Predictability: the wrong way

Andrea Buraschi and Andrea Carnelli

First draft: November 2011
This version: May 2012

ABSTRACT

We study the predictability of S&P500 returns using short term expected risk premia as a conditioning variable. We construct short term expected risk premia by combining dividend prices implied by futures markets with a simple model of dividend expectations. Regression results for forecasting horizons from 1 to 4 quarters show that time variation in risk premia captures time variation in realized excess returns, albeit with the wrong sign. Counter to the intuition that a high price of risk commands high returns, we find that high expected returns predict low returns. The economic significance is strong: a one standard deviation rise in expected risk premia decreases realized excess returns by 0.16 (0.34) standard deviations at quarterly (annual) horizons. No asset pricing model is able to generate these patterns of predictability.

Keywords: equity risk premium, predictability, dividend prices, asset pricing models, term structure.
JEL Codes: G10, G12, G13.

Andrea Buraschi (andrea.buraschi@chicagobooth.edu) is at The University of Chicago Booth School of Business. Andrea Carnelli (andrea.carnelli08@imperial.ac.uk) is at Imperial College London Business School.

I Introduction

In this paper we construct an empirical proxy for short term expected risk premia and study its empirical properties. Understanding the properties of risk premia is one of the most important yet challenging tasks in asset pricing. One of the challenges is related to the fact that, in general, it is difficult to identify risk premia demanded by investors: they cannot be directly observed when the timing and magnitudes of dividends are unknown. A stock index pays uncertain dividends at uncertain times over a life of uncertain duration. Investors form expectations about future cash flows and use a set of discount rates to determine the value. All information is then aggregated by a tatonnement and the econometrician can only observe a time series of stock prices. This "loss" of information is a binding constraint for our understanding of asset prices; its limits are laid bare when contrasted to our understanding of discount rates in fixed income markets. In the fixed income literature, the term structure of interest rates is an observable entity with two important roles. First, it encodes information about the dynamics of the stochastic discount factor (yields are risk adjusted expectations of future short rates). Second, it provides a host of restrictions that allow to better discriminate among alternative models. If equity discount rates were observable, we could learn about their dynamics and generate additional restrictions that asset pricing models have to satisfy.

How can we construct a measure of expected returns? We follow the pioneering work of Van Binsbergen et al. (2011) and address this challenge by using observations on dividend prices. Dividend prices can be obtained by imposing a simple no-arbitrage pricing restriction on cash and derivatives markets. The prices of dividends carry information on both future expected dividend and risk premia. By modeling expectations about short term dividends, we are able to identify the short term risk premium component embedded in dividend prices. We investigate the properties of these market-implied short term risk premia (erx_t^{ST} hence). In particular, we are interested in the relationship between erx_t^{ST} and longer maturity discount rates. We explore this relationship indirectly by running predictive regressions of realized excess returns on lagged short term risk premia. Since realized excess returns reveal changes in equity discount rates at all maturities, these regressions allow us to uncover the dynamic link among short- and long-term discount rates. We ask to what extent the dynamics of risk premia on short term dividends help to explain index returns and which asset pricing model can be consistent with some of its basic properties.

We test the null hypothesis of no predictability by projecting realized excess returns on short term expected risk premia. Our results can be summarized as follows. We find that short term expected returns predict realized returns over short horizons. The economic significance is strong: a one standard deviation rise in expected risk premia decreases realized excess returns by 0.16 (0.34) standard deviations at quarterly (annual) horizons. The predictive power of expected returns over short horizon is over and above that of dividend yields: expected returns detect high frequency predictability components in stock returns that dividend yields do not capture. Even more interesting, the sign of the slope coefficient is negative: higher market-implied risk premia at time t are followed by lower excess returns at time $t + h$. We control for the overlapping nature of the regressions, and verify that the results are robust to potential mis-specification of the model for expected dividends. The results are difficult to be reconciled with leading asset pricing models, at least in their traditional specifications. These models generate risk premia dynamics and term structures of expected equity returns that are too restrictive. We identify the two ingredients that are necessary for a model to be able to explain this type of predictability: shocks to dividends must be priced, and there must be a multivariate state vector driving the term structure of expected risk premia.

Our first empirical exercise is a regression of excess returns on expected one year risk premia: can theory say something about the sign and magnitude of the slope coefficient? The problem can be best understood by means of an analogy to fixed income securities. We can think of an equity index as a long maturity bond (say, 30 years) that pays coupons (i.e. dividends), and we can think of short term expected risk premia as the yield on a short maturity (say, 1 year) zero coupon bond. The issue we address is akin to the question: can one year yields predict one-year holding period returns on a 30 years maturity bond? A rigorous answer to the question involves specifying a term structure model that: i) derives the price of a bond as a function of state variables, and ii) determines expected risk premia by modeling the covariance between shocks to bond prices and shocks to the stochastic discount factor. To the extent that the yield on the 1 year bond reveals the state variable, it is also possible to form expectations about the 1 year holding period return of the 30 year maturity bond. An equity index is like a coupon-bearing bond, except that is has infinite maturity and cash flow risk. Just like a coupon-bearing bond can be thought of as a portfolio of zero coupons, an equity index is a portfolio of "zero coupon" dividend strips, i.e. claims to dividends payable at future dates. The term structure model of dividend strips we adopt is the one developed in Lettau and Wachter (2007). Lettau and Wachter (2007) show that short term expected risk premia (our regressor) identify the price of risk up to a scaling factor, and that expected returns on dividend strips with longer maturities are positive functions of the price of risk. Since the return on the index is the sum of returns on dividend strips of all maturities, the model provides a tight link between the left and right hand side variables of our regression, and motivates its predictive power.

This work is especially related to Van Binsbergen et al. (2011) study of equity yields. Equity yields are the ratio of the futures price of dividend futures to the current dividend. Time variation in equity yields reflects time variation in dividend growth expectations and/or risk premia. The authors test the predictive power of equity yields, and use a VAR to separate the dividend growth from the risk premia components. They find that equity yields are good forecasters of dividend and consumption growth, and that risk premia are counter cyclical. Our analysis is different from Van Binsbergen et al. (2011) in two respects. First, instead of looking at the predictability of dividend growth, we focus on the predictability of returns. Second, we exploit the predictability of returns conditional on short term risk premia to uncover the dynamics and time dependence of the term structure of equity risk premia.

II A Motivating Example

As a motivation for the regressions that follow, we use the term structure model of dividend prices developed by Lettau and Wachter (2007) to shed light on the relationship between short term expected risk premia and index returns. The logarithm of dividend growth $d_{t+1} = \log(D_{t+1}/D_t)$ and its conditional mean z_{t+1} are assumed to follow

$$\begin{aligned} \Delta d_{t+1} &= g + z_t + \sigma_d \epsilon_{t+1} \\ z_{t+1} &= \phi_z z_t + \sigma_z \epsilon_{t+1}. \end{aligned} \tag{1}$$

The price of risk is driven by a state variable x_t that evolves according to

$$x_{t+1} = (1 - \phi_x)\bar{x} + \phi_x x_t + \sigma_x \epsilon_{t+1},$$

where ϵ_{t+1} is a 3×1 vector of independent standard normal shocks, so that the "shock loadings" σ_d, σ_z, σ_x are 1×3 row vectors. Note that the conditional volatility of each variable is given by the norm of the vector of shock loadings, while the conditional covariance between any two variables is given by the inner product of the respective vectors of shock loadings. Since only dividend risk is assumed to be priced, and the (log) risk free rate $r^f = \log R^f$ is assumed to be constant, the stochastic discount factor equals

$$M_{t+1} = \exp\left\{ - r^f - \frac{1}{2}x_t^2 - x_t \frac{\sigma_d}{||\sigma_d||}\epsilon_{t+1} \right\}.$$

Let $P_{t,n}$ be the price of a claim that pays the aggregate dividend n periods from now, and let $R_{n,t+1} = P_{t+1,n-1}/P_{t,n}$ denote the return on holding the dividend claim for one period. Note that by no arbitrage, it must be that $P_{t,0} = D_t$. Lettau and Wachter (2007) show that under these assumptions both the price-to-dividend ratio and the expected excess return on dividend claims are exponential-affine functions of the state variables:

$$\log P_{t,n} = \log D_t + A(n) + B_x(n)x_t + B_z(n)z_t$$

$$\log E_t\left[R_{n,t+1}/R^f\right] = \left(\sigma_d + B_x(n-1)\sigma_x + B_z(n-1)\sigma_z\right)\frac{\sigma_d'}{||\sigma_d||}x_t, \tag{2}$$

where the values of the coefficients $A(n)$, $B_z(n)$, and $B_x(n)$ are given by the boundary condition for $n = 0$

$$A(0) = B_z(0) = B_x(0) = 0,$$

and, for $n > 0$, by the recursions

$$A(n) = A(n-1) - r^f + g + B_x(n-1)(1-\phi_x)\bar{x} + \frac{1}{2}V_{n-1}V_{n-1}'$$

$$B_z(n) = \frac{1-\phi_z^n}{1-\phi_z}$$

$$B_x(n) = B_x(n-1)\left(\phi_x - \sigma_x\frac{\sigma_d'}{||\sigma_d||}\right) - (\sigma_d + B_z(n-1)\sigma_z)\frac{\sigma_d'}{||\sigma_d||}.$$

Since an index can be thought of as a portfolio of dividend claims, its value today is equal to the sum of the prices of the dividend strips

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

and the expected excess return it generates is a weighted average of dividend risk premia across the range of maturities

$$\log E_t\left[R_{t+1}/R^f\right] = \log\left(\sum_{n=s}^{\infty} w_n \exp\left\{(\sigma_d + B_x(n-1)\sigma_x + B_z(n-1)\sigma_z)\frac{\sigma_d'}{||\sigma_d||}x_t\right\}\right), \tag{3}$$

where $w_n = P_{n,t}/P_t$ and $s = 1$ ($s = 2$) for cum-dividend (ex-dividend) returns.

Equations (2) and (3) allow us to gain insight into the relationship between short term expected risk premia (the right hand side of our regression) and index returns (the left hand side of our regression). Short term expected risk premia are the empirical counterpart to (2) for $n = 1$

$$\log E_t \left[R_{1,t+1}/R^f \right] = ||\sigma_d|| x_t. \tag{4}$$

Equation (4) says that expected risk premia backed out of short term dividend expectations and prices are a constant multiple of the latent factor x_t. Since the expected return on the index is also a function of x_t, its derivative with respect to $||\sigma_d|| x_t$ is different from zero

$$\beta_1 = \frac{\partial \log E_t \left[R_{t+1}/R^f \right]}{\partial ||\sigma_d|| x_t}. \tag{5}$$

Defining $\beta_0 = \log E_t \left[R_{t+1}/R^f \right] - \beta_1 ||\sigma_d|| x_t$, and letting $\{u_t\}$ denote a white noise process orthogonal to x_t, one can see that the model imposes a restriction on the coefficients of the regression

$$\log(R_{t+1}/R^f) = \beta_0 + \beta_1 ||\sigma_d|| x_t + u_{t+1}, \tag{6}$$

and suggests that expected risk premia on short term dividends can predict future realized returns. Equations (**??**) is the core regression of our empirical exercise. The relationship lends itself to a simple interpretation. The only risk involved in buying a claim to one-year forward dividends is dividend volatility, so that short term expected risk premia are proportional to the latent factor that drives the price of risk. Since the price of risk is a state variable for the entire term structure of dividend strips, variation in short term expected risk premia captures variation in realized returns.

III Data

This section describes the data set. All data are at quarterly frequency and span the period from 1986 : 1 to 2011 : 4. The reason behind picking 1986 as a start year is data availability: since we construct short term expected risk premia from dividend prices implied by index futures, we can only study the period over which S&P500 futures are traded. The use of futures data also motivates the choice of sampling frequency: since contract maturities are in the March quarterly cycle, quarterly observations dispense us with the need to obtain constant maturity contracts. As an alternative to backing out dividend prices indirectly from index futures via the cost of carry, we could have obtained direct observations from dividend swaps, as in Binsbergen et al. (2011). Dividend swaps have the advantage of being traded for a large range of maturities, but, at the same time, they are OTC contracts with counter-party risk and have gained in popularity quite some time after index futures. Also, dividend prices implied by dividend swaps are potentially subject to price pressure owing to institutional hedging.[1] Since our intention is to construct short term expected risk premia, the trade-off between the two datasets is biased in favor of futures data.

A Derivation of Expected Risk Premia

The price of a stock is the sum of the prices of the dividends it will pay in the future. By combining dividend prices with dividend expectations of different maturities, it is possible to back out a term structure of discount rates. Dividend expectations can be recovered either directly (from analysts' forecasts) or indirectly (via an econometric model). Dividend prices on the other hand are implicit in the spread between cash and (discounted) futures prices. More formally, consider a dividend-paying financial asset with time t price S_t and a forward contract written on the asset with delivery n periods in the future. Define $P_{t,n}$ as the value that makes the cost of carry formula true:

$$F_{t,n} = [S_t - P_{t,n}]\, e^{nr_{t,n}},$$

where $r_{t,n}$ is a (continuously compounded) risk-free rate appropriate to discount $t+n$ cash flows back to t. The expression can be re-written in terms of $P_{t,n}$

$$P_{t,n} = S_t - F_{t,n} e^{-nr_{t,n}}. \tag{7}$$

[1] Investment banks typically sell derivatives whose payoffs are ex-dividend; hedging these derivatives requires a position in the underlying, which gives rise to a long exposure to dividend risk. If dividend risk is too high, the investment bank might have the incentive to sell dividend derivatives at particularly attractive prices.

By no arbitrage $P_{t,n}$ must be the price that markets are willing to pay for the dividend stream between t and $t+n$. Equation (6) thus provides a means to infer the price of dividend claims from spot and futures market data. Let D_{t+n} denote the dividend that the asset will pay at $t+n$. Since D_{t+n} is a random variable at t, its price today is equal to its expected value discounted by a risk free rate and an appropriate risk premium $\mu_{t,n}$

$$P_{t,n} = E_t\left[D_{t+n}\right]e^{-n(r_{t,n}+\mu_{t,n})}.$$

Re-arranging terms, time t expected risk premia for $t+n$ cash flows can be expressed as

$$\mu_{t,n} = \frac{\log E_t[D_{t+n}] - \log P_{t,n}}{n} - r_{t,n}. \tag{8}$$

The following sections explore the data counterparts to equations (6) and (7).

B S&P500 Dividend Prices

In order to construct term structures of dividend prices using (6) we need the time series of spot prices S_t, futures prices $F_{t,n}$, and interest rates $r_{t,n}$ that are appropriate to discount time $t+n$ cash flows to t. S&P500 futures trade on the Chicago Mercantile Exchange (CME). Available maturities vary within the sample. When S&P500 futures started trading in 1982, available maturities included the first 4 months in the March cycle. Ever since, the number of available contracts has increased progressively.[2] All open positions are settled in cash at close of the last day of trading (3:15 p.m. on Thursday prior to third Friday of the contract month) to the Special Opening Quotation (SOQ). The SPQ is calculated using normal index procedures except that the prices used are the actual opening values of the constituents. Since securities might start trading minutes or hours after the opening bell of the NYSE, the SPQ can be significantly different than the published opening value of the index. The SPQ is used because it represents actionable values and thus ensures convergence of cash and futures markets.

[2] As of March 2012, available maturities include 11 contracts in the March cycle.

Both futures and spot data are taken from Thomson Datastream. There is a slight mismatch in the timing of index and futures close prices. CME sets settlement prices for the day equal to midpoint of the closing range based on pit trading activity between 15:14:30-15:15:00 Central Time. S&P indices on the other hand are computed using New York Stock Exchange (NYSE) close prices at 16:00:00, Eastern Time. The implicit assumption is that the 15 minutes mismatch in quotes can be treated as white noise. To compute dividend prices, we use equation (6) by combining underlying and futures prices with interest rates data from IvyDB (Optionmetrics). The database contains (annualized, continuously compounded) zero curves derived from LIBOR rates from the British Bankers Association (BBA) and settlement prices of Chicago Mercantile Exchange (CME) Eurodollar futures. For a given futures contract, the appropriate interest rate corresponds to the zero-coupon rate that has a maturity equal to the futures' expiration date, and is obtained by linearly interpolating between the two closest zero-coupon rates on the zero curve. [3] At the end of each quarter, we obtain the prices of cumulative dividends for different maturities. Since we use quarterly data, the maturities of the dividend prices we obtain are constant over time; this allows us to dispense with the need of interpolation techniques to obtain constant maturity prices, which would be necessary with higher sampling frequency. The number of maturities depends on the set of available contracts. Ideally, we would like to exploit the information of the entire term structure of dividend prices, and back out a term structure of expected risk premia. This approach, however, has three limitations. First, the depth of the term structure varies over time, as argued above. Second, even though the use of quarterly data virtually eliminates the possibility of stale prices, dividend prices might still reflect, to some extent, the different levels of liquidity of the futures contracts used to construct them. Third, the prices of short-maturity dividends may inherit the seasonality of dividend payments. Bearing these considerations in mind, we construct dividend prices using the third available contract, which is the longest maturity contract available over the entire sample.

[3] IvyDB data starts in 1996. Between 1986 and 1996, we use the term structure of monthly LIBOR rates (obtained from Datastream).

C S&P500 Dividend Expectations

In order to construct dividend expectations, we work under the null hypothesis that (log) dividend growth is equal to a constant plus a white noise process. This assumption is well defended by a large body of empirical literature. Many authors (see Cochrane, 2005, for a review) have found that time variation in dividend yields can forecast future excess returns. It can be easily shown that, in rational economies without bubbles, this empirical finding, coupled with an accounting identity that relates current prices to future cash flows and discount rates, implies the absence of dividend growth predictability (see, for instance, Cochrane 2006 and references therein). Absence of dividend growth predictability can be understood as the consequence of dividend smoothing corporate policies; importantly, it implies that time variation in the value of dividend claims is dominated by revisions in the "quantity" and "price" of risk. If (log) dividend growth is equal to a constant plus white noise, $\Delta d_{t+1} = g + \sigma_d \epsilon_{t+1}$, the level of future dividends is $D_{t+1} = D_t \exp(g + \sigma_d \epsilon_{t+1})$, and expected future dividends are equal to $E_t[D_{t+1}] = D_t \exp\left(g + \frac{1}{2}||\sigma_d||\right)$. We use current dividends D_t as a proxy of dividend expectations. The term $g + \frac{1}{2}||\sigma_d||$ is a constant that is absorbed by the intercept of our regression since we take logs of the RHS. IBES dividend forecasts would be an ideal measure of model-free, forward-looking dividend expectations. Unfortunately, dividend per share forecast data are well populated only since 2004. In the robustness section, however, we rerun predictive regressions using expected risk premia constructed from IBES consensus forecasts, and show that the results are robust. The assumption that dividend growth is a constant plus a white noise process seems not to be rejected in our sample: the plot of Δd_{t+1} in figure 1 resembles a white noise process.

D S&P500 expected risk premia, realized returns, and dividend yields

In order to construct excess expected and realized returns, we need data on risk free rates. We take the 3 month US Treasury rate from Datastream. Rates are quoted on a discount basis, annualized using the 30/360-day-count convention, and expressed in percentage points. We convert all rates to their annualized, continuously compounded equivalents using the *Actual*/360-day-count convention. [4] We compute one-year forward expected risk premia erx_t^{ST} by combining one-year forward dividend expectations and prices with 12 month risk-free rates via (7). Figure 2 plots the time series of erx_t^{ST}. The plot peaks in December 2000 at a level that is economically unreasonable, so in the analysis that follows we exclude this observation (the inclusion of the observation has the only effect of strengthening all results). Table I reports the sample statistics: expected risk premia are generally high (the mean is 31%) and volatile (26% standard deviation). Consistent with the findings of Van Binsbergen et al. (2010), our measure of expected returns documents that risk premia on short term cash flows are considerably higher than the risk premium earned historically on the $S\&P500$. Finally, realized excess returns $rx_{t,t+h}$ are constructed as the (annualized) cum-dividend (log) return on the index between t and $t+h$, minus the (annualized) 3 month Treasury rate prevailing at time t.

IV Forecasting regressions

A Baseline model

The reduced-form model presented in the first section suggests that one year expected returns predict realized returns with a positive sign. To investigate this link empirically we run regressions of realized excess returns on expected one-year risk premia over different forecasting horizons h:

$$rx_{t+h} = \beta_0^{(h)} + \beta_1^{(h)} erx_t^{ST} + \epsilon_{t+h}^{(h)}. \tag{9}$$

When data sampling is finer than forecasting horizons, predictive regressions are overlapping. The overlap causes regression errors to be serially correlated even under the null hypothesis of no predictability: estimates of the covariance matrix of coefficients based on the classic assumption of homoskedasticity and no autocorrelation overstate the statistical significance of the estimates. In order to rule out spurious statistical significance, we estimate the covariance matrix of coefficients

[4] Let R be the discount yield on a bond due in τ days. The annualized, continuously compounded, rate associated with R is equal to $r = (365/\tau)\log\left(1 + (R/100) \times (\tau/360)\right)$.

in two ways: Newey and West (1987), and Hansen and Hodrick (1980). The two methods can be understood as instances of the GMM framework for different choices of the long run covariance matrix of sample moments. To illustrate, let β denote a $(K \times 1)$ vector of OLS coefficients in the forecasting regression $y_{t+h} = x_t \beta_t' + \epsilon_{t+h}$ for $t = 1, \ldots, T+h$. Since the OLS estimate of β satisfies $E[\epsilon_{t+h} \otimes x_t] = 0$, $E[\epsilon_{t+h} \otimes x_t] = 0$ is the equivalent GMM moment condition. The benefit of working with GMM is that it only assumes that errors are orthogonal to the regressors. The covariance matrix of OLS estimates is

$$Var\left[\beta\right] = E[x_t x_t']^{-1} S E[x_t x_t']^{-1},$$

where

$$S = \sum_{j=-\infty}^{\infty} E\left[\epsilon_{t+h} x_t x_{t-j}' \epsilon_{t-j+h}\right].$$

Henceforth, we will denote Newey and West (1987) and Hansen and Hodrick (1980) errors by NW and HH, respectively. The difference is based on alternative methods to estimate the (asymptotic) covariance matrix of coefficients in regression (8), taking the overlaps into account.[5] We consider four forecasting horizons: $1, 2, 3, 4$ quarters. For each forecasting horizon, we test for the marginal significance of slope coefficients using NW and HH errors.

Table II contains the output of the regressions. With the exception of very short horizons, the statistical significance of the loadings of realized excess returns on expected risk premia is high. While the significance of the slope coefficient for one quarter horizons is just below 10%, all other t-statistics, computed using both NW and HH errors, are significant at the 5% level. Statistical significance tends to be increasing in forecasting horizon, and peaks at forecasting horizons of three quarters, achieving -3.45 (NW errors) and -3.14 (HH errors). Also the adjusted R^2 are

[5] The two methods differ in the estimation of S. Newey and West (1987) propose to underweight higher order sample autocorrelations: $S = \sum_{j=-l}^{l} \left(\frac{h-|j|}{h}\right) \left(\frac{1}{T} \sum_{h}^{T+h} \epsilon_{t+h} x_t x_{t+h-j} \epsilon_{t+h-j}\right)$, where the maximum number of lags l is constrained by sample size $l_{max} = T + h - 1$. Hansen and Hodrick (1980) suggest to correct only for the $MA(h-1)$ structure introduced by the overlap, and to weight autocorrelations equally: $S = \sum_{j=-h+1}^{h-1} \frac{1}{T}$ $\sum_{h}^{T+h} \epsilon_{t+h} x_t x_{t+h-j} \epsilon_{t+h-j}$.

increasing in maturity: they range from 1.37% to 10.91%. Given the simplicity of the econometric model and the high noise in realized returns, the degree of predictability we find is interesting.

The most striking feature of the results relates to its economic significance. Since both regressands and regressors are standardized, the coefficients can be interpreted as standard deviation changes in realized excess returns caused by a one standard deviation change in expected risk premia.The estimated coefficients over the four forecasting horizons are $-0.16, -0.23, -0.30, -0.34$. In order to facilitate comparability across forecasting horizon, consider the annualized slope coefficients: $-0.64, -0.46, -0.23, -0.34$. The magnitudes are economically significant. Importantly, contrary to the predictions of the model of Lettau and Wachter (2007), the estimated slope coefficients are negative: higher levels of expected risk premia are associated with lower subsequent realizations of excess returns.

B Horse races

A natural question is whether expected risk premia feature explanatory power over and above that of traditional predictive variables studied by the empirical literature. A popular strand of literature, for instance, looks at the predictive power of dividend yields. These regressions have a strong theoretical motivation: under the assumption of rationality and non bubbles, it can be easily shown that dividend yields can vary over time only if either dividend growth or excess returns are predictable. Many authors find evidence that the dividend yield is a predictor of future excess returns, and that the predictive power rises in the forecasting maturity (see, for instance, Campbell 1991 and Cochrane 1992). Examining the joint predictive power of expected risk premia and dividend yields is interesting for two reasons. First, since the two variables capture time variation of future excess returns at different frequencies (we find that expected risk premia forecast excess returns over short horizons, while dividend yields are known to have predictive power over long horizons), a model that includes both variables can lead to better forecasting performance. Second, by controlling for time varying expectations about all future excess returns, the inclusion of dividend yields allows to isolate the role of short term expected risk premia.

Table III summarizes the evidence about dividend yield predictability in our sample, by showing the output of forecasting regressions of realized excess returns on lagged dividend yields. (We also include 3-month treasury rates because Ang and Bekaert (2006) find that it improves the short term forecasting power of dividend yields). The results are consistent with the empirical literature. Dividend yields predict excess returns with a positive sign (low prices relative to dividends capture high discount rates, rather than low dividend growth), and forecasting power builds over longer maturities: both the t-statistics and adjusted R^2 rise monotonically with the forecasting horizon. On the other hand, contrary to the finding of Ang and Bekaert (2006), we find that 3-month Treasury rates do not have marginal significance.

Having run univariate regressions on erx_t^{ST} and dy_t separately, we analyze their combined predictive power in bivariate regressions:

$$rx_{t+h} = \beta_0^{(h)} + \beta_1^{(h)} erx_t + \beta_2^{(h)} dy_t + \beta_3^{(h)} rf_t + \epsilon_{t+h}^{(h)}. \tag{10}$$

Table IV summarizes the regression results. While the inclusion of dividend yields seem to absorb some of the forecasting power of erx_t^{ST}, the statistical significance of expected risk premia over horizons of three and four quarters remains high: slope coefficients are significant at least at the 5% level. Controlling for dividend yields, the economic significance is only marginally affected: the slope coefficients remain negative. The estimated coefficients over the four forecasting horizons are $-0.12, -0.18, -0.23, -0.27$, or, expressed on a per annum basis, $-0.48, -0.36, -0.17, -0.27$: relative to the baseline case, we still find that the annualized loadings are decreasing (in absolute value) in the forecasting horizons, but this effect is slightly accentuated by the inclusion of dividend yields. The loadings on dividend yields, on the other hand, still feature a statistical significance that increases in the forecasting horizon; erx_t^{ST} seems to absorb part of its forecasting power: the null hypothesis that the estimated slope coefficient on dy_t is equal to zero can be rejected only for annual horizons. The loadings on dy_t are increasing in the horizon $0.09, 0.15, 0.20, 0.26$.

There are three main lessons to be learned from the results of regressions 8 and 9. First, the finding that higher expected risk premia predict lower subsequent excess returns features high statistical and economic significance. Importantly, the finding is robust to the inclusion of dividend yields and interest rates. Second, while expected risk premia and dividend yields feature some commonality, marginal significance statistics suggest that they carry different information. Third, differences in the patterns of loadings across maturities suggest that time variation in erx_t^{ST} and dy_t captures high (low) frequency components of future realized returns.

C Bond predictability

To the extent that short term risk premia on short term dividends capture the compensation that investors require on short-term investments, it is reasonable to conjecture that they possess forecasting ability for other asset classes. We test this conjecture by looking at the ability of short term risk premia to predict excess returns on Treasury securities. We consider Treasury securities for two reasons. First, bonds have a fixed maturity, so the decomposition of their returns into short- and long-term shocks to discount factors is easier than for equities. Second, the absence of cash flow risk allows us to focus on the "discount rate risk" component of short term risk premia.

We follow the empirical fixed income literature and run predictive regressions of annual bond excess returns constructed from the Fama-Bliss dataset (from WRDS). Unfortunately, data limitations do not allow us to consider shorter forecasting horizons. The annual excess return on a bond with maturity n is defined as $p_{t+1Y}^{(n-1)} - p_t^{(n)} - y_t^{(1)}$, where $p_t^{(n)}$ denotes the time t log-price of a zero discount bond with maturity in n years, and $y_t^{(n)}$ is the yield associated with maturity n at time t, i.e. $y_t^{(n)} = -(1/n)p_t^{(n)}$. Since n -year forward spreads are known to have predictive power for the annual returns of n-year bonds (Fama and Bliss, 1987), we include forward spreads as controls. The n-year forward spread is defined as $f_t^{(n)} - y_t^{(1)}$, where $f_t^{(n)} = p_t^{(n-1)} - p_t^{(n)}$.

Table V contains the results of the predictive regressions. The loadings on short term risk premia feature high statistical significance for short maturity bonds. For $2-$ and $3-$ years maturity bonds the null hypothesis of no predictability can be rejected at the 5% level of significance; for $4-$ and $5-$ years, on the other hand, erx_t^{ST} seems to lose its predictive power. This pattern of statistical significance is flipped for forward-spot spreads, the classic forecasting variable in the empirical fixed income literature: its loadings are not significant, except for long maturity bonds. Importantly, contrary to what we find in the equity forecasting regressions, the loadings on short term expected risk premia are *positive*. The size of the loadings is economically significant, and decreasing in the bond maturity. Overall, we find that short term risk premia predict bond returns with a positive sign, a result which is consistent with the intuition that high ex-ante risk premia should command, on average, high realized excess returns; also, we find that this result is weaker, both statistically and economically, as we increase the maturity of the bond. These results, coupled with the findings about equity return predictability, suggest that the negative sign in equity forecasting regressions might be traced back to the dynamics of cash flow expectations.

D Robustness

We use current dividends as an instrument for expectations of future dividends. In order to assess the robustness of our results, we use construct short term risk premia by using consensus dividend forecasts from IBES. IBES data are ideal instruments of dividend expectations: they are model-free, forward-looking, and surmount the classic problems features by econometric models (omitted variables and time varying parameters, to name a few). Unfortunately, IBES database is well populated with dividend forecasts on S&P500 firms only from 2004 onwards. Forecasting regressions with quarterly data and 7 years of data are clearly hopeless. In order to increase the number of observations, we sample data at the daily frequency and run regressions for up to 6 monthly horizons. We aggregate firm level one-year forward dividend forecasts to reflect the procedure followed by Standard & Poor's: we use float-adjusted market capitalization weights and standardize the resulting value by the index divisor. In order to obtain dividend prices with constant time to maturity, we interpolate linearly between available contracts. Table VI summarizes the results: the slope coefficients on erx_t^{ST} are statistically significant and negative. Bearing in mind the short span of the sample and the high frequency of the regressions, also these regressions confirm the robustness of our results to potential misspecification of the model of dividend expectations.

V Understanding predictability

The finding that expected risk premia predict realized returns with a negative sign is puzzling as it suggests that higher levels of the price of risk tend to be accompanied by lower realized returns. In this section we make this observation precise by investigating whether any of the leading asset pricing models explored by the literature imply negative forecasting power of short term expected risk premia. We evaluate the implications of three asset pricing models: the reduced form model of Lettau and Wachter (2007), consumption habit (Campbell and Cochrane, 1999), and long run risk (Bansal and Yaron, 2004). The main conclusion of the section is that none of these models is able to generate the predictability patterns that we find in the data.

A Time varying expected dividends

We start by considering the implications of the model by Lettau and Wachter (2007), introduced above as a motivating example. What range of regression coefficients does the model imply? Equation (5) provides a means to answer this question. Unfortunately, the expression requires to compute the derivative of a complex function involving non-linearities and an infinite series, so an analytical expression for β_1 cannot be found. Even though we cannot pin down the magnitude of β_1, we can say something about the loading of expected returns of individual dividend strips on $||\sigma_d||x_t$ via (2). Computation of (2) requires a parametrization of the model. We follow Lettau and Wachter (2007) in part, by considering annual frequency and setting $r^f = 1.93\%$, $g = 2.28\%$, $\phi_z = 0.91$, $\phi_x = 0.87$, $\bar{x} = 0.625$, $||\sigma_d|| = 0.145$, $||\sigma_z|| = 0.0032$, $||\sigma_x|| = 0.24$. We depart from Lettau and Wachter (2007) in the parametrization of σ_d, σ_z, and σ_x. Lettau and Wachter (2007) assume that dividend shocks are correlated with shocks to dividend expectations, but not with shocks to the price of risk. These assumptions, together with the estimates of the conditional volatilities, imply the following specification for $\Sigma = [\sigma_d' \quad \sigma_z' \quad \sigma_x']'$

$$\begin{pmatrix} \sigma_{dd} & 0 & 0 \\ \sigma_{zd} & \sigma_{zz} & 0 \\ 0 & 0 & \sigma_{xx} \end{pmatrix},$$

where

$$\begin{aligned} \sigma_{dd} &= ||\sigma_d|| \\ \sigma_{zd} &= \rho_{z,d}||\sigma_z|| \\ \sigma_{zz} &= \sqrt{||\sigma_z||^2 - \rho_{z,d}^2||\sigma_z||^2} \\ \sigma_{xx} &= ||\sigma_x||, \end{aligned}$$

and

$$\rho_{z,d} = \frac{\sigma_z \sigma_d'}{||\sigma_z||||\sigma_d||}.$$

As Lettau and Wachter (2007) note, the assumption of no conditional correlation between dividend and price of risk shocks is not innocuous and implies assumptions about more primitive features of the economy. Instead of choosing a particular specification, we explore model predictions for a variety of combinations of ρ_{zd} and ρ_{xd}. The characterization differs from Lettau and Wachter (2007) only in the choice of σ_x

$$\sigma_x = [\sigma_{xd} \qquad 0 \qquad \sigma_{xd}],$$

where

$$\begin{aligned} \sigma_{xd} &= \rho_{x,d}||\sigma_z|| \\ \sigma_{xx} &= \sqrt{||\sigma_x||^2 - \rho_{x,d}^2||\sigma_x||^2}. \end{aligned}$$

Not all combinations of ρ_{zd} and ρ_{xd} are allowed: restrictions have to be imposed to ensure that P_t converges for all z_t and x_t and that σ_{zz} and σ_{xx} are real. The appendix explores the derivation of these restrictions in detail. Figure 3 shows which combinations of correlation coefficients ensure the convergence for our choice of parameters. It is interesting to note that for this specific choice of parameter values, the “binding constraint” comes in the form of a lower bound for $\rho_{x,d}$ around -0.3. In figure 4, each panel shows the loadings on x_t of expected returns (equation (2)) for a fixed maturity and different combinations of ρ_{zd} and ρ_{xd}. The figures exhibit clear patterns. The key “driver” of the loadings is the correlation between shocks to the price of risk and to realized dividends. For a fixed combination of ρ_{xd} and ρ_{zd}, the loadings barely change across maturities. Fixing maturity, the loadings decrease from approximately 0.8 for high values of ρ_{xd}, to approximately 0.20 for low values of ρ_{xd}. The sign of the loadings is positive for all maturities: an increase in the price of risk raises expected excess returns at all horizons. Since the expected return on the index is a weighted sum of the expected returns on its dividend strips, the graphs suggest that a regression of realized excess returns on expected risk premia should yield a positive slope coefficient.

B Long run risk

In the appendix we show that the model explored by Bansal and Yaron (2004) implies that the risk premium on short term dividends is equal to zero: long run risk models rule out the possibility that erx_t has any predictive power with regards to future returns. The intuition of this result is simple. In long run risk models, agents are fearful about states of the world in which expected growth prospects are low and economic uncertainty is high. A short maturity dividend claim is exposed to the risk that realized dividends undershoot expectations, but it is virtually not affected by long run growth prospects and economic uncertainty. Since shocks to realized dividends are not priced in this economy, agents demand zero risk premia to hold short maturity dividend claims.

C Habit formation

Also a simple habit formation model will find it difficult to explain the data. As habit implies that dividend shocks are negatively correlated with innovations to the price of risk, the model can be (somewhat) casted in the context of Lettau and Wachter (2007) specification with $\rho_{xd} < 0$. In this case, however, the lower the correlation between realized dividends and the price of risk, the higher the implied slope coefficient.

The asset pricing models presented above fail to capture the predictability we find in the data. The mechanics of the failures are precious because they shed light on the direction of improvement that future models will have to pursue to match the empirical evidence of short term predictability. Bansal and Yaron (2004) on one side, and Lettau and Wachter (2007) and Campbell and Cochrane (1999) on the other, fail to capture predictability for two different reasons. Long run risk models deny the existence of predictability conditional on erx_t^{ST} because short term dividends are not exposed to the two sources of priced risk of the economy, i.e. long run economic uncertainty and growth prospects. The reduced form model by Lettau and Wachter (2007), on the other hand, does allow for predictability conditional on erx_t^{ST}, but implies a positive slope coefficient. The intuition behind this result is simple. Expected returns are increasing in the price of risk, which is revealed, up to a scaling factor, by the expected returns on the short term asset (via equation (2)). If the price of risk is the only state variable driving the term structure of expected returns and follows a persistent process, an increase in short term expected risk premia (i.e. the price of risk) will raise expected excess returns at all horizons, yielding a positive slope coefficient in regressions of realized returns on erx_t^{ST}. Hence, we learn that two ingredients have to co-exist in a model for it have a chance to rationalize our empirical findings. First, shocks to realized dividends must be priced. Second, the term structure of expected risk premia must be driven by a multivariate state vector.

D Epilogue

Having shown that leading asset pricing are unable to explain our empirical findings, we study the implications for the dynamics of agents' subjective expectations. In running the predictive regressions above

$$\frac{P_{t+1} + D_{t+1}}{P_t} - R_t^f = \alpha + \beta \left(E_t[R_{1,t+1}] - R_t^f \right) + \epsilon_{t+1},$$

we have implicitly assumed that the subjective measure of the representative agent used in the computation of $E_t[R_{1,t+1}]$ coincides with the objective measure. Consider now what happens if the subjective and objective measures differ. In particular, let $E_t^*[\cdot]$ denote the expectation operator associated with the subjective probability measure of the representative agent. Expectations under the subjective and physical measure are related by the equation

$$E_t^*[R_{1,t+1}] = E_t[\eta_{t+1} R_{1,t+1}],$$

where η_{t+1} is the Radon-Nikodym derivative (dQ^*/dQ) of the subjective measure with respect to the physical measure. If we allow for subjective probabilities to differ from physical probabilities, the regression we run should be written as

$$\frac{P_{t+1} + D_{t+1}}{P_t} - R_t^f = \alpha + \beta \left(E_t[\eta_{t+1} R_{1,t+1}] - R_t^f \right) + \epsilon_{t+1}.$$

Under the assumption that realized returns are i.i.d. under the objective measure, and that agents are overly optimistic about short term expected returns in good times, i.e., when stock excess return are high ($\frac{\partial \eta_{t+1}}{\partial rx_t} > 0$), the regression can yield a negative coefficient. To see the intuition behind this claim, let us consider the case where expected returns are zero. Suppose that realized excess return at time t are positive. Since expected excess returns are zero, this realization is just a "lucky draw". The agent, however, upon observing $rx_t > 0$, assigns more probability weight to states with high returns (i.e. $\frac{\partial \eta_{t+1}}{\partial rx_t} > 0$), thus raising the expected risk premium via $erx_t = \log(R_{1,t+1}) - \log(R_t^f)$. Since the average return realization at time $t+1$ is zero, regressions of realized returns on expected risk premia can generate a negative coefficient under these assumptions.

VI Conclusion

We construct short horizon expected risk premia by combining dividend prices with dividends expectations, and find evidence that expected risk premia capture short horizon predictability in $S\&P500$ realized excess returns. The results are robust to the inclusion of dividend yields in the conditioning information set. The slope coefficients of the predictive regressions are difficult to reconcile with the predictions of leading asset pricing models. The failure of these models may be the consequence of their inability to produce a multi-factor structure for the term structure of equity risk premia and/or to the fact that they fail to capture the slow moving dynamics of risk premia.

VII Appendix

A Parameter restrictions for Lettau and Wachter (2007)

We impose two restrictions in order to derive the set of admissible combinations of $\rho_{z,d}$ and $\rho_{x,d}$: P_t must converge for all z_t and x_t, and σ_{zz} and σ_{xx} must be real. Lettau and Wachter (2007) show that the price of equity converges to a finite value if and only if the following three conditions are satisfied:

1. $|\phi_z| < 1$.

2. $|\phi_x - \sigma_x \frac{\sigma_d}{||\sigma_d||}| < 1$.

3. $-r + g + \bar{B}_x(1-\phi_x)\bar{x} + \frac{1}{2}\bar{V}\bar{V}' < 0$.

where

$$\bar{B}_x = -\frac{\left(\sigma_d + \frac{\sigma_z}{1-\phi_z}\right)\frac{\sigma_d}{||\sigma_d||}}{1-\left(\phi_x - \frac{\sigma_x\sigma_d}{||\sigma_d||}\right)},$$

and

$$\bar{V} = \sigma_d + \frac{\sigma_z}{1-\phi_z} + \bar{B}_x\sigma_x.$$

Condition 1 is satisfied because we set $\phi_z = 0.91$. Condition 2 can be rewritten as follows

$$|\phi_x - \rho_{xd}||\sigma_x||| < 1 \rightarrow \left|\rho_{xd} - \frac{\phi_x}{||\sigma_x||}\right| < \frac{1}{||\sigma_d||},$$

which implies the following condition for ρ_{xd}

$$\rho_{xd} = \left(\frac{\phi_x}{||\sigma_x||} - \frac{1}{||\sigma_d||}, \quad \frac{\phi_x}{||\sigma_x||} + \frac{1}{||\sigma_d||}\right).$$

For the set of admissible ρ_{xd} derived above, we evaluate condition 3 numerically and rule out the ρ_{zd} that do not satisfy it.

We assume that $\Sigma = [\sigma_d' \quad \sigma_z' \quad \sigma_x']'$ equals:

$$\begin{pmatrix} \sigma_{dd} & 0 & 0 \\ \sigma_{zd} & \sigma_{zz} & 0 \\ \sigma_{xd} & 0 & \sigma_{xx} \end{pmatrix}.$$

The unknown parameters can be computed from the knowledge of the conditional volatilities and pairwise correlations. We have:

1. $\sigma_{dd} = ||\sigma_d||$.
2. Since $\rho_{zd} = \frac{\sigma_z \sigma_d'}{||\sigma_z|| \cdot ||\sigma_d||} = \frac{\sigma_{zd}}{||\sigma_z||}$, we have $\sigma_{zd} = \rho_{zd}||\sigma_z||$.
3. $\sigma_{zz} = \sqrt{||\sigma_z||^2 - \sigma_{zd}^2}$.
4. The same argument as point 2 leads to: $\sigma_{xd} = \rho_{xd}||\sigma_x||$.
5. $\sigma_{xx} = \sqrt{||\sigma_x||^2 - \sigma_{xd}^2}$.

Points 3 and 5 show that ρ_{zd} and ρ_{xd} must be such that the argument of the square roots is positive: we rule out all combinations of correlations that do not satisfy this condition.

B Short horizon expected risk premia in Bansal and Yaron (2004)

The model specifies a process for (log) dividend growth, (log) consumption growth, the conditional expected growth rate, and economic uncertainty:

$$\begin{aligned}
\Delta d_{t+1} &= \mu_d + \phi z_t + w_d \sigma_t \epsilon_{d,t+1} \\
\Delta c_{t+1} &= \mu_c + z_t + w_c \sigma_t \epsilon_{c,t+1} \\
z_{t+1} &= \phi_z z_t + w_z \sigma_t \epsilon_{z,t+1} \\
\sigma_{t+1}^2 &= \mu_\sigma + v\left(\sigma_t^2 - \sigma^2\right) + w_\sigma + \epsilon_{\sigma,t+1},
\end{aligned}$$

where $\epsilon_{d,t+1}$, $\epsilon_{c,t+1}$, $\epsilon_{z,t+1}$, and $\epsilon_{\sigma,t+1}$ are i.i.d. $N(0,1)$. The equilibrium log return on the market asset is given by

$$r_{t+1} = \Delta d_{t+1} + k_1 A_{1,m} z_{t+1} - A_{1,m} z_t + k_{1,m} A_{2,m} \sigma_{t+1}^2 - A_{2,m} \sigma_t^2, \tag{11}$$

where a complete characterization of $A_{1,m}$, $A_{2,m}$, k_1 and $k_{1,m}$ is provided by Bansal and Yaron (2004). The expected risk premium on the first dividend strip erx_t^{ST} is given by $\log E_t[D_{t+1}] - \log P_{t,1} - r_t^f$. Let M_{t+1} denote the stochastic discount factor between t and $t+1$, i.e. $P_{t,1} = E_t[M_{t+1} D_{t+1}]$. By the properties of normal variables, we have

$$\begin{aligned} \log E_t[D_{t+1}] =& \log D_t E_t[\exp(\Delta d_{t+1})] \\ =& d_t + E_t[\Delta d_{t+1}] + 0.5 Var_t[\Delta d_{t+1}] \end{aligned}$$

$$\begin{aligned} \log P_{t,1} =& \log D_t E_t[\exp(m_{t+1} + \Delta d_{t+1})] \\ =& d_t + E_t[m_{t+1}] + E_t[\Delta d_{t+1}] \\ & + 0.5\big(Var_t[m_{t+1}] + Var_t[\Delta d_{t+1}] + 2Cov_t[m_{t+1}, \Delta d_{t+1}]\big) \end{aligned}$$

$$\begin{aligned} r_t^f =& -\log E_t[\exp(m_{t+1})] \\ =& -E_t[m_{t+1}] - 0.5 Var_t[m_{t+1}]. \end{aligned}$$

Hence, the expression for erx_t^{ST} simplifies to

$$\log E_t[D_{t+1}] - \log P_{t,1} - r_t^f = -2Cov_t[m_{t+1}, \Delta d_{t+1}],$$

i.e. the expected risk premium depends on the covariance between m_{t+1} and Δd_{t+1}. The innovations to the SDF are equal to

$$m_{t+1} - E_t[m_{t+1}] = \lambda_{mc}\sigma_t\epsilon_{c,t+1} - \lambda_{mz}\sigma_t\epsilon_{z,t+1} - \lambda_{m\sigma}\sigma_t\epsilon_{\sigma,t+1},$$

where expressions for the prices of risk λ_{mc}, λ_{mz}, $\lambda_{m\sigma}$ are given in the original paper. Since innovations to Δd_{t+1} are equal to

$$\Delta d_{t+1} - E_t[\Delta d_{t+1}] = w_d\sigma_t\epsilon_{d,t+1},$$

it follows that $Cov_t[m_{t+1}, \Delta d_{t+1}] = 0$ and $erx_t^{ST} = 0$.

References

[1] A. Ang and G. Bekaert. 2007. Stock Return Predictability: Is it There? Review of Financial Studies 20(3), pages 651-707.

[2] M. Baker and J. Wurgler. 2006. Investor Sentiment and the Cross-Section of Stock Returns. Journal of Finance, vol. 61(4), pages 1645-1680.

[3] R. Bansal and A. Yaron. 2004. Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles. Journal of Finance, vol. 59(4), pages 1481-1509.

[4] J. H. Van Binsbergen, M. W. Brandt, R. S. J. Koijen. 2010. On the Timing and Pricing of Dividends. American Economic Review, Forthcoming.

[5] J. H. Van Binsbergen, W. Hueskes, R. S. J. Koijen, and E. B. Vrugt, 2011. Equity Yields. NBER Working Papers 17416, National Bureau of Economic Research, Inc.

[6] J. Y. Campbell. 1991. A Variance Decomposition for Stock Returns. The Economic Journal, Vol. 101, pages 157-179.

[8] J. Y. Campbell and J. H. Cochrane. 1999. By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior. Journal of Political Economy, Vol. 107, No. 2: pp. 205-251.

[9] J. Y. Campbell, and R. J. Shiller. 1988. The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors. Review of Financial Studies, Oxford University Press for Society for Financial Studies, vol. 1(3), pages 195-228.

[7] J.H. Cochrane. 1992. Explaining the variance of price-dividend ratios. Review of Financial Studies, 5(2), pages 243-280.

[10] J. H. Cochrane. 2005. Asset pricing, revised edition. Princeton University Press.

[18] Cochrane, John H., 2006. The Dog That Did Not Bark: A Defense of Return Predictability. Review of Financial Studies 21, 1533-1575.

[11] Fama, E.F., and R.R. Bliss, 1987. The information in long-maturity forward rates. The American Economic Review, 77, p. 680?692.

[12] Lars Peter Hansen and Robert J. Hodrick. 1980. Forward Exchange Rates as Optimal Predictors of Future Spot Rates: An Econometric Analysis. Journal of Political Economy, 1980, 88(5), pp. 829-53.

[13] R. J. Hodrick. 1992. Dividend yields and expected stock returns: alternative procedures for inference and measurement. Review of Financial Studies (1992) 5 (3): 357-386.

[14] M. Lettau and J. A. Wachter. 2007. Why is long-horizon equity less risky? a duration-based explanation of the value premium. Journal of Finance, 62(1):55-92.

[15] Newey, Whitney K; West, Kenneth D. 1987. A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. Econometrica 55 (3): 703–708.

[16] Phillips, P.C.B and P. Perron. 1988. Testing for a Unit Root in Time Series Regression. Biometrika, 75, 335–346.

[17] Stambaugh, Robert F. 1999. Predictive regressions. Journal of Financial Economics, vol. 54(3), pages 375-421.

Table I – Expected risk premia: sample statistics

Mean	Stdev	AR(1)
0.31	0.26	0.55

Table II – Univariate forecasting regressions

This table reports estimates from OLS univariate regressions of $S\&P500$ $(t \rightarrow t+h)$ excess returns on time t expected risk premia: $rx_{t+h} = \beta_0^{(h)} + \beta_1^{(h)} erx_t^{ST} + \epsilon_{t+h}^{(h)}$. The holding period h ranges from one to four quarters. For each horizon considered, the first row (in grey) contains OLS estimates of the coefficients and the adjusted R^2. The rows below feature the t-statistics of the coefficients, computed with NW and HH errors respectively. All errors are estimated with h lags. Two (one) stars indicate significance at the 5% (10% significance level.

$Horizon(h)$	$const$	erx_t^{ST}	$\bar{R}^2$
1Q	-0.00	-0.16	1.37%
	-0.00	-1.61	
	-0.00	-1.61	
2Q	0.00	-0.23	4.31%
	0.00	-2.19**	
	0.00	-1.98**	
3Q	0.00	-0.30	7.82%
	0.00	-3.45**	
	0.00	-3.14**	
4Q	-0.00	-0.34	10.91%
	-0.00	-2.90**	
	-0.00	-2.49**	

Table III – Bivariate forecasting regressions without expected risk premia

This table reports estimates from OLS univariate regressions of $S\&P500$ $(t \rightarrow t+h)$ excess returns on time t (log) dividend yield and the 3-month (log) Treasury rate: $rx_{t+h} = \beta_0^{(h)} + \beta_1^{(h)} dy_t + \beta_2^{(h)} rf_t + \epsilon_{t+h}^{(h)}$. The holding period h ranges from one to four quarters. For each horizon considered, the first row (in grey) contains OLS estimates of the coefficients and the adjusted R^2. The rows below feature the t-statistics of the coefficients, computed with NW and HH errors respectively. All errors are estimated with h lags. Two (one) stars indicate significance at the 5% (10% significance level.

$Horizon(h)$	$const$	dy_t	rf_t	$\bar{R}^2$
1Q	0.00	0.15	0.01	0.28%
	0.00	1.40	0.06	
	0.00	1.40	0.06	
2Q	0.00	0.23	-0.02	3.01%
	0.00	1.90*	-0.12	
	0.00	1.66*	-0.11	
3Q	0.00	0.30	-0.05	6.41%
	0.00	2.22**	-0.29	
	0.00	1.85*	-0.26	
4Q	0.00	0.36	-0.09	10.08%
	0.00	2.42**	-0.58	
	0.00	1.99**	-0.53	

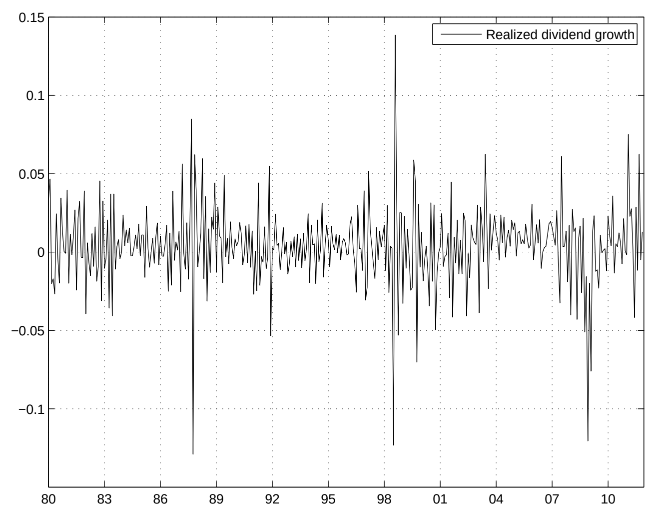

Figure 1 – Log Dividend Growth

Table IV – Bivariate forecasting regressions with expected risk premia

This table reports estimates from OLS univariate regressions of $S\&P500$ $(t \to t+h)$ excess returns on time t expected risk premia, (log) dividend yield, and 3-month Treasury (log) rate: $rx_{t+h} = \beta_0^{(h)} + \beta_1^{(h)} erx_t^{ST} + \beta_2^{(h)} dy_t + \beta_3^{(h)} rf_t + \epsilon_{t+h}^{(h)}$. The holding period h ranges from one to four quarters. For each horizon considered, the first row (in grey) contains OLS estimates of the coefficients and the adjusted R^2. The rows below feature the t-statistics of the coefficients, computed with NW and HH errors respectively. All errors are estimated with h lags. One and two stars indicate significance at the 10% and 5% significance level respectively.

h	$const$	erx_t^{ST}	dy_t	rf_t	$\bar{R}^2$
1Q	0.00	-0.12	0.09	0.04	0.25%
	0.00	-1.21	0.78	0.32	
	0.00	-1.21	0.78	0.32	
2Q	0.00	-0.18	0.15	0.02	4.36%
	0.00	-1.75*	1.25	0.13	
	0.00	-1.61	1.11	0.11	
3Q	0.00	-0.23	0.20	-0.02	8.98%
	0.00	-2.54**	1.41	-0.15	
	0.00	-2.40**	1.24	-0.14	
4Q	0.00	-0.27	0.26	-0.10	14.43%
	0.00	-2.50**	1.85*	-0.70	
	0.00	-2.30**	1.64*	-0.66	

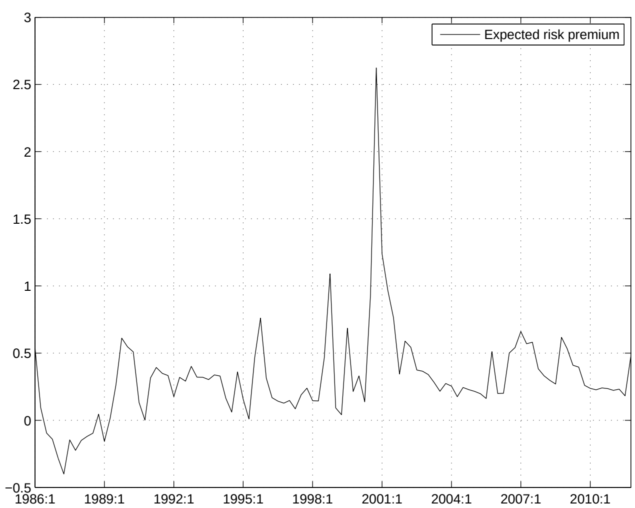

Figure 2 – Expected risk premia

Table V – Bond forecasting regressions

This table reports estimates from OLS univariate regressions of n-year Treasury securities annual ($(t \rightarrow t+4)$) excess returns on time t expected risk premia and the n-year forward spot spread $\left(f_t^{(n)} - y_t^{(1)}\right)$: $rx_{t+4} = \beta_0^{(n)} + \beta_1^{(n)} erx_t^{ST} + \beta_2^{(n)} \left(f_t^{(n)} - y_t^{(1)}\right) + \epsilon_{t+h}^{(h)}$. The bond maturity n ranges from two to five years. For each maturity considered, the first row (in grey) contains OLS estimates of the coefficients and the adjusted R^2. The rows below feature the t-statistics of the coefficients, computed with NW and HH errors respectively. All errors are estimated with 6 lags (equivalent to 1.5 years). One and two stars indicate significance at the 10% and 5% significance level, respectively.

$Bondmaturity(n)$	$const$	erx_t^{ST}	$\left(f_t^{(n)} - y_t^{(1)}\right)$	$\bar{R}^2$
2 years	0.00	0.30	0.18	7.88%
	0.00	2.57**	1.22	
	0.00	2.22**	1.48	
3 years	-0.00	0.26	0.17	6.96%
	-0.00	2.18**	1.18	
	-0.00	1.85*	1.61	
4 years	0.00	0.22	0.22	6.78%
	0.00	1.84*	1.55	
	0.00	1.58	2.22**	
5 years	0.00	0.17	0.18	3.73%
	0.00	1.44	1.38	
	0.00	1.24	1.98**	

Table VI – Univariate forecasting regressions (daily frequency)

This table reports estimates from OLS univariate regressions of $S\&P500$ $(t \rightarrow t+h)$ excess returns on time t one-year-forward expected risk premia: $rx_{t+h} = \beta_0^{(h)} + \beta_1^{(h)} erx_t + \epsilon_{t+h}^{(h)}$. The holding period h ranges from one to six months. For each horizon considered, the first row (in grey) contains OLS estimates of the coefficients and the adjusted R^2. The rows below feature the t-statistics of the coefficients, computed with NW and HH errors respectively. All errors are estimated with h lags. One (two) stars indicate significance at the 0.05 (0.01) level.

$Horizon$ (h)	$Constant$	erx_t^{ST}	$\bar{R}^2$
1 month	0.00	-0.03	1.17%
	0.59	-1.26	
	0.51	-1.06	
2 months	0.01	-0.07	4.37%
	1.25	-1.88	
	1.01	-1.60	
3 months	0.02	-0.11	7.37%
	1.30	-2.16*	
	1.05	-1.83	
4 months	0.03	-0.14	8.02%
	1.21	-2.25*	
	1.00	-2.00*	
5 months	3.32	-0.17	8.01%
	1.22	-2.37*	
	1.08	-2.41*	
6 months	0.04	-0.21	8.30%
	1.27	-2.59**	
	1.21	-3.16**	

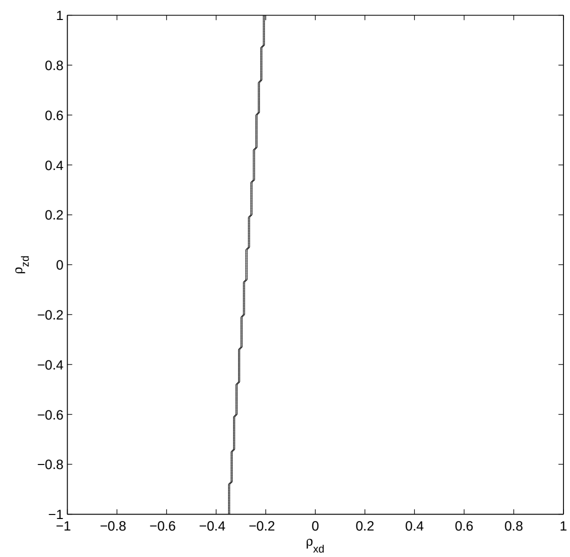

Figure 3 – Admissible correlations

The region on the right of the line is the set of $(\rho_{x,d}, \rho_{z,d}) = \left(\frac{\sigma_x \sigma_d'}{||\sigma_x||||\sigma_d||}, \frac{\sigma_z \sigma_d'}{||\sigma_z||||\sigma_d||}\right)$ that ensure that: i) the price to dividend ratio of dividend strips converges for any maturity, and ii) the values of σ_{zz} and σ_{xx} are real.

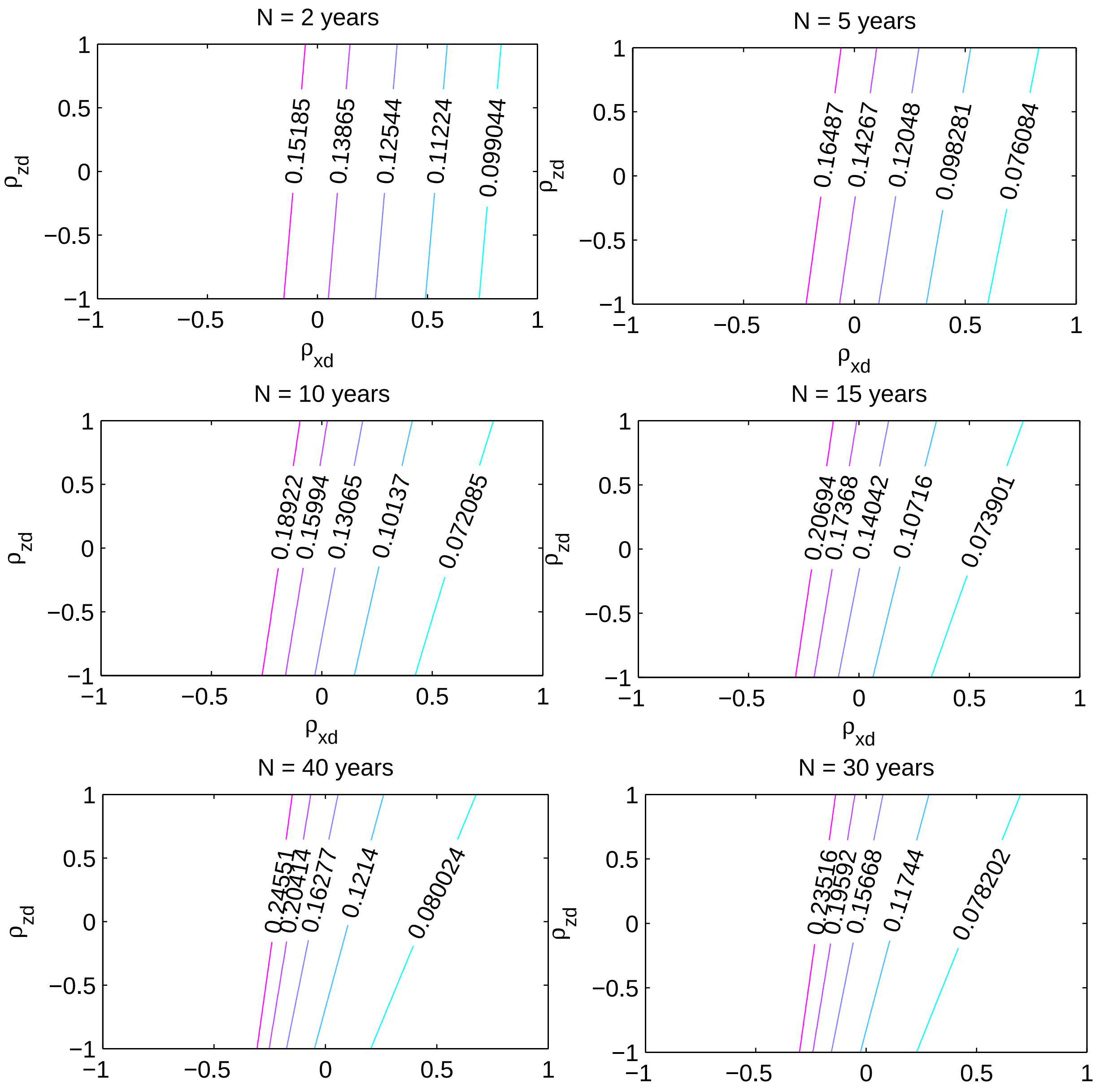

Figure 4 – Loadings of expected returns on x_t

The graphs illustrate the loadings of expected returns of dividend strips of maturity N on the price of risk x_t for a variety of $(\rho_{x,d}, \rho_{z,d})$ combinations. The lines are contours connecting the $(\rho_{x,d}, \rho_{z,d})$ pairs with the same loading x_t; the number on each contour indicates its level.