations from means.

Derivation. This is just the definition of return:

$$\begin{aligned}
R_{t+1} &= \frac{P_{t+1} + D_{t+1}}{P_t} = \frac{\left(1 + \frac{P_{t+1}}{D_{t+1}}\right)\frac{D_{t+1}}{D_t}}{\frac{P_t}{D_t}} \\
r_{t+1} &= \log\left(1 + e^{pd_{t+1}}\right) + \Delta d_{t+1} - pd_t \\
r_{t+1} &\approx \log\left(1 + e^{pd}\right) + \frac{e^{pd}}{\left(1 + e^{pd}\right)}(pd_{t+1} - pd) + \Delta d_{t+1} - pd_t
\end{aligned}$$

$$\rho = \frac{1}{1 + D/P} \approx 0.96 \text{ (Annual, with } D/P = 0.04; \text{ P/D=20)}$$

2. Solve forward to present value formula.

$$pd_t \approx \rho pd_{t+1} + \Delta d_{t+1} - r_{t+1}$$

$$pd_t \approx \sum_{j=1}^{k} \rho^{j-1}\Delta d_{t+j} - \sum_{j=1}^{k} \rho^{j-1} r_{t+j} + \rho^k\left(pd_{t+k}\right)$$

when $\rho^k\left(pd_{t+k}\right) \to 0$ (it does, pd can't explode, "transversality condition"),

$$pd_t \approx \sum_{j=1}^{\infty} \rho^{j-1}\Delta d_{t+j} - \sum_{j=1}^{\infty} \rho^{j-1} r_{t+j} = \text{long run } \Delta d \text{ - long run } r$$

   (a) Stare at this. This is true ex-post. It's just the definition of long horizon return,

$$\sum_{j=1}^{k} \rho^{j-1} r_{t+j} \approx \sum_{j=1}^{k} \rho^{j-1}\Delta d_{t+j} + \rho^k\left(pd_{t+k}\right) - pd_t$$

   return comes from lots of dividends, a low initial price, or a high final price.

   (b) If it's true ex post, then also ex-ante. Apply E_t to both sides,

$$pd_t \approx E_t \sum_{j=1}^{\infty} \rho^{j-1}\Delta d_{t+j} - E_t \sum_{j=1}^{\infty} \rho^{j-1} r_{t+j}$$

   A present value formula. higher expected dividend growth or lower expected return → higher price

3. The nature of forecasting regressions

$$R_{t+1} = a + b \times \frac{D_t}{P_t} + \varepsilon_{t+1}$$

   (a) Cause and Effect

$$\text{actual temperature at } t+1 = a + b \times (\text{prediction made at } t) + \text{forecast error}_{t+1}$$

   (b) Reverse causality: $E_t(R_{t+1})$ rises and this pushes P_t down!

   (c) Errors are forecast errors, so uncorrelated with dp. That's why they go on the right.

   (d) We learn about prices with this regression!

   (e) R^2 is not answering an interesting question!

1.3 Volatility and Bubbles

1. Volatility

$$pd_t \approx E_t \sum_{j=1}^{\infty} \rho^{j-1} \Delta d_{t+j} - E_t \sum_{j=1}^{\infty} \rho^{j-1} r_{t+j}$$

   (a) *The P/D ratio moves if and only if there is news about long run dividend growth or returns. If $E_t(r_{t+j})$ and $E_t(\Delta d_{t+j})$ are constant, then $p_t - d_t$ must be a constant! P/D varies so we know we don't live in an iid world.*

   (b) Run both sides of

$$dp_t \approx \sum_{j=1}^{\infty} \rho^{j-1} r_{t+j} - E_t \sum_{j=1}^{\infty} \rho^{j-1} \Delta d_{t+j}$$

   on $d_t - p_t$. Result?

$$1 \approx b_r^{lr} - b_d^{lr}$$

   where b^{lr} means

$$\begin{aligned} \sum_{j=1}^{\infty} \rho^{j-1} \Delta d_{t+j} &= b_d^{lr} dp_t + \varepsilon^d \\ \sum_{j=1}^{\infty} \rho^{j-1} r_{t+j} &= b_r^{lr} dp_t + \varepsilon^r \end{aligned}$$

   (c) *Long-run* return forecast and dividend forecast must add to one. These are the right regressions to run in order to understand prices, not one-period regressions. This is just like the one-period case now. *If dividend yields vary, they must forecast long-run returns or dividend growth..*

   (d) Or, $b = cov(x, y)/var(x)$, so

$$var\left(p_t - d_t\right) \approx cov\left\{p_t - d_t, \sum_{j=1}^{\infty} \rho^{j-1} \Delta d_{t+j}\right\} - cov\left\{p_t - d_t, \sum_{j=1}^{\infty} \rho^{j-1} r_{t+j}\right\}$$

   (e) *$p - d$ varies if and only if it* **forecasts** *long run dividend growth or long run returns. Which* is it?

   (f) Volatility Facts: Summary **Table II** from "Discount rates" (See also *Asset Pricing* Table 20.3)

	Coefficient		
Method and horizon	$b_r^{(k)}$	$b_{\Delta d}^{(k)}$	$\rho^k b_{dp}^{(k)}$
Direct regression , $k = 15$	1.01	-0.11	-0.11
Implied by VAR, $k = 15$	1.05	0.27	0.22
VAR, $k = \infty$	1.35	0.35	0.00

$$\sum_{j=1}^{k} \rho^{j-1} r_{t+j} = a + b_r^{(k)} dp_t + \varepsilon_{t+k}^r$$

   - $p - d$ variation is almost all due to expected returns. It has nothing to do with expected dividend growth.
   - 100% / 0% has become 0%/100%!
   - Note – *on average. This* high P/D might be due to D news. *On average, in the past,* high P/D has meant low returns. *Period!*

2. Bubbles

(a) Use the k-year identity

$$dp_t = \sum_{j=1}^{k} \rho^{j-1} r_{t+j} - \sum_{j=1}^{k} \rho^{j-1} \Delta d_{t+j} + \rho^k dp_{t+k}$$

$$1 = b_r^{(k)} - b_d^{(k)} + \rho^k b_{pd}^{(k)}$$

(b) It's possible prices vary with no news of future returns, or dividends, but only on news of future prices. We would see this by running the regression backwards, future dp on today's dp.

(c) At short horizons it's all prices, $b_{pd}^{(k)}$ is large.

(d) Table II: *No rational bubbles!*

(e) So what's the big Fama / Shiller debate all about? Volatility = pd forecasts returns, pd does not forecast dividend growth. pd varies "because" $E(R)$ varies. Fama: $E(R)$ varies sensibly as the business cycle goes up or down. Shiller: E(R) varies more than "it should" even over the business cycle. It's fads of psychology. Resolution: write down macroeconomic or psychological models of $E(R)$ variation. No amount of arguing will settle this. Empirical work and identities really help to narrow down this argument!

3. Return decomposition in the VAR

$$(E_{t+1} - E_t) : pd_t \approx E_t \sum_{j=1}^{\infty} \rho^{j-1} \Delta d_{t+j} - E_t \sum_{j=1}^{\infty} \rho^{j-1} r_{t+j}$$

$$(E_{t+1} - E_t)\, r_{t+1} = (E_{t+1} - E_t)\, \Delta d_{t+1} + (E_{t+1} - E_t) \sum_{j=1}^{\infty} \rho^j \Delta d_{t+1+j} - (E_{t+1} - E_t) \sum_{j=1}^{\infty} \rho^j r_{t+1+j}$$

(a) Return shocks come from *current* dividend shocks, news to expected future dividends, and news to expected future returns. It will turn out about half is current dividends, half is expected future returns, and none is expected future dividends.

(b) This is not inconsistent with the finding that all dp variance comes from expected returns: When d moves, p moves one for one and dp does not change. So half of return variance is due to dividends but none of pd variance is due to dividends. But don't fall into the trap of confusing volatility of pd facts with volatility of p and volatility of r facts! The latter have current dividends in them.

1.4 Vector autoregression

1. The basic VAR

$$\begin{aligned} r_{t+1} &= b_r dp_t + \varepsilon^r_{t+1} \\ \Delta d_{t+1} &= b_d dp_t + \varepsilon^d_{t+1} \\ dp_{t+1} &= \phi dp_t + \varepsilon^{dp}_{t+1} \end{aligned}$$

   (a) It's just like the AR(1) but with a vector/matrix

$$\begin{aligned} \begin{bmatrix} r_{t+1} \\ \Delta d_{t+1} \\ dp_{t+1} \end{bmatrix} &= \begin{bmatrix} 0 & 0 & b_r \\ 0 & 0 & b_d \\ 0 & 0 & \phi \end{bmatrix} \begin{bmatrix} r_t \\ \Delta d_t \\ dp_t \end{bmatrix} + \begin{bmatrix} \varepsilon^r_{t+1} \\ \varepsilon^d_{t+1} \\ \varepsilon^{dp}_{t+1} \end{bmatrix} \\ \mathbf{x}_{t+1} &= \mathbf{A}\mathbf{x}_t + \boldsymbol{\varepsilon}_{t+1} \end{aligned}$$

   (b) Estimates, round numbers

	Estimates $\hat{b}, \hat{\phi}$	ε s. d. (diagonal) and correlation. r	Δd	dp
r	0.1	20	+big	-big
Δd	0		14	0
dp	0.94			15

   It is a cool and useful happenstance that dividend growth and dp shocks are basically uncorrelated.

2. Identity constrains coefficients and shocks:

$$r_{t+1} \approx -\rho dp_{t+1} + \Delta d_{t+1} + dp_t.$$

   There are really only two independent variables, and shocks. Returns come from price changes and dividend changes! Hence (E_t of the identity)

$$\begin{aligned} b_r &= 1 - \rho\phi + b_d \\ 0.1 &= 1 - 0.96 \times 0.94 + 0 \end{aligned}$$

   (($E_{t+1} - E_t$) of the identity)

$$\varepsilon^r_{t+1} \approx -\rho\varepsilon^{dp}_{t+1} + \varepsilon^d_{t+1}.$$

   (a) Restrictions on the variance / covariance follow

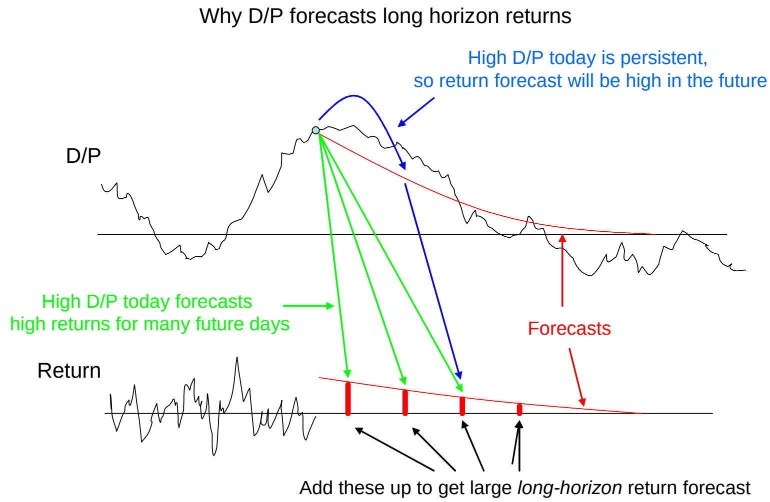
$$\begin{aligned} \sigma^2\left(\varepsilon^r_{t+1}\right) &\approx \rho^2\sigma^2\left(\varepsilon^{dp}_{t+1}\right) + \sigma^2\left(\varepsilon^d_{t+1}\right). \\ 0.20^2 &= 0.96^2 \times 0.15^2 + 0.14^2. \end{aligned}$$

$$cov(\varepsilon^r_{t+1}, \varepsilon^{dp}_{t+1}) = -\rho\sigma^2\left(\varepsilon^{dp}_{t+1}\right) << 0$$

   A positive dp shock usually has no news about dividends, so means a negative p shock and a negative r shock. ε^{dp} and ε^r are strongly negatively correlated. This negative correlation gives rise to "Stambaugh bias" and lots of other interesting phenomena

3. Use 1: Connecting long and short horizons.

   - The rise of coefficients and R^2 with horizon is not a separate phenomenon. It is a mechanical result of a small short horizon b and R^2 and a persistent (ϕ large) forecasting variable. Equivalently, it is the result of a forecasting variable that forecasts returns many periods in the future.



- Equations:

$$\begin{aligned} r_{t+1} &= b_r dp_t + \varepsilon^r_{t+1} \\ dp_{t+1} &= \phi dp_t + \varepsilon^{dp}_{t+1}. \end{aligned}$$

$$\begin{aligned} &\Leftrightarrow \quad r_{t+1} + r_{t+2} = b_r(1+\phi)dp_t + (error) \\ &\Leftrightarrow \quad r_{t+1} + r_{t+2} + r_{t+3} = b_r(1+\phi+\phi^2)dp_t + (error) \end{aligned}$$

$$\Leftrightarrow r_{t+2} = b_r \phi dp_t + (error)\,;\; r_{t+3} = b_r \phi^2 dp_t + (error)$$

- Long horizon b = short horizon b + persistent forecasting variable
- Long horizon $b = x_t$ predicts one-year returns far in the future
- Long horizon R^2

$$\begin{aligned} R^2_{k=1} &= \frac{b_r^2 \sigma^2(dp_t)}{\sigma^2\left(r_{t+1}\right)} \\ R^2_{k=2} &= \frac{b_r^2(1+\phi)^2\sigma^2(dp_t)}{\sigma^2\left(r_{t+1}+r_{t+2}\right)} \approx \frac{b_r^2(1+\phi)^2\sigma^2(dp_t)}{2\sigma^2\left(r_{t+1}\right)} = \frac{(1+\phi)^2}{2} R^2_{k=1} \end{aligned}$$

4. Volatility tests in the VAR

(a) Recall

$$\begin{aligned} \sum_{j=1}^{\infty} \rho^{j-1} \Delta d_{t+j} &= b_d^{lr} dp_t + \varepsilon^d \\ \sum_{j=1}^{\infty} \rho^{j-1} r_{t+j} &= b_r^{lr} dp_t + \varepsilon^r \end{aligned}$$

and we derived

$$1 = b_r^{lr} - b_d^{lr}.$$

We interpreted b_d^{lr}, b_r^{lr} as "fraction of d-p variance accounted for by cashflow/expected return variation" (I standardized here on d-p rather than p-d as the right hand variable)

(b) In the VAR

$$b_r^{lr} = \sum_{j=1}^{\infty} \rho^{j-1}\phi^{j-1} b_r = \frac{b_r}{1-\rho\phi}$$

Thus, the volatility decomposition is simply

$$\frac{b_r}{1-\rho\phi} - \frac{b_d}{1-\rho\phi} = b_r^{lr} - b_d^{lr} = 1$$

You can get here much more quickly just from the identity

$$\begin{aligned} b_r &= 1-\rho\phi + b_d \\ 1 &= \frac{b_r}{1-\rho\phi} - \frac{b_d}{1-\rho\phi} \end{aligned}$$

but then the interpretation isn't so clear.

(c) Simplified numbers

$$\begin{aligned} b_r^{lr} &= \frac{0.1}{1-0.94\times 0.96} = 1 \\ b_d^{lr} &= 0 \end{aligned}$$

Again, the return coefficient is just enough. The dividend coefficient is zero. No bubbles needed. These units are nice, because $b_r^{lr}=1$ is easier to remember than $b_r = 0.1$

1.5 Impulse-Response Function

1. There are only two shocks really. Recall

$$r_{t+1} \approx -\rho dp_{t+1} + \Delta d_{t+1} + dp_t.$$

$$\varepsilon^r_{t+1} \approx -\rho \varepsilon^{dp}_{t+1} + \varepsilon^d_{t+1}.$$

It makes no sense to move returns without moving both dp and d – it can't happen. So choose two shocks.

2. My choice.

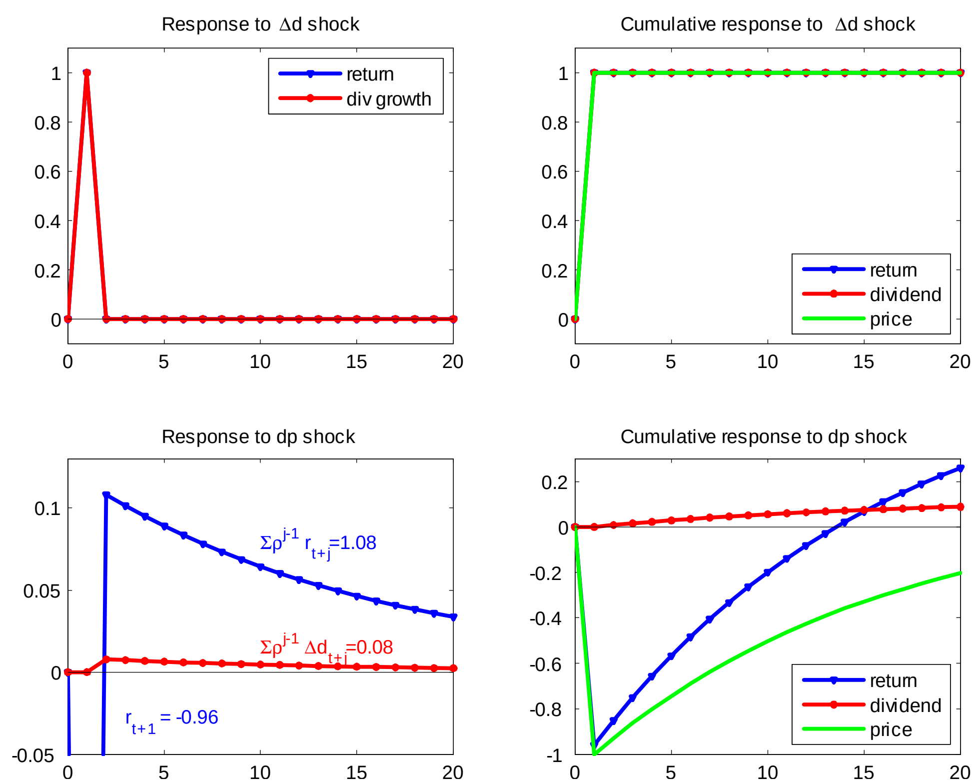
$$\begin{array}{rl} \text{Shock 1 ("}\Delta d\text{ shock"):} & \left[\begin{array}{ccc} \varepsilon^r_1 & \varepsilon^d_1 & \varepsilon^{dp}_1 \end{array}\right] = \left[\begin{array}{ccc} 1 & 1 & 0 \end{array}\right] \\ \text{Shock 2 ("}dp\text{ shock" or "}Er\text{" shock)} \quad : & \left[\begin{array}{ccc} \varepsilon^r_1 & \varepsilon^d_1 & \varepsilon^{dp}_1 \end{array}\right] = \left[\begin{array}{ccc} -\rho & 0 & 1 \end{array}\right] \end{array}$$

   (a) Shock 1: returns go up 1% because dividends go up 1%, with no change in dividend yield
   (b) Shock 2: no change to dividends, expected returns go up, so actual returns (prices) go down.

3. Impulse-response function: Simulate forward

$$\begin{aligned} r_{t+1} &= 0.108 \times dp_t + \varepsilon^r_{t+1} \\ \Delta d_{t+1} &= 0.015 \times dp_t + \varepsilon^d_{t+1} \\ dp_{t+1} &= 0.0937 \times dp_t + \varepsilon^{dp}_{t+1} \end{aligned}$$

starting with the above two shocks, and leaving all other shocks to zero.



4. Interpretation: how news about the future changes prices today.

   (a) ε^d, dividend shock with no dp change is a pure expected-cashflow shock with no change in expected returns

(b) ε^{dp}, dp shock with no change in dividends is (almost) a pure discount-rate shock with no change in expected cashflows.

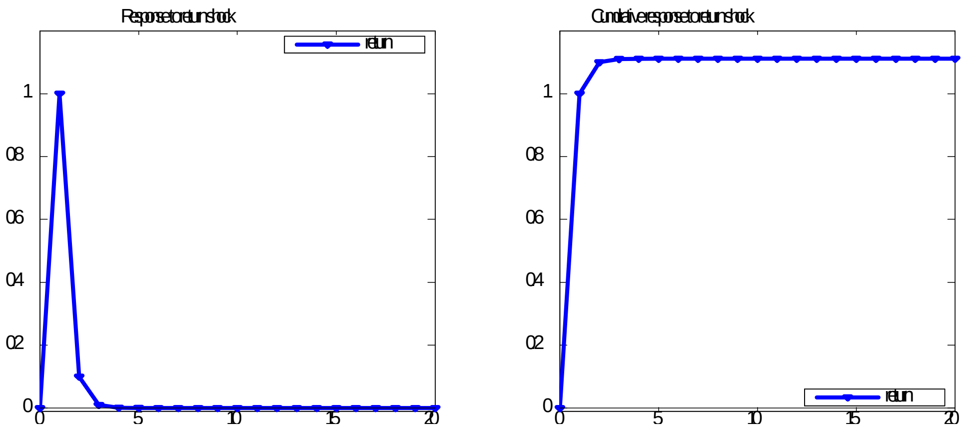
5. *There is a "temporary component" to stock prices. You need to look at* **both** *prices and dividends to see it.*

1.6 Univariate vs multivariate responses

1. Compare to the response to all return shocks lumped together: If you run

$$r_{t+1} = 0.1 \times r_t + \varepsilon_{t+1}$$

The response looks like this:



Response function $r_{t+1} = 0.1 \times r_t + \varepsilon_{t+1}$

2. Similarly, as shown in *Asset Pricing*, $\sigma^2(r_{t+1} + r_{t+2} + .. + r_{t+k}) \approx k\sigma^2(r)$. Stocks really are not "safer in the long run."

3. How can stocks be predictable from DP, but nearly a random walk on their own – not "safer in the long run?"

   (a) Temperature forecast story

   (b) *The univariate return process implied by the VAR is very close to uncorrelated over time.*

$$\begin{aligned} r_{t+1} &= b_r dp_t + \varepsilon^r_{t+1} \\ r_{t+2} &= b_r \left(\phi dp_t + \varepsilon^{dp}_{t+1}\right) + \varepsilon^r_{t+2} \end{aligned}$$

   so

$$\begin{aligned} cov(r_{t+1}, r_{t+2}) &= cov\left[b_r dp_t + \varepsilon^r_{t+1}, b_r \left(\phi dp_t + \varepsilon^{dp}_{t+1}\right) + \varepsilon^r_{t+2}\right] \\ cov(r_{t+1}, r_{t+2}) &= b_r^2 \phi \sigma^2(dp_t) + b_r cov(\varepsilon^r_{t+1}, \varepsilon^{dp}_{t+1}) \end{aligned}$$

      i. Intuition: First term: A slow moving predictor = momentum. Second term: like bonds, a lower price means higher expected returns = mean reversion. Fact: these terms almost exactly offset.

      ii. Algebra: recall

$$b_r = 1 - \rho\phi + 0$$

$$\varepsilon^r_{t+1} \approx -\rho\varepsilon^{dp}_{t+1} + \varepsilon^d_{t+1};\ cov(\varepsilon^{dp}_{t+1}, \varepsilon^d_{t+1}) \approx 0$$

      so

$$cov(\varepsilon^r_{t+1}, \varepsilon^{dp}_{t+1}) = -\rho\sigma^2\left(\varepsilon^{dp}_{t+1}\right).$$

      then

$$\begin{aligned} cov(r_{t+1}, r_{t+2}) &= b_r^2 \phi \sigma^2(dp_t) - b_r \rho \sigma^2\left(\varepsilon^{dp}_{t+1}\right) \\ &= b_r^2 \phi \frac{\sigma^2\left(\varepsilon^{dp}_{t+1}\right)}{1-\phi^2} - b_r \rho \sigma^2\left(\varepsilon^{dp}_{t+1}\right) \\ &= b_r\left(\phi \frac{1-\rho\phi}{1-\phi^2} - \rho\right)\sigma^2\left(\varepsilon^{dp}_{t+1}\right) \end{aligned}$$

If we had $\phi = \rho = 0.96 = 0$, we get the result,

$$cov(r_{t+1}, r_{t+2}) = 0.$$

(c) Bottom line: *Predictability from dp does not affect the safety of stocks in the long run!* Predictability does affect long-run investment decisions, because it becomes a "Merton state variable."

1.7 More variables

- Should we add more variables, including lags

$$R_{t+1} = a + b\,(D/P_t) + cx_t + \varepsilon_{t+1}?$$

Yes!

- Common confusion. We have shown "dividend growth is not predictable from dp." We have not shown "dividend growth is not predictable" from other variables. It can be. And it is!

- Example ("Discount rates") cay helps to forecast returns!

Table IV

Forecasting Regressions with the Consumption-wealth Ratio

	Coefficients		t-statistics		Other statistics	
Left-hand Variable	dp_t	cay_t	dp_t	cay_t	R^2	$\sigma\left[E_t(y_{t+1})\right]\%$
r_{t+1}	0.12	0.071	(2.14)	(3.19)	0.26	8.99
Δd_{t+1}	0.024	0.025	(0.46)	(1.69)	0.05	2.80
dp_{t+1}	0.94	-0.047	(20.4)	(-3.05)	0.91	
cay_{t+1}	0.15	0.65	(0.63)	(5.95)	0.43	
$r_t^{lr} = \sum_{j=1}^{\infty} \rho^{j-1} r_{t+j}$	1.29	0.033				0.51
$\Delta d_t^{lr} = \sum_{j=1}^{\infty} \rho^{j-1} \Delta d_{t+j}$	0.29	0.033				0.12

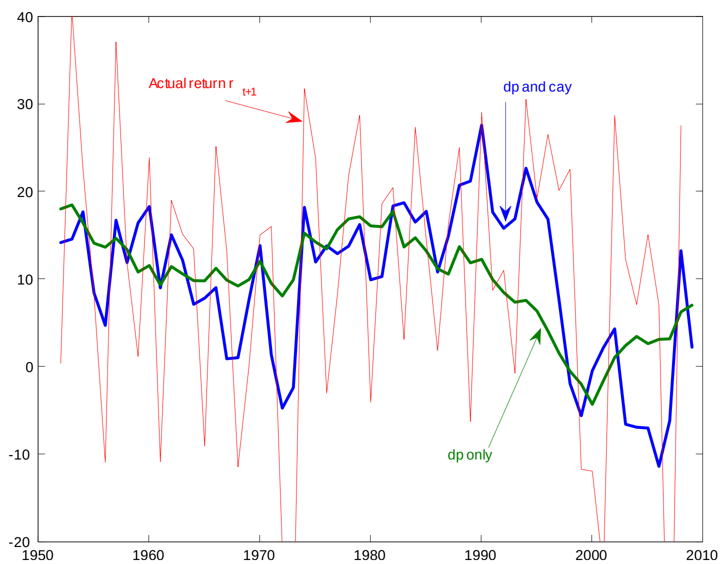


Figure 1: **Forecast and actual one-year returns.** The forecasts are fitted values of regressions of returns on dividend yield and *cay*. Actual returns r_{t+1} are plotted on the same date as their forecast, $a + b \times dp_t$.

- Identities, variance, etc?

$$d_t - p_t \approx E_t \sum_{j=1}^{\infty} \rho^{j-1} \left(r_{t+j} - \Delta d_{t+j}\right).$$

How can anything help to forecast r_{t+1}? A: by *also* forecasting Δd_{t+j} or r_{t+j}! *variables that help to forecast cashflows must also help to forecast returns. DP is affected by cashflow and return forecasts, so other cashflow forecasts "clean up" DP as a return forecaster. Deep point. This is why accounting ratios that forecast cashflows help to forecast returns.*

1.

$$\sum_{j=1}^{\infty} \rho^{j-1} r_{t+j} = a_r + b_r^{lr} \times dp_t + c_r^{lr} \times z_t + \varepsilon_t^r$$
$$\sum_{j=1}^{\infty} \rho^{j-1} \Delta d_{t+j} = a_d + b_d^{lr} \times dp_t + c_d^{lr} \times z_t + \varepsilon_t^d$$

$$b_r^{lr} - b_d^{lr} = 1$$
$$c_r^{lr} - c_d^{lr} = 0$$

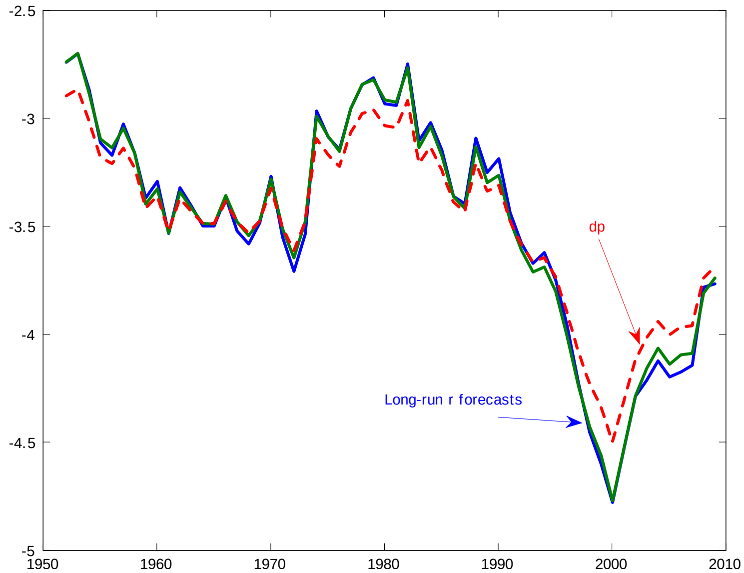
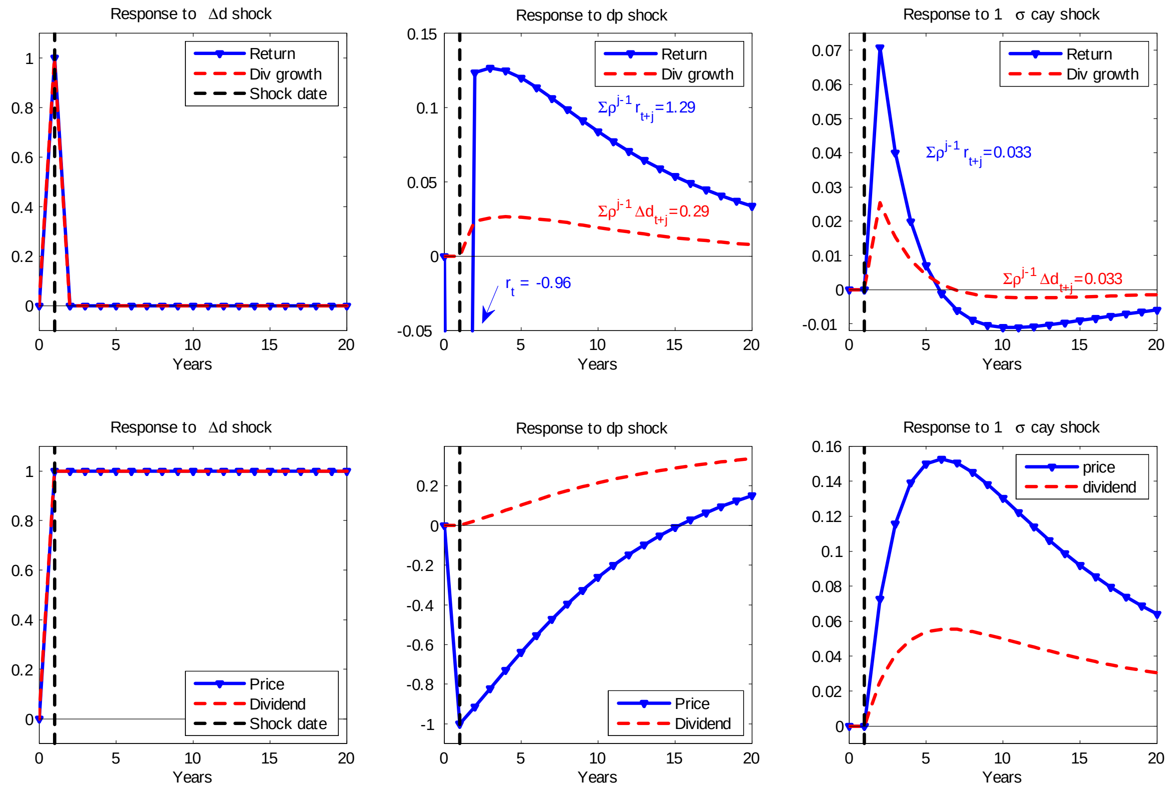


Figure 2: **Log dividend yield** dp **and forecasts of long-run returns** $\sum_{j=1}^{\infty} \rho^{j-1} r_{t+j}$. Return forecasts are computed from a VAR including dp, and a VAR including dp and cay.

2. Plot of $d_t - p_t$, $E_t \sum_{j=1}^{\infty} \rho^{j-1} r_{t+j}$. and $E_t \sum_{j=1}^{\infty} \rho^{j-1} \Delta d_{t+j}$. cay dies out, makes very little difference to pd decomposition. pd variance is still almost all expected returns

3.



Impulse-response Functions. Response functions to dividend growth, dividend yield, and *cay* shocks. Calculations are based on the VAR of Table IV. Each shock changes the indicated variable without changing the others, and includes a contemporaneous return shock from the identity $r_{t+1} = \Delta d_{t+1} - \rho dp_{t+1} + dp_t$. The vertical dashed line indicates the period of the shock.

- Summary: more variables can make returns *even more predictable,* because they *can* forecast dividends, or the term structure of risk premiums.
- Fast moving forecasters do little however to alter our view of the source of price variation.

1.8 Pervasive predictability: a preview

Preview: Predictability beyond d/p and stock market returns. (Variety on left hand side as well as on right hand side)

1. Us:

$$\begin{aligned} r_{t+1} &= a + 0.1 \times (d_t - p_t) + \varepsilon_{t+1} \\ \Delta d_{t+1} &= a + 0 \times (d_t - p_t) + \varepsilon_{t+1} \end{aligned}$$

2. More variables:

$$R_{t+1} + a + b(D/P)_t + c \times \text{term}_t + d \times \text{def}_t + f \times \text{I/K}_t + g \times \text{cay}_t + h \times \pi_t + \text{volatility}_t + m \times \text{VIX}_t + n \times \text{vol}_t ... + \varepsilon_{t+1}$$

In identity?

$$dp_t \approx E_t \sum_{j=1}^{\infty} \rho^{j-1} (r_{t+j} - \Delta d_{t+j})$$

Other variables can make both dividend growth and returns more predictable.

3. Individual stocks? (Pay attention)

$$\begin{aligned} R^i_{t+1} &= a + bx^i_t + \varepsilon^i_{t+1}? \\ &\leftrightarrow E(R^i_{t+1}) \text{ is higher when } x^i_t \text{ is higher} \end{aligned}$$

B/M, size, momentum, etc.

4. Bonds

   (a) "Expectations hypothesis." $y^{long} = 5\%$, $y^{short} = 2\%$ Implication?

   (b) Facts:

$$\begin{aligned} R^{bond}_{t+1} - R^f_t &= a + 1 \times (y^{long}_t - y^{short}_t) + \varepsilon_{t+1} \\ R^f_{t+1} - R^f_t &= a^f + 0 \times (y^{long}_t - y^{short}_t) + \varepsilon^f_{t+1} \end{aligned}$$

5. Foreign exchange. "Forward premium anomaly"

   (a) Expectations. $r^{US} = 1\%$, $r^{Eu} = 5\%$ Implication?

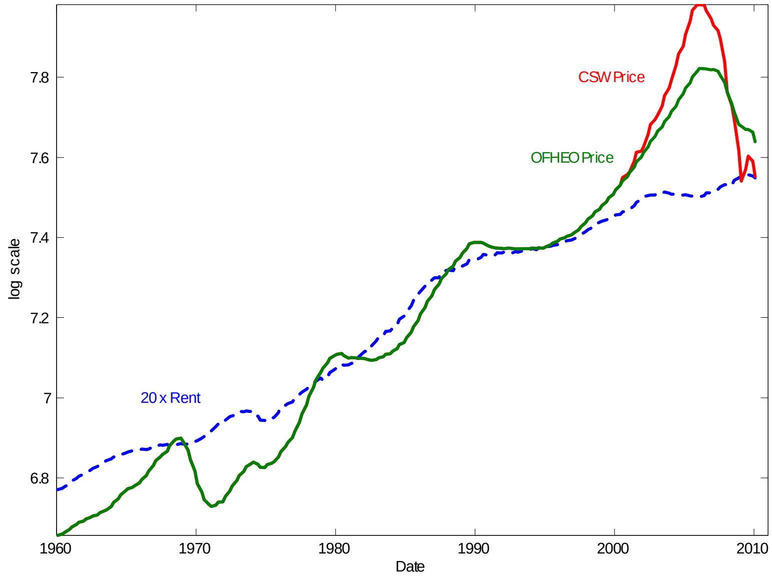
   (b) Regression

$$\begin{aligned} R^{Eu}_{t+1} - r^{\$}_{t+1} &= a + 1 \times (r^{Eu}_t - r^{\$}_t) + \varepsilon_{t+1} \\ \Delta e^{Eu/\$}_{t+1} &= a^e + 0 \times (r^{Eu}_t - r^{\$}_t) + \varepsilon^e_{t+1} \end{aligned}$$

   R^{Eu} = dollar return to holding Euro bonds for a year, unhedged (actually >1, <0), e = exchange rate.

6. Credit spreads do not mean (much) higher chance of default, do mean higher expected return.

7. Houses



Houses:	b	t	R^2	Stocks:	b	t	R^2
r_{t+1}	0.12	(2.52)	0.15		0.13	(2.61)	0.10
Δd_{t+1}	0.03	(2.22)	0.07		0.04	(0.92)	0.02
dp_{t+1}	0.90	(16.2)	0.90		0.94	(23.8)	0.91

8. Many questions! Do the variables that forecast one thing forecast another? What is the *factor structure of expected returns* across markets?