

# Section 1: Return Predictability and the Term Structure of Returns

Ralph S.J. Koijen      Stijn Van Nieuwerburgh\*

March 17, 2024

---

\*Koijen: University of Chicago, Booth School of Business, NBER, and CEPR. Van Nieuwerburgh: Columbia Business School, CEPR, and NBER. If you find typos, or have any comments or suggestions, then please let us know via [ralph.koijen@chicagobooth.edu](mailto:ralph.koijen@chicagobooth.edu) or [svnieuwe@gsb.columbia.edu](mailto:svnieuwe@gsb.columbia.edu).

## 1. Basic structure of the notes

- High-level summary of theoretical frameworks to interpret empirical facts.
- Per asset class, we will discuss:
  1. Key empirical facts in terms of prices (unconditional and conditional risk premia) and asset ownership.
  2. Interpret the facts using the theoretical frameworks.
  3. Facts and theories linking financial markets and the real economy.
  4. Active areas of research and some potentially interesting directions for future research.
- The notes cover the following asset classes:
  1. Equities (weeks 1-5).
    - Predictability and the term structure of risk (week 1)
    - Cross-section and the factor zoo (week 2)
    - Intermediary-based asset pricing (week 3)
    - Production-based asset pricing (week 4)
    - Asset pricing via demand systems (week 5)
  2. Mutual Funds and Hedge Funds (week 6).
  3. Options and volatility (week 7).
  4. Government bonds (week 8).
  5. Corporate bonds and CDS (week 9).
  6. Currencies and international finance (week 10).
  7. Commodities (week 11).
  8. Real estate (week 12).

## 2. Stock Return Predictability

### 2.1. The equity premium and stock market volatility

- The average returns on stocks is higher than the returns on short-term nominal bonds.
- Data source: [Ken French](#), using data from CRSP and Bloomberg.
- Annualized estimates of equity risk premium based on monthly returns:

---

1990.7-2021.07	U.S.	N-America	Europe	Asia Pac, ex-Japan	Japan
Mean	9.1	9.0	6.5	8.3	1.5
Stand. dev.	15.0	15.0	17.1	20.1	19.3
Sharpe Ratio	0.60	0.60	0.38	0.41	0.08

---

- The equity premium and Sharpe ratio for the U.S. is robust across samples. For a long sample from 1926.7-2021.7, the equity risk premium in the U.S. is 8.3%. Return volatility is 18.5%. The Sharpe Ratio is 0.45.
- Equity risk premium is similarly large for Europe<sup>1</sup> and Asia Pacific, excluding Japan.
- Japan is a surprising “outlier” with no equity risk premium whatsoever during a 30-year period. How plausible is it that investors were negatively surprised 30 years in a row?

---

<sup>1</sup>Europe includes Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, the Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, and the United Kingdom.

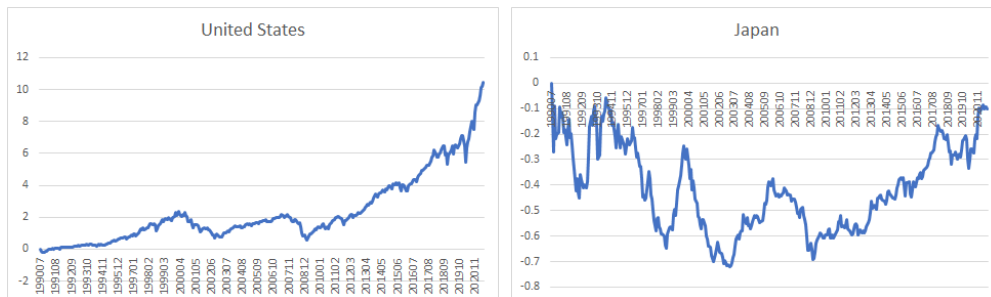
- Equity returns are volatile, which makes it challenging to measure the equity premium precisely. The standard error over the long sample, which contains 95 years of data is  $18.5\%/\sqrt{95} = 2\%$ . Hence a 95%-confidence interval ranges from 4% to 12%!
- [Avdis and Wachter \(2017\)](#) provide unconditional maximum likelihood estimators of the equity risk premium ( $\mu_r$ ) using systems of the form

$$r_{t+1} = \mu_r + \beta(x_t - \mu_x) + \epsilon_{r,t+1},$$

$$x_{t+1} = \mu_x + \phi(x_t - \mu_x) + \epsilon_{x,t+1},$$

Estimates of  $\mu_r$  via this system of equations are more precise when  $\phi$  is high and when the innovations are correlated.

- Obviously, stock markets tend to decline in bad economic times: cumulative excess market return -66% from Oct 2007-Feb 2009, -22% from Jan -March 2020, and in 2022.



## 2.2. Time-series predictability and excess volatility

- **Campbell and Shiller (1988)** develop a log-linear approximation of returns that results in a useful accounting identity to understand the link between stock prices, fundamentals (that is, dividends) and expected returns.
- This relationship starts from the definition of log stock returns:

$$r_{t+1} = \log\left(\frac{P_{t+1} + D_{t+1}}{P_t}\right) = \Delta d_{t+1} - pd_t + \log\left(1 + \frac{P_{t+1}}{D_{t+1}}\right)$$

where  $pd_t = \log(P_t/D_t)$  and  $\Delta d_{t+1} = \log(D_{t+1}/D_t)$ .

- Apply a first-order Taylor approximation to the last term

$$\begin{aligned} \log\left(1 + \frac{P_{t+1}}{D_{t+1}}\right) &\approx \kappa_0 + \kappa_1 pd_{t+1}, \\ \kappa_1 &= \frac{e^{\bar{pd}}}{1 + e^{\bar{pd}}}, \quad \kappa_0 = \log\left(1 + e^{\bar{pd}}\right) - \kappa_1 \bar{pd} \\ r_{t+1} &\approx \kappa_0 + \Delta d_{t+1} + \kappa_1 pd_{t+1} - pd_t \end{aligned}$$

- Iterate forward on this equation to obtain:

$$pd_t = \frac{\kappa_0}{1 - \kappa_1} + \sum_{j=0}^{\infty} \kappa_1^j \Delta d_{t+1+j} - \sum_{j=0}^{\infty} \kappa_1^j r_{t+1+j}.$$

- After imposing the transversality condition, which is a no-bubbles condition

$$\lim_{j \rightarrow \infty} \kappa_1^j E_t[pd_{t+j}] = 0.$$

- As an aside, Giglio, Maggiori, and Stroebel (2016) test the no-bubble condition in housing markets by comparing very long-term (700+ years!) leases and freeholds in the UK and Singapore. They find no evidence of bubbles.

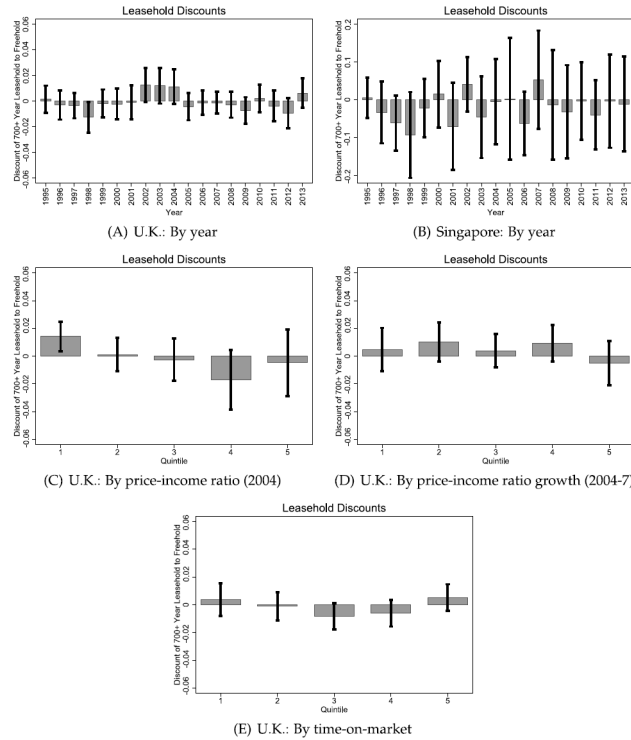


FIGURE 2.—Time-series and cross-section of bubble claim. *Note:* Figure reports estimates of the discount between 700+ year leaseholds and freeholds from regression (3), dividing the sample along time-series and cross-sectional dimensions. Panels A and B show the coefficients of the 700+ leasehold discount year-by-year for the U.K. and Singapore, respectively. Panels C through E report the coefficients of the 700+ leasehold discount, splitting Middle Layer Super Output Areas by quintiles of measures of the potential for a bubble: the price-income ratio in 2004 (Panel C), the growth of the price-income ratio between 2004 and 2007 (Panel D), and the time-on-market (Panel E). These measures of bubble potential are constructed as in columns 4–6 of Table III, Panel B. The bars indicate the 95% confidence interval of the estimate using standard errors clustered at the 3-digit postcode level in the U.K., and at the 5-digit postcode level in Singapore.

- The present-value relationship holds ex-post as well as ex-ante:

$$pd_t = \frac{\kappa_0}{1 - \kappa_1} + E_t \left[ \underbrace{\sum_{j=0}^{\infty} \kappa_1^j \Delta d_{t+1+j}}_{\Delta d_t^H} \right] - E_t \left[ \underbrace{\sum_{j=0}^{\infty} \kappa_1^j r_{t+1+j}}_{r_t^H} \right]. \quad (1)$$

- Hence, movements in prices can be attributed to fluctuations in expected growth rates ( $\Delta d_t^H$ ), expected returns ( $r_t^H$ ), or both.
- Expected discounted future dividend growth rates or returns have to be volatile or they have to be negatively correlated if prices are to be volatile:

$$V[pd_t] = V[\Delta d_t^H] + V[r_t^H] - 2Cov[\Delta d_t^H, r_t^H].$$

- [Shiller \(1981\)](#) provides the first evidence that prices appear to move more than what is implied by expected dividends, even *realized* dividends. This is the celebrated [excess volatility puzzle](#). The classic figure from Shiller's paper:

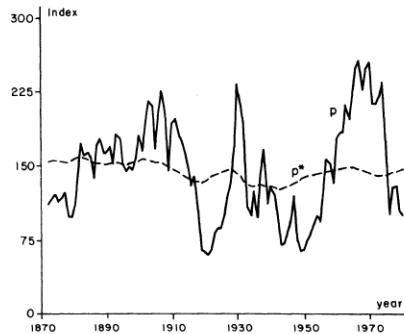


FIGURE 1

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.

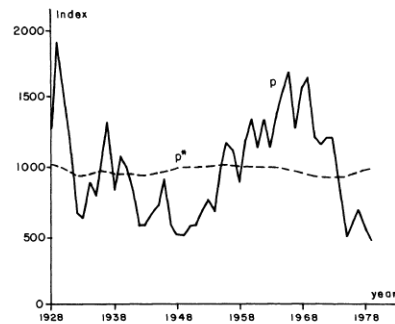


FIGURE 2

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

- As prices are more volatile than realized dividends, equation (1) implies that discount rates must move over time.
- Time-varying expected returns means that returns are predictable. The natural candidate predictor variable is the price-dividend ratio.
- Rewrite (1) in terms of covariances:

$$\begin{aligned}
 V[pd_t] &= Cov[\Delta d_t^H, pd_t] - Cov[r_t^H, pd_t] \\
 1 &= \frac{Cov[\Delta d_t^H, pd_t]}{V[pd_t]} - \frac{Cov[r_t^H, pd_t]}{V[pd_t]}
 \end{aligned}$$

- First term is the slope of a regression predicting future dividend growth rates with  $pd_t$
- Second term is the slope of a regression predicting future returns with  $pd_t$
- There is an adding-up constraint on the two long-horizon predictability slope coefficients
- The dog that did not bark (Lettau and Van Nieuwerburgh, 2008 and Cochrane, 2008)



### 2.3. Empirical Evidence

- Typical empirical framework:

$$\Delta d_{t+1} = a_d + \kappa_d dp_t + e_{d,t+1}, \quad (2)$$

$$r_{t+1} = a_r + \kappa_r dp_t + e_{r,t+1}, \quad (3)$$

$$dp_{t+1} = a_{dp} + \phi dp_t + e_{pd,t+1}, \quad (4)$$

where the present-value identity implies a coefficient restriction  $1 - \kappa_1 \phi = \kappa_r - \kappa_d$

- Summary of the evidence ([Kojien and Van Nieuwerburgh, 2011](#))

Panel A: Return Predictability						
	Div. Reinv. at $R^f$			Div. Reinv. at $R^m$		
	$\kappa_r$	$t - stat$	$R^2$	$\kappa_r$	$t - stat$	$R^2$
1926-2009	0.077	1.31	2.90	0.104	2.08	4.82
1945-2009	0.130	2.56	10.84	0.126	2.58	10.02
Panel B: Dividend Growth Predictability						
	Div. Reinv. at $R^f$			Div. Reinv. at $R^m$		
	$\kappa_d$	$t - stat$	$R^2$	$\kappa_d$	$t - stat$	$R^2$
1926-2009	-0.078	-1.48	7.64	0.008	0.20	0.05
1945-2009	0.017	0.68	1.13	0.044	1.10	2.03

Source: Kojien and Van Nieuwerburgh (2011), Table 1

- Findings:
  - Evidence of return predictability in the post-war sample period, but weaker before the second world war.
  - The reinvestment strategy of dividends during the year matters ([Binsbergen and Kojien, 2010](#)).
  - Dividend growth is predictable by the price-dividend ratio before the second world war, not thereafter. Potential explanation: changes in dividend smoothing ([Chen, 2009](#)).

See also historical evidence from Amsterdam and London stock exchanges (Golez and Koudijs, 2018).

- Return predictability tends to be stronger at longer horizons, see Cochrane (2011):

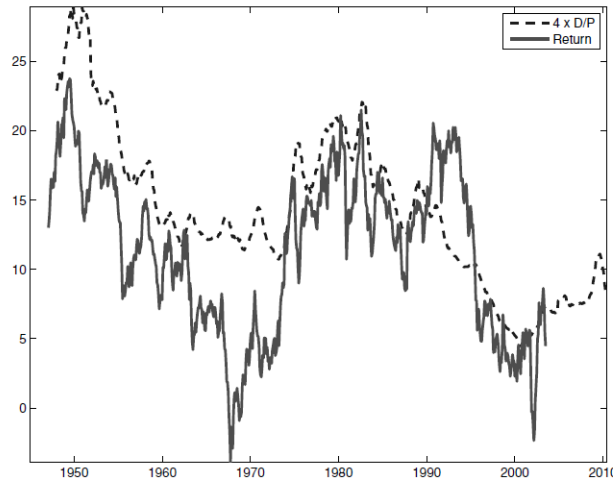


Figure 1. Dividend yield and following 7-year return. The dividend yield is multiplied by four. Both series use the CRSP value-weighted market index.

- Stock return predictability literature can be divided into:
  1. Better statistical methods to infer expected returns or expected dividend growth rates given the persistence of the  $pd$  ratio, see for instance
    - Structural breaks (Lettau and Van Nieuwerburgh, 2008).
    - Filtering methods (Binsbergen and Koijen, 2010).
    - Near-unit root inference (Campbell and Yogo, 2006)
  2. Use additional variables besides  $pd_t$  to predict returns, see for instance
    - CAY (Lettau and Ludvigson, 2001).
    - The cross-section of valuation ratios (Kelly and Pruitt, 2013).

- The variance risk premium (Bollerslev and Zhou, 2009). More on this later.
  - Many more predictors have been proposed, the predictive qualities of many of which were called into question by Goyal and Welch (2008) and (2021).
- Lettau and Van Nieuwerburgh (2008): Break-adjusting  $dp$  strengthens evidence for return predictability considerably, but also the evidence for dividend growth predictability

Panel A: Return Predictability						
	Div. Reinv. at $R^f$			Div. Reinv. at $R^m$		
	$\kappa_r$	$t - stat$	$R^2$	$\kappa_r$	$t - stat$	$R^2$
1926-2009	0.212	2.32	6.20	0.393	4.29	14.91
1945-2009	0.322	4.47	17.25	0.357	4.17	17.72

Panel B: Dividend Growth Predictability						
	Div. Reinv. at $R^f$			Div. Reinv. at $R^m$		
	$\kappa_d$	$t - stat$	$R^2$	$\kappa_d$	$t - stat$	$R^2$
1926-2009	-0.240	-2.53	20.52	0.107	1.37	2.15
1945-2009	-0.021	-0.33	0.42	0.133	1.86	4.08

Source: Kojien and Van Nieuwerburgh (2011), Table 2

- Clarifies that a permanent or at least highly persistent component in the pd ratio obscures return and dividend growth predictability.
- This is useful input for theoretical asset pricing models which must possess both return and dividend growth predictability.

## 2.4. Extracting expected returns and dividend growth rates

### 2.4.1. Gaussian Setting

- Follows [Binsbergen and Kojen \(2010\)](#).
- Rather than pre-specifying that a variable  $x_t$  predicts returns or dividend growth, we can model expected returns ( $\mu_t$ ) and expected growth ( $g_t$ ) rates as [latent variables](#).
- The assumptions are about the time-series dynamics, which we assume to be an AR(1) for both

$$\begin{aligned}\mu_{t+1} &= \delta_0 + \delta_1(\mu_t - \delta_0) + \epsilon_{t+1}^\mu, \\ g_{t+1} &= \gamma_0 + \gamma_1(g_t - \gamma_0) + \epsilon_{t+1}^g,\end{aligned}$$

combined with the model for realized dividend growth

$$\Delta d_{t+1} = g_t + \epsilon_{t+1}^d.$$

- We assume that the shocks are normally distributed

$$\epsilon_t \equiv (\epsilon_t^d, \epsilon_t^g, \epsilon_t^\mu)' \sim N(0, \Sigma).$$

- The log price-dividend ratio as implied by the Campbell and Shiller identity.

$$\begin{aligned}pd_t &= \frac{\kappa_0}{1 - \kappa_1} + \sum_{s=1}^{\infty} \kappa_1^{s-1} E_t[\Delta d_{t+s}] - \sum_{s=0}^{\infty} \kappa_1^{s-1} E_t[r_{t+s}] \\ &= A - B_1(\mu_t - \delta_0) + B_2(g_t - \gamma_0),\end{aligned}$$

where  $A = \kappa_0(1 - \kappa_1)^{-1} + (\gamma_0 - \delta_0)(1 - \kappa_1)^{-1}$ ,  $B_1 = (1 - \delta_1\kappa_1)^{-1}$ , and  $B_2 = (1 - \gamma_1\kappa_1)^{-1}$ .

- **Note 1:** If expected returns and expected growth are an AR(1), then the price-dividend ratio is an AR(1) if only if expected returns and expected growth rates are *equally* persistent, that is,  $\delta_1 = \gamma_1$ .
- **Note 2:** The equation for the price-dividend ratio has no error in it. This means that instead of having two latent variables, we only have one.
- Denoting the demeaned expected growth rate of dividends by  $\hat{g}_t = g_t - \gamma_0$ , we arrive at the final system

$$\begin{aligned}\Delta d_{t+1} &= \gamma_0 + \hat{g}_t + \epsilon_{t+1}^d, \\ pd_{t+1} &= (1 - \delta_1)A + B_2(\gamma_1 - \delta_1)\hat{g}_t + \delta_1 pd_t - B_1\epsilon_{t+1}^\mu + B_2\epsilon_{t+1}^g, \\ \hat{g}_{t+1} &= \gamma_1\hat{g}_t + \epsilon_{t+1}^g.\end{aligned}$$

The first two equations are **measurement equations**. The third equation is the **transition equation** of the latent variable.

- One can estimate this model via maximum likelihood, using the Kalman filter to construct the likelihood. The appendix contains the derivations.
- The Kalman filter effectively introduces moving average terms of returns and dividend growth rates to predict future returns and future dividend growth rates.

- Estimation results:

	benchmark 1926-2009	benchmark 1945-2009	break-adjusted 1926-2009	break-adjusted 1945-2009
AC exp ret	0.93	0.92	0.66	0.64
AC exp div gr	0.26	0.38	0.29	0.35
<i>Std[exp ret]</i>	4.2%	4.6%	7.8%	8.5%
<i>Std[exp div gr]</i>	12.2%	6.9%	12.3%	6.8%
$R^2$ returns	3.0%	9.1%	6.7%	14.1%
$R^2$ div gr	46.8%	18.9%	46.5%	19.9%
% <i>DR</i>	93%	103%	79%	107%
% <i>CF</i>	13%	7%	50%	22%
$-2Cov(CF, DR)$	-6%	-10%	-29%	-30%

Source: Kojien and Van Nieuwerburgh (2011), Table 3

- Notice the much higher persistence in expected returns than in expected dividend growth rates
- Also notice that dividend growth rates are strongly predictable (but not by the pd ratio as we saw earlier)
- **Most of the variation in the pd ratio comes from discount rates** (see also [Cochrane, 2011](#))

## 2.4.2. Beyond the Kalman Filter

- We need a linear-Normal model to apply the Kalman filter.
- In non-linear or non-Gaussian models, the updating steps are not always known analytically.
- However, there has been a lot of work on non-linear filters

– Fast and simple non-linear filters:

- \* **Extended Kalman filter**: The conditional mean can be a non-linear function, but the innovations are additive and normally distributed, e.g.,

$$X_t = h(X_{t-1}) + \epsilon_t.$$

- \* **Unscented Kalman filter**: The model can be fully non-linear and numerical integration is done using Gaussian quadrature. For a “finance-oriented” introduction, see [Zoeter, Ypma, and Heskes \(2004\)](#).
- General approach, but numerically much more challenging is through [particle filtering](#), see for introductions the lecture notes by [Jesus Fernandez-Villaverde](#) and for a more formal treatment, see [Doucet, de Freitas, and Gordon \(2001\)](#). For an application to estimating dynamic stochastic general equilibrium models, see [Fernandez-Villaverde and Rubio-Ramirez \(2007\)](#).

## 2.5. Frequencies in expected returns

- The expected returns extracted as above are highly persistent; they move at generational frequencies
- Alternative methods and additional data tend to uncover a business-cycle frequency in expected returns. From [Cochrane \(2011\)](#):

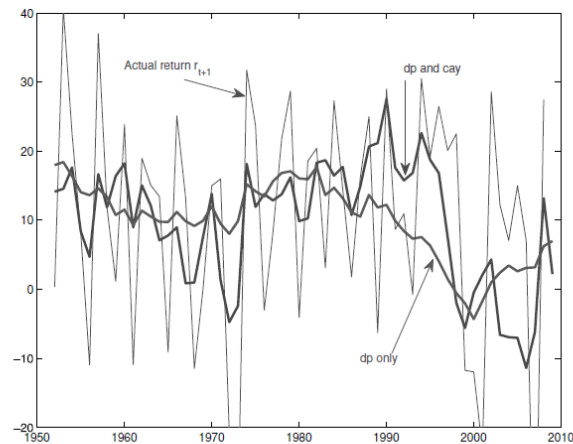


Figure 3. Forecast and actual 1-year returns. The forecasts are fitted values of regressions of returns on dividend yield and  $cay$ . Actual returns  $r_{t+1}$  are plotted on the same date as their forecast,  $a + b \times dp_t$ .

- Hence, the persistence in the price-dividend ratio suggests a highly persistent component. CAY from Lettau and Ludvigson or the cross-section of valuation ratios from Kelly and Pruitt point to a higher-frequency component.
- Evidence from the variance risk premium points to predictability that disappears after weeks or months, rather than years or decades. This is a third frequency component in expected returns.



## 2.6. *Econometric issues in return predictability*

- A large econometric literature is concerned with correct inference as many variables, including the price-dividend ratio, are highly persistent:
  - Bias and correct test statistics if predictors are persistent (Mankiw and Shapiro (1986), Stambaugh (1999), Lewellen (2004), and Campbell and Yogo (2006)).
  - Correct inference in case of long-horizon regressions (Boudoukh, Richardson, and Whitelaw, 2008).
  - Poor out-of-sample performance (Goyal and Welch, 2008 and Ferreira and Santa-Clara, 2011).
- In response to Goyal and Welch (2008), it is common practice to include a section on the out-of-sample predictability of a new predictor variable or a new method.
- However, we are repeatedly studying the same out-of-sample period, which turns out-of-sample into in-sample tests again.

- Illustration of the [Mankiw-Shapiro / Stambaugh bias](#) (omitting means)

$$\begin{aligned} r_{t+1} &= \beta dp_t + \epsilon_{t+1}, \\ dp_{t+1} &= \phi dp_t + u_{t+1}. \end{aligned}$$

In this system,  $dp_t$  is highly persistent ( $\phi \simeq 1$ ),  $\beta > 0$ , and  $Cov(\epsilon_{t+1}, u_{t+1}) < 0$  (why?).

- In small samples,  $\hat{\phi}$  tends to be downward biased (standard issue in OLS).
- This implies for the bias in the predictive coefficient,  $\beta$

$$E(\hat{\beta} - \beta) = \frac{Cov(\epsilon_{t+1}, u_{t+1})}{Var(u_{t+1})} E(\hat{\phi} - \phi).$$

- Hence,  $\hat{\beta}$  is upward biased, which means that we reject the null of no predictability too often.
- The upward bias is larger when (i) the predictor is more persistent and (ii) the innovations of the predictor and returns are more negatively correlated.
- This problem arises in other areas of (financial) economics as well and is just a basic property of VAR models.

- Jiang, Lustig, Van Nieuwerburgh, and Xiaolan (2022) apply the insights of the equity return predictability to the portfolio of outstanding government debt.
  - Campbell-Shiller decomposition of debt/output ratio into:
    - \* expected future cash flows on the government bond portfolio, primary surpluses, over the next  $T$  years,
    - \* expected future real, growth-adjusted returns on the government bond portfolio over the next  $T$  years
    - \* the debt/output ratio  $T$  years from now
  - The debt/output ratio plays the role of the price-dividend ratio; it is highly persistent.
  - Its innovations are negatively correlated with the innovations in fiscal cash flows (surplus) and positively with the innovations in government bond return
  - So there is a severe Stambaugh bias in both cash-flow and return predictability regressions.
  - After correcting for the bias, there is evidence for neither cash flow nor return predictability: the dogs that did not bark.
  - Instead, the debt/output ratio at  $t$  forecasts the debt/output ratio at  $t + T$ .
  - This is evidence of fiscal unsustainability.
  - This lack of predictability is to a large extent due to the presence of a permanent component in the debt/output ratio; removing it restores some role for cash flow growth predictability.

## 2.7. Expectations and information sets

- We often write  $E_t(\cdot)$  in the equations so far.
- But whose expectations do we measure?
- Standard assumption in empirical asset pricing: Investors know more than the econometrician and we can apply the law of iterated expectations.

$$E_t(M_{t+1}R_{t+1}^e) = 0 \quad \Rightarrow \quad E(M_{t+1}R_{t+1}^e) = 0.$$

In many cases, conditioning down solves the problem of testing models as long as we assume that we condition on a smaller information set than the information set of investors.

- Alternatively, we use [survey expectations](#) to predict future returns.
- Survey expectations exist for households, CFOs, analysts, ...
- Data sources:
  - Gallup: Individual investors.
  - Graham-Harvey: CFOs.
  - American Association of Individual Investors.
  - Investor Intelligence: Summary of newsletters.
  - Shiller: Individual investors.
  - Michigan Survey Research Center: Consumers.
  - [New York Fed Survey of Consumer Expectations](#)

- Greenwood and Shleifer (2014) suggest that there is quite some co-movement between different surveys of returns expectations. The average correlation is 43%.

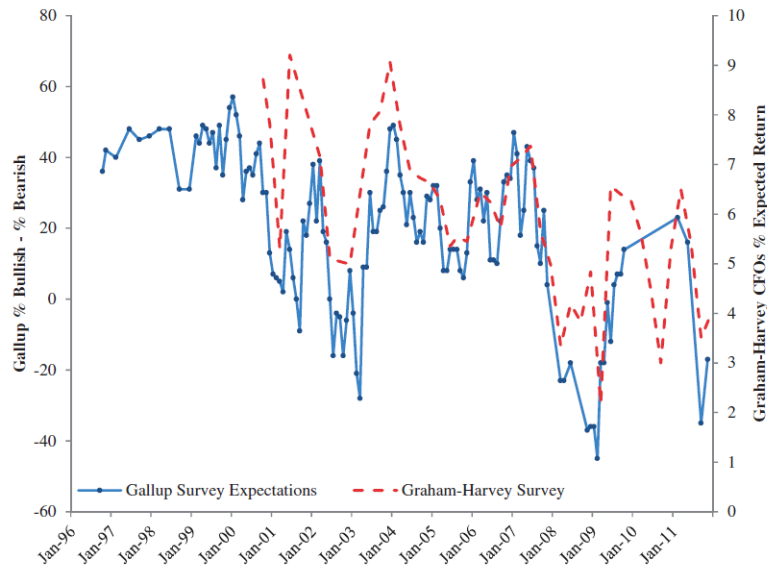
Table 2  
Correlations between different measures of investor expectations

	Gallup (N = 135)	Graham- Harvey (N = 42)	American Association (N = 294)	Investor Intelligence (N = 588)	Shiller (N = 132)	Michigan (N = 22)	Index (N = 294)
Graham-Harvey	0.77 [0.000]						
American Association	0.64 [0.000]	0.56 [0.000]					
Investor Intelligence	0.60 [0.000]	0.64 [0.000]	0.55 [0.000]				
Shiller	0.39 [0.000]	0.66 [0.000]	0.51 [0.000]	0.43 [0.000]			
Michigan	0.61 [0.003]	-0.12 [0.922]	0.60 [0.003]	0.19 [0.395]	-0.55 [0.020]		
Expectations Index	0.87 [0.000]	0.58 [0.000]	0.87 [0.000]	0.81 [0.000]	0.52 [0.000]	0.55 [0.008]	
Fund flow	0.69 [0.000]	0.71 [0.000]	0.42 [0.000]	0.20 [0.002]	0.51 [0.001]	0.40 [0.068]	0.45 [0.000]

This table shows partial correlation coefficients; that is, it uses the full sample of overlapping data for each series. The Expectations Index combines data in the Gallup, American Association, and Investor Intelligence series. Numbers in brackets denote  $p$ -values on the hypothesis that the correlation between the two series is zero.

- **Striking fact:** Survey expectations of returns are *low in bad times*.
- This is inconsistent with most (rational) theories of asset pricing.

- Overview of the evidence in Greenwood and Shleifer (2014):



**Figure 2**  
**Comparing the Gallup survey with Graham-Harvey CFO expectations**  
 The main Gallup series, marked with a solid line (left axis), is computed as the fraction of investors who are bullish (optimistic or very optimistic) minus the fraction of investors who are bearish. The dashed line denotes forecasts of nominal returns made by CFOs in John Graham and Campbell Harvey's quarterly surveys (right axis).

- Potential explanations
  1. Investors confound fundamentals and prices (= do not understand that discount rates fluctuate a lot).
  2. Investors extrapolate returns.
- Importantly, incorrect expectations of a group of investors can be a source of excess volatility.

- In recent work, [Wang \(2021\)](#) argues that there are actually two clusters of agents:
  - the ones identified by Greenwood and Shleifer: CFOs and retail investors
  - a second group: buy-side and sell-side **analysts**
- Turns out that the second group has subjective expectations of stock returns that are contrarian (high in recessions), and hence negatively correlated with the beliefs of the first group of agents (who are like momentum traders).
- He shows that *both* sets of beliefs can be reconciled with the data on returns, dividend growth rates, and pd ratios using separate Kalman Filter estimation for each group (rational learning).
- He estimates different values of **correlation between the subjective expected returns and expected growth rates**: the first group has a positive correlation while the second group has a negative correlation.

### 3. Term Structure of Risk and Returns

#### 3.1. What is it and why do we care?

- **Definition:** The **term structure of returns** refers to returns on assets with the same underlying cash flows, where the return is measured over the *same holding period*, but for *different maturities*.
- E.g., the 1-month return on a 3-year and a 5-year Treasury bond.
- We will see evidence for Treasuries, corporate bonds, variance swaps, and housing later in the course. We now discuss evidence from equity markets.
- Why do we care?
  - Expected returns and risk over different horizons are important for real and financial investment decisions.
  - Short-maturity asset prices are informative about future growth, even in the presence of the ZLB.
  - Informative about the cross-section of expected returns.
  - Powerful test of theoretical asset pricing models.



- We focus on the term structure of equity returns, and will re-visit this topic later when we discuss other asset classes.
- To fix ideas, it is useful to start from the dividend discount model.
- The price of a stock or equity index  $S_t$  is given by the discounted value of its dividends  $D_t$ :

$$S_t = \sum_{n=1}^{\infty} E_t (M_{t:t+n} D_{t+n}),$$

$M_{t:t+n} = \prod_{j=1}^n M_{t+j}$  is the product of one-period stochastic discount factors

- Alternative notation:

$$S_t = \sum_{n=1}^{\infty} \frac{E_t (D_{t+n})}{(1 + \mu_{t,n})^n}$$

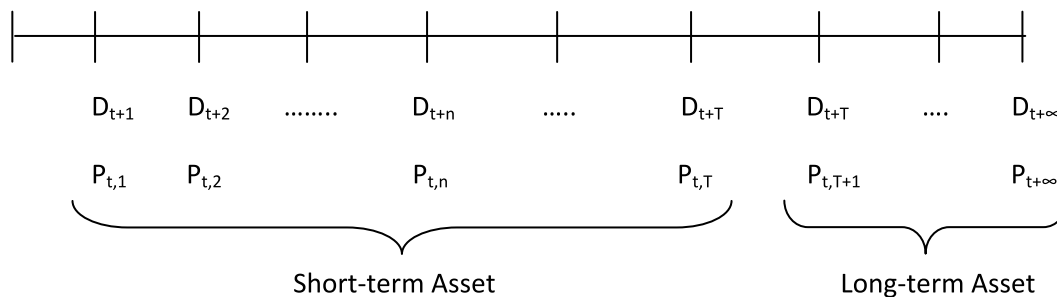
$\mu_{t,n}$  is appropriate per-period discount rate for period  $t + n$ .

- Decompose the stock index as:

$$\begin{aligned}
 S_t &= \sum_{n=1}^{\infty} E_t (M_{t:t+n} D_{t+n}) \\
 &= \underbrace{\sum_{n=1}^T E_t (M_{t:t+n} D_{t+n})}_{\text{Short-term asset}} + \underbrace{\sum_{i=T+1}^{\infty} E_t (M_{t:t+n} D_{t+n})}_{\text{Long-term asset}}.
 \end{aligned}$$

- We call  $P_{t,n} = E_t (M_{t:t+n} D_{t+n})$  the price of the  $n^{\text{th}}$  dividend strip, see [Brennan \(1998\)](#). The equity index price is the sum of all strip prices (value additivity):

$$S_t = \sum_{n=1}^{\infty} P_{t,n}.$$



- Properties of the aggregate stock market that have been challenging to explain:
  1. Equity premium puzzle
  2. Excess volatility puzzle
  3. Return predictability
- We want to “strip” down the stock market index and study the pricing of “short-term” and “long-term” dividend payments.
- **Big picture question:**  
Are facts (1) - (3) a “long-term” or a “short-term” phenomenon?
- What do leading macro-finance models predict regarding the term structure of equity returns?

- Let's start with the basic consumption CAPM.

- Preferences:

$$\max \sum_{s=0}^{\infty} E_t (\beta^s u(C_{t+s})),$$

where  $u(x) = x^{1-\gamma}/(1-\gamma)$ .

- Consumption growth is assumed to be i.i.d.

$$\Delta C_{t+1} = \mu_c + \sigma_c \epsilon_{c,t+1}.$$

- The price of dividend strips in this case is given by:

$$P_{t,n} = E_t (M_{t:t+n} D_{t+n}) = \phi_n D_t,$$

where  $M_{t:t+n} = \beta^n (C_{t+n}/C_t)^{-\gamma}$  denotes the  $n$ -period stochastic discount factor and  $\phi_n$  a constant that depends on maturity.

- The expected geometric return for strips of all maturities is constant.
- In the most basic consumption CAPM, the term structure of risk premia and volatility is **constant across maturities**.
- However, this model fails to reproduce the **level and volatility of both the risk-free rate and the equity risk premium**.

- Models that are successful at matching moments of the risk-free rate and the equity risk premium:
  - Campbell and Cochrane (1999) external habit formation model.
  - Bansal and Yaron (2004) long-run risk model.
  - Gabaix (2012) and Wachter (2014) variable rare disasters model.
- Let's use the external habit model to illustrate the main predictions.
- In this model, the only modification relative to the consumption CAPM are the preferences.
- The stochastic discount factor changes to:

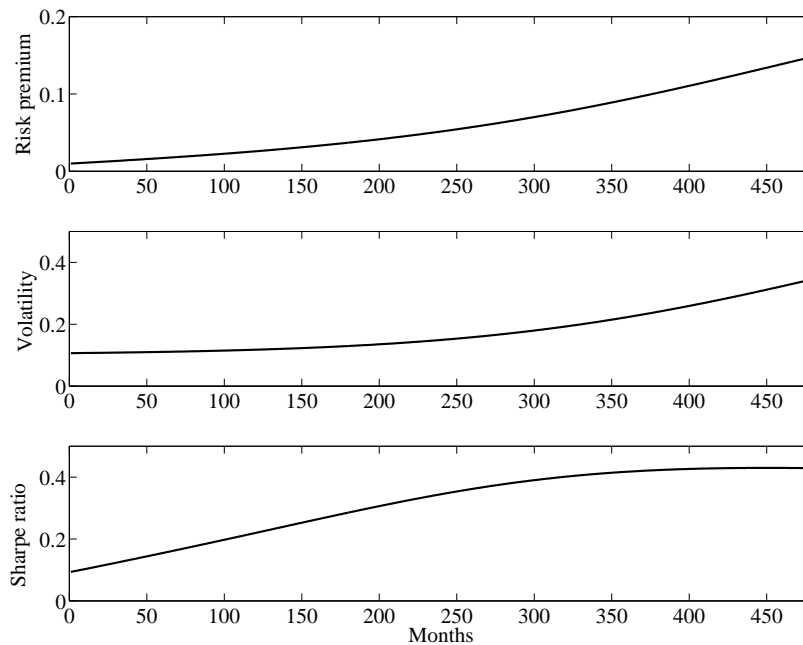
$$M_{t+1} = \beta e^{-\gamma \mu_c} e^{-\gamma(s_{t+1} - s_t + \epsilon_{c,t+1})},$$

where  $s_t$  denotes the surplus consumption ratio with dynamics:

$$s_{t+1} = (1 - \phi)\bar{s} + \phi s_t + \lambda(s_t)v_{t+1},$$

where  $\lambda(s_t)$  is the sensitivity function which is chosen so that the risk-free rate is constant.

- Here is what the external habit model (standard calibration) implies for the term structure of dividend strips: expected returns (first panel), return volatility (second panel), Sharpe ratio (third panel)



- Overview of theoretical benchmarks:

	Expected returns	Volatility	Sharpe ratios
Data	Down	Down	Down
Campbell and Cochrane (1999)	Up	Up	Up
Bansal and Yaron (2004)	Up	Up	Up
Gabaix (2012)	Flat	Up	Down

- Despite different economic mechanisms, the external habit and LRR model make similar predictions for the term structure of equity.
- In the variable rare disaster model, volatilities still increase with maturity, but expected returns are flat, leading to downward-sloping Sharpe ratios across maturities.

### 3.2. *Extracting the term structure of equity risk using the cross-section of stocks*

- **Intuition:** If different firms have different cash flow structures across maturity, then differences in average returns across stocks are informative about risk premia across maturities.
- **Note:** This is not about differences in *average* growth rates (Chen, 2017), but it is about differences in *risk exposures across maturities*, see also Hansen, Heaton, and Li (2008).
- Differences in average growth rates will generate differences in risk premia only due to the term premium.
- See Cornell (1999), Dechow, Sloan, and Soliman (2004), Bansal, Dittmar, and Lundblad (2005), and Da (2009) for early contributions.
- Weber (2018) is a more recent example. Finds that low-duration stocks outperform high-duration stocks by 1.1% per month, but have lower betas. Favors behavioral explanation.

- Hansen , Heaton, and Li (2008) measure the term structure of expected returns for value and growth firms.
- Large differences in risk premia, for a fixed holding period, on value and growth cash flows at longer horizons (Figure 2.B).
- Solid = Value, Dotted = Market, Dash-dotted = Growth.

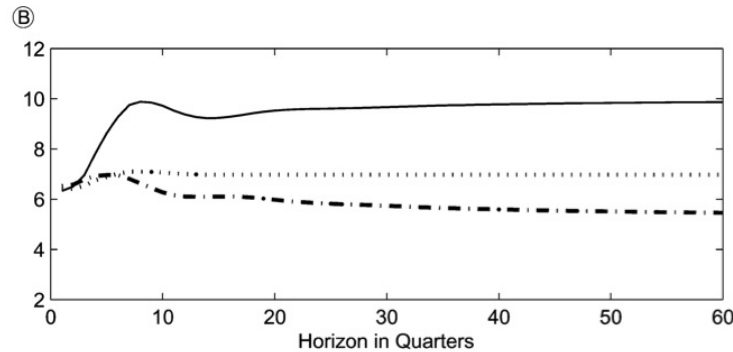


FIG. 2.—A, Expected returns by cash flow horizon. B, Holding period returns. Rates of return are given in annual percentage rates. The expected rates of return in panel A are computed over the investment horizon, and the returns in panel B are one-quarter holding period returns on payoffs at the respective horizons.

- To construct this figure, Hansen, Heaton, and Li combine a statistical model for the dynamics of consumption with recursive preferences to obtain a SDF (i.e., the risk prices).
- Shocks are identified via joint VAR of consumption growth and earnings.
- A similar statistical model for dividends of value and growth portfolios provides the risk exposures of the cash flows.
- By combining risk prices and exposures, they can compute risk premia across horizons.
- **Note:** Interesting variation for value and growth portfolios across horizons, but not for the aggregate stock market.



### 3.3. *Extracting the term structure of equity risk from options*

- Binsbergen, Brandt, and Koijen (2012) use the **put-call parity** relationship for a European option on a dividend paying stock to measure dividend strip prices directly

$$c_{t,T} + Xe^{-r_{t,T}(T-t)} = p_{t,T} + S_t - \mathcal{P}_{t,T},$$

where  $p_{t,T}$  and  $c_{t,T}$  are the prices of a European put and call options at time  $t$ , with maturity  $T$  and strike price  $X$ .

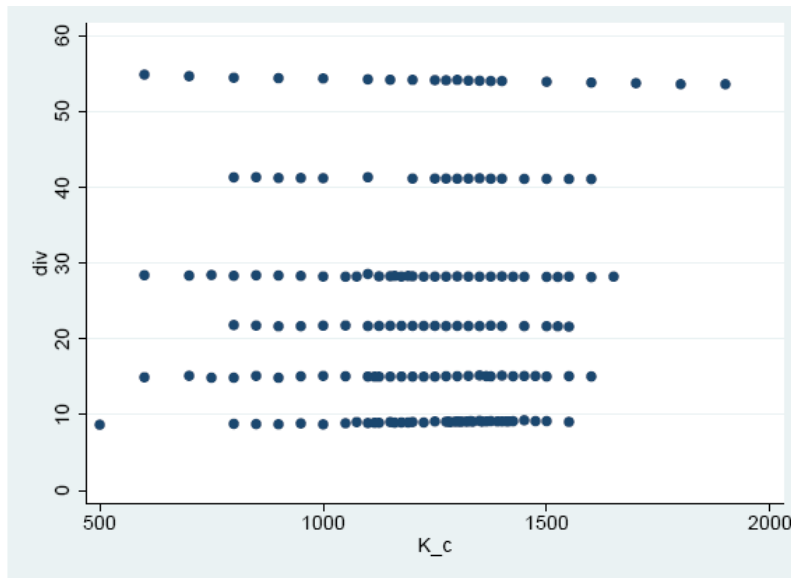
- $\mathcal{P}_{t,T}$  is the price of an asset that pays the dividends on the stock between periods  $t$  and  $t+T$ , i.e., the short-term dividend asset.
- We compute the price of the short-term asset by rearranging the equation above:

$$\mathcal{P}_{t,T} = p_{t,T} - c_{t,T} + S_t - Xe^{-r_{t,T}(T-t)}.$$

- Data set from the CBOE containing TAQ data on S&P 500 index options.
- S&P500 index options are European-style options.
- Index data from Tick Data Inc.
- Futures data from Tick Data Inc.
- Interest rates from Option Metrics based on BBA LIBOR rates.
- Sample period: January 1996-October 2009.

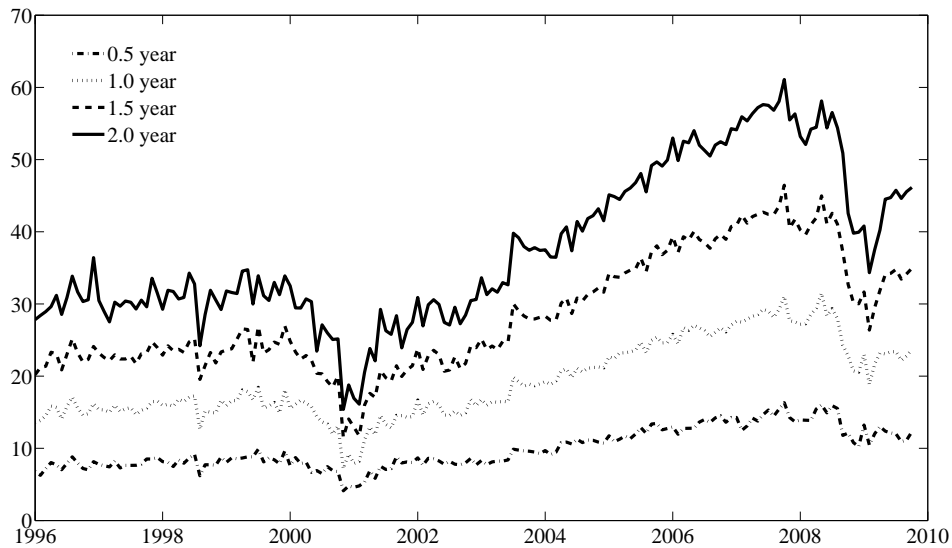
- Selecting the sample:
  - Find pairs of put and call quotes with the same strike and maturity that are closest together in time between 10am and 2pm for the last trading day of each month.
  - Pick the pair with the smallest time difference.
    - ⇒ Typically, many matches within the same second.
  - If multiple matches exist, take the median of all dividend prices for a given maturity.
    - ⇒ Designed to **minimize measurement error** and issues related to **microstructure noise**.
  - Pick the maximum maturity under 2 years and follow it until another contract closer to 2 years is introduced.

- Dividend prices in November 2006. Maturities: 0.31, 0.55, 0.81, 1.06, 1.56, and 2.05 years. S&P Value: 1397.92.

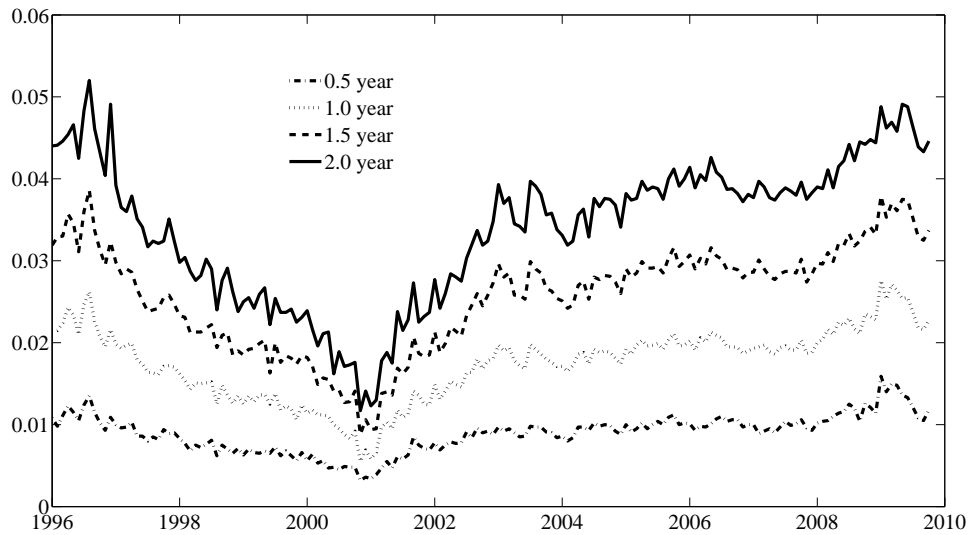


- Note that:
  - If the **wrong interest rate had been used**, the lines would not be flat. Indeed, one can recover the interest rate used in markets by ensuring these lines are flat.
  - If there had been a **lot of microstructure noise or liquidity effects**, the “lines” would be “clouds” instead.

- Here is what the procedure implies for the dividend strip prices:

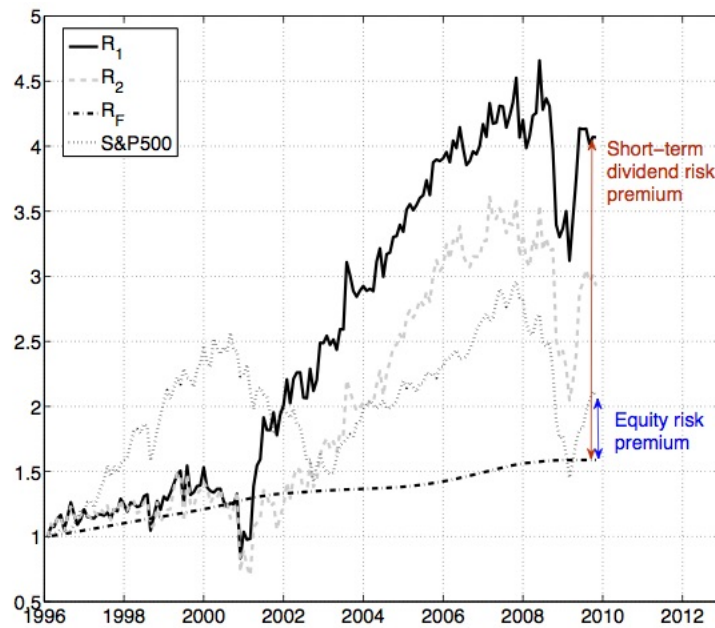


- Cumulative dividend prices *as a share* of the index:



- The first two years of dividends represent about 4% of the total index value. Much less in 2001: recession expected to be short.

- Two dividend strategies:
  - Buy two years of dividends ( $R_{1,t}$ ).
  - Buy two years of dividends and sell the first six months ( $R_{2,t}$ ).
- The second strategy is tax neutral and hence dividend taxation does not explain these results.



**Figure 1. Cumulative returns of dividend strategies, the S&P500, and 30-day T-bills.**

The figure shows the cumulative returns of investing \$1 in January 1996 until October 2009 for four strategies: (i) dividend strategy 1 ( $R_1$ ), dividend strategy 2 ( $R_2$ ), (iii) the S&P500, and (iv) 30-day T-bills ( $R_F$ ).

- Summary of results:

TABLE 1—SUMMARY STATISTICS

	$R_{1,t}$	$R_{1,t} - R_{f,t}$	$R_{2,t}$	$R_{2,t} - R_{f,t}$	$R_{SP500,t}$	$R_{SP500,t} - R_{f,t}$
Mean	0.0116 (0.0044)	0.0088 (0.0044)	0.0112 (0.0044)	0.0084 (0.0045)	0.0056 (0.0047)	0.0027 (0.0047)
Standard deviation	0.0780 (0.0136)	0.0781 (0.0136)	0.0965 (0.0171)	0.0966 (0.0171)	0.0469 (0.0050)	0.0468 (0.0050)
Sharpe ratio	0.1124 (0.0520)	— —	0.0872 (0.0494)	— —	0.0586 (0.1058)	— —
Observations	165	165	165	165	165	165

*Notes:* The table presents descriptive statistics of the monthly returns on the two trading strategies described in the main text. Block bootstrapped standard errors (blocks of 15 observations) of each of the moments are in parentheses. Sample period is February 1996 through October 2010.

- Three puzzling findings compared to the benchmark models:
  1. Average risk premia of short-maturity assets are large and positive, while theoretical benchmarks predict near-zero risk premia.
  2. High volatility of short-maturity assets.
  3. Sharpe ratios decline with maturity.
- **Note:** Because dividend strips are volatile, the risk premium estimates based on this sample are insignificant or borderline significant.

- What matters is the comparison to the S&P500.

TABLE 5—MONTHLY RETURNS ON THE TWO TRADING STRATEGIES AND THE THREE S&P 500 FACTORS

Dependent variable	$R_{1,t+1} - R_{f,t}$	$R_{2,t+1} - R_{f,t}$
<i>c</i>	0.0061 (0.0047)	0.0066 (0.0056)
sp500rf	0.3972 (0.1824)	0.4137 (0.2058)
hml-sp500	0.1526 (0.1752)	0.5668 (0.1994)
smb-sp500	0.3043 (0.3117)	-0.0528 (0.3614)
$R^2$	0.1000	0.1011

*Notes:* The table presents OLS regressions of the returns on trading strategies 1 and 2 (dependent variables) on the Fama-French three factor model, where the three factors are constructed using S&P 500 firms only. Newey-West standard errors in parentheses.

- Short-maturity assets have a market beta well below one.
- Consistent with the theory of [Lettau and Wachter \(2007\)](#), short-maturity assets have a positive HML beta, although the exposure is small.
- Three-factor alpha is 66bp per month or 8% per year.
- There seems to be excess compensation for bearing short-term risk in the stock market.

- Recall the excess volatility figure of Shiller.
- In Shiller's calculations, one may worry about dividends far out in the future. Using short-maturity assets, there is **direct evidence** of excess volatility.

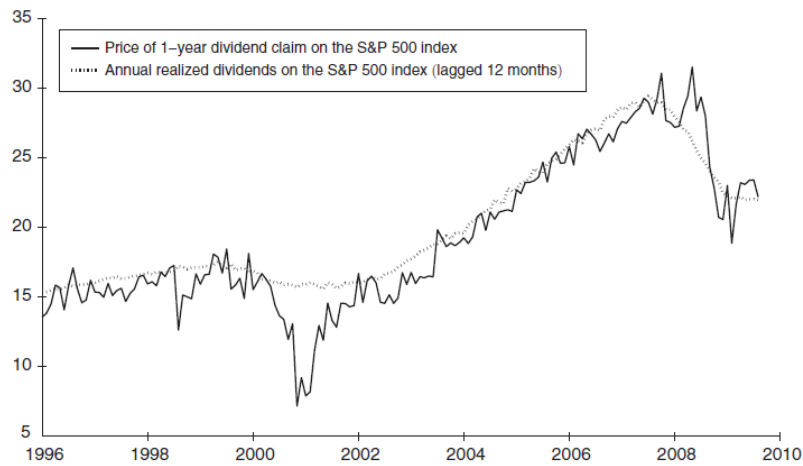


FIGURE 4. PRICES AND REALIZATIONS OF DIVIDEND CLAIMS: 1996:1–2009:10



- Summary so far:
  1. Expected returns and Sharpe ratios on the short-term asset are higher than on the aggregate market, although statistical significance is weak **because of:**
  2. The return volatility of the short-term asset is higher than on the aggregate market.
  3. The beta with respect to the aggregate stock market is 0.5.
  4. The alpha with respect to the aggregate stock market is about 8% per annum.
  5. The prices of short-term dividends are more volatile than their realizations, pointing to excess volatility on the short end of the equity curve.
  6. The returns on the short-term asset are predictable.
- These properties hard to explain using the leading macro-finance models, where long-run risks tends to receive higher compensation than short-term risks.

### 3.4. Extracting the term structure of equity risk from futures

- Instead of using option prices, one can use direct evidence from dividend futures.
- We use dividend futures to define equity yields.
- We start from the price of an  $n$ -period dividend strip (recall Campbell-Shiller):

$$P_{t,n} = D_t \exp(n(g_{t,n} - \mu_{t,n})).$$

- We define the **per-period expected growth rate**  $g_{t,n}$  as:

$$g_{t,n} = \frac{1}{n} E_t \left[ \log \left( \frac{D_{t+n}}{D_t} \right) \right],$$

- We decompose expected returns,  $\mu_{t,n}$ , into a risk premium,  $\theta_{t,n}$ , and a Treasury yield,  $y_{t,n}$ :

$$\mu_{t,n} = \theta_{t,n} + y_{t,n}.$$

- This implies for the price of an  $n$ -period dividend strip:

$$P_{t,n} = D_t \exp(-n(y_{t,n} + \theta_{t,n} - g_{t,n})).$$

- Binsbergen, Hueskes, Koijen, and Vrugt (2013) define the log dividend yield on an equity strip, the **equity yield**, as:

$$e_{t,n} \equiv \frac{1}{n} \log \left( \frac{D_t}{P_{t,n}} \right) = y_{t,n} + \theta_{t,n} - g_{t,n}.$$

- We do not observe the spot price  $P_{t,n}$  but its futures price:

$$F_{t,n} = P_{t,n} \exp (ny_{t,n}).$$

- Define the **forward equity yield** as:

$$e_{t,n}^f \equiv \frac{1}{n} \log \left( \frac{D_t}{F_{t,n}} \right) = \theta_{t,n} - g_{t,n}.$$

- How can you earn the risk premium  $\theta_{t,n}$ ?
- Buy the  $n$ -period futures contract at time  $t$  (known payment at  $t$ , due at  $t + n$ ), hold till maturity  $t + n$ , receive risky realized dividends in period  $t + n$ .
- The  $n$ -period return is:

$$r_{t+n}^D = \log \left( \frac{D_{t+n}}{F_{t,n}} \right) = \log \left( \frac{D_{t+n}}{D_t} \right) + \log \left( \frac{D_t}{F_{t,n}} \right).$$

Because the forward price is known at time  $t$ , but paid at time  $t + n$ , this is a zero-cost strategy, and no money is exchanged at time  $t$ . The expected return on this strategy is given by:

$$E_t [r_{t+n}^D] = n\theta_{t,n}.$$

- So this is a  $n$ -period risk premium, net of the bond risk premium (term premium).

- [van Binsbergen and Koijen \(2017\)](#) use prices of dividend futures with maturities up to 10 years for 2002-2014 from four major regions:
  1. U.S.: S&P500.
  2. Europe: Eurostoxx 50.
  3. Japan: Nikkei 225.
  4. U.K.: FTSE 100.
- Natural players in the market: derivatives desks, pension funds, ...
- Before 2008, these contracts are traded in over-the-counter markets, but exchange-traded products available now.
- Pricing data from Goldman Sachs (to mark their internal trading books to the market). Data verified with the prices from BNP Paribas and the data from exchange-traded options and futures (Bloomberg).

- The one-period return on a futures contract is given by:

$$R_{t,n}^F = F_{t,n-1}/F_{t-1,n} - 1.$$

- Up to a first-order approximation, the return on the index,  $R_t^M$ , can be written as the return on a portfolio of dividend futures returns plus the return on a portfolio of bonds:

$$R_t^M \approx \sum_{n=1}^{\infty} w_{t-1,n} R_{t,n}^F + \sum_{n=1}^{\infty} w_{t-1,n} R_{t,n}^B,$$

where the weights  $w_{t,n}$  are given by  $w_{t,n} = P_{t,n}/S_t$  and  $S_t$  is the index level.

- To compare expected returns on short-term dividend futures to those on the overall stock market, we compute the 10-year-bond-adjusted market return,  $R_{B,t}^M$ , as:

$$R_{B,t}^M \equiv \frac{1 + R_t^M}{1 + R_{t,120}^B} - 1.$$

- Alternatively, we can convert the dividend futures contracts to spot contracts using the cost-of-carry formula:

$$F_{t,n} = P_{t,n} \exp(ny_{t,n}).$$

- Then the no-arbitrage relationship implies that the dividend spot return  $R_{t,n}^S$  can be computed as:

$$R_{t,n}^S = \frac{P_{t,n-1}}{P_{t-1,n}} - 1 = (1 + R_{t,n}^F)(1 + R_{t,n}^B) - 1.$$

- This return can be compared directly to the market return.
- Cumulative performance dividend futures contracts:

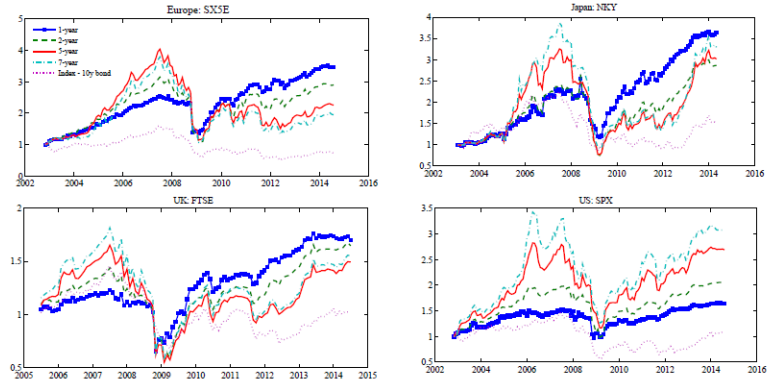


Figure 1: Cumulative Performance Dividend Futures Contracts  
 The four panels in the figure display the cumulative performance of constant maturity dividend futures contracts with maturities of 1,2,5 and 7 years between 2002 and 2014 for the Europe and the US, between 2003 and 2014 for Japan and between 2006 and 2014 for the UK. Dividend futures returns are excess returns in excess of their corresponding constant maturity zero-coupon bond return. As a comparison we also plot the performance of the index in excess of the long-term (10-year) bond return.

Maturity in years	1	2	3	4	5	6	7	Index - 10y bonds
Europe: SX5E (Nov 2002 - Jul 2014)								
mean	0.0101	0.0097	0.0085	0.0082	0.0085	0.0079	0.0075	-0.0004
stdev	0.0458	0.0624	0.0689	0.0708	0.0713	0.0719	0.0721	0.0608
Sharpe	0.2214	0.1556	0.1239	0.1152	0.1200	0.1096	0.1046	-0.0064
Japan: NKY (Jan 2003 - Jul 2014)								
mean	0.0112	0.0107	0.0107	0.0115	0.0115	0.0119	0.0124	0.0048
stdev	0.0521	0.0701	0.0738	0.0772	0.0771	0.0809	0.0821	0.0614
Sharpe	0.2160	0.1526	0.1450	0.1489	0.1490	0.1473	0.1511	0.0781
UK: FTSE (Jan 2005 - Jul 2014)								
mean	0.0061	0.0065	0.0062	0.0063	0.0065	0.0067	0.0072	0.0015
stdev	0.0499	0.0610	0.0675	0.0703	0.0729	0.0746	0.0775	0.0505
Sharpe	0.1222	0.1059	0.0916	0.0893	0.0894	0.0900	0.0933	0.0289
US: SPX (Nov 2002 - Jul 2014)								
mean	0.0041	0.0059	0.0067	0.0072	0.0084	0.0090	0.0095	0.0020
stdev	0.0337	0.0405	0.0482	0.0496	0.0513	0.0544	0.0559	0.0553
Sharpe	0.1224	0.1464	0.1390	0.1453	0.1633	0.1645	0.1707	0.0368

Table 1: Properties of Dividend Strip Returns  
 We summarize the average return, standard deviation and Sharpe ratio of dividend strip returns across maturities of 1-7 years, and regions: Europe, Japan, the United Kingdom, and the United States.

- International evidence on CAPM betas across maturities:

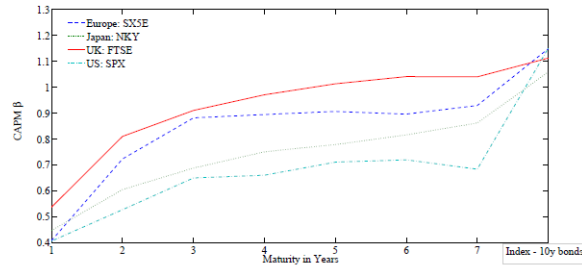


Figure 4: CAPM Betas  
The graph plots the unconditional CAPM betas for dividend futures for all 7 maturities, as well as for the strategy that goes long in the index and short in the 10-year bond. Each line represents one of the four regions.

- International evidence on excess volatility:

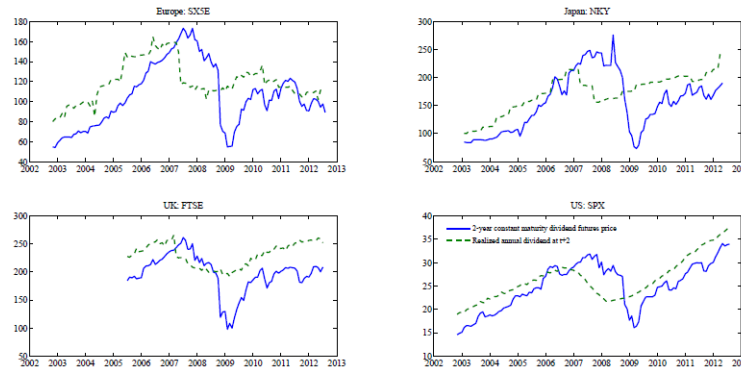


Figure 3: Excess Volatility  
The four panels in the figure display the 2-year constant maturity dividend futures price and plots it against the realized annual dividend two years later for all four regions.

- Short-maturity assets have *significantly* higher returns than the market once we form international portfolios.
- One obtains more powerful statistical test as a result of international diversification.

Asset class	Sharpe ratios		Test statistic	Sharpe ratios		Test statistic
	Short	Long		Short	Long	
	Dividend					
Equity	futures	Index		1y	7y	
Europe	0.50	-0.02	2.27	0.77	0.36	1.78
Japan	0.55	0.27	1.07	0.75	0.52	1.12
UK	0.35	0.10	1.37	0.42	0.32	0.60
US	0.54	0.13	1.74	0.42	0.59	-0.76
Global	0.62	0.12	2.00	0.81	0.58	1.05



- Forward equity yields not only predict returns on dividend futures but also dividend growth

$$\Delta d_{t+1} = \alpha_n - \beta_n e_{t,n}^f + \epsilon_{t,n}.$$

**Table 3**

Dividend growth predictability.

The table reports the predictive coefficient ( $\beta_n$ ), the  $t$ -statistic, and the  $R$ -squared value of a predictive regression of annual dividend growth using monthly overlapping data by forward equity yields of maturities 1, ..., 5 years using univariate regressions with one forward equity yield on the right-hand side of the regression, see Eq. (10). The  $t$ -statistics are computed using Hodrick (1992) standard errors. The sample period is from October 2002 until March 2011 for the US and Europe, and from January 2003 until March 2011 for Japan.

Index		S&P 500			
Maturity	$\beta$	OLS $t$ -statistic	Hodrick $t$ -statistic	$R$ -squared	
1	0.84	20.16	9.60	80%	
2	1.13	18.23	6.59	77%	
3	1.47	14.02	5.63	66%	
4	1.76	12.18	4.86	60%	
5	1.97	10.82	4.20	54%	
Index		EuroStoxx 50			
Maturity	$\beta$	OLS $t$ -statistic	Hodrick $t$ -statistic	$R$ -squared	
1	0.88	17.43	10.29	75%	
2	0.99	15.00	7.22	69%	
3	1.40	14.28	6.93	67%	
4	1.83	13.20	6.41	64%	
5	2.24	12.33	6.00	60%	
Index		Nikkei 225			
Maturity	$\beta$	OLS $t$ -statistic	Hodrick $t$ -statistic	$R$ -squared	
1	0.545	12.82	4.33	63%	
2	0.720	12.71	4.84	62%	
3	0.979	12.28	4.83	64%	
4	1.220	11.90	4.64	59%	
5	1.433	11.42	4.43	57%	

- Forward equity yields not only predict dividend growth but also economic growth more broadly, such as consumption growth

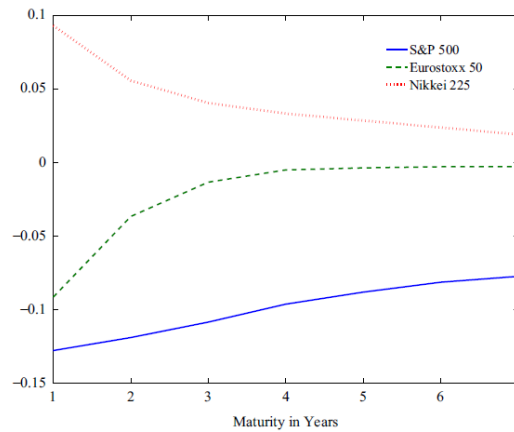
**Table 8**

Predictability of annual consumption growth.

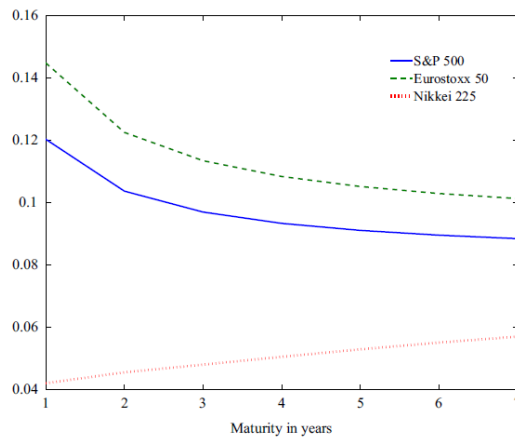
The table reports the results of a predictive regression at a quarterly frequency of US consumption growth by S&P 500 forward equity yields (Panel A), US nominal bond yields (Panel B), and US real bond yields (Panel C) in Eq. (22). The table reports the predictive coefficient ( $\beta_n$ ), the  $t$ -statistic, and the  $R$ -squared value. The  $t$ -statistics are computed using Hodrick (1992) standard errors. The sample period is from December 2002 and March 2011.

Panel A: Consumption growth predictability by equity yields			
Maturity	$\beta_n$	$t$ -Statistic	$R$ -squared
1	0.09	3.46	48%
2	0.11	3.80	47%
3	0.15	3.05	38%
4	0.18	2.53	32%
5	0.19	2.20	27%
Panel B: Consumption growth predictability by nominal bond yields			
Maturity	$\beta_n$	$t$ -Statistic	$R$ -squared
1	0.20	0.52	6%
5	0.49	0.89	12%
5y-1y	-0.11	-0.13	0%
Panel C: Consumption growth predictability by real bond yields			
Maturity	$\beta_n$	$t$ -Statistic	$R$ -squared
1	-0.16	-0.44	2%
5	-0.21	-0.31	1%
5y-1y	0.62	0.93	5%

- Forward equity yields are therefore useful indicators of risk premia (+) and growth expectations (-), for instance around the tsunami in Japan.
- Second graph isolates expected dividend growth.

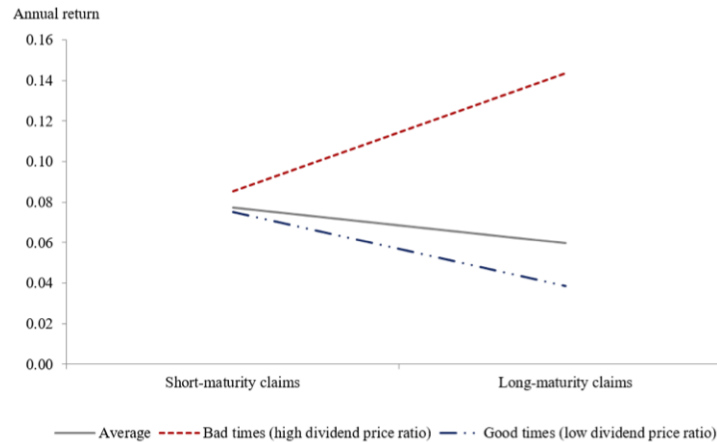


**Fig. 12.** Term structure of forward equity yields on March 31st, 2011. The graph displays the forward equity yields at the end of our sample period on March 31st, 2011 for maturities up to 7 years for the US, Europe, and Japan.



**Fig. 13.** Term structure of dividend growth expectations on March 31st, 2011. The graph displays the dividend growth expectations at the end of our sample period on March 31st, 2011 for maturities up to 7 years for the US, Europe, and Japan. We use a VAR model as described in Eqs. (16) and (19) to decompose forward equity yields into dividend growth expectations and risk premia by maturity. The sample period is from January 2003 to March 2011 to estimate the VAR model.

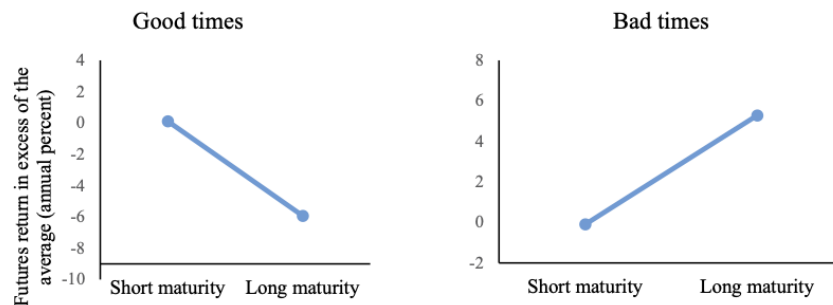
- Gormsen (2021) extends these facts and focuses on the time variation in equity premia across maturities (for a fixed holding period), conditional on good and bad times.
- Bad times are periods where the ex-ante dividend-price ratio is in the top quintile of the time series.



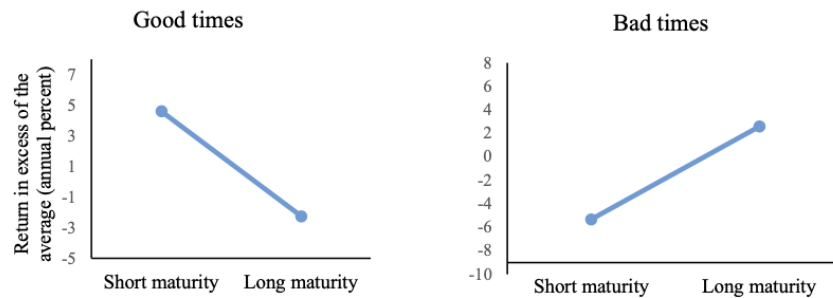
- Unconditionally, the equity premia decline with maturity, also for the longer sample from 1996 to 2019.
- However, the term structure of risk premia is upward sloping in bad times and more steeply downward sloping in good times.

- This result is robust across different ways of measuring equity term premia (using options, dividend futures, or using the cross-section of stocks). These results are also robust across sample periods.

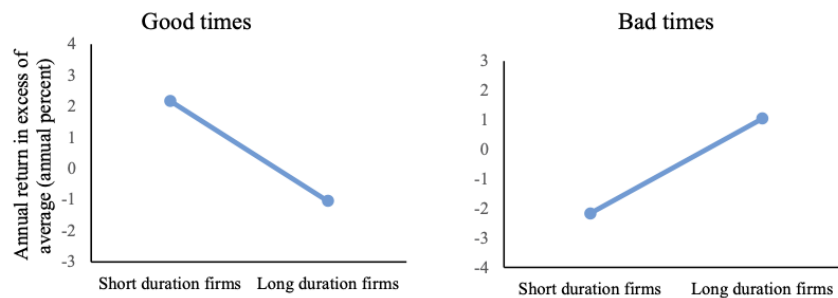
Panel A: Estimated using dividend futures (2003-2019)



Panel B: Estimated using option implied dividend prices (1996-2009)

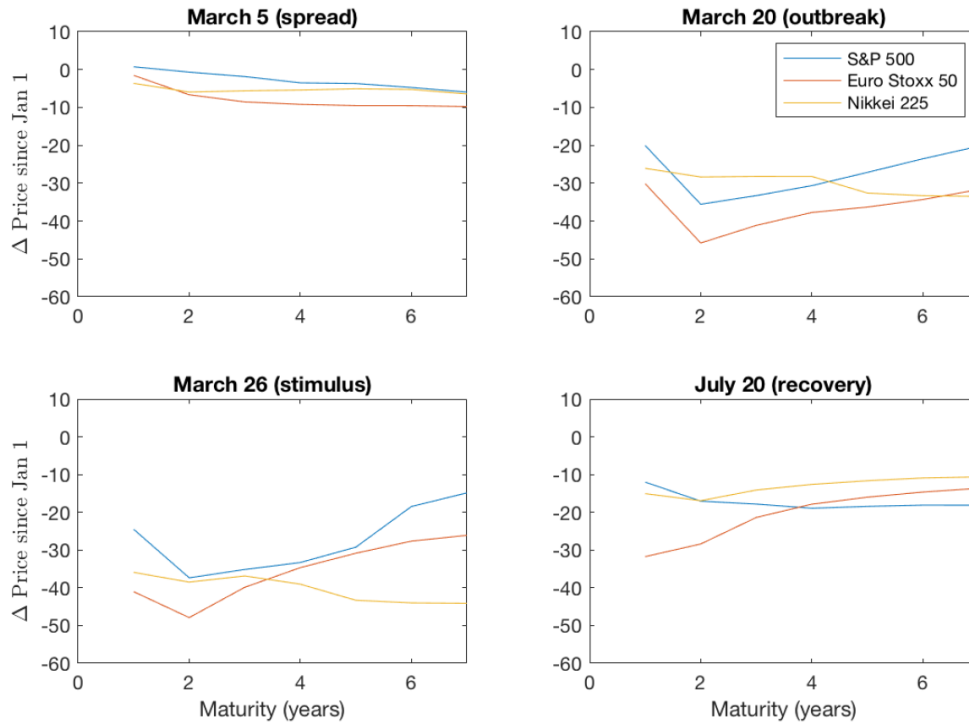


Panel C: Estimated using the cross-section of stocks (1963-2015)



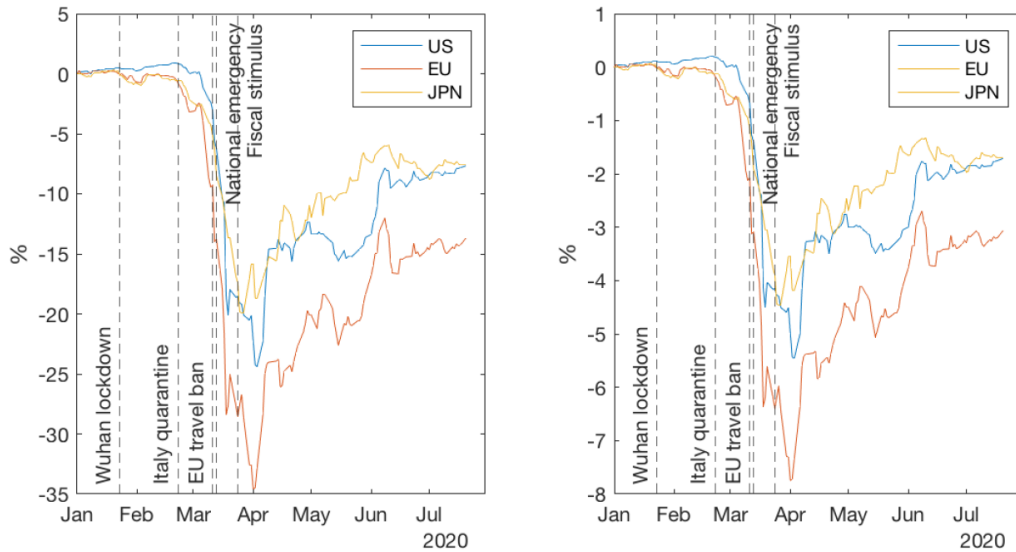
- Gomsen and Koijen (2020) study dividend futures during the onset of the COVID-19 crisis in 2020.

Figure 2: The Development of the Dividend Term Structure over the COVID-19 Outbreak



- They then extract the market’s expectation about future dividend growth and GDP growth by regressing dividend (GDP) growth on the lagged forward equity yield (at 2-year maturity) in the pre-covid era and use the fitted value and daily equity yields during the covid era to form expected div (GDP) growth.

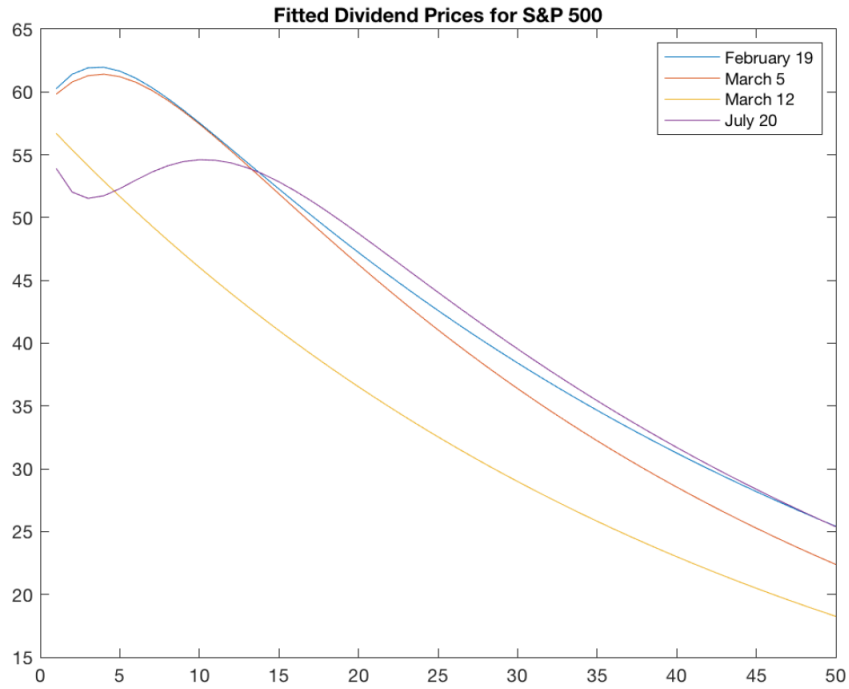
Figure 5: Expected Dividend and GDP Growth from Dividend Futures  
**Change in one-year expected dividend growth**      **Change in one-year expected GDP growth**



This figure shows the change in expected dividend and GDP growth relative to expected value at January 1, 2020. The figure shows expected growth in the US in blue, the EU in red, and Japan in yellow. Key events are indicated by the vertical dashed lines. The expected dividend growth is revised slowly in response to the outbreak, particularly in the US where it was revised down by less than 5% at March 11. By July 20, expected dividend growth is down by 8% in the US and Japan and 14% in the EU. Expected GDP growth over the next year is down by 2% in the US and Japan and 3% in the EU. Details of the estimation are in Section 3.3.

- Covid-19 had initially far-out dividend strips fall in price, then both short-end and long-end prices fell. The recovery was exclusively in long-dated strips.

Figure 9: Reconciling the stock market and dividend price responses





### 3.5. Revisiting the structural asset pricing models

- One can test the theoretical asset pricing models directly. If we simulate from the model, how likely is it to draw a sample that looks like the data?

Panel A: Campbell Cochrane (1999) Habit formation Model								Data			
Percentile	0.5	1	2.5	50	97.5	99	99.5	Eur	Japan	UK	US
$E[R_{t,P}^F]$	-0.0056	-0.0047	-0.0040	0.0013	0.0060	0.0071	0.0083	0.0090	0.0111	0.0063	0.0065
$E[R_{t,P}^F - R_{t,B}^M]$	-0.0066	-0.0062	-0.0058	-0.0031	-0.0006	0.0002	0.0004	0.0094	0.0081	0.0049	0.0044
$E[R_{t,P}^S - R_{t,B}^M]$	-0.0066	-0.0062	-0.0058	-0.0031	-0.0006	0.0002	0.0004	0.0060	0.0064	0.0026	0.0011

Panel B: Bansal and Yaron (2004) Long-run Risk Model								Data			
Percentile	0.5	1	2.5	50	97.5	99	99.5	Eur	Japan	UK	US
$E[R_{t,P}^S - R_{t,B}^M]$	-0.0051	-0.0047	-0.0043	-0.0020	0.0002	0.0006	0.0009	0.0060	0.0064	0.0026	0.0011

Panel C: Bansal, Kiku, and Yaron (2011) Long-run Risk Model								Data			
Percentile	0.5	1	2.5	50	97.5	99	99.5	Eur	Japan	UK	US
$E[R_{t,P}^S - R_{t,B}^M]$	-0.0064	-0.0060	-0.0054	-0.0026	-0.0006	-0.0002	0.0001	0.0060	0.0064	0.0026	0.0011

Table 3: The Simulated Average Return of Dividend Strips in the Habit Formation Model  
The table provides the percentiles of the simulated mean of a portfolio of dividend futures and dividend spot returns (in absolute terms as well as in excess of the market) in the habit formation model, where the number of monthly simulations corresponds to the number of observations in the data (146 months). The last four columns report the estimated means in the data for the four geographic regions.

- Simulate 1,000 samples of 146 months from the external habit model and find a minuscule likelihood that short-maturity assets beat the index. Same for the long-run risk model.
- However, using expected returns as moments is not the most powerful test of asset pricing models.
- Excess volatility on the short end of the equity curve leads to much more powerful volatility tests.
- Recall that  $e_{t,n}^f = \theta_{t,n} - g_{t,n}$ .
- We can compute the volatility on forward equity yields in the models and compare it to that in the data.
- As before, we use the external habit model as a test.
- Equity yields are much too smooth in the habit model.

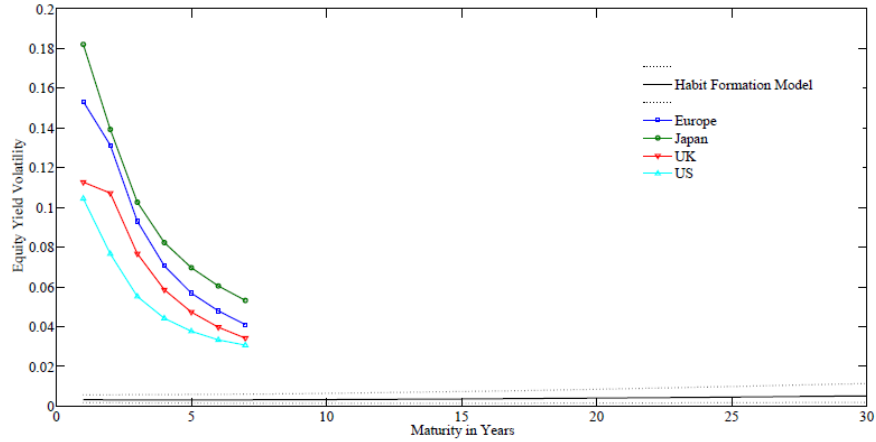


Figure 5: Equity Yield Volatility

The graph shows the estimated yield volatility in the data and contrasts it with the estimated yield volatility in the simulated Campbell and Cochrane (1999) habit formation model. We simulate the model 1000 times for 146 months, and for each simulation compute the volatility of the equity yields. The solid black line reports the median across 1000 simulations, and the dotted lines indicate the 95% confidence bound.

- The dotted lines indicate the confidence interval, which points to a powerful rejection of the model.

- New theories have been proposed to address these facts on the term structure of risk. They can be classified as:
  - Alternative models of preferences.
  - Alternative models of technology.
  - Alternative models of beliefs.
  - Heterogeneous agent models.
  - Pricing models with an exogenous SDF.
    - \* See for instance [Lettau and Wachter \(2007\)](#) and [Lynch and Randall \(2014\)](#).
- We briefly discuss some of the main mechanisms.
- Few models (so far) are able to explain:
  - Facts about average returns, Sharpe ratios, volatilities, and equity yields jointly.
  - Facts across asset classes.

- Alternative models of preferences:
  - Eisenbach and Schmalz (2016) and Andries, Eisenbach, and Schmalz (2019) consider a model in which the representative agent is more risk averse over imminent risks than risks in the distant future.
    - \* The model matches facts of the term structure of equity and variance risk.
  
- Alternative models of technology:
  - Nakamura, Steinsson, Barro, and Ursua (2013) consider a model with disasters **and recoveries** (see also Gourio, 2008).
    - \* Long-term dividend strips are less exposed to disaster risk due to economic recoveries.

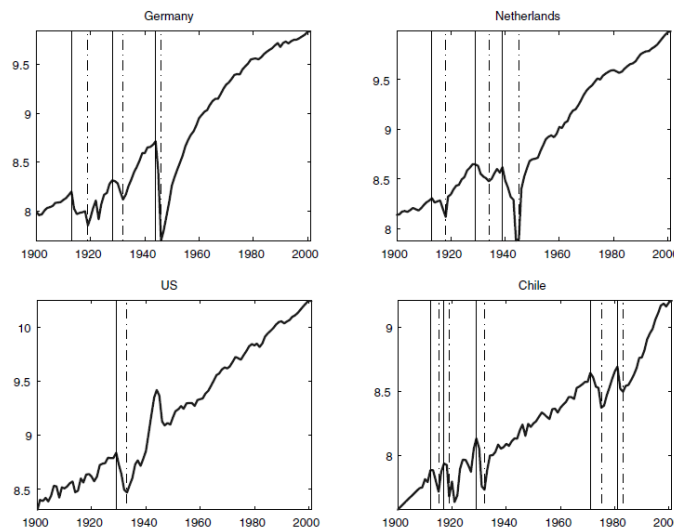


FIGURE 1. LOG GDP PER CAPITA (IN 1990 DOLLARS) FOR FOUR COUNTRIES: GERMANY, NETHERLANDS, THE US, AND CHILE  
 Note: The disaster start (end) dates are taken from Barro (2006), and are shown with a vertical full (dashed) line.

- [Belo, Colin-Dufresne, and Goldstein \(2015\)](#) propose a model to modify the dividend process.
  - Start by modeling firms' earnings, not dividends
  - Assume that leverage ratios are stationary
  - Shareholders are being forced to divest (invest) when leverage is low (high), which shifts long-horizon growth risk of earnings to short-horizon dividends.
  - As a result, dividends are more volatile than earnings over short horizons, but equally volatile over long horizons as dividends and earnings are co-integrated.

Alternative models of beliefs:

- Croce, Lettau, and Ludvigson (2014) consider a model with short-term and long-run shocks to consumption.
  - The representative decision maker optimizes based on a cash-flow model that is *sparse* in the sense that it ignores cross-equation restrictions that are difficult (if not impossible) to infer in finite samples.
  - Assets that have small exposure to long-run consumption risk, but are highly exposed to short-run (even i.i.d.) consumption risk, can command high risk premiums in the bounded rationality limited information case.
  - As a result, the term structure of equity risk premia can be downward sloping under the bounded-rational model, while it is upward sloping under full-information model.

## Heterogeneous agent models:

- All the models so far are representative agent models.
- [Lustig and Van Nieuwerburgh \(2006\)](#) are the first to show that a heterogeneous-agent model, where agents differ in their histories of income shocks, can produce a downward-sloping term structure of equity.
  - Risk sharing of income shocks is limited by the amount of housing collateral that agents have.
  - Agents face both shocks to the wealth distribution, which fluctuates at business cycle frequency, and shocks to housing collateral, which fluctuates at lower frequencies.
  - A negative consumption shock temporarily increases discount rates, but it does not affect housing collateral, which governs discount rates in the long run.
  - As a result, the price of consumption strips of longer maturity is insulated from bad consumption shocks today, which do affect short-maturity consumption strips.
- [Hodor and Zapatero \(2023\)](#) model investors that are heterogeneous in their investment horizon. Mutual fund managers are focused on the short term, due to their compensation incentives, while long-term value investors have longer investment horizons. This structure can lead to downward sloping term structure of expected returns when mutual funds' appetite for risk is strong.

### 3.6. *Applications and Open Questions*

- Real excess volatility:
  - Hiring depends on the present value of marginal product of labor minus wages.
  - In the data, hiring is too volatile.
  - [Hall \(2017\)](#) shows that variation in short-term discount rates could explain the variation in hiring.
- The argument extends to investment as well, providing a potential link between asset prices and both investment and hiring decisions.
- Indeed, it would be interesting to see whether we can use data on various term structures to come up with discount rates that can be used to understand hiring, investment, and the valuation of both listed equity and private companies.



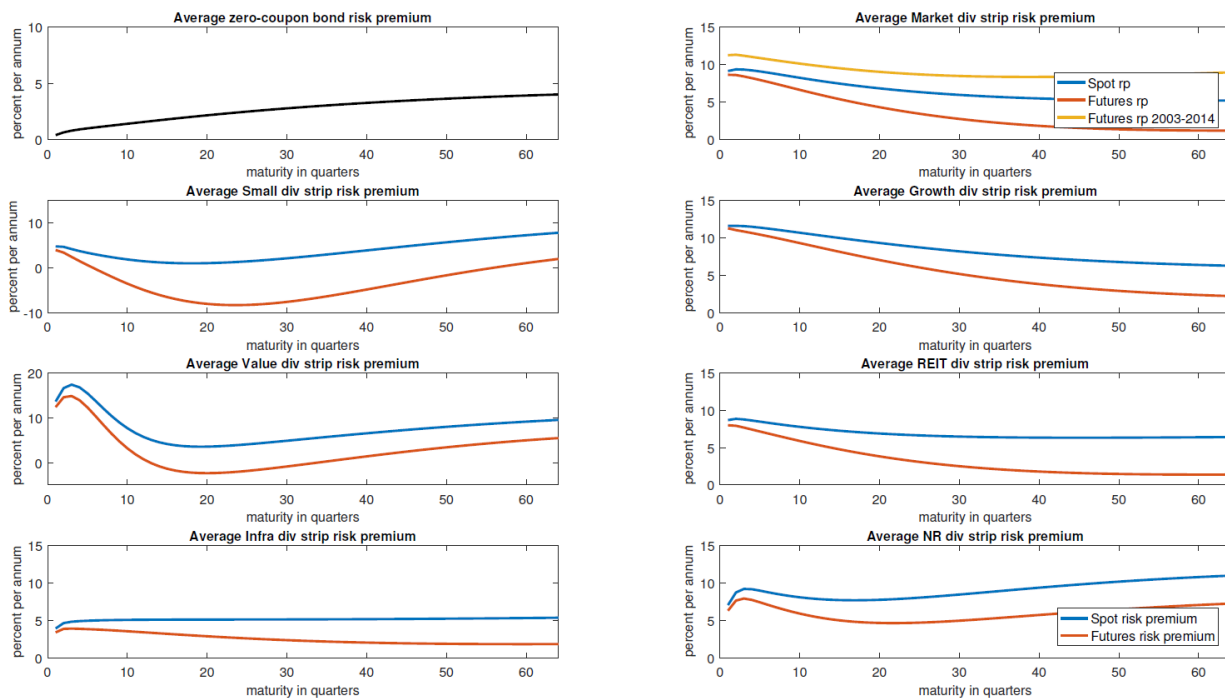


Figure 8. Strip expected returns by horizon.

- [Gupta and Van Nieuwerburgh \(2021\)](#) use the term structure of risk in stock and bond markets to value [private equity](#).
  - In the process they extract term structures of returns for value, growth, small, real estate, infrastructure, and natural resource stocks for the full post-war sample. Have different shapes unconditionally, as well as dependence on state of economy.
  - Their asset pricing model matches the evidence from dividend strips from options and futures for the subsample when data is available. Futures risk premium curve is not particularly steeply downward sloping in the data sample.

## APPENDIX: Extracting expected returns and dividend growth rates using the Kalman Filter

- Follows [van Binsbergen and Koijen, 2010](#)
- Denoting the demeaned expected growth rate of dividends by  $\hat{g}_t = g_t - \gamma_0$ , we arrive at the final system

$$\begin{aligned}\Delta d_{t+1} &= \gamma_0 + \hat{g}_t + \epsilon_{t+1}^d, \\ pd_{t+1} &= (1 - \delta_1)A + B_2(\gamma_1 - \delta_1)\hat{g}_t + \delta_1 pd_t - B_1\epsilon_{t+1}^\mu + B_2\epsilon_{t+1}^g, \\ \hat{g}_{t+1} &= \gamma_1\hat{g}_t + \epsilon_{t+1}^g.\end{aligned}$$

The first two equations are [measurement equations](#). The third equation is the [transition equation](#) of the latent variable.

- We estimate the model via maximum likelihood, where we use the Kalman filter to construct the likelihood.
- We write the state and observation vectors in general form

$$X_t = \begin{bmatrix} \hat{g}_{t-1} \\ \epsilon_t^d \\ \epsilon_t^g \\ \epsilon_t^\mu \end{bmatrix}, \quad Y_t = \begin{bmatrix} \Delta d_t \\ pd_t \end{bmatrix}.$$

- We can write the dynamics of the state vector and observation vectors as

$$\begin{aligned}X_t &= FX_{t-1} + \Gamma\epsilon_t, \\ Y_t &= M_0 + M_1Y_{t-1} + M_2X_t,\end{aligned}$$

where the coefficient matrices  $F$ ,  $\Gamma$ ,  $M_0$ ,  $M_1$ , and  $M_2$  follow from the earlier equations.

- In the Kalman filter, we recursively update our estimate of the state.
- Define  $X_{t|s} = E_s[X_t]$  and  $P_{t|s} = E_s[X_t X_t']$ . These are our best estimates of the latent state and covariance matrix, conditional on the information until time  $s$ .
- In the procedure below, we use  $s = t - 1$  and  $s = t$ . However, you can do similar calculations for  $s = T$ , which is our best estimate of the latent state using the full sample. This is called the [Kalman smoother](#).
- We can now compute the likelihood. We initialize the filter using the unconditional distribution

$$\begin{aligned} X_{0|0} &= E[X_0] = 0_{4 \times 1}, \\ P_{0|0} &= E[X_0 X_0']. \end{aligned}$$

- Next, we construct predictions for time  $t$  using time- $(t - 1)$  information:

$$\begin{aligned} X_{t|t-1} &= F X_{t-1|t-1}, \\ P_{t|t-1} &= F P_{t-1|t-1} F' + \Gamma \Sigma \Gamma'. \end{aligned}$$

- Based on these predictions, we can compute the residuals of the observation equation and their covariance matrix

$$\begin{aligned} \eta_t &= Y_t - M_0 - M_1 Y_{t-1} - M_2 X_{t|t-1}, \\ S_t &= M_2 P_{t|t-1} M_2', \end{aligned}$$

where  $S_t = E_{t-1}[\eta_t \eta_t']$ . We use this to construct the log likelihood

$$\mathcal{L} = - \sum_{t=1}^T \log(\det(S_t)) - \sum_{t=1}^T \eta_t' S_t^{-1} \eta_t.$$

- To complete the iteration, we need to update  $X_t$  and  $P_t$  with the new time- $t$  observation

$$\begin{aligned} K_t &= P_{t|t-1} M_2' S_t^{-1}, \\ X_{t|t} &= X_{t|t-1} + K_t \eta_t, \\ P_{t|t} &= (I - K_t M_2) P_{t|t-1}, \end{aligned}$$

where  $K_t$  is called the Kalman gain and measures the revision of the latent state based on the innovations,  $\eta_t$ .

- It is easy to show (see the appendix of [van Binsbergen and Kojen, 2010](#)) that the Kalman filter effectively introduces moving average terms of returns and dividend growth rates to predict future returns and future dividend growth rates.