2 Week 1 Returns are predictable over time.

2.1 Return forecasting regressions

1. Method.

   (a) *Are excess returns (stock - bond) predictable, or not – like coin flips*

      i. Test/meaning. *Run a regression*

      $$R_{t+1} = a + bx_t + \varepsilon_{t+1}$$

      ii. ↔Expected return *at time t* is

      $$E_t(R_{t+1}) = a + bx_t$$

      Regression measures whether expected returns (risk premium) varies over time.

      iii. Meaning of the word "predictable."

2. Classic view: returns should not be predictable.

   (a) Definition of efficiency: Information is reflected in prices. *Period.* A result of competition.

   (b) Efficiency. → Returns are unpredictable.

3. See if you can debunk these:

   (a) "The market declined temporarily because of profit-taking. It will bounce back next week."

   (b) "The stock price rises slowly after an announcement as new information diffuses through the market."

   (c) "The internet is the wave of the future. You should put your money in internet stocks."

   (d) "Buy stocks of strong companies, with good earnings and good earnings growth. They will be more profitable and give better returns to stockholders."

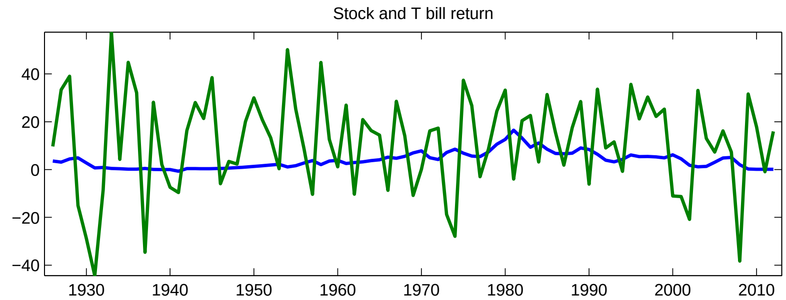
   (e) "The demand curve slopes down;" "Big trades have a lot of price impact." "Stock prices fell today under a lot of temporary selling pressure," "Some stocks fell too far in the crash because mutual funds and hedge funds had to unload them to meet redemptions" "Small losers fall in December as dentists harvest tax losses."

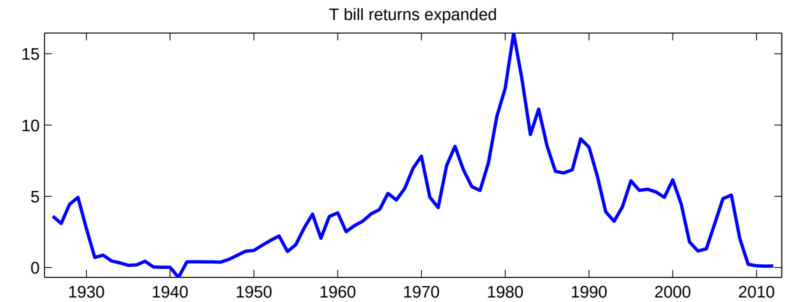
   (f) "Some stocks fell too far in the crash because mutual funds and hedge funds had to unload them to meet redemptions"

4. Facts, first generation: (Problem set 1 updates data)

Regression of returns on lagged returns
Annual data 1927-2008
$R_{t+1} = a + bR_t + \varepsilon_{t+1}$

	b	t(b)	R^2	E(R)	$\sigma(E_t(R_{t+1}))$
Stock	0.04	0.33	0.002	11.4	0.77
T bill	0.91	19.5	0.83	4.1	3.12
Excess	0.04	0.39	0.00	7.25	0.91





(a) "Efficiency," saving vs. risk-bearing. Why we use excess returns

5. New *Facts:* new variables (D/P), longer horizons.

(a) "Discount rates" Table 1

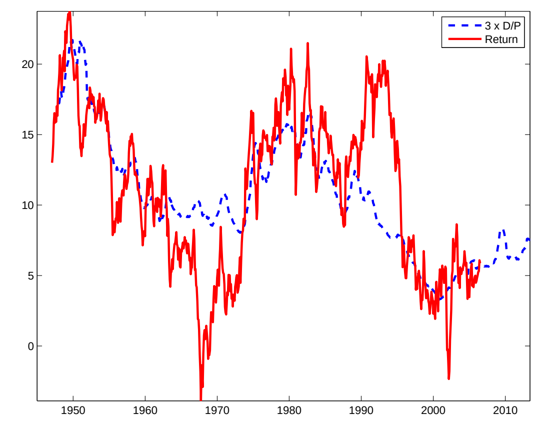
Horizon k	b	t(b)	R^2	$\sigma\left[E_t(R^e)\right]$	$\frac{\sigma[E_t(R^e)]}{E(R^e)}$
1 year	3.8	(2.6)	0.09	5.46	0.76
5 years	20.6	(3.4)	0.28	29.3	0.62

- Return Coefficient is huge. > 1!

$$R_{t+1} = \frac{P_{t+1} + D_{t+1}}{P_t} = \frac{D_t}{P_t}\left(\frac{D_{t+1}}{D_t}\right) + \frac{P_{t+1}}{P_t}$$

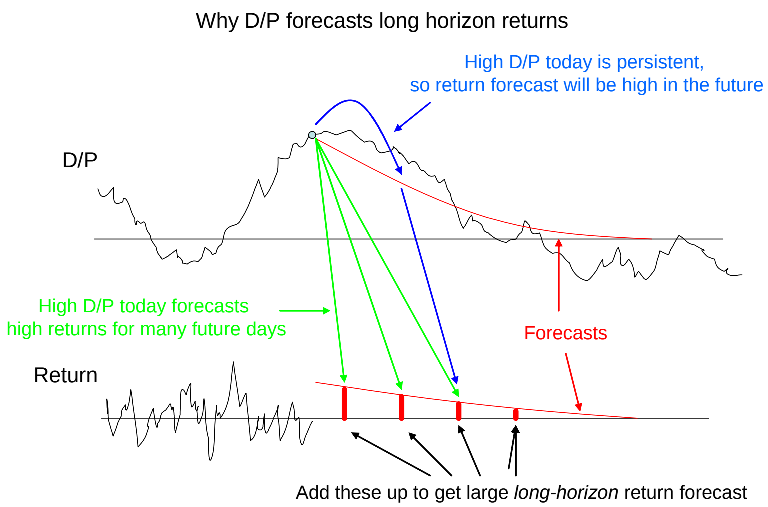
- t is significant, though not dramatic.
- R^2 is big at long horizons *long-horizon returns are much more predictable.*

(b) Discount rates graph



(c) Time-varying expected return is huge. (Relative to $E(R)$) see D/P graph

6. Connecting long and short horizons.



Equations:

$$\begin{aligned} r_{t+1} &= bx_t + \varepsilon_{t+1} \\ x_{t+1} &= \phi x_t + \delta_{t+1}. \end{aligned}$$

$$\Leftrightarrow \quad r_{t+1} + r_{t+2} = b(1+\phi)x_t + (error)$$
$$\Leftrightarrow \quad r_{t+1} + r_{t+2} + r_{t+3} = b(1+\phi+\phi^2)x_t + (error)$$

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

- 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

$$R^2 = \frac{b^2(1+\phi+\phi^2+..)^2\sigma^2(x_t)}{\sigma^2\left(r_{t+1}+r_{t+2}+r_{t+3}\right)} \approx \frac{k^2}{k}\frac{b^2\sigma^2(x_t)}{\sigma^2(r)} \approx kR^2_{k=1}$$

- Your first "vector autoregression." Imply long-run dynamics from vector AR(1).

7. Does predictability mean that markets are "inefficient?"

   (a) Efficient view: high risk aversion in bad economic times. "Time-varying risk premium" is at least possible, plausible.

   (b) Alternative, "bubbles," "irrational exuberance" – people think $E(R)$ is constant, get expected cashflows wrong.

   (c) Fact: High prices = low subsequent returns and vice versa.

   (d) *>>This is the central (only) fact in the bubbles/irrationality debate.<<*

8. 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.

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

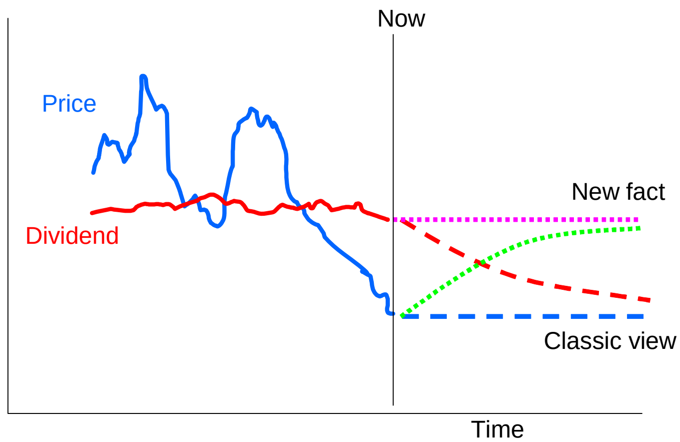
2.2 Dividends

1.

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
2	7.9	3.0	0.12	-0.42	-0.22	0.001
3	12.6	3.0	0.20	0.16	0.13	0.0001
5	20.6	2.6	0.22	2.42	1.11	0.02

(Source: "Financial markets and the real economy")

2. Dividends *should* be forecastable.



2.3 Present value identities, volatility, bubbles

1. Motivation:

   (a) "Higher P/D must forecast either dividends or returns."

   (b) "Required returns rise, people try to sell, prices go down, then we see the subsequent higher returns (on average)"

   (c) "What about bubbles?"

   (d) Need to start with price = present value of dividends.

   (e) Major new tool: linear, dynamic present value formula.

2. Volatility background.

   (a) Shiller

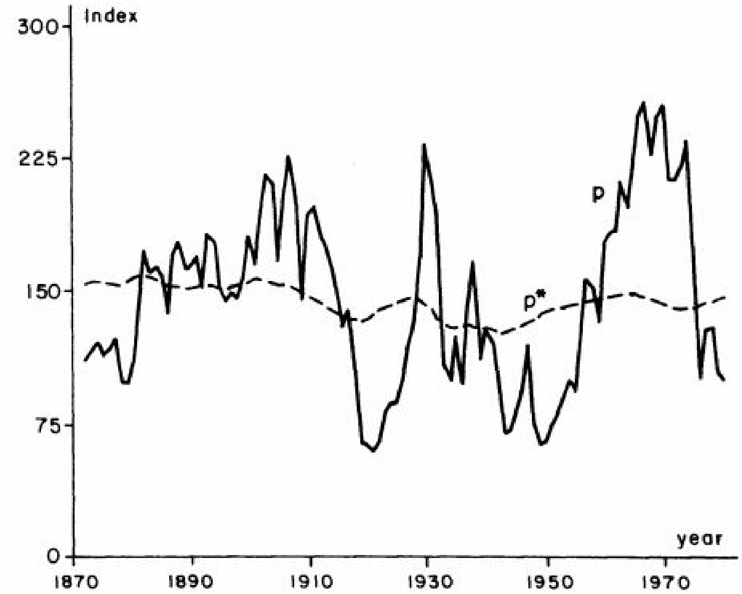
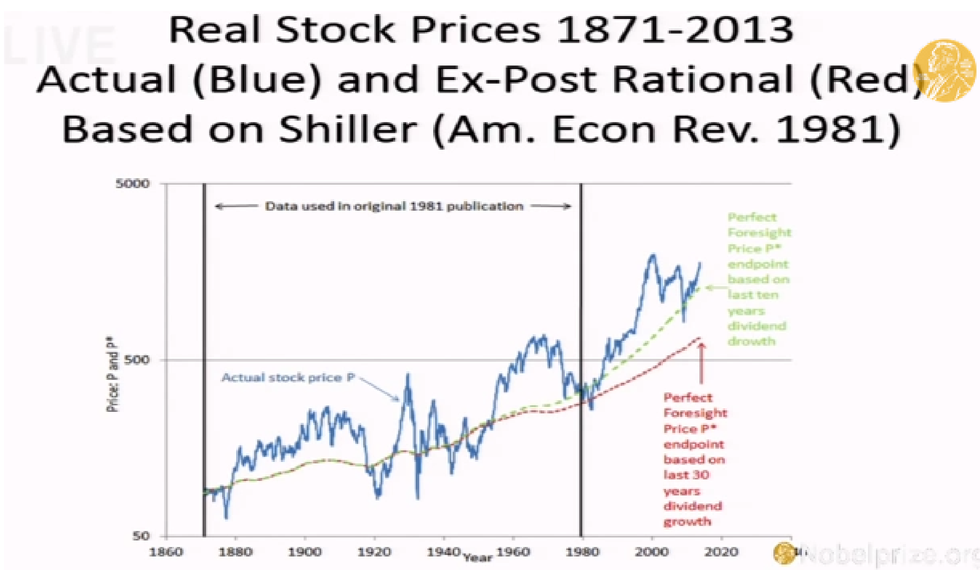


FIGURE 1
Note: Real Standard and Poor's Composite Stock Price Index (solid line p) and *ex post* rational price (dotted line p^*), 1871–1979, both detrended by dividing a long-run exponential growth factor. The variable p^* is the present value of actual subsequent real detrended dividends, subject to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

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



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

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

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

(b) A new and different test of efficiency having nothing to do with predictions?

(c) It is the same as prices don't predict dividends!

(d) It is the same as prices do predict returns!

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

4. 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) 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!

5. Tie volatility to predictability. 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)

$$\begin{aligned} r_{t+1} &= b_r(d_t - p_t) + \varepsilon_{t+1}^r \\ \Delta d_{t+1} &= b_d(d_t - p_t) + \varepsilon_{t+1}^d \end{aligned}$$

Then

$$d_t - p_t = \left[b_r(d_t - p_t) + \varepsilon^r_{t+1}\right] - \left[b_d(d_t - p_t) + \varepsilon^d_{t+1}\right]$$

Result:

$$1 = b_r - b_d.$$

(Also $0 = \varepsilon^r_{t+1} - \varepsilon^d_{t+1}$)

$$b_r = \frac{cov(r_{t+1}, d_t - p_t)}{var(d_t - p_t)}$$

$$var(d_t - p_t) = cov(r_{t+1}, d_t - p_t) - cov(\Delta d_{t+1}, d_t - p_t)$$

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.*

6. Linearized dynamic present value formula

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

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

   (a) Stare at this.

   (b) Source: Definition of return (like 1-period model)

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

   solve

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

   Iterate forward. Return identity derivation:

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

   (c) *NOTICE I am ignoring the constant term – treat all variables $p - d, r, \Delta d$ as deviations from their means (so k=0) or put the constant back in.*

7. Volatility

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

   (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!*

(b) Run both sides of

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

on $d_t - p_t$. Result?

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

where b^{lr} means

$$\begin{aligned} \sum_{j=1}^{\infty} \rho^{j-1} \Delta d_{t+j} &= b_d^{lr} \left(d_t - p_t \right) + \varepsilon \\ \sum_{j=1}^{\infty} \rho^{j-1} r_{t+j} &= b_r^{lr} \left(d_t - p_t \right) + \varepsilon \end{aligned}$$

(c) *Long-run* return forecast and dividend forecast must add to one. *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!*

(g) Another interpretation

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

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

$$dp_t \approx b_r^{lr} \times dp_t - b_d^{lr} \times dp_t$$

Our long-run regression splits up each day's dp into expected future returns and expected future dividend growth. Since $b_r^{lr} \approx 1$, the variance of the first term is almost entirely the same as the variance of dividend yields.

8. Volatility and bubbles

   (a) Volatility background and puzzle.

   (b) Volatility = predictable return, it's not a different kind of efficiency test.

   (c) Bubbles?

      i. Time-varying expected returns? Yes, the unanswered question.

      ii. Rational bubbles?

   (d) "Rational bubbles" $D = 0$, "buy and sell to the greater fool"

$$\begin{aligned} P_{t+1} &= RP_t + \varepsilon_{t+1}. \\ E_t(R_{t+1}) &= E_t\left(\frac{P_{t+1}}{P_t}\right) = \frac{1}{R} \\ P_t &= \frac{1}{R} E_t\left(P_{t+1}\right) \end{aligned}$$

   (e) Doing it right. Use the k-year identity

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

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

   - Table II: *No rational bubbles!*

9. Q: How "big" is return forecastability? t stats and R^2 don't look great. A:

   (a) $\sigma(E_t(R_{t+1}))$ is large (6%) compared to $E(R_{t+1})$

   (b) R^2 grows quite large at long horizons

   (c) Long horizon return forecast + long horizon dividend forecast = 1. Dividend forecast = 0! 0/1 expected became 1/0

   (d) Long horizon return forecasts are just enough to account for huge price volatility.

10. Not answered here: why do $E_t(R_{t+1})$ vary *so much* over time?

2.4 Vector autoregression and impulse-response function

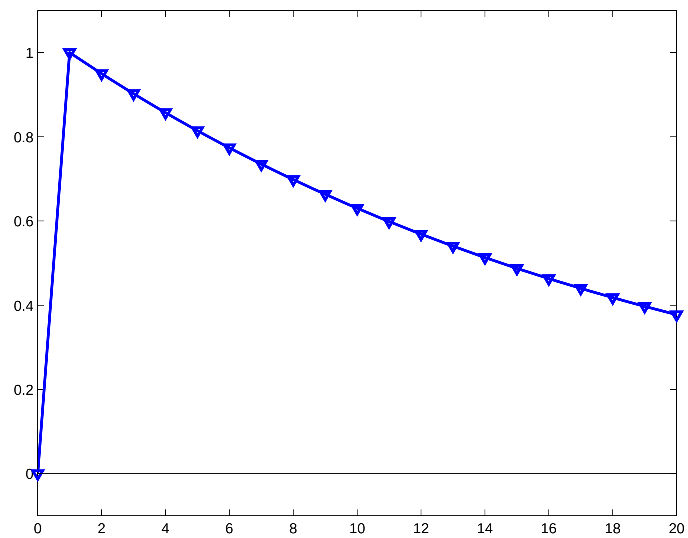
1. "Vector autoregression" (VAR) and "impulse-response function."

(a) Example

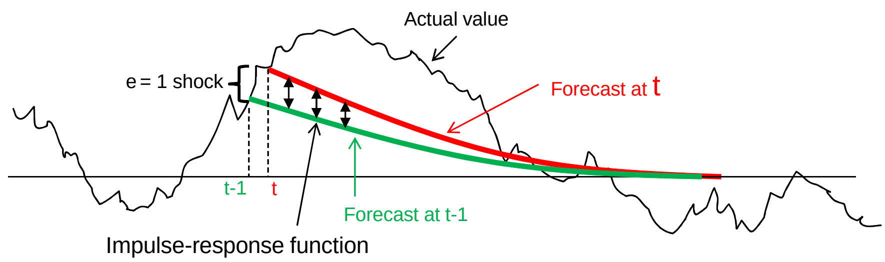
$$x_t = \phi x_{t-1} + \varepsilon_t$$

$x_0 = 0$, $\varepsilon_1 = 1$, $\varepsilon_2 = \varepsilon_3 = ... = 0$,simulate

$$\begin{aligned} x_1 &= 1 \\ x_2 &= \phi x_1 + 0 = \phi \\ x_3 &= \phi x_2 = \phi^2 \\ x_4 &= \phi x_3 = \phi^3 \end{aligned}$$



(b) Idea: "How does a shock today correspond to changes in expected future values?" (Warning: not *cause!*)



2. Us: Same idea, more variables

(a) Regressions

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

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) Shocks don't move independently. Identity:

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

Hence $((E_{t+1} - E_t))$

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

If returns go up, either prices or dividends must also go up! Only 2 shocks.

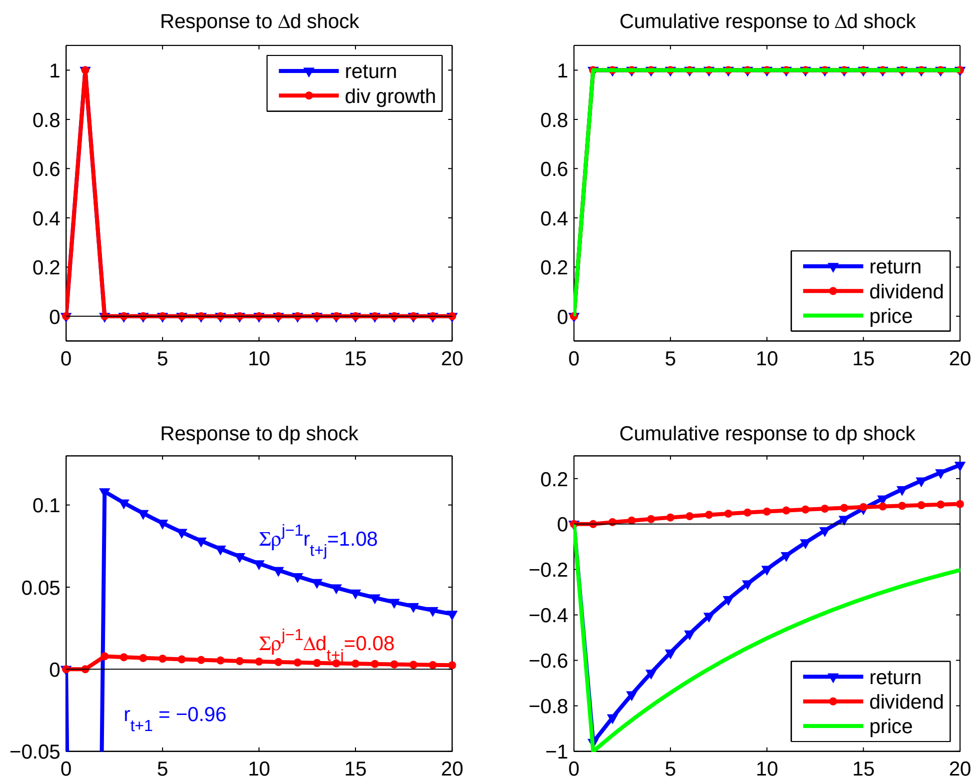
(c) My choice.

$$\begin{aligned}\text{Shock 1 (“}\Delta d\text{ shock”):} \quad & \begin{bmatrix} \varepsilon^r_1 & \varepsilon^d_1 & \varepsilon^{dp}_1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \end{bmatrix} \\ \text{Shock 2 (“}dp\text{ shock” or “}Er\text{” shock)} \quad : \quad & \begin{bmatrix} \varepsilon^r_1 & \varepsilon^d_1 & \varepsilon^{dp}_1 \end{bmatrix} = \begin{bmatrix} -\rho & 0 & 1 \end{bmatrix}\end{aligned}$$

i. Shock 1: returns go up 1% because dividends go up 1%, with no change in dividend yield

ii. Shock 2: no change to dividends, expected returns go up, so actual returns (prices) go down.

(d) Impulse-response Results



Response to ε^{d} and ε^{dp} shocks from

$$\begin{aligned}
\Delta d_{t+1} &= a_d + b_d dp_t + \varepsilon^{d}_{t+1} \\
r_{t+1} &= a_r + b_r dp_t + \varepsilon^{r}_{t+1} \\
dp_{t+1} &= a_{dp} + \phi dp_t + \varepsilon^{dp}_{t+1} \\
\text{with } \varepsilon^{r}_{t} &= \varepsilon^{d}_{t} - \rho \varepsilon^{dp}_{t}
\end{aligned}$$

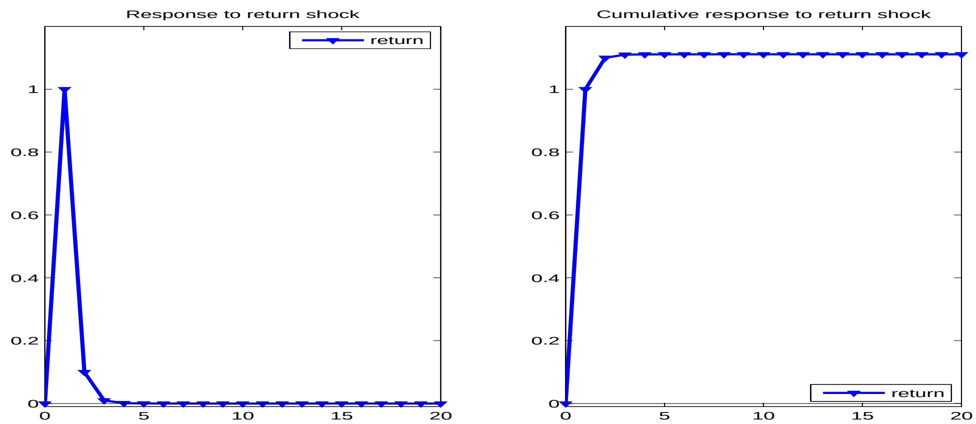
(e) Interpretation: how news about the future changes prices today.

  i. ε^{d}, dividend shock with no dp change is a pure expected-cashflow shock with no change in expected returns

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

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

(g) Compare to the response to all return shocks lumped together:



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

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

i. Temperature forecast story

ii. *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

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

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

2.5 More variables

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

  Yes!

- Example ("discount rates")

Table IV

Forecasting Regressions with the Consumption-wealth Ratio

Left-hand Variable	Coefficients dp_t	cay_t	t-statistics dp_t	cay_t	Other statistics 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

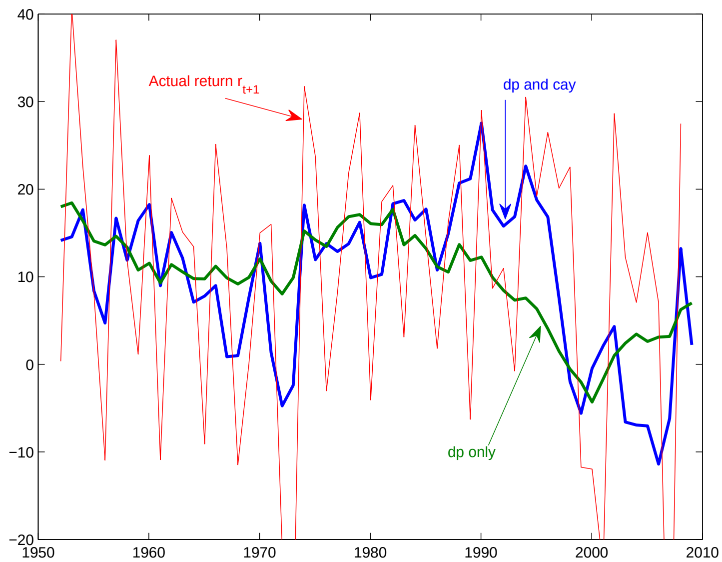


Figure 1: **Forecast and actual one-year returns.** The forecasts are fitted values of regressions of returns on dividend yield and *cay*. Actual returns r_{t+1} are plotted on the same date as their forecast, $a + b \times dp_t$.

- 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 foreast 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 foreast cashflows help to forecast returns.*

1.

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

$$\begin{aligned} b_r - b_d &= 1 \\ c_r - c_d &= 0 \end{aligned}$$

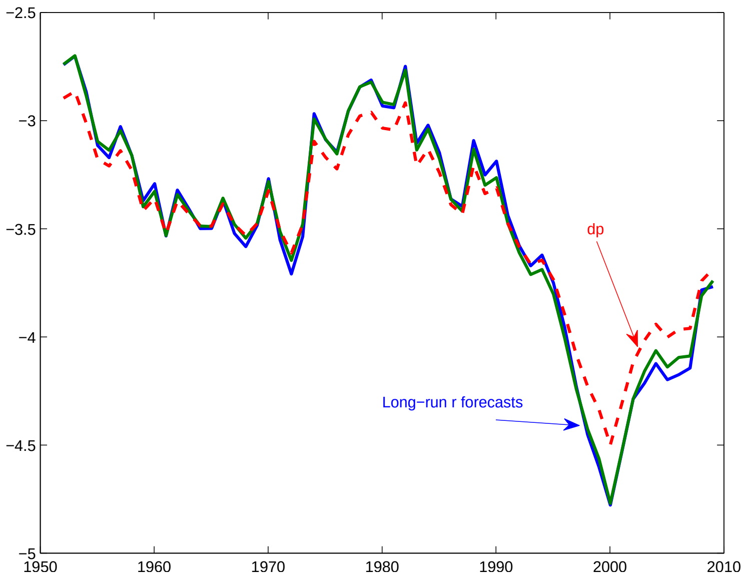
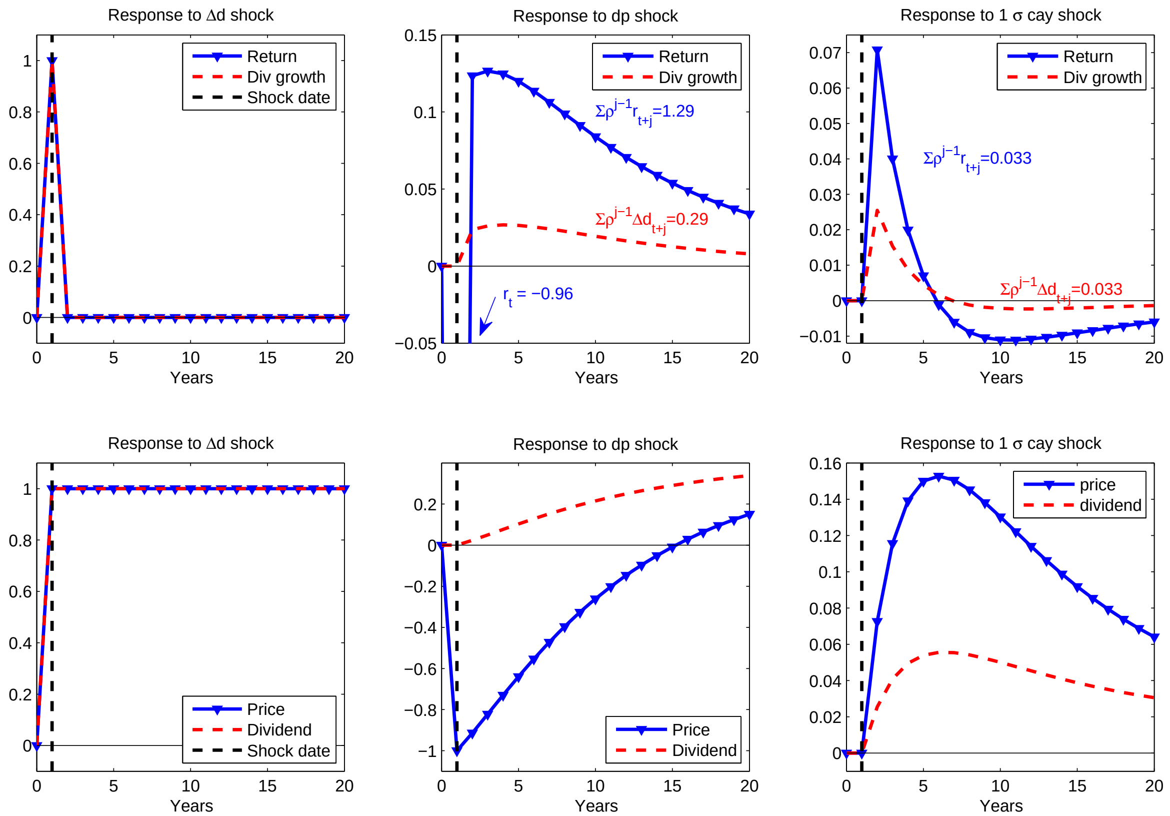


Figure 2: **Log dividend yield** dp **and forecasts of long-run returns** $\sum_{j=1}^{\infty} \rho^{j-1} r_{t+j}$. Return forecasts are computed from a VAR including dp, and a VAR including dp and cay.

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.

2.6 Pervasive predictability: a preview

Preview: Predictability beyond d/p and market returns.

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?

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

Other variables make both dividend growth and returns more predictable.

3. Individual stocks? (Pay attention)

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

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?

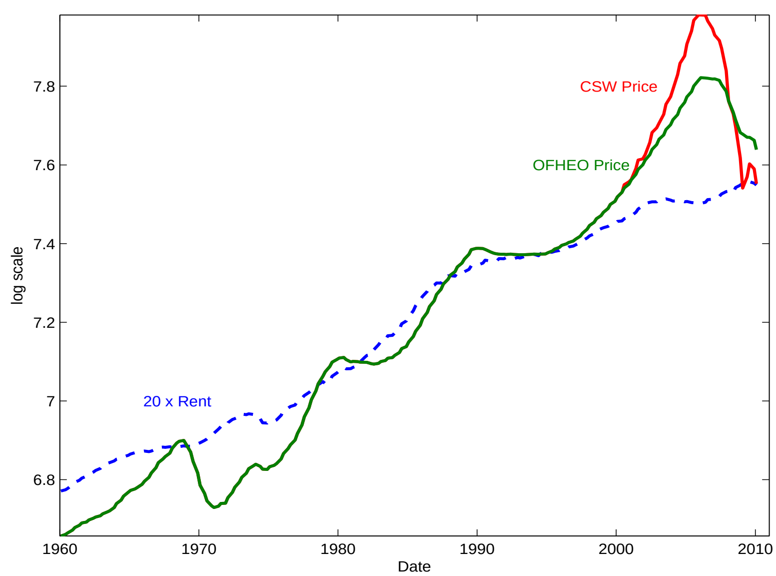
   (b) Regression

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

   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

2.7 Forecastability Review and Interpretation

1. Big picture: expected returns vary *over time* in ways not described by the classic random walk/CAPM. (Later, *across assets.*)

2. Tool 1: *Forecasting regression*

$$R_{t+1} = a + bx_t + \varepsilon_{t+1}$$

3. 1st generation: $b \approx 0$, R^2 small.

4. New view: you can forecast returns..

   (a) Evidence: $R_{t+1} = a + b(D/P)_t + \varepsilon_{t+1}$.

   (b) Innovation: "1/p variables", long horizons

   (c) $b \approx 2 - 5$; b, R^2 grow with horizon.

5. Inefficiency/quick profit? No, alas

   (a) R^2 is only large at long horizons; x_t moves slowly

   (b) To profit, you have to buy in bad times.

   (c) $\rightarrow$ *Risk premium varies over time.*

6. Long horizon predictability.

   (a) R^2 is big! Wow!

   (b) long horizon = small short horizon predictability plus persistent x_t.

7. Tool 2: Useful linearized version of $R_{t+1} = \frac{P_{t+1}+D_{t+1}}{P_t}$.

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

8. Tool 3: linearized present value formula

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

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

   p-d *reveals to us* changes in market risk premia.

9. Variance decomposition/coefficient identity

$$\begin{aligned} 1 &= b_r^{lr} - b_d^{lr} \\ var\left(dp_t\right) &= var\left(p_t - d_t\right) \approx \sum_{j=1}^{\infty} \rho^{j-1} cov\left[dp_t, r_{t+j}\right] - \sum_{j=1}^{\infty} \rho^{j-1} cov\left[dp_t, \Delta d_{t+j}\right] \end{aligned}$$

   (a) Results: $E_t r_{t+j}$ is the dominant cause of market price movements. (Alas).

   (b) D/P does *not* forecast Δd. It "should"

(c) If D/P varies, it *must* forecast one of dividend growth or returns. Long-run regression coefficients *must* add to one.

(d) Evidence on bubbles. Long run expected return variation *is* enough. Really

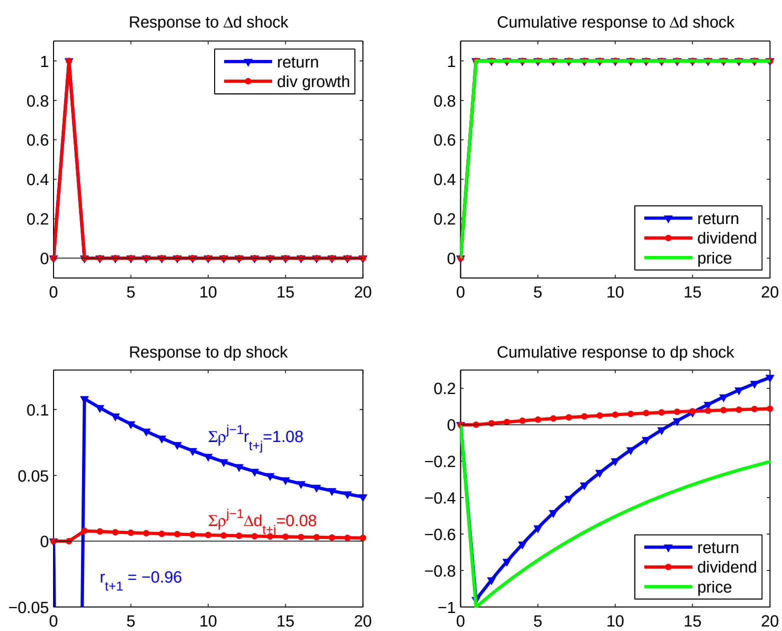
$$p_t - d_t \approx \kappa + E_t \sum_{j=1}^{k} \rho^{j-1} (\Delta d_{t+j} - r_{t+j}) + \rho^k (p_{t+k} - d_{t+k})$$

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

$$\begin{aligned} 1 &= b_r^{(k)} - b_d^{(k)} + \rho^k b_{dp}^{(k)} \\ var(dp_t) &= var(p_t - d_t) \approx \sum_{j=1}^{k} \rho^{j-1} cov[dp_t, r_{t+j}] - \sum_{j=1}^{k} \rho^{j-1} cov[dp_t, \Delta d_{t+j}] + \rho^k cov(dp_t, dp_{t+k}) \end{aligned}$$

and the r terms are enough; no "rational bubbles."

10. Tool 4: VAR, I-R, and "expected return" vs. "cashflow" shocks



11. Bonds and fx show the same pattern.

(a) Yield spread forecasts returns, not future rates; coefficients add to one.

(b) Interest rate spreads forecast fx returns, not exchange rates; coefficients add to one.

12. Real world: Uses many more variables. Example: cay

$$R_{t+1} = a + b \times (D/P)_t + c \times cay_t + \varepsilon_{t+1}$$

Present value identity: variables help to forecast returns if they *also* help to forecast dividends or term structure of risk premia.

13. Warnings/Practical use
    (a) Fishing; Peso problems
    (b) Imprecise estimates.
    (c) Not useful for high speed trading!
    (d) "Buy in bad times," and miss bull markets.
    (e) *But* the point estimates are big.
    (f) View of the world is 100% different – *Expected return news not dividend news drives markets.*
14. Many other implications
    (a) CAPM. Assumes iid returns.
    (b) Cost-of-capital. $V = E(CF)/E(R) = E(CF)/\left[R^f + \beta E(R^m - R^f)\right]$
    (c) Marking to market assumes a random walk. "Accounting for future value" makes sense.
15. Some concept benchmarks
    (a) *Definition of Market Efficiency*
    (b) *Average investor theorem*
    (c) *Time-varying market risk premium – thinking about a market in equilibrium*
    (d) "*Forecast*" and use of regressions to forecast without causal interpretation