Lecture 6: Excessive Volatility in Prices and Returns

Raul Riva

FGV EPGE

February, 2026

Intro

Intro

- Last class: equity premium and risk-free rate puzzles
- Consumption too smooth: $\sigma(\Delta c) \ll \sigma(r^e)$

Intro

- Last class: equity premium and risk-free rate puzzles
- Consumption too smooth: $\sigma(\Delta c) \ll \sigma(r^e)$

Today: Excessive Volatility

- Prices move too much for constant discount rates (Shiller (1981))
- Campbell-Shiller: decompose into expected returns vs dividend growth
- Discount-rate news dominates (Campbell (1991), Cochrane (2011))
- Subjective beliefs revive cash-flow news (DeLaO and Myers (2021))

Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends?

Motivation

- Shiller (1981) is a provocative paper
- Equity prices move around a lot... can dividend expectation changes justify this?

Motivation

- Shiller (1981) is a provocative paper
- Equity prices move around a lot... can dividend expectation changes justify this?
- Typical model to price stocks (recall your analyst friend from Lecture 1):

$$P_t = \mathbb{E}_t \left[\sum_{k=1}^{\infty} \frac{D_{t+k}}{(1+r)^k} \right]$$

- D_{t+k} = dividend at time $t + k$
- r = constant discount rate

Motivation

- Shiller (1981) is a provocative paper
- Equity prices move around a lot... can dividend expectation changes justify this?
- Typical model to price stocks (recall your analyst friend from Lecture 1):

$$P_t = \mathbb{E}_t \left[\sum_{k=1}^{\infty} \frac{D_{t+k}}{(1+r)^k} \right]$$

- D_{t+k} = dividend at time $t+k$
- r = constant discount rate
- With constant r, can this model generate the same volatility in prices as seen in the data?
- Price changes **must** be driven by news about future dividends!

Realized vs Forecasted Dividends

- The "efficient market model" implies:

$$P_t = \mathbb{E}_t\left[\sum_{k=1}^{\infty} \frac{D_{t+k}}{(1+r)^k}\right] = \sum_{k=1}^{\infty} \gamma^k \mathbb{E}_t[D_{t+k}], \qquad \gamma \equiv \frac{1}{1+r}$$

Realized vs Forecasted Dividends

- The "efficient market model" implies:

$$P_t = \mathbb{E}_t\left[\sum_{k=1}^{\infty} \frac{D_{t+k}}{(1+r)^k}\right] = \sum_{k=1}^{\infty} \gamma^k \mathbb{E}_t[D_{t+k}], \qquad \gamma \equiv \frac{1}{1+r}$$

- Realized dividends are the sum of expected dividends and an "innovation" (forecast error):

$$D_{t+k} = \mathbb{E}_t[D_{t+k}] + \varepsilon_{t+k}, \quad \mathbb{E}_t[\varepsilon_{t+k}] = 0$$

Realized vs Forecasted Dividends

- The "efficient market model" implies:

$$P_t = \mathbb{E}_t\left[\sum_{k=1}^{\infty} \frac{D_{t+k}}{(1+r)^k}\right] = \sum_{k=1}^{\infty} \gamma^k \mathbb{E}_t[D_{t+k}], \qquad \gamma \equiv \frac{1}{1+r}$$

- Realized dividends are the sum of expected dividends and an "innovation" (forecast error):

$$D_{t+k} = \mathbb{E}_t[D_{t+k}] + \varepsilon_{t+k}, \quad \mathbb{E}_t[\varepsilon_{t+k}] = 0$$

- The realized series cannot be smoother than its projection:

$$\sigma^2(D_{t+k}) = \sigma^2(\mathbb{E}_t[D_{t+k}] + \varepsilon_{t+k}) = \sigma^2(\mathbb{E}_t[D_{t+k}]) + \sigma^2(\varepsilon_{t+k}) \geq \sigma^2(\mathbb{E}_t[D_{t+k}])$$

Shiller's Key Insight

- Let $P_t^* \equiv \sum_{k=1}^{\infty} \gamma^k D_{t+k}$. How should $\sigma(P_t^*)$ and $\sigma(P_t)$ compare?

Shiller's Key Insight

- Let $P_t^* \equiv \sum_{k=1}^{\infty} \gamma^k D_{t+k}$. How should $\sigma(P_t^*)$ and $\sigma(P_t)$ compare?
- Let Π_t be the GDP deflator and $A \equiv 1 + g$, where g a growth rate for real dividends
- Shiller deflated the index prices and dividends by inflation and the growth rate:

$$p_t \equiv \frac{P_t}{\Pi_t A^{t-T}}, \qquad d_t \equiv \frac{D_t}{\Pi_t A^{t+1-T}}$$

where $T = 1978$ is the last year of the sample and g is estimated as the constant log-growth rate of real dividends.

Shiller's Key Insight

- Let $P_t^* \equiv \sum_{k=1}^{\infty} \gamma^k D_{t+k}$. How should $\sigma(P_t^*)$ and $\sigma(P_t)$ compare?
- Let Π_t be the GDP deflator and $A \equiv 1 + g$, where g a growth rate for real dividends
- Shiller deflated the index prices and dividends by inflation and the growth rate:

$$p_t \equiv \frac{P_t}{\Pi_t A^{t-T}}, \qquad d_t \equiv \frac{D_t}{\Pi_t A^{t+1-T}}$$

where $T = 1978$ is the last year of the sample and g is estimated as the constant log-growth rate of real dividends.

- Then we have $p_t = \sum_{k=1}^{\infty} \tilde{\gamma}^k \mathbb{E}_t[d_{t+k}]$, where $\tilde{\gamma} \equiv \frac{A}{1+r} = A\gamma$

Shiller's Key Insight

- Let $P_t^* \equiv \sum_{k=1}^{\infty} \gamma^k D_{t+k}$. How should $\sigma(P_t^*)$ and $\sigma(P_t)$ compare?
- Let Π_t be the GDP deflator and $A \equiv 1 + g$, where g a growth rate for real dividends
- Shiller deflated the index prices and dividends by inflation and the growth rate:

$$p_t \equiv \frac{P_t}{\Pi_t A^{t-T}}, \qquad d_t \equiv \frac{D_t}{\Pi_t A^{t+1-T}}$$

where $T = 1978$ is the last year of the sample and g is estimated as the constant log-growth rate of real dividends.

- Then we have $p_t = \sum_{k=1}^{\infty} \tilde{\gamma}^k \mathbb{E}_t[d_{t+k}]$, where $\tilde{\gamma} \equiv \frac{A}{1+r} = A\gamma$
- Similarly: $p_t^* = \sum_{k=1}^{\infty} \tilde{\gamma}^k d_{t+k}$

Shiller's Key Insight

- Crucially, we have the relationship:

$$p_t = \mathbb{E}_t[p_t^*] \implies \sigma^2(p_t) = \sigma^2(\mathbb{E}_t[p_t^*]) \leq \sigma^2(p_t^*)$$

Shiller's Key Insight

- Crucially, we have the relationship:

$$p_t = \mathbb{E}_t[p_t^*] \implies \sigma^2(p_t) = \sigma^2(\mathbb{E}_t[p_t^*]) \leq \sigma^2(p_t^*)$$

Implementation matters:

- Dividends after 1978 are unknown at time t
- Shiller assumed that they would grow at rate g forever
- This assumption likely understates $\sigma^2(p_t^*)$

Shiller's Key Insight

- Crucially, we have the relationship:

$$p_t = \mathbb{E}_t[p_t^*] \implies \sigma^2(p_t) = \sigma^2(\mathbb{E}_t[p_t^*]) \leq \sigma^2(p_t^*)$$

Implementation matters:

- Dividends after 1978 are unknown at time t
- Shiller assumed that they would grow at rate g forever
- This assumption likely understates $\sigma^2(p_t^*)$
- p_{1978}^* is computed with "fake dividends" and then you use the recursion backward:

$$p_t^* = \tilde{\gamma}(d_{t+1} + p_{t+1}^*)$$

- The discount rate r is such that $r = \frac{\mathbb{E}[d]}{\mathbb{E}[p]}$ (why was deflating necessary for this step?)

The Main Plot

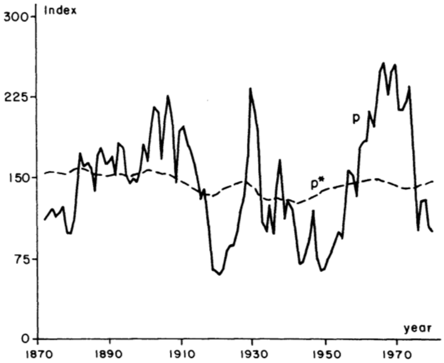


FIGURE 1

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.

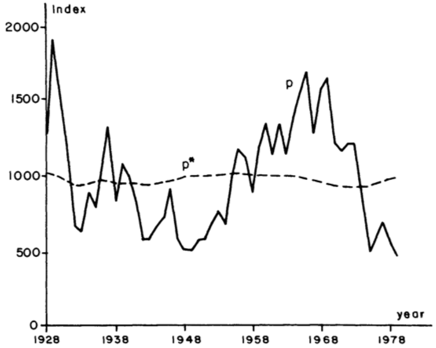


FIGURE 2

Note: Real modified Dow Jones Industrial Average (solid line p) and *ex post* rational price (dotted line p^*), 1928-1979, both detrended by dividing by 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 2, Appendix.

Main Numerical Result

- Strange things: inequalities upside down!
- You see many rows because he derived many bounds with different assumptions
- The data is rejecting theory by an order of magnitude
- $50 > 8$ and $355.9 > 26.8$

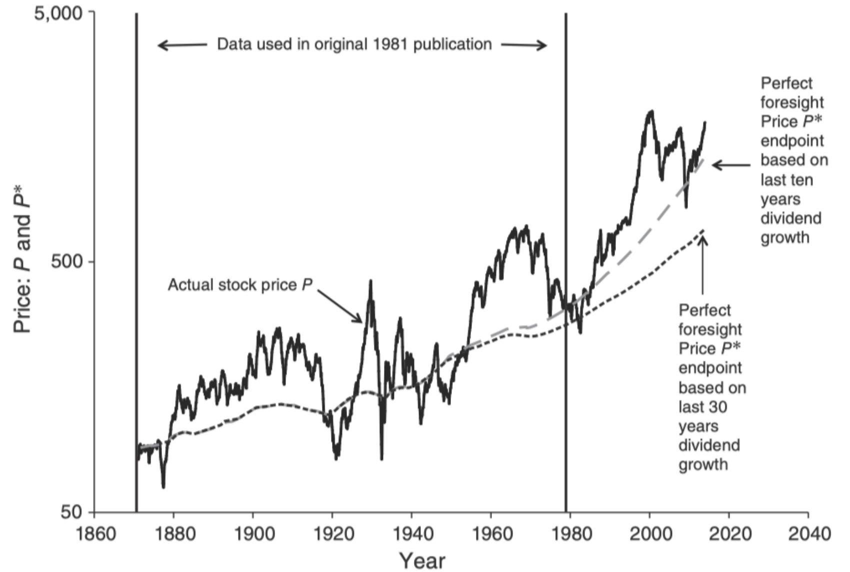
Historical evidence:

- During the Great Depression, (real) dividends were below average only on 1933-1935, and 1939
- But returns were pretty bad!

TABLE 2—SAMPLE STATISTICS FOR PRICE AND DIVIDEND SERIES

	Data Set 1: Standard and Poor's	Data Set 2: Modified Dow Industrial
Sample Period:	1871–1979	1928–1979
1) $E(p)$	145.5	982.6
$E(d)$	6.989	44.76
2) $\bar{r}$	.0480	0.456
$\bar{r}_2$	.0984	.0932
3) $b = ln\lambda$	.0148	.0188
$\hat{\sigma}(b)$	(.0011)	(1.0035)
4) $cor(p, p^*)$	.3918	.1626
$\sigma(d)$	1.481	9.828
Elements of Inequalities:		
Inequality (1)		
5) $\sigma(p)$	50.12	355.9
6) $\sigma(p^*)$	8.968	26.80
Inequality (11)		
7) $\sigma(\Delta p + d_{-1} - \bar{r}p_{-1})$	25.57	242.1
$min(\sigma)$	23.01	209.0
8) $\sigma(d)/\sqrt{\bar{r}_2}$	4.721	32.20
Inequality (13)		
9) $\sigma(\Delta p)$	25.24	239.5
$min(\sigma)$	22.71	206.4
10) $\sigma(d)/\sqrt{2\bar{r}}$	4.777	32.56

Newer Data



5,000
500
50
Price: P and P*
Data used in original 1981 publication
Actual stock price P
Perfect foresight Price P* endpoint based on last ten years dividend growth
Perfect foresight Price P* endpoint based on last 30 years dividend growth
1860
1880
1900
1920
1940
1960
1980
2000
2020
2040
Year

What if Discount Rates Vary?

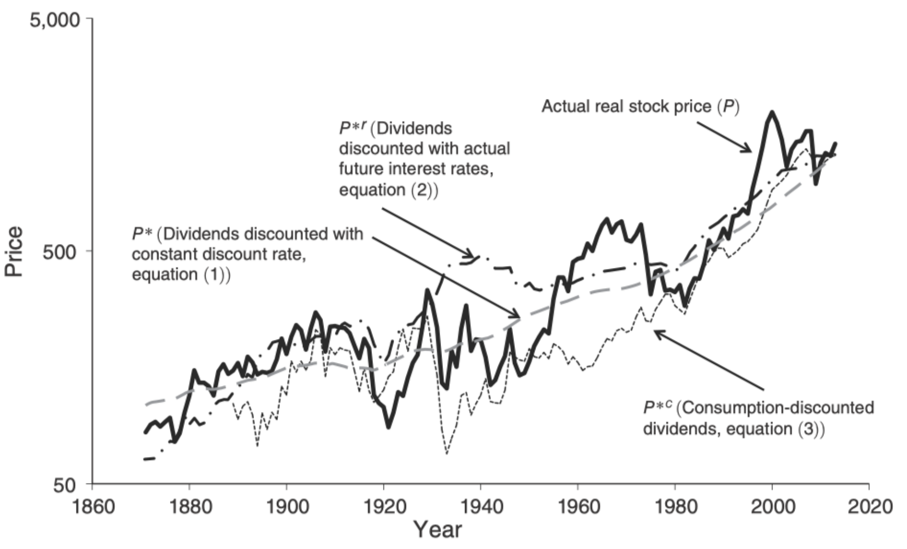
- We can always write:

$$P_t = \mathbb{E}_t \left[\sum_{k=1}^{\infty} D_{t+k} \prod_{j=1}^{k} \frac{1}{1 + r_{t+j-1}} \right]$$

- If we impose no discipline, this will always be true for some series of returns
- Shiller (1981) used 6-month prime commercial paper + a constant
- Shiller (2014) used our good and old friend:

$$r_{t+1} = -\log(\beta) + \delta \cdot \Delta c_{t+1}$$

Newer Data and Time-Varying Interest Rates



Where The Literature Took This

Takeaway message:

- Stock prices move too much to be justified by subsequent changes in dividends
- The natural way to reconcile this is allowing for rich time-variation in discount rates

Two Paths Forward:

Where The Literature Took This

Takeaway message:

- Stock prices move too much to be justified by subsequent changes in dividends
- The natural way to reconcile this is allowing for rich time-variation in discount rates

Two Paths Forward:

1. **Return decompositions**: how much of price movements is due to changes in dividend expectations vs expected returns? (Campbell and Shiller 1988; Campbell 1991; Cochrane 2011)

Where The Literature Took This

Takeaway message:

- Stock prices move too much to be justified by subsequent changes in dividends
- The natural way to reconcile this is allowing for rich time-variation in discount rates

Two Paths Forward:

1. **Return decompositions**: how much of price movements is due to changes in dividend expectations vs expected returns? (Campbell and Shiller 1988; Campbell 1991; Cochrane 2011)
2. **Behavioral finance and limits to arbitrage**: prices deviate from fundamentals due to investor psychology and limits to arbitrage (De Long et al. 1990; Shleifer and Vishny 1997; Shiller 2014; Barberis and Thaler 2003; Barberis 2018)

Behavioral Finance

- Excess volatility could reflect investor psychology, not rational risk premia
- Noise traders and limits to arbitrage (De Long et al. 1990; Shleifer and Vishny 1997)
- Shiller (2014): irrational exuberance and social dynamics drive speculative prices

Behavioral Finance

- Excess volatility could reflect investor psychology, not rational risk premia
- Noise traders and limits to arbitrage (De Long et al. 1990; Shleifer and Vishny 1997)
- Shiller (2014): irrational exuberance and social dynamics drive speculative prices
- Surveys: Barberis and Thaler (2003) and Barberis (2018) (extrapolation, overconfidence, prospect theory)
- We will not pursue this path, but it is an important and active research program

Questions?

Discount Rates: They Like to Move It!

The Campbell-Shiller Decomposition

- Campbell and Shiller (1988) made a simple, crazy powerful point
- Prices are *discounted expected cash-flows*: either return or dividends expectations must move to **generate** returns!

The Campbell-Shiller Decomposition

- Campbell and Shiller (1988) made a simple, crazy powerful point
- Prices are *discounted expected cash-flows*: either return or dividends expectations must move to **generate** returns!
- *The dividend-price ratio can only move if either* ***expected returns*** *or* ***expected dividend growth*** *moves, regardless of your mambo-jambo model!*
- They introduced the "Campbell-Shiller" decomposition, based on a Taylor expansion
- An almost model-free playground to test theories of time-varying discount rates

Campbell–Shiller Log-Linearization (Setup)

- Exact log return identity:

$$h_t \equiv \log(P_{t+1} + D_t) - \log(P_t)$$

Campbell–Shiller Log-Linearization (Setup)

- Exact log return identity:

$$h_t \equiv \log(P_{t+1} + D_t) - \log(P_t)$$

- Rewrite the nonlinear term:

$$\log(P_{t+1} + D_t) = p_{t+1} + \log(1 + \exp(d_t - p_{t+1}))$$

- Define the log dividend–price ratio:

$$\delta_{t+1} \equiv d_t - p_{t+1}$$

- Dividend D_t is paid at time $t+1$ in this notation (keeping it close to the paper)

Campbell–Shiller Log-Linearization (Taylor Step)

- Let $g(\delta) \equiv \log(1 + e^{\delta})$ and expand around $\bar{\delta}$ (the long-run average):

$$g(\delta_{t+1}) \approx g(\bar{\delta}) + g'(\bar{\delta})(\delta_{t+1} - \bar{\delta})$$

Campbell–Shiller Log-Linearization (Taylor Step)

- Let $g(\delta) \equiv \log(1 + e^{\delta})$ and expand around $\bar{\delta}$ (the long-run average):

$$g(\delta_{t+1}) \approx g(\bar{\delta}) + g'(\bar{\delta})(\delta_{t+1} - \bar{\delta})$$

- Derivative and the key constant ρ:

$$g'(\delta) = \frac{e^{\delta}}{1 + e^{\delta}}, \qquad 1 - \rho \equiv g'(\bar{\delta}), \qquad \rho = \frac{1}{1 + e^{\bar{\delta}}}$$

Campbell–Shiller Log-Linearization (Taylor Step)

- Let $g(\delta) \equiv \log(1 + e^{\delta})$ and expand around $\bar{\delta}$ (the long-run average):

$$g(\delta_{t+1}) \approx g(\bar{\delta}) + g'(\bar{\delta})(\delta_{t+1} - \bar{\delta})$$

- Derivative and the key constant ρ:

$$g'(\delta) = \frac{e^{\delta}}{1 + e^{\delta}}, \qquad 1 - \rho \equiv g'(\bar{\delta}), \qquad \rho = \frac{1}{1 + e^{\bar{\delta}}}$$

- Collect constants into k:

$$\log(P_{t+1} + D_t) \approx k + \rho\, p_{t+1} + (1 - \rho)\, d_t$$

Campbell–Shiller Log-Linear Return + Forward Solution

- Substitute into the return identity:

$$h_t \approx k + \rho\, p_{t+1} + (1-\rho)\, d_t - p_t$$

Campbell–Shiller Log-Linear Return + Forward Solution

- Substitute into the return identity:

$$h_t \approx k + \rho\, p_{t+1} + (1-\rho)\, d_t - p_t$$

- Express in terms of $\delta_t \equiv d_{t-1} - p_t$:

$$h_t \approx k + \Delta d_t + \delta_t - \rho\, \delta_{t+1}$$

Campbell–Shiller Log-Linear Return + Forward Solution

- Substitute into the return identity:

$$h_t \approx k + \rho\, p_{t+1} + (1-\rho)\, d_t - p_t$$

- Express in terms of $\delta_t \equiv d_{t-1} - p_t$:

$$h_t \approx k + \Delta d_t + \delta_t - \rho\, \delta_{t+1}$$

- Rearranged recursion and forward solution:

$$\delta_t = (h_t - \Delta d_t - k) + \rho\, \delta_{t+1} \;\Rightarrow\; \boxed{\delta_t = \sum_{j=0}^{\infty} \rho^j (h_{t+j} - \Delta d_{t+j}) - \frac{k}{1-\rho}}$$

- This equation has no economic content: it is an (approximate) **identity**. Intuition?

How to Add Empirical Content

- Notice that we can take expectations conditional on information at time t:

$$\delta_t = \mathbb{E}_t\left[\sum_{j=0}^{\infty} \rho^j (h_{t+j} - \Delta d_{t+j})\right] - \frac{k}{1-\rho} = \sum_{j=0}^{\infty} \rho^j (\mathbb{E}_t[h_{t+j}] - \mathbb{E}_t[\Delta d_{t+j}]) - \frac{k}{1-\rho}$$

- An Asset Pricing model will imply structure about $\mathbb{E}_t[h_{t+j}]$ and $\mathbb{E}_t[\Delta d_{t+j}]$

How to Add Empirical Content

- Notice that we can take expectations conditional on information at time t:

$$\delta_t = \mathbb{E}_t\left[\sum_{j=0}^{\infty} \rho^j (h_{t+j} - \Delta d_{t+j})\right] - \frac{k}{1-\rho} = \sum_{j=0}^{\infty} \rho^j (\mathbb{E}_t[h_{t+j}] - \mathbb{E}_t[\Delta d_{t+j}]) - \frac{k}{1-\rho}$$

- An Asset Pricing model will imply structure about $\mathbb{E}_t[h_{t+j}]$ and $\mathbb{E}_t[\Delta d_{t+j}]$
- Example: suppose that for some r_{t+j} and some c we have $\mathbb{E}_t[h_{t+j}] = \mathbb{E}_t[r_{t+j}] + c$
- We can write:

$$\delta_t = \sum_{j=0}^{\infty} \rho^j (\mathbb{E}_t[r_{t+j}] - \mathbb{E}_t[\Delta d_{t+j}]) + \frac{c-k}{1-\rho}$$

- This will become testable once we specify a model for returns and dividends

How to Add Empirical Content

- Notice that we can take expectations conditional on information at time t:

$$\delta_t = \mathbb{E}_t\left[\sum_{j=0}^{\infty} \rho^j (h_{t+j} - \Delta d_{t+j})\right] - \frac{k}{1-\rho} = \sum_{j=0}^{\infty} \rho^j (\mathbb{E}_t[h_{t+j}] - \mathbb{E}_t[\Delta d_{t+j}]) - \frac{k}{1-\rho}$$

- An Asset Pricing model will imply structure about $\mathbb{E}_t[h_{t+j}]$ and $\mathbb{E}_t[\Delta d_{t+j}]$
- Example: suppose that for some r_{t+j} and some c we have $\mathbb{E}_t[h_{t+j}] = \mathbb{E}_t[r_{t+j}] + c$
- We can write:

$$\delta_t = \sum_{j=0}^{\infty} \rho^j (\mathbb{E}_t[r_{t+j}] - \mathbb{E}_t[\Delta d_{t+j}]) + \frac{c-k}{1-\rho}$$

- This will become testable once we specify a model for returns and dividends
- Example: give me $\tilde{r}_{t,j} \equiv \mathbb{E}_t[r_{t+j}]$ and $\tilde{\Delta} d_{t,j} \equiv \mathbb{E}_t[\Delta d_{t+j}]$ and generate $\tilde{\delta}_t$
- We should have $\delta_t \approx \tilde{\delta}_t$

Questions?

Campbell and Shiller Meet the Data

There is hope in the approximation!

Table 2
Summary statistics for stock market data

Statistic	Data set and sample period: Cowles/S&P, 1871–1986	Cowles/S&P, 1926–1986	NYSE, 1926–1986	Correlation of Cowles/S&P and NYSE, 1926–1986
Δp_t:				
Mean	0.032	0.044	0.042	
Standard deviation	0.178	0.200	0.208	0.972
Δd_t:				
Mean	0.030	0.041	0.040	
Standard deviation	0.132	0.131	0.134	0.958
δ_t:				
Mean	−3.053	−3.121	−3.143	
Standard deviation	0.277	0.294	0.290	0.985

All variables in this table are nominal and measured annually. p_t is the log stock price, d_t is the log dividend, and δ_t is the log dividend-price ratio $d_{t-1} - p_t$.

Setup

- Assume that market participants use state variables y_t to do forecasts
- y_t evolves linearly (think about an MA(∞) process)
- The econometrician observes (with abuse of notation) $x_t \equiv (\delta_t, h_{t-1}, \Delta d_{t-1})^\top \subset y_t$

Setup

- Assume that market participants use state variables y_t to do forecasts
- y_t evolves linearly (think about an MA(∞) process)
- The econometrician observes (with abuse of notation) $x_t \equiv (\delta_t, h_{t-1}, \Delta d_{t-1})^\top \subset y_t$
- Assume that x_t (already demeaned) follows a VAR(p):

$$x_t = C_1 x_{t-1} + ... + C_p x_{t-p} + u_t$$

where all variables are demeaned and u_t is a white noise process each C_i is a 3×3 matrix

Setup

- Assume that market participants use state variables y_t to do forecasts
- y_t evolves linearly (think about an MA(∞) process)
- The econometrician observes (with abuse of notation) $x_t \equiv (\delta_t, h_{t-1}, \Delta d_{t-1})^\top \subset y_t$
- Assume that x_t (already demeaned) follows a VAR(p):

$$x_t = C_1 x_{t-1} + ... + C_p x_{t-p} + u_t$$

where all variables are demeaned and u_t is a white noise process each C_i is a 3×3 matrix

- Rewrite the VAR in companion form:

$$z_t \equiv \begin{bmatrix} x_t \\ x_{t-1} \\ \vdots \\ x_{t-p+1} \end{bmatrix} = \begin{bmatrix} C_1 & C_2 & ... & C_p \\ I & 0 & ... & 0 \\ 0 & I & ... & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & ... & 0 \end{bmatrix} \begin{bmatrix} x_{t-1} \\ x_{t-2} \\ \vdots \\ x_{t-p} \end{bmatrix} + \begin{bmatrix} u_t \\ 0 \\ \vdots \\ 0 \end{bmatrix} = A z_{t-1} + v_t$$

The VAR Approach

- The companion form implies: $\mathbb{E}_t[z_{t+j}] = A^j z_t$
- The VAR is a forecasting machine to generate discount rate and dividend forecasts
- Given A, we can compute:

$$\tilde{\delta}_t = \sum_{j=0}^{\infty} \rho^j \left(\mathbb{E}[h_{t+j} | \{x_t, x_{t-1}, ...\}] - \mathbb{E}[\Delta d_{t+j} | \{x_t, x_{t-1}, ...\}] \right)$$

The VAR Approach

- The companion form implies: $\mathbb{E}_t[z_{t+j}] = A^j z_t$
- The VAR is a forecasting machine to generate discount rate and dividend forecasts
- Given A, we can compute:

$$\tilde{\delta}_t = \sum_{j=0}^{\infty} \rho^j \left(\mathbb{E}[h_{t+j} | \{x_t, x_{t-1}, ...\}] - \mathbb{E}[\Delta d_{t+j} | \{x_t, x_{t-1}, ...\}]\right)$$

- Let $e_1 = (1, 0, 0, \dots, 0)^\top$ and $e_2 = (0, 1, 0, \dots, 0)^\top$ and $e_3 = (0, 0, 1, 0, \dots, 0)^\top$. Then:

$$\tilde{\delta}_t = \sum_{j=0}^{\infty} \rho^j (e_2^\top A^{j+1} z_t - e_3^\top A^{j+1} z_t) = \sum_{j=0}^{\infty} \rho^j ((e_2 - e_3)^\top A^{j+1} z_t)$$

The VAR Approach

- The companion form implies: $\mathbb{E}_t[z_{t+j}] = A^j z_t$
- The VAR is a forecasting machine to generate discount rate and dividend forecasts
- Given A, we can compute:

$$\tilde{\delta}_t = \sum_{j=0}^{\infty} \rho^j \left(\mathbb{E}[h_{t+j} | \{x_t, x_{t-1}, \dots\}] - \mathbb{E}[\Delta d_{t+j} | \{x_t, x_{t-1}, \dots\}]\right)$$

- Let $e_1 = (1, 0, 0, \dots, 0)^\top$ and $e_2 = (0, 1, 0, \dots, 0)^\top$ and $e_3 = (0, 0, 1, 0, \dots, 0)^\top$. Then:

$$\tilde{\delta}_t = \sum_{j=0}^{\infty} \rho^j (e_2^\top A^{j+1} z_t - e_3^\top A^{j+1} z_t) = \sum_{j=0}^{\infty} \rho^j ((e_2 - e_3)^\top A^{j+1} z_t)$$

- What condition would make the observed δ_t and the synthetic $\tilde{\delta}_t$ match?

Rational Pricing Implies Objective Forecasts

- If we impose $\delta_t = \tilde{\delta}_t$, then:

$$e_1^\top z_t = \delta_t = \tilde{\delta}_t = \sum_{j=0}^{\infty} \rho^j \left((e_2 - e_3)^\top A^{j+1} z_t\right) = \left(\sum_{j=0}^{\infty} \rho^j (e_2 - e_3)^\top A^{j+1}\right) z_t, \forall z_t$$

Rational Pricing Implies Objective Forecasts

- If we impose $\delta_t = \tilde{\delta}_t$, then:

$$e_1^\top z_t = \delta_t = \tilde{\delta}_t = \sum_{j=0}^{\infty} \rho^j \left((e_2 - e_3)^\top A^{j+1} z_t \right) = \left(\sum_{j=0}^{\infty} \rho^j (e_2 - e_3)^\top A^{j+1} \right) z_t, \forall z_t$$

- This would imply the following restriction on the VAR coefficients:

$$\boxed{e_1^\top = \left(\sum_{j=0}^{\infty} \rho^j (e_2 - e_3)^\top A^{j+1} \right) = (e_2 - e_3)^\top A(I - \rho A)^{-1}}$$

Rational Pricing Implies Objective Forecasts

- If we impose $\delta_t = \tilde{\delta}_t$, then:

$$e_1^\top z_t = \delta_t = \tilde{\delta}_t = \sum_{j=0}^{\infty} \rho^j \left((e_2 - e_3)^\top A^{j+1} z_t \right) = \left(\sum_{j=0}^{\infty} \rho^j (e_2 - e_3)^\top A^{j+1} \right) z_t, \forall z_t$$

- This would imply the following restriction on the VAR coefficients:

$$\boxed{e_1^\top = \left(\sum_{j=0}^{\infty} \rho^j (e_2 - e_3)^\top A^{j+1} \right) = (e_2 - e_3)^\top A (I - \rho A)^{-1}}$$

- We can do a Wald test! This is just a (non-linear) restriction on the VAR coefficients!
- A rejection would imply that whatever theory we imposed to generate the forecasts is not consistent with the data
- $\rho = 0.93$ to match average dividend–price ratio in the data

Tested Theory 1

- They assume a constant conditional expected return $\mathbb{E}_t[h_{t+j}] = \bar{r}$
- They can use $x_t = (\delta_t, \Delta d_{t-1})$ here since $\bar{r}$ gets absorbed

Tested Theory 1

- They assume a constant conditional expected return $\mathbb{E}_t[h_{t+j}] = \bar{r}$
- They can use $x_t = (\delta_t, \Delta d_{t-1})$ here since $\bar{r}$ gets absorbed

VAR estimation:

δ_{t+1}	0.706	0.259	0.515	0.000
	(0.066)	(0.139)		
Δd_t	−0.197	0.231	0.227	0.000
	(0.039)	(0.083)		

Implications of VAR estimates:

$$\delta'_t = \underset{(0.123)}{0.636}\delta_t - \underset{(0.106)}{0.097}\Delta d_{t-1}$$

$$\sigma(\delta'_t)/\sigma(\delta_t) = \underset{(0.124)}{0.637} \qquad \text{corr}\,(\delta'_t, \delta_t) = \underset{(0.006)}{0.997}$$

Significance level for Wald test that $\delta'_t = \delta_t$: 0.005

- The idea of constant expected returns is easily rejected!

Tested Theory 2

- They impose a constant *premium* over the risk-free rate: $\mathbb{E}_t[h_{t+j}] = \mathbb{E}_t[r^f_{t+j}] + c$
- Use $x_t = (\delta_t, r^f_{t-1} - \Delta d_{t-1})$ with nominal variables – inflation cancels out!

Tested Theory 2

- They impose a constant *premium* over the risk-free rate: $\mathbb{E}_t[h_{t+j}] = \mathbb{E}_t[r^f_{t+j}] + c$
- Use $x_t = (\delta_t, r^f_{t-1} - \Delta d_{t-1})$ with nominal variables – inflation cancels out!

Table 5
Testing constant expected real and excess returns

	Model version 1 (constant expected real returns)			Model version 2 (constant expected excess returns)		
	Lag length			Lag length		
	1	3	5	1	3	5
Cowles/S&P, 1871–1986						
h_t regression test	0.078	0.056	0.083	0.100	0.007	0.014
ξ_t regression test	0.045	0.035	0.055	0.073	0.005	0.009
Test that $\delta'_t = \delta_t$	0.005	0.000	0.000	0.009	0.000	0.000
$\sigma(\delta'_t)/\sigma(\delta_t)$	0.637	0.370	0.382	0.674	0.434	0.402
	(0.124)	(0.129)	(0.073)	(0.111)	(0.143)	(0.060)
corr (δ'_t, δ_t)	0.997	0.837	0.326	0.999	0.856	0.431
	(0.006)	(0.246)	(0.565)	(0.004)	(0.200)	(0.583)

Tested Theory 3 and 4

- The third one: $\mathbb{E}_t[h_{t+j}] = \alpha \cdot \mathbb{E}_t[\Delta c_{t+j}] +$ constant, where α is an estimated RRA parameter
- The fourth one: $\mathbb{E}_t[h_{t+j}] = \alpha \cdot V_t +$ constant terms, where $V_t = h_t^2$ is a proxy for volatility
- This is due to Mehra and Prescott (1985) and Pindyck (1984)

Tested Theory 3 and 4

Table 7
Testing consumption- and volatility-based models of the discount rate

	Model version 3 (consumption)		Model version 4 (volatility)	
	Lag length		Lag length	
	1	3	1	3
Cowles/S&P, 1871–1986				
Estimate of α	−2.191	−0.423	2.552	0.960
(standard error)	(1.399)	(0.818)	(2.549)	(0.985)
ξ_t regression test	0.072	0.038	0.170	0.096
Test that $\delta'_t = \delta_t$	0.012	0.000	0.039	0.007
$\sigma(\delta_t - \delta'_{dt})/\sigma(\delta_t)$	0.484	0.711	0.368	0.724
	(0.100)	(0.216)	(0.122)	(0.211)
$\sigma(\delta'_{rt})/\sigma(\delta_t)$	0.227	0.038	0.077	0.095
	(0.149)	(0.081)	(0.116)	(0.148)
corr $(\delta_t - \delta'_{dt}, \delta'_{rt})$	0.910	−0.024	0.339	0.660
	(0.073)	(0.788)	(1.415)	(0.525)

Main Takeaways from C-S 1988

- Campbell and Shiller (1988) called it: the dividend-price ratio can only move if either expected returns or expected dividend growth moves
- Their decomposition imposes very little structure! *You give me some theory, my VAR gives you a test.*
- Constant expected returns cannot rationalize the data!
- Allowing the discount rates to depend on consumption and volatility was *not* enough
- Expected discount rates *have* to move, and they have to move a lot!

Questions?

The Variance Decomposition of Returns

From C-S Identity to News Decomposition

- Campbell (1991) took the C-S log-linearization one step further
- Recall the approximate identity:

$$h_t \approx k + \Delta d_t + \delta_t - \rho\,\delta_{t+1}$$

From C-S Identity to News Decomposition

- Campbell (1991) took the C-S log-linearization one step further
- Recall the approximate identity:

$$h_t \approx k + \Delta d_t + \delta_t - \rho\, \delta_{t+1}$$

- The forward solution gives:

$$h_t \approx -\sum_{j=1}^{\infty} \rho^j \Delta d_{t+j} + \sum_{j=1}^{\infty} \rho^j h_{t+j} + \Delta d_t + (1-\rho)\delta_t + k'$$

where $k' \equiv k/(1-\rho)$ collects constants

From C-S Identity to News Decomposition

- Campbell (1991) took the C-S log-linearization one step further
- Recall the approximate identity:

$$h_t \approx k + \Delta d_t + \delta_t - \rho\,\delta_{t+1}$$

- The forward solution gives:

$$h_t \approx -\sum_{j=1}^{\infty} \rho^j \Delta d_{t+j} + \sum_{j=1}^{\infty} \rho^j h_{t+j} + \Delta d_t + (1-\rho)\delta_t + k'$$

  where $k' \equiv k/(1-\rho)$ collects constants

- This is still just an identity – but now let's take *surprises*

Taking Surprises

- Define the "surprise operator": $(\mathbb{E}_t - \mathbb{E}_{t-1})[\cdot]$
- Apply it to both sides of the identity. Constants and $t-1$-known quantities vanish:

$$h_t - \mathbb{E}_{t-1}[h_t] = (\mathbb{E}_t - \mathbb{E}_{t-1})\left[\sum_{j=0}^{\infty} \rho^j \Delta d_{t+j}\right] - (\mathbb{E}_t - \mathbb{E}_{t-1})\left[\sum_{j=1}^{\infty} \rho^j h_{t+j}\right]$$

Taking Surprises

- Define the "surprise operator": $(\mathbb{E}_t - \mathbb{E}_{t-1})[\cdot]$
- Apply it to both sides of the identity. Constants and $t-1$-known quantities vanish:

$$h_t - \mathbb{E}_{t-1}[h_t] = (\mathbb{E}_t - \mathbb{E}_{t-1})\left[\sum_{j=0}^{\infty} \rho^j \Delta d_{t+j}\right] - (\mathbb{E}_t - \mathbb{E}_{t-1})\left[\sum_{j=1}^{\infty} \rho^j h_{t+j}\right]$$

- Define **cash-flow news** and **discount-rate news**:

$$\eta_{d,t} \equiv (\mathbb{E}_t - \mathbb{E}_{t-1})\sum_{j=0}^{\infty} \rho^j \Delta d_{t+j}, \qquad \eta_{h,t} \equiv (\mathbb{E}_t - \mathbb{E}_{t-1})\sum_{j=1}^{\infty} \rho^j h_{t+j}$$

The Campbell (1991) Decomposition

$$\boxed{h_t - \mathbb{E}_{t-1}[h_t] = \eta_{d,t} - \eta_{h,t}}$$

- Unexpected returns = cash-flow news − discount-rate news
- Still an approximate identity – no model content yet!

The Campbell (1991) Decomposition

$$\boxed{h_t - \mathbb{E}_{t-1}[h_t] = \eta_{d,t} - \eta_{h,t}}$$

- Unexpected returns = cash-flow news − discount-rate news
- Still an approximate identity – no model content yet!
- Intuition for the signs:
  - Good cash-flow news ($\eta_{d,t} > 0$): future dividends revised upward $\implies$ positive surprise return
  - Positive discount-rate news ($\eta_{h,t} > 0$): expected future returns revised upward $\implies$ current price drops $\implies$ negative surprise return

The Campbell (1991) Decomposition

$$\boxed{h_t - \mathbb{E}_{t-1}[h_t] = \eta_{d,t} - \eta_{h,t}}$$

- Unexpected returns = cash-flow news − discount-rate news
- Still an approximate identity – no model content yet!
- Intuition for the signs:
  - Good cash-flow news ($\eta_{d,t} > 0$): future dividends revised upward $\implies$ positive surprise return
  - Positive discount-rate news ($\eta_{h,t} > 0$): expected future returns revised upward $\implies$ current price drops $\implies$ negative surprise return
- This is the key equation: how much of the variance of unexpected returns comes from each component?

VAR Implementation for News

- Use the same VAR machinery from the C-S section. With $z_t = Az_{t-1} + v_t$:

$$(\mathbb{E}_t - \mathbb{E}_{t-1})[z_{t+j}] = A^j z_t - A^j \mathbb{E}_{t-1}[z_t] = A^j \left(z_t - \mathbb{E}_{t-1}[z_t]\right) = A^j v_t$$

VAR Implementation for News

- Use the same VAR machinery from the C-S section. With $z_t = Az_{t-1} + v_t$:

$$(\mathbb{E}_t - \mathbb{E}_{t-1})[z_{t+j}] = A^j z_t - A^j \mathbb{E}_{t-1}[z_t] = A^j \left(z_t - \mathbb{E}_{t-1}[z_t]\right) = A^j v_t$$

- Discount-rate news (let e_h select the return row from z_t):

$$\eta_{h,t} = e_h^\top \sum_{j=1}^{\infty} \rho^j A^j \, v_t = e_h^\top \rho A (I - \rho A)^{-1} v_t$$

VAR Implementation for News

- Use the same VAR machinery from the C-S section. With $z_t = Az_{t-1} + v_t$:

$$(\mathbb{E}_t - \mathbb{E}_{t-1})[z_{t+j}] = A^j z_t - A^j \mathbb{E}_{t-1}[z_t] = A^j \left(z_t - \mathbb{E}_{t-1}[z_t]\right) = A^j v_t$$

- Discount-rate news (let e_h select the return row from z_t):

$$\eta_{h,t} = e_h^\top \sum_{j=1}^{\infty} \rho^j A^j \, v_t = e_h^\top \rho A (I - \rho A)^{-1} v_t$$

- Cash-flow news (backed out from the identity):

$$\eta_{d,t} = (h_t - \mathbb{E}_{t-1}[h_t]) + \eta_{h,t} = e_h^\top v_t + e_h^\top \rho A (I - \rho A)^{-1} v_t$$

VAR Implementation for News

- Use the same VAR machinery from the C-S section. With $z_t = Az_{t-1} + v_t$:

$$(\mathbb{E}_t - \mathbb{E}_{t-1})[z_{t+j}] = A^j z_t - A^j \mathbb{E}_{t-1}[z_t] = A^j \left(z_t - \mathbb{E}_{t-1}[z_t]\right) = A^j v_t$$

- Discount-rate news (let e_h select the return row from z_t):

$$\eta_{h,t} = e_h^\top \sum_{j=1}^{\infty} \rho^j A^j \, v_t = e_h^\top \rho A (I - \rho A)^{-1} v_t$$

- Cash-flow news (backed out from the identity):

$$\eta_{d,t} = (h_t - \mathbb{E}_{t-1}[h_t]) + \eta_{h,t} = e_h^\top v_t + e_h^\top \rho A (I - \rho A)^{-1} v_t$$

- Both $\eta_{d,t}$ and $\eta_{h,t}$ are linear functions of the VAR innovations v_t
- $\text{Var}(\eta_{h,t}) = \lambda' \Sigma_v \lambda$ where $\lambda \equiv e_h^\top \rho A (I - \rho A)^{-1}$ and Σ_v is the covariance matrix of v_t

The Variance Decomposition

- Take variances of the decomposition $h_t - \mathbb{E}_{t-1}[h_t] = \eta_{d,t} - \eta_{h,t}$:

$$\mathsf{Var}(h_t - \mathbb{E}_{t-1}[h_t]) = \mathsf{Var}(\eta_{d,t}) + \mathsf{Var}(\eta_{h,t}) - 2\,\mathsf{Cov}(\eta_{d,t}, \eta_{h,t})$$

The Variance Decomposition

- Take variances of the decomposition $h_t - \mathbb{E}_{t-1}[h_t] = \eta_{d,t} - \eta_{h,t}$:

$$\mathsf{Var}(h_t - \mathbb{E}_{t-1}[h_t]) = \mathsf{Var}(\eta_{d,t}) + \mathsf{Var}(\eta_{h,t}) - 2\,\mathsf{Cov}(\eta_{d,t}, \eta_{h,t})$$

- Divide both sides by $\mathsf{Var}(h_t - \mathbb{E}_{t-1}[h_t])$:

$$1 = \frac{\mathsf{Var}(\eta_{d,t})}{\mathsf{Var}(h_t - \mathbb{E}_{t-1}[h_t])} + \frac{\mathsf{Var}(\eta_{h,t})}{\mathsf{Var}(h_t - \mathbb{E}_{t-1}[h_t])} - \frac{2\,\mathsf{Cov}(\eta_{d,t}, \eta_{h,t})}{\mathsf{Var}(h_t - \mathbb{E}_{t-1}[h_t])}$$

The Variance Decomposition

- Take variances of the decomposition $h_t - \mathbb{E}_{t-1}[h_t] = \eta_{d,t} - \eta_{h,t}$:

$$\mathrm{Var}(h_t - \mathbb{E}_{t-1}[h_t]) = \mathrm{Var}(\eta_{d,t}) + \mathrm{Var}(\eta_{h,t}) - 2\,\mathrm{Cov}(\eta_{d,t}, \eta_{h,t})$$

- Divide both sides by $\mathrm{Var}(h_t - \mathbb{E}_{t-1}[h_t])$:

$$1 = \frac{\mathrm{Var}(\eta_{d,t})}{\mathrm{Var}(h_t - \mathbb{E}_{t-1}[h_t])} + \frac{\mathrm{Var}(\eta_{h,t})}{\mathrm{Var}(h_t - \mathbb{E}_{t-1}[h_t])} - \frac{2\,\mathrm{Cov}(\eta_{d,t}, \eta_{h,t})}{\mathrm{Var}(h_t - \mathbb{E}_{t-1}[h_t])}$$

- Notice that $\mathrm{Var}(h_t - \mathbb{E}_{t-1}[h_t]) = e_h^\top \Sigma_v e_h$
- The question: which share is larger?
- This is all computable from the VAR estimates (A and Σ_v)

Another Interesting Measure

- If expected returns are moving around, these expectations might have some persistence
- Let $u_t \equiv [\mathbb{E}_t - \mathbb{E}_{t-1}]h_{t+1}$.
- This is news about tomorrow's return, which came in today
- Notice that $u_t = e_h^\top A v_t$

Another Interesting Measure

- If expected returns are moving around, these expectations might have some persistence
- Let $u_t \equiv [\mathbb{E}_t - \mathbb{E}_{t-1}]h_{t+1}$.
- This is news about tomorrow's return, which came in today
- Notice that $u_t = e_h^\top A v_t$
- Define P_h as the following ratio:

$$P_h \equiv \frac{\sigma(\eta_{h,t})}{\sigma(u_t)} = \frac{\lambda^\top \Sigma_v \lambda}{e_h^\top A \Sigma_v A^\top e_h}$$

- Intuition: if u_t increases, by how much does $\eta_{h,t}$ increase?
- Important exercise: show that if $\mathbb{E}_t[h_{t+1}]$ follows an AR(1), $\eta_{h,t}$ and u_t are proportional

Another Interesting Measure

- If expected returns are moving around, these expectations might have some persistence
- Let $u_t \equiv [\mathbb{E}_t - \mathbb{E}_{t-1}]h_{t+1}$.
- This is news about tomorrow's return, which came in today
- Notice that $u_t = e_h^\top A v_t$
- Define P_h as the following ratio:

$$P_h \equiv \frac{\sigma(\eta_{h,t})}{\sigma(u_t)} = \frac{\lambda^\top \Sigma_v \lambda}{e_h^\top A \Sigma_v A^\top e_h}$$

- Intuition: if u_t increases, by how much does $\eta_{h,t}$ increase?
- Important exercise: show that if $\mathbb{E}_t[h_{t+1}]$ follows an AR(1), $\eta_{h,t}$ and u_t are proportional

Notes on implementation:

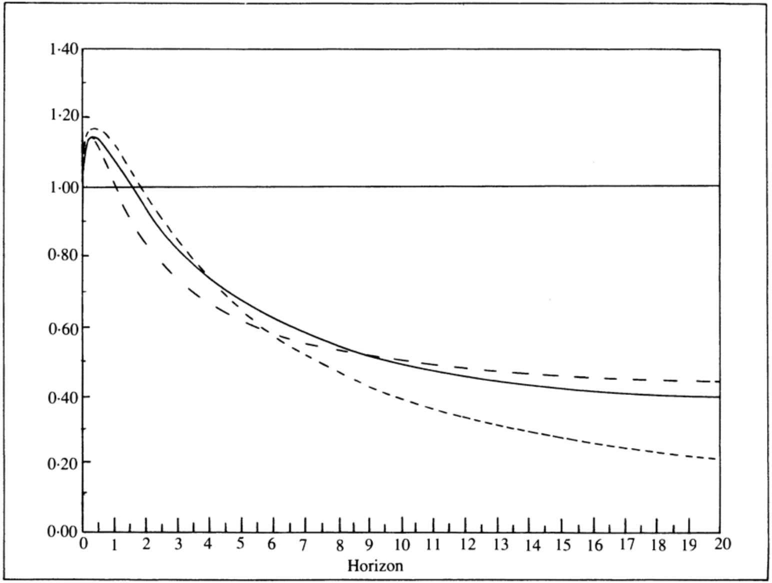
- Everything is done at the monthly frequency, unless specified otherwise

Variance Decomposition

Variance Decomposition for Real Stock Returns

VAR specification and time period	R_h^2 (Sig.)	Var (η_d) (SE)	Var (η_h) (SE)	-2Cov(η_d, η_h) (SE)	Corr(η_d, η_h) (SE)	P_h (SE)
h, D/P, rrel						
1 lag, monthly						
A: 1927:1–1988:12	0·024	0·369	0·285	0·346	−0·534	4·772
	(0·018)	(0·119)	(0·145)	(0·046)	(0·127)	(2·247)
B: 1927:1–1951:12	0·028	0·437	0·185	0·378	−0·664	3·258
	(0·183)	(0·226)	(0·182)	(0·053)	(0·118)	(2·414)
C: 1952:1–1988:12	0·065	0·127	0·772	0·101	−0·161	5·794
	(0·000)	(0·016)	(0·164)	(0·153)	(0·256)	(1·469)
h, D/P, rrel						
6 lags, monthly						
A: 1927:1–1988:12	0·087	0·538	0·265	0·197	−0·261	3·972
	(0·004)	(0·181)	(0·162)	(0·121)	(0·203)	(2·253)
B: 1927:1–1951:12	0·129	0·661	0·118	0·222	−0·398	1·909
	(0·083)	(0·363)	(0·142)	(0·288)	(0·565)	(1·515)
C: 1952:1–1988:12	0·118	0·127	0·797	0·075	−0·118	4·100
	(0·000)	(0·035)	(0·175)	(0·165)	(0·269)	(1·112)
h, D/P, rrel						
4 lags, quarterly						
A: 1927:1–1988:4	0·162	0·334	0·497	0·170	−0·208	2·726
	(0·045)	(0·096)	(0·193)	(0·186)	(0·269)	(1·435)
B: 1927:1–1951:4	0·307	0·428	0·476	0·096	−0·106	1·856
	(0·024)	(0·195)	(0·166)	(0·236)	(0·290)	(0·820)
C: 1952:1–1988:4	0·213	0·158	0·916	−0·074	0·097	7·289
	(0·000)	(0·067)	(0·184)	(0·211)	(0·257)	(5·437)

Implied Variance Ratios



Main Takeaways from Campbell (1991)

- News about *expected returns* explain at least half of the variation in unexpected *realized returns*
- Expected returns are persistent, and show negative autocorrelation at longer horizons
- This is consistent with the idea that expected returns are time-varying and that they are a key driver of stock price movements
- In the post-war period, it's almost all discount-rate news and very little cash-flow news

Questions?

Taking Stock

- Shiller (1981) showed that stocks prices are too volatile to be consistent by constant discount rates
- Campbell and Shiller (1988) built the decomposition framework and tested theories of time-varying discount rates
- Strong rejection of constant premia and of premia that depend on consumption growth and volatility
- Campbell (1991): *You guys have been paying attention to wrong thing!*
- Most of the variation in stock prices comes from variation in expected returns!

Taking Stock

- Shiller (1981) showed that stocks prices are too volatile to be consistent by constant discount rates
- Campbell and Shiller (1988) built the decomposition framework and tested theories of time-varying discount rates
- Strong rejection of constant premia and of premia that depend on consumption growth and volatility
- Campbell (1991): *You guys have been paying attention to wrong thing!*
- Most of the variation in stock prices comes from variation in expected returns!

Implications for Asset Pricing: - The bar is high: models should have time-varying discount rates - From Mehra and Prescott (1985): rates should be higher than what a Lucas-economy can deliver - Oh, and by the way, they should move a lot!

Implications and Road Ahead

- Shiller showed that prices are too volatile under constant discount rates
- Campbell and Shiller (1988) built the decomposition framework and tested theories
- Campbell (1991) showed that most price variation is **discount-rate news**

Implications and Road Ahead

- Shiller showed that prices are too volatile under constant discount rates
- Campbell and Shiller (1988) built the decomposition framework and tested theories
- Campbell (1991) showed that most price variation is **discount-rate news**
- The message is clear: we need models that generate rich time-variation in discount rates!
- Cochrane (2011): this is *the* central finding of modern empirical asset pricing

Questions?

Some More Modern Stuff

Newer Data, Same Message

- Cochrane (2011) revisited Campbell and Shiller (1988) with data up to 2010
- Recall the decomposition:

$$\delta_t = \sum_{j=0}^{\infty} \rho^j h_{t+1+j} - \sum_{j=0}^{\infty} \rho^j \Delta d_{t+j} + \text{constants}$$

Newer Data, Same Message

- Cochrane (2011) revisited Campbell and Shiller (1988) with data up to 2010
- Recall the decomposition:

$$\delta_t = \sum_{j=0}^{\infty} \rho^j h_{t+1+j} - \sum_{j=0}^{\infty} \rho^j \Delta d_{t+j} + \text{constants}$$

Take the covariance on both sides with δ_t:

$$\text{Var}(\delta_t, \delta_t) = \text{Cov}(\delta_t, \delta_t) = \text{Cov}\left(\delta_t, \sum_{j=1}^{\infty} \rho^j h_{t+1+j}\right) - \text{Cov}\left(\delta_t, \sum_{j=1}^{\infty} \rho^j \Delta d_{t+j}\right)$$

Newer Data, Same Message

- Cochrane (2011) revisited Campbell and Shiller (1988) with data up to 2010
- Recall the decomposition:

$$\delta_t = \sum_{j=0}^{\infty} \rho^j h_{t+1+j} - \sum_{j=0}^{\infty} \rho^j \Delta d_{t+j} + \text{constants}$$

Take the covariance on both sides with δ_t:

$$\text{Var}(\delta_t, \delta_t) = \text{Cov}(\delta_t, \delta_t) = \text{Cov}\left(\delta_t, \sum_{j=1}^{\infty} \rho^j h_{t+1+j}\right) - \text{Cov}\left(\delta_t, \sum_{j=1}^{\infty} \rho^j \Delta d_{t+j}\right)$$

- We get an identity in terms of regression coefficients:

$$1 = \frac{\text{Cov}\left(\delta_t, \sum_{j=1}^{\infty} \rho^j h_{t+1+j}\right)}{\text{Var}(\delta_t)} - \frac{\text{Cov}\left(\delta_t, \sum_{j=1}^{\infty} \rho^j \Delta d_{t+j}\right)}{\text{Var}(\delta_t)} \equiv b_r - b_{\Delta d}$$

It's All Discount-Rate News

Method and Horizon	Coefficient $b_r^{(k)}$	$b^{(k)}_{\Delta d}$	$\rho^k b^{(k)}_{dp}$
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

- Essentially **all** variation in the dividend-price ratio is explained by variation in expected returns
- The game in town is explaining time-variation of expected returns!
- Given early lectures: this is tightly connected to explaining time-variation of the SDF!

30 years later... Dividends Strike Back!

- DeLaO and Myers (2021) brought *subjective beliefs* into the picture
- Under rational expectations, we have that

$$\text{Cov}(\mathbb{E}_t[\Delta d_{t+1}], \delta_t) = \text{Cov}(\Delta d_{t+1}, \delta_t)$$

- Under subjective beliefs, it does not need to be the case!

30 years later... Dividends Strike Back!

- DeLaO and Myers (2021) brought *subjective beliefs* into the picture
- Under rational expectations, we have that

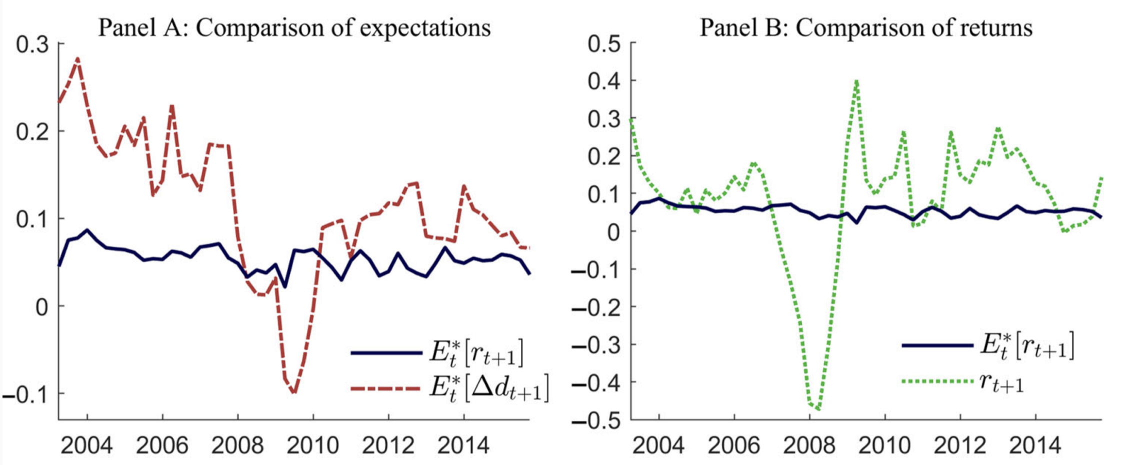
$$\text{Cov}(\mathbb{E}_t[\Delta d_{t+1}], \delta_t) = \text{Cov}(\Delta d_{t+1}, \delta_t)$$

- Under subjective beliefs, it does not need to be the case!

Their idea:

- Measure return and dividend growth expectation using survey data (IBES and Graham-Harvey)!
- Instead of inferring what people think, just ask them!
- This is a *very* influential paper

One-year ahead expectations



The Subjective Belief Campbell-Shiller Decomposition

- If $\mathbb{E}_t^*$ is a *subjective* expectation operator, we can write:

$$\delta_t = \sum_{j=0}^{\infty} \rho^j \mathbb{E}_t^*[h_{t+1+j}] - \sum_{j=0}^{\infty} \rho^j \mathbb{E}_t^*[\Delta d_{t+j}] + \text{constants}$$

- This has to hold because it's just an accounting identity!

The Subjective Belief Campbell-Shiller Decomposition

- If $\mathbb{E}_t^*$ is a *subjective* expectation operator, we can write:

$$\delta_t = \sum_{j=0}^{\infty} \rho^j \mathbb{E}_t^*[h_{t+1+j}] - \sum_{j=0}^{\infty} \rho^j \mathbb{E}_t^*[\Delta d_{t+j}] + \text{constants}$$

- This has to hold because it's just an accounting identity!
- They don't observe, at time t, the full path of expectations. They focus on:

$$1 = \underbrace{\frac{\operatorname{cov}(E_t^*[\Delta d_{t+1}],\, \delta_t)}{\operatorname{var}(\delta_t)}}_{CF_1} + \underbrace{\frac{-\operatorname{cov}(E_t^*[r_{t+1}],\, \delta_t)}{\operatorname{var}(\delta_t)}}_{DR_1} + \underbrace{\rho\frac{\operatorname{cov}(E_t^*[\delta_{t+1}],\, \delta_t)}{\operatorname{var}(\delta_t)}}_{LT}$$

- They can **directly** measure CF_1 and DR_1 in the data

Empirical Results

	Sample (1)	CF_1 (2)	DR_1 (3)	LT (4)
Price-dividend ratio	2003 to 2015	0.390	−0.049	0.659
		(0.034)	(0.013)	(0.037)
Price-earnings ratio	2003 to 2015	0.937	−0.004	0.064
		(0.095)	(0.008)	(0.105)
	1976 to 2015	0.417		
		(0.105)		

- You'd expect very high DR_1 and very low CF_1... not what happens!
- Most of the variation in price-dividends and price-earnings is steaming from dividends!
- This is exactly the *opposite* to Campbell (1991)

Empirical Results

	Sample (1)	CF_2 (2)	DR_{10} (3)
Price-dividend ratio	2003 to 2015	0.65	−0.07
		(0.12)	(0.11)
Price-earnings ratio	2003 to 2015	0.98	−0.08
		(0.10)	(0.04)
	1985 to 2015	0.64	
		(0.15)	

$$CF_2 \equiv \frac{\operatorname{cov}\left(\sum_{j=1}^{2} \rho^{j-1} E_t^*[\Delta d_{t+j}],\ \delta_t\right)}{\operatorname{var}(\delta_t)} \qquad DR_{10} \equiv \frac{\operatorname{cov}\left(-\sum_{j=1}^{10} \rho^{j-1} E_t^*[r_{t+j}],\ \delta_t\right)}{\operatorname{var}(\delta_t)}$$

Benchmarks

Panel A: Price-Dividend Ratio				
	CF (1)	DR (2)	CF_1 (3)	CF_2 (4)
Data 2003 to 2015	1.09	−0.09	0.39	0.65
	(0.04)	(0.04)	(0.03)	(0.12)
Habit formation	0.00	1.00	0.00	0.00
	n/a	n/a	n/a	n/a
Long-run risk	0.38	0.62	0.11	0.19
	(0.19)	(0.19)	(0.06)	(0.09)
Learning	0.07	0.93	0.00	0.00
	(0.16)	(0.16)	(0.01)	(0.01)
Return extrapolation	∞	$-\infty$	0.87	1.74
	n/a	n/a	n/a	n/a
Earnings growth reversal	1.09	−0.09	0.39	0.64
	n/a	n/a	n/a	n/a
Panel B: Price-Earnings Ratio				
Data 2003 to 2015	1.01	−0.01	0.94	0.98
	(0.01)	(0.01)	(0.09)	(0.10)
Earnings growth reversal	0.95	0.05	0.95	0.95
	(0.01)	(0.01)	(0.01)	(0.01)

Questions?

AI-Generated Summary

- Shiller (1981): with constant r, prices are too volatile
- Takeaway: big time-variation in expected returns / discount rates (SDF)
- Campbell and Shiller (1988): dividend–price ratio reflects expected returns vs dividend growth
- VAR tests reject constant expected returns (and simple consumption/volatility stories)
- Campbell (1991) / Cochrane (2011): discount-rate news dominates
- DeLaO and Myers (2021): surveys + subjective beliefs revive cash-flow news

Readings and References i

Barberis, Nicholas. 2018. "Psychology-Based Models of Asset Prices and Trading Volume." In *Handbook of Behavioral Economics: Applications and Foundations*, 1:79–175. Elsevier. https://doi.org/10.1016/bs.hesbe.2018.07.001.

Barberis, Nicholas, and Richard Thaler. 2003. "A Survey of Behavioral Finance." In *Handbook of the Economics of Finance*, 1:1053–1128. Elsevier. https://doi.org/10.1016/S1574-0102(03)01027-6.

Campbell, John Y. 1991. "A Variance Decomposition for Stock Returns." *The Economic Journal* 101 (405): 157–79. https://doi.org/10.2307/2233809.

Campbell, John Y., and Robert J. Shiller. 1988. "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors." *The Review of Financial Studies* 1 (3): 195–228. https://doi.org/10.1093/rfs/1.3.195.

Cochrane, John H. 2011. "Presidential Address: Discount Rates." *The Journal of Finance* 66 (4): 1047–1108. https://doi.org/10.1111/j.1540-6261.2011.01671.x.

De Long, J. Bradford, Andrei Shleifer, Lawrence H. Summers, and Robert J. Waldmann. 1990. "Noise Trader Risk in Financial Markets." *Journal of Political Economy* 98 (4): 703–38. https://doi.org/10.1086/261703.

Readings and References ii

DeLaO, Ricardo, and Sean Myers. 2021. "Subjective Cash Flow and Discount Rate Expectations." *The Journal of Finance* 76 (3): 1339–87. https://doi.org/10.1111/jofi.13016.

Mehra, Rajnish, and Edward C. Prescott. 1985. "The Equity Premium: A Puzzle." *Journal of Monetary Economics* 15 (2): 145–61. https://doi.org/10.1016/0304-3932(85)90061-3.

Pindyck, Robert S. 1984. "Risk, Inflation, and the Stock Market." *The American Economic Review* 74 (3): 335–51. http://www.jstor.org/stable/1804011.

Shiller, Robert J. 1981. "Do Stock Prices Move Too Much to Be Justified by Subsequent Changes in Dividends?" *The American Economic Review* 71 (3): 421–36. http://www.jstor.org/stable/1802789.

———. 2014. "Speculative Asset Prices." *American Economic Review* 104 (6): 1486–1517. https://doi.org/10.1257/aer.104.6.1486.

Shleifer, Andrei, and Robert W. Vishny. 1997. "The Limits of Arbitrage." *The Journal of Finance* 52 (1): 35–55. https://doi.org/10.1111/j.1540-6261.1997.tb03807.x.