Should Investors Learn about the Timing of Equity Risk?*

Michael Hasler† Mariana Khapko‡ Roberto Marfè§

August 4, 2017

Accepted in the Journal of Financial Economics

Abstract

The term structure of equity risk has been documented to be downward-sloping. We capture this feature using return dynamics driven by both a transitory and a permanent component. We study the asset allocation and portfolio performance when transitory and permanent components cannot be observed, and therefore need to be estimated. Strategies that account for the observed timing of equity risk outperform those that do not, particularly so out-of-sample. Indeed, the mean (median) certainty equivalent return increases from about 13% (12%) to about 21% (15%) because properly modeling the timing of equity risk implies surges in portfolio returns.

*We would like to thank an anonymous referee and G. William Schwert (editor) for providing insightful comments that helped significantly improve the quality of the paper. We are also grateful for suggestions received from Pat Akey, Daniel Andrei, Tomas Björk, Olivier Dessaint, Elisa Luciano, Chayawat Ornthanalai, Julien Penasse, Mike Simutin, and Andrea Tamoni. Hasler thanks the University of Toronto and especially the Montreal Institute of Structured Finance and Derivatives (IFSID) for its generous financial support. Marfè and Khapko thank the Collegio Carlo Alberto and the University of Toronto, respectively, for their financial support.

†University of Toronto, 105 St. George Street, Toronto, ON, M5S 3E6, Canada, Michael.Hasler@rotman.utoronto.ca; www.utoronto.hasler

‡University of Toronto, 105 St. George Street, Toronto, ON, M5S 3E6, Canada, Mariana.Khapko@rotman.utoronto.ca; www.utoronto.khapko

§Collegio Carlo Alberto, Via Real Collegio, 30, 10024 Moncalieri (Torino), Italy, roberto.marfe@carloalberto.org; http://robertomarfe.altervista.org/

1 Introduction

Empirical research has recently provided evidence of a downward-sloping term structure of *equity risk* (van Binsbergen, Brandt, and Koijen, 2012). That is, short-horizon equity returns are riskier than long-horizon ones. This result is groundbreaking because it challenges most models in the asset pricing and portfolio selection literature. Indeed, asset-pricing models typically consider a *permanent component* with stochastic drift and stochastic volatility in consumption growth (Bansal and Yaron, 2004), whereas portfolio selection models consider the same features for equity returns (Kim and Omberg, 1996; Wachter, 2002; Chacko and Viceira, 2005). These two features imply an upward-sloping term structure of equity risk.

In this paper, we do not investigate the foundations of downward-sloping equity risk. Instead, we focus on its implications for investors. Namely, we provide a simple and flexible model of equity returns that allows the slope of the term structure of equity risk to be either positive or negative. The key ingredient generating the downward-sloping effect is the *transitory component* of equity returns, which is ignored by most models.[1] This transitory component significantly increases equity risk in the short term and has little influence on it in the long term because it is mean-reverting. When the transitory component reverts sufficiently rapidly, its downward-sloping effect dominates the upward-sloping effect implied by both the stochastic drift and volatility in the permanent component, generating a downward-sloping term structure of equity risk.

We show that properly modeling the shape of the term structure of equity risk is important to investors because it improves the performance of their investment strategies. Indeed, the *out-of-sample* mean and median certainty equivalent returns of a strategy that considers the transitory component are about 21% and 15%, respectively, whereas those of a strategy that ignores the transitory component are about 13% and 12%. Accounting for the transitory component has no significant impact on the left tail of the distribution of portfolio returns, whereas it markedly lengthens and fattens its right tail thereby providing investors with more attractive returns on their wealth. Importantly, we show that the performance advantages of learning about the timing of equity risk are larger in market downturns.

To investigate the implications of learning about the timing of equity risk, we consider a dynamic portfolio choice problem of an investor who has CRRA preferences and maximizes the expected utility of terminal wealth. The investor can invest in two securities: one stock and one risk-free asset. Stock returns feature two standard ingredients, a *permanent component* with stochastic drift (Kim and Omberg, 1996) and stochastic volatility (Heston, 1993). In addition, returns are driven by a *transitory component* that allows the term structure of equity risk to be potentially downward-sloping. Since the investor observes neither the permanent nor the transitory component, she filters them out by combining Bayesian updating techniques with the observation of past returns. Learning dynamically about the relative importance of the two components allows the investor to

[1] A few exceptions that consider a transitory component in either equity returns or dividend growth are Fama and French (1988) and, more recently, Menzly, Santos, and Veronesi (2004), Bansal, Kiku, and Yaron (2010), Greenwald, Lettau, and Ludvigson (2014), and Marfè (2015a, 2016).

account for changes in horizon-specific risk levels.

To understand the determinants of the timing of risk, we first discuss the general relation between the term structure of holding period return (HPR) volatility and the term structure of equity yield volatility, both of which are measures of the timing of equity risk.[2] We then provide model-specific conditions that generate downward-sloping term structures of HPR volatility and equity yield volatility. We show that downward-sloping term structures of equity risk can only be obtained in the presence of a sufficiently volatile transitory component.

We fit the model to historical monthly S&P 500 returns. Estimated parameter values show that the transitory component is persistent, that the stochastic drift of the permanent component is a quickly reverting and highly volatile process, and that the model accurately replicates the observed downward-sloping term structures of S&P 500 HPR volatility and equity yield volatility. Furthermore, we provide evidence that the transitory component strongly negatively covaries with the S&P 500 dividend yield (65% R^2), whereas the stochastic drift does so to only a limited extent (5% R^2). Since the dividend yield is a persistent forward-looking measure of stock returns and dividend growth (Cochrane, 2008; van Binsbergen and Koijen, 2010), so is the transitory component. As a result, models accounting for the transitory component can replicate the observed predictive power of the dividend yield, whereas models abstracting from the transitory component cannot.

Using estimated parameter values, we focus on portfolio allocation. We start by investigating the properties of investor's hedging demands. Since returns are negatively correlated with volatility and positively correlated with the stochastic drift of the permanent component, both variables command a negative hedging term. The reason is that, in bad times (i.e. when returns are low), volatility is high and the expected return tends to be low because the stochastic drift is low. Since our investor is more risk averse than a log-utility investor, she hedges these two risks by reducing her investment in the stock. In contrast, the transitory component commands a positive hedging term because it is positively correlated with returns and it negatively impacts expected returns. Indeed, the transitory component is small in bad times and therefore implies a high expected return, which our investor exploits by increasing her position in the stock.

As the investor's investment horizon increases, the size of the hedging terms tends to increase. The reason is that the model-implied term structure of equity risk features two upward-sloping components and one downward-sloping component. The upward slope is implied by the stochastic drift and stochastic volatility, whereas the downward slope is implied by the transitory component. Since the stochastic drift and volatility generate more risk in the long term than in the short term, the investor hedges these two risks more when her horizon is long. At the same time, the investor exploits the small amount of long-term risk implied by the transitory component and therefore has the incentive to increase her risky position when her horizon increases.

We compare the performance of investment strategies that account for the transitory component of returns to that of those ignoring it. We show that modeling the transitory component of returns

[2]The former and latter measures have been empirically analyzed by Campbell and Viceira (2005) and van Binsbergen, Hueskes, Koijen, and Vrugt (2013), respectively.

has a minor impact on the Sharpe ratio of portfolio returns but significantly increases their skewness, good relative to bad volatility, and good relative to bad kurtosis[3]. In addition, we provide evidence that modeling the transitory component is particularly beneficial to the investor who takes into account stochastic return volatility. These results hold both in-sample and out-of-sample and are robust to changes in the investment horizon and risk aversion.

To provide a utility-based measure of the benefits of modeling the transitory component, we compare the mean certainty equivalent return (CER)[4] of strategies considering the transitory component to those ignoring it. Both in-sample and out-of-sample, accounting for the transitory component weakly impacts the CER when return volatility is assumed to be constant. However, when return volatility is assumed to be stochastic, including the transitory component increases the median CER by about 10% and 20% in-sample and out-of-sample, respectively. Since time-varying return volatility is a well-known feature of the data (Engle, 1982; Bollerslev, 1986), our results show that ignoring the transitory component of returns is particularly costly for investors. Furthermore, we show that accounting for the transitory component increases the CER more when the coefficient of relative risk aversion is small and the investment horizon is long. The reason is that the investor's position in the risky asset is large when her risk aversion is small, but also when her horizon is long because there is less return volatility in the long run than in the short run.

Finally, we provide evidence for and explore the implications of the fact that the observed term structure of equity risk is more downward-sloping in bad times than in good times, which is consistent with the findings by van Binsbergen et al. (2013) and Aït-Sahalia, Karaman, and Mancini (2015). In our out-of-sample analysis, the investor re-estimates the parameters of the model every year and therefore captures this feature of the data when considering the transitory component in her model but not when ignoring it. As a result, the improvement in portfolio performance implied by modeling the transitory component is expected to be concentrated in bad times. We confirm this statement by showing that the unconditional benefits of considering the transitory component, in terms of both portfolio return moments and certainty equivalent returns, mostly come from the significantly better performance of such strategies in recessions.

This paper builds on the influential empirical studies of van Binsbergen et al. (2012) and van Binsbergen et al. (2013), which show that the term structures of equity risk and return are downward-sloping. Similarly, Dechow, Sloan, and Soliman (2004) provide evidence of a positive (short) duration premium in the cross-section and, more recently, Weber (2016) and Marfè (2015b) offer additional empirical evidence. Since the most successful asset-pricing models (e.g. Campbell and Cochrane, 1999; Bansal and Yaron, 2004; Gabaix, 2012; Wachter, 2013) imply markedly upward-sloping term structures of equity risk and return, the findings of van Binsbergen, Koijen, and co-authors have highlighted new asset-pricing puzzles that researchers have focused on since

[3]Good and bad volatility (resp. good and bad kurtosis) is defined as the volatility (resp. kurtosis) conditional on returns being larger than their mean (Segal, Shaliastovich, and Yaron, 2015).

[4]We follow Johannes, Korteweg, and Polson (2014) and define the certainty equivalent return as the yield of a fictitious riskless asset that makes the investor indifferent between implementing her optimal risky strategy and investing her entire wealth in this riskless asset.

then. Among these researchers, Belo, Collin-Dufresne, and Goldstein (2015), Croce, Lettau, and Ludvigson (2015), Hasler and Marfè (2016), and Marfè (2016) provide foundations for the observed shape of the term structures of equity risk and return based on financial leverage, learning, economic recoveries, and labour rigidity, respectively. Refer to van Binsbergen and Koijen (2017) for an excellent survey of this research stream.

The aforementioned literature on the term structure of equity is silent about the implications of the timing of risk on asset allocation and portfolio performance. The goal of our paper is therefore to investigate the latter. In a related paper, Campbell and Viceira (2005) use a VAR model to study asset returns across different investment horizons. They show that the term structure of holding period bond return volatility is less downward-sloping than that of equity return volatility, whereas the term structure of holding period T-bill return volatility is upward-sloping. Based on this evidence, Campbell and Viceira point out that the myopic mean-variance portfolio (Markowitz, 1952) is not optimal for long-horizon investors. We depart from the static mean-variance problem discussed in Campbell and Viceira (2005) along a number of important dimensions. In our paper, we propose a model of stock returns that allows for a flexible shape of the term structure of equity risk. We then study a full-fledged dynamic learning and portfolio choice problem of a CRRA investor. This setting allows us to discuss in- and out-of sample higher order moments and utility-based measures for a comprehensive examination of the effects of the timing of equity risk on portfolio performance.

Our paper is also related to the dynamic asset allocation literature with learning. Xia (2001) investigates the optimal portfolio allocation of an investor who faces uncertainty about return predictability. That is, the investor knows that some observable variable predicts future stock returns but is uncertain about the strength of the relation and therefore learns about it. Johannes et al. (2014) extend the setting of Xia (2001) to allow for parameter uncertainty and show that the out-of-sample performance of the associated strategy can be significantly better than that obtained with a model assuming constant mean and constant volatility of returns. We contribute to this literature by first showing that simple learning models without predictive variables generate significantly better out-of-sample portfolio performance than that of the constant mean and constant volatility model. More importantly, we provide evidence that learning about the shape of the term structure of equity risk significantly improves portfolio performance in terms of both portfolio return moments and certainty equivalent returns. These benefits turn out to be particularly large in bad times and under the assumption of stochastic return volatility.

The remainder of the paper is organized as follows. Section 2 provides empirical support for the assumptions of the model; Section 3 describes the model; Section 4 estimates the parameters of the model; Section 5 discusses the results; and Section 6 concludes.

2 Empirical Support

This section investigates equity risk across different investment horizons. Time-variation in stock returns leads to horizon-specific risk levels. This timing of equity risk is informative about the role of the transitory and permanent components of equity returns. We document movements in the slope of the timing of equity risk and their relation with the dividend yield.

We measure the timing of equity risk by computing the term structure of holding period stock return variance ratios (VRs hereafter), in other words, the ratio of the annualized return variance at horizon τ relative to the annualized return variance at the reference horizon τ_0:

$$\mathrm{VR}(\tau) = \frac{\mathrm{Var}(R_{t\to t+\tau})/\tau}{\mathrm{Var}(R_{t\to t+\tau_0})/\tau_0}.$$

VRs capture whether or not the variance of a given variable increases linearly with the observation interval. Therefore, upward-sloping (downward-sloping) VRs that are above (below) unity imply that risk concentrates at long (short) horizons. If the stock return variation is mostly due to shocks that have a permanent (transitory) impact, we expect increasing (decreasing) VRs. If stock returns are either close to i.i.d. or the permanent and transitory components offset each other, we expect approximatively flat VRs.

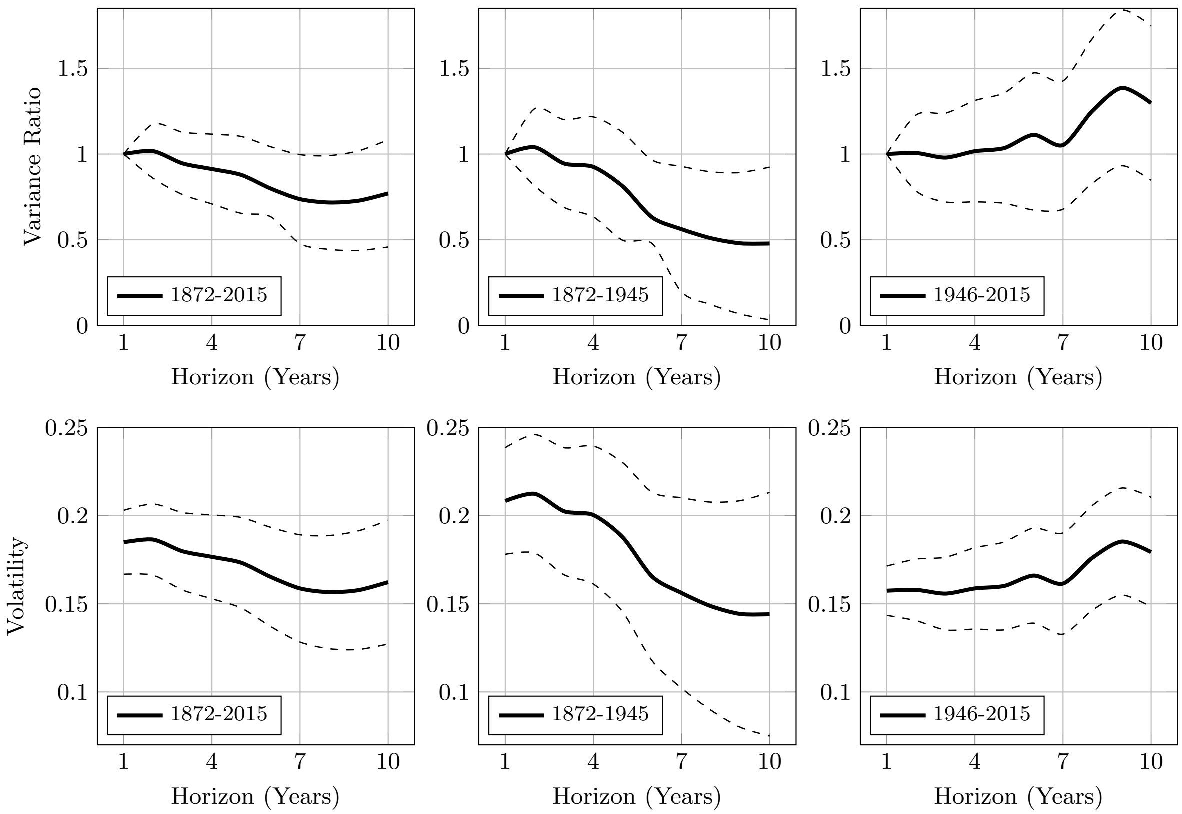
We consider monthly S&P 500 nominal log-returns from 01/1872 to 12/2015. Data are from Robert Shiller's website. We compute variance ratios from monthly data and use the 1-year variance as the reference level.[5] The upper panels of Figure 1 depict the term structures of VRs for horizons up to 10 years. We compute VRs for the full sample (1872-2015) as well as for the pre-war and post-war sub-samples.

We observe three key aspects. First, the full sample VRs are downward-sloping, the slope is particularly steep between the 3-year and 8-year horizons, and the VRs are about 0.75 at longer horizons. Second, pre-war VRs are markedly downward-sloping, the negative slope starts at the 3-year horizon and persists, and the 10-year VR is about 0.50. Third, post-war VRs are mostly flat up to the 7-year horizon, they are slightly upward-sloping at longer horizons, and they reach a level of about 1.25 at the 10-year horizon. The 90% confidence intervals suggest that equity risk is decreasing with the holding period. In both the full and pre-war samples, the confidence intervals belong approximatively to the interval $(0, 1)$ at any horizon; in the post-war sample the confidence interval includes unity at any horizon. Thus, although the point estimates of variance ratios in the post-war sample are larger than one, we cannot reject the null hypothesis that they are smaller than one.[6]

[5] Similar results are obtained when computing variance ratios using yearly returns aggregated from monthly data.

[6] Note that our variance ratio estimates of *realized* returns over different holding periods can be interpreted as *predictive* variance ratios subject to some degree of measurement error. Campbell and Viceira (2005) document that equity risk decreases with the holding period using a VAR approach. By accounting for estimation risk, Pástor and Stambaugh (2012) reach the opposite conclusion for very long horizons. Recently, Jondeau and Rockinger (2016) recover downward-sloping equity risk from the Bayesian estimation of a structural DSGE model. While *predictive* variance ratios rely on the assumption of a specific predictive model and information set, we work with realized returns and therefore interpret the confidence intervals as bounds for the measurement error.

Figure 1: Variance Ratios and Volatilities of S&P 500 Holding Period Returns.
The figure displays the S&P 500 holding period return variance ratios (upper panels) and volatilities (lower panels) over three different samples. Holding periods are between 1 year and 10 years. Dashed lines denote 90% confidence intervals.



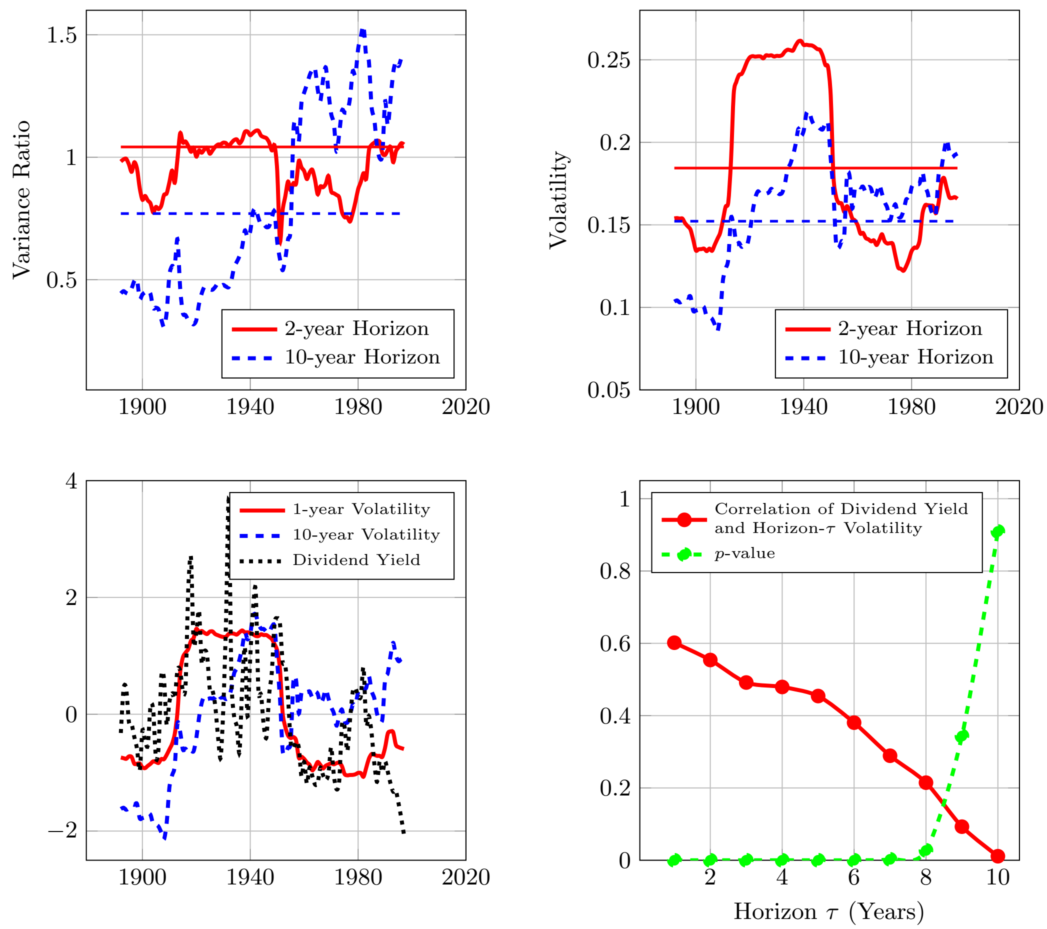
In both the full sample and the sub-samples, the 2-year VRs are close to unity. This comes from the fact that stock returns exhibit short-term time series momentum (Moskowitz, Ooi, and Pedersen, 2012), whereas long-term reversal implies decreasing VRs between the 2-year and the 4-year horizons. Long-horizon VRs are larger than zero indicating that prices follow an integrated process, but they are below or close to unity indicating that the transitory component of equity risk either offsets or dominates the permanent component.

The lower panels of Figure 1 report the term structures of the annualized volatility of holding period returns. In the full sample, return volatility decreases from about 18.5% at the 1-year horizon to about 16% at the 10-year horizon. In the pre-war sample, volatility decreases from 21% to 14%, whereas it increases from 16% to 18% in the post-war sample. Therefore, the increase in the slope of VRs from pre-war to post-war data is due to both a decrease in short-horizon risk and an increase in long-horizon risk.

We also investigate the time variation of stock return VRs and volatilities. Using the full sample, we build time series at yearly frequency from a 40-year rolling window of monthly returns. This procedure gives 106 yearly observations of VRs and volatilities with horizons ranging from 1 to 10 years. The left and right panels of Figure 2 depict the time series of VRs and volatilities,

Figure 2: S&P 500 Holding Period Return Variance Ratio, Volatility, and Dividend Yield in the 20th Century.

The upper panels display the S&P 500 holding period return variance ratios (left panel) and volatilities (right panel) with 2-year and 10-year horizons. Statistics are computed through a 40-year rolling window of monthly returns. Horizontal lines denote the full sample averages. The lower left panel displays the standardized dividend yield, 1-year holding period return volatility, and 10-year holding period return volatility of the S&P 500. The dividend yield is computed by averaging monthly observations over each year and volatilities are computed through a 40-year rolling window of monthly returns. The lower right panel displays the correlations (and p-values) between the dividend yield and the holding period return volatilities.



respectively, at the 2-year and 10-year horizons. While the 2-year VRs fluctuate around the full sample average, the 10-year VRs show a secular positive trend and lie significantly below and above the full sample average in the pre-war and post-war samples, respectively. This positive trend is mostly due to the behaviour of the 10-year return volatility. In contrast, short-horizon volatility increases in the first two decades of the 20th century and then decreases in the 50's and 60's.

The lower left panel of Figure 2 displays the standardized time series of 1-year and 10-year return volatilities together with the standardized S&P 500 dividend yield. The latter is a forward-looking measure of equity compensation. Hence, it is important to verify how the dividend yield relates with short-horizon and long-horizon equity risk. While the positive trend in long-horizon volatility seems

Table 1: Correlation between Holding Period Return Variance Ratio and Dividend Yield of the S&P 500.

This table reports the correlation between the S&P 500 holding period return variance ratios with 5-, 7-, and 10-year horizons and both the contemporaneous and the future average S&P 500 dividend yield. $\widetilde{Y}$ stands for the dividend yield. Variance ratios are computed through a 40-year rolling window of monthly returns and using a reference horizon τ_0 of one year. p-values are reported in parentheses.

	$\widetilde{Y}_t$	$\frac{1}{n}\sum_{i=1}^{n}\widetilde{Y}_{t+i}$			
		1	3	5	7
VR(5)	-0.203	-0.228	-0.283	-0.319	-0.340
p-value	(0.037)	(0.019)	(0.003)	(0.001)	(0.000)
VR(7)	-0.502	-0.526	-0.596	-0.633	-0.656
p-value	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
VR(10)	-0.543	-0.559	-0.626	-0.664	-0.690
p-value	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

unrelated to time variation in the dividend yield, we observe that short-horizon volatility captures both the hump and the decline of the dividend yield in pre-war and post-war samples, respectively. The lower right panel of Figure 2 shows the correlations (and p-values) between the dividend yield and return volatility for horizons ranging from 1 to 10 years. While the correlations with short-horizon volatility are large and significant, those with long-horizon volatility decline toward zero and become insignificant. This result highlights the cyclicality of the slope of equity risk. While the negative slope appears as an unconditional property of the data, we can understand how the slope behaves dynamically by observing the contemporaneous and intertemporal relation between the long-horizon VRs and the dividend yield. The former represent a quantitative measure of the slope of equity risk and the latter represents a (counter-cyclical) measure of economic conditions. Since short-term return volatility is highly positively correlated with the dividend yield, we expect VRs to be negatively correlated with it.

Confirming our expectation, Table 1 shows that the contemporaneous correlations between the dividend yield and the 5-, 7-, and 10-year VRs are negative and large, ranging between -20% and -55%. This suggests that the slope of the term structure of equity risk is pro-cyclical, consistent with the empirical findings of van Binsbergen et al. (2013) and Aït-Sahalia et al. (2015). Table 1 also shows that long-horizon VRs are highly negatively correlated with future levels of the dividend yield. This implies that the slope of the term structure of equity risk forecasts fluctuations in economic conditions and expected market compensation.

The analysis in this section documents that the timing of equity risk is markedly downward-sloping, confirming former results in Campbell and Viceira (2005).[7] This pattern of the timing of equity risk is robust, but its magnitude has decreased over the 20th century. Analysis of the relation between the dividend yield and the holding period return volatility suggests that short-run equity risk is much more closely linked to the dividend yield—a proxy of the equity risk premium—than is long-run equity risk. Importantly, the slope of equity risk is pro-cyclical (i.e. short-run risk increases in bad times) and forecasts the dividend yield. Thus, disentangling the permanent and the transitory components of stock returns is key to understand, model, and estimate equity risk.

3 The Model

This section presents the economic setting in which stock returns feature permanent and transitory components that are unobservable to the investor. First, we describe the investment opportunities and solve the investor's learning problem. This allows us to characterize the state-price density, capital gains, and the dividend yield. Then, we solve for dividend strip prices and discuss the connection between holding period stock returns and equity yields. Finally, we solve the investor's portfolio choice problem.

3.1 Investment Opportunities

We consider an investor who can invest in one risky asset—*a stock*—and one riskless asset paying a constant risk-free interest rate r_f. The stock price with dividends reinvested (Xia, 2001),[8] S_t^c, is modeled as the product of two components:

$$S_t^c \equiv Y_t Z_t = S_0 e^{\int_0^t \left(\mu + x_u - \frac{1}{2}\sigma_u^2\right)du + \sigma_u dB_u^y} \times e^{z_t},$$

where the dynamics of Y_t and Z_t satisfy

$$d\log Y_t = dy_t = \left(\mu + x_t - \sigma_t^2/2\right)dt + \sigma_t dB_t^y,$$

$$dx_t = -\kappa_x x_t dt + \sigma_x dB_t^x, \tag{2}$$

$$d\log Z_t = dz_t = -\kappa_z z_t dt + \sigma_z dB_t^z. \tag{3}$$

The investor observes only the price S_t^c. Through Bayesian updating, she therefore needs to filter out the unobservable individual components of the price before solving a portfolio choice problem

[7] Downward-sloping equity risk is a robust characteristic of the U.S. stock market: a broader index than the S&P 500 such as the value weighted CRSP index (including NYSE, Nasdaq, and Amex stocks) features VRs declining from one to 0.48, and volatilities declining from 20% to 13.8% between the 1-year and 10-year horizons.

[8] S^c denotes the stock price with dividends reinvested, whereas S denotes the ex-dividend value. Note that S_t^c and S_t are related via $S_t^c = S_t \exp\left(\int_0^t \frac{D_\tau}{S_\tau} d\tau\right)$. Applying Ito's lemma yields the stock return, which consists of the capital gain plus the dividend yield:

$$\frac{dS_t^c}{S_t^c} = \frac{dS_t + D_t dt}{S_t}. \tag{1}$$

(Detemple, 1986). The Brownian motions B^y, B^x, and B^z are independent[9] and the full filtration generated by observing all three Brownian shocks is denoted by $\mathbf{F}$.

The aim of considering the stock price as a product of the two components is to introduce some flexibility in modelling the timing of equity risk (Marfè, 2015a, 2016). The first term, Y, is an integrated process that depends on the time integral of x. Shocks in x accumulate and therefore permanently affect future prices. For this reason, we call x the *stochastic drift* of the stock return *permanent component*. If $Z_t \equiv 1$ we have the standard stock price dynamics with mean-reverting returns (Kim and Omberg, 1996; Wachter, 2002), which, as we will show, give rise to an upward-sloping term structure of equity risk because of the accumulation of x shocks in Y. The second term, Z_t, depends on the current value of z_t. Since the process z is mean-reverting, shocks in z dissipate as time passes and therefore have a transitory impact on prices. For this reason, we call z the *transitory component* of stock returns. By definition, the transitory component generates risk in the short term that dissipates in the longer term. In other words, the transitory component helps to reconcile the timing of risk in the model with that observed empirically.

Note that the transitory component defined here is different from the transitory component discussed in Alvarez and Jermann (2005) and Hansen and Scheinkman (2009). They show that any integrated process can be decomposed into a martingale and a transitory component. For instance, a geometric Brownian motion with mean-reverting drift can be written as the product of a martingale and a transitory process that is driven by the mean-reverting drift and is negatively correlated with the martingale. However, since the drift is embedded in an integrated process, its variation accumulates over time and leads to upward-sloping risk. Instead, in our paper we refer to a transitory component that is not embedded in an integrated process, such that its variation does not accumulate over time and therefore leads to downward-sloping risk. We model a transitory component that is orthogonal to (instead of negatively correlated with) the integrated component of the cumulative return (and its martingale).

By an application of Ito's lemma, the stock return satisfies

$$\frac{dS_t^c}{S_t^c} = (m + x_t - \kappa_z z_t)dt + \sqrt{v_t}dB_t, \tag{4}$$

where $m \equiv \mu + \sigma_z^2/2$ is the long-run expected rate of return, $v_t \equiv \sigma_t^2 + \sigma_z^2$ is the conditional variance of stock returns (assumed to be either constant, $v_t \equiv \bar{v}$, or allowed to vary stochastically over time), and $dB_t \equiv \left(\sigma_t dB_t^y + \sigma_z dB_t^z\right)/\sqrt{v_t}$ is an increment of a standard Brownian motion.

[9]Note that assuming non-zero correlations among unobservable state variables would imply additional parameters to be estimated in Section 4.1, but would not qualitatively change our results in Section 5. The reason is that the investor solves her portfolio choice problem by observing stock returns only. In turn, the filtered state variables are perfectly correlated irrespective of the correlations among the unobservable variables. Therefore, we keep the model simple and assume that the unobservable variables are uncorrelated.

3.2 Learning

The expected rate of return on the stock varies over time due to shocks that come from two sources: the drift of the permanent component x_t defined in (2) and the transitory component z_t defined in (3). The investor only has access to information generated by observing the realized return defined in (4), and thus does not have access to the full information contained in the filtration $\mathbf{F}$. Therefore, all her actions must be adapted to her observation filtration $\mathbf{F}^S = \left\{\mathcal{F}_t^S\right\}_{t\geq 0}$, defined as the flow of information generated by the stock return process dS_t^c/S_t^c. In other words, the investor needs to filter out through Bayesian updating the *unobservable* factors x_t and z_t by observing the history of stock returns only.

Since the stock return variance v_t can be inferred from the quadratic variation of returns, it is also observable. We assume, however, that it does not provide more information on x_t and z_t than does the stock return dS_t^c/S_t^c. That is, the return process dS_t^c/S_t^c is assumed to be perfectly (positively or negatively) correlated with the variance process dv_t. This simplifying assumption implies that the dynamics of v_t do not enter the learning problem, and therefore that the investor faces a single source of risk under the observation filtration (see Proposition 1 below).

Assuming that the variance v_t does not provide more information than the return dS_t^c/S_t^c is reasonable if and only if stock returns and their variance are either highly positively or highly negatively correlated. Andersen, Bondarenko, and Gonzalez-Perez (2015), Andersen, Fusari, and Todorov (2015), and Babaoglu, Christoffersen, Heston, and Jacobs (2016) have recently provided empirical evidence that the mean correlation between returns and their variance lies between -0.85 and -0.95. This implies that filtering out x_t and z_t using only the history of returns is almost equivalent to filtering them out using the history of both returns and their variance. In other words, the informational loss incurred by ignoring the history of the return variance is small. Overall, these results show that assuming a perfect correlation between dS_t^c/S_t^c and dv_t is reasonable.

Proposition 1 provides the dynamics of the filtered state variables.

Proposition 1. *With respect to the investor's observation filtration, the stock return dS_t^c/S_t^c and the filtered state variables $\widehat{z}_t$ and $\widehat{x}_t$ satisfy*

$$\frac{dS_t^c}{S_t^c} = \left(m + \widehat{x}_t - \kappa_z \widehat{z}_t\right) dt + \sqrt{v_t} d\widehat{B}_t, \tag{5}$$

$$d\widehat{x}_t = -\kappa_x \widehat{x}_t dt + \frac{1}{\sqrt{v_t}} \left(\gamma_{x,t} - \kappa_z \gamma_{zx,t}\right) d\widehat{B}_t, \tag{6}$$

$$d\widehat{z}_t = -\kappa_z \widehat{z}_t dt + \frac{1}{\sqrt{v_t}} \left(\gamma_{zx,t} - \kappa_z \gamma_{z,t} + \sigma_z^2\right) d\widehat{B}_t, \tag{7}$$

where $\widehat{x}_t \equiv E_t(x_t)$, $\widehat{z}_t \equiv E_t(z_t)$, and $\widehat{B}_t$ is a $\mathcal{F}_t^S$-Brownian motion process. The posterior variance-covariance matrix Γ_t is defined as follows:

$$\Gamma_t \equiv \begin{pmatrix} \gamma_{z,t} & \gamma_{zx,t} \\ \gamma_{zx,t} & \gamma_{x,t} \end{pmatrix} = \begin{pmatrix} \mathrm{Var}_t(z_t) & \mathrm{Cov}_t(z_t, x_t) \\ \mathrm{Cov}_t(z_t, x_t) & \mathrm{Var}_t(x_t) \end{pmatrix} \tag{8}$$

and the posterior variance-covariances $\gamma_{z,t}$, $\gamma_{x,t}$, $\gamma_{zx,t}$ are given by

$$\frac{d\gamma_{z,t}}{dt} = \sigma_z^2 - 2\kappa_z\gamma_{z,t} - \frac{1}{v_t}\left(\gamma_{zx,t} - \gamma_{z,t}\kappa_z + \sigma_z^2\right)^2, \tag{9}$$

$$\frac{d\gamma_{x,t}}{dt} = \sigma_x^2 - 2\kappa_x\gamma_{x,t} - \frac{1}{v_t}\left(\gamma_{x,t} - \gamma_{zx,t}\kappa_z\right)^2, \tag{10}$$

$$\frac{d\gamma_{zx,t}}{dt} = -\left(\kappa_x + \kappa_z\right)\gamma_{zx,t} - \frac{1}{v_t}\left(\gamma_{x,t} - \gamma_{zx,t}\kappa_z\right)\left(\gamma_{zx,t} - \gamma_{z,t}\kappa_z + \sigma_z^2\right). \tag{11}$$

Note that, for notational convenience, we denote by $E_t(.)$, $Cov_t(.,.)$, and $Var_t(.)$ the moments conditional on the time t observation filtration, $\mathcal{F}_t^S$.

Proof. See Appendix A.1.

Equation (5) gives the return dynamics projected on the observable filtration, while Equations (6) and (7) describe the investor's updating rule of the expectation of the latent state variables x_t and z_t. We refer to $\widehat{x}_t$ and $\widehat{z}_t$ as the filters. Equations (9)–(11) provide the dynamics of the posterior variance-covariance matrix (8) and hence capture the evolution of uncertainty associated with the estimation of the unobserved factors.

3.3 State Prices, Capital Gains, and Dividends

The dynamics provided in Proposition 1 and the assumption that stock returns are perfectly correlated with their variance implies that markets are complete;[10] the investor can replicate any payoff by trading the stock and the riskless asset because she faces a single source of risk $\widehat{B}_t$. The Second Fundamental Theorem of Asset Pricing (see, e.g., Duffie, 2001) therefore implies that there exists a unique risk-neutral measure characterized by a unique market price of risk.

The stock return dynamics in (5) and the constant risk-free rate r_f imply that the market price of risk ψ_t is given by

$$\psi_t = \frac{m + \widehat{x}_t - \kappa_z\widehat{z}_t - r_f}{\sqrt{v_t}},$$

which is the expected excess stock return relative to the stock return volatility. By definition, the

[10] Note that if we had assumed that dS_t^c/S_t^c were not perfectly correlated to dv_t, then the filters $\widehat{x}_t$ and $\widehat{z}_t$ provided in Proposition 1 would have been driven by two sources of risk. Therefore, markets would have been incomplete and one of the two components of the market price of risk could have been chosen arbitrarily. In this case, it would have been unclear whether the downward-sloping timing of equity risk discussed in the next sections were driven by the choice of the market price of risk or by the transitory component *per se*. Such ambiguity is resolved when markets are complete because the market price of risk is unique, motivating our choice that dS_t^c/S_t^c and dv_t are perfectly correlated.

state-price density ξ_t and the (ex-dividend) price of the stock S_t satisfy

$$\frac{d\xi_t}{\xi_t} = -r_f dt - \psi_t d\widehat{B}_t, \qquad (12)$$
$$S_t = E_t\left[\int_t^\infty \frac{\xi_s}{\xi_t} D_s ds\right],$$

where D_t is the dividend process paid by the stock. We specify the dynamics of the ex-dividend stock price in accordance with the model for stock returns in (5). Specifically,

$$\frac{dS_t}{S_t} = (m_s + a_x\widehat{x}_t - a_z\kappa_z\widehat{z}_t)dt + \sqrt{v_t}d\widehat{B}_t, \qquad (13)$$

where m_s, a_x, and a_z are constants. The parameter m_s captures the long-term expected stock price growth, while the parameters a_x and a_z are the stock price growth sensitivities to $\widehat{x}_t$ and $\widehat{z}_t$, respectively. Comparing (5) and (13), we note that in line with the requirement (1), the diffusions of returns dS_t^c/S_t^c and capital gains dS_t/S_t coincide. Moreover, from the difference between the two drift terms in (5) and (13), we can recover the dividend yield $\widetilde{Y}_t$:

$$\widetilde{Y}_t \equiv D_t/S_t = m - m_s + (1 - a_x)\widehat{x}_t - (1 - a_z)\kappa_z\widehat{z}_t. \qquad (14)$$

The dynamics of the dividend process

$$\frac{dD_t}{D_t} = \mu_{D,t}dt + \sigma_{D,t}d\widehat{B}_t \qquad (15)$$

are readily obtained by applying Ito's lemma to $D_t = \widetilde{Y}_t S_t$ using the dynamics of the stock price S_t and the dividend yield $\widetilde{Y}_t$ defined in (13) and (14), respectively.

3.4 Equity Risk

In this subsection we study the measures and sources of equity risk over different horizons. One approach is to look at the volatility of holding period stock returns. Holding the stock from time t to $T = t + \tau$ provides the investor with a return consisting of capital gains and dividends over a horizon τ. Therefore, the holding period return (HPR hereafter) is defined as

$$R(t, T) \equiv S_T^c/S_t^c,$$

where S_t^c is the stock price with dividends reinvested. The variance or volatility of the HPR captures how the risk profile of the stock changes as the holding period increases.

Alternatively, we can describe the timing of equity risk by studying the volatility of equity yields (van Binsbergen et al., 2013; van Binsbergen and Koijen, 2017). The equity yield $e(t, T)$ is defined by

$$e(t,T) = -\frac{1}{T-t}\log\left(\frac{P(t,T)}{D_t}\right),$$

where $P(t,T)$ is the price of a dividend strip with maturity date T[11]

$$P(t,T) = E_t\left[\frac{\xi_T}{\xi_t}D_T\right]. \tag{16}$$

The equity yield reveals information about the internal rate of return on the dividend strip, and therefore its volatility is a measure of equity risk that complements the HPR volatility.

Before investigating the properties of the two measures of equity risk in our model, it is important to study the relation among HPRs, dividend strip returns, and equity yields as well as between their respective risks in a general framework. Indeed, this allows to better understand the components of the two measures of equity risk and shows how they are connected.

By definition, the stock price can be represented as the integral of dividend strip prices over the continuum of maturities in $[t,\infty)$

$$S_t = \int_t^\infty P(t,i)di.$$

Using this definition, Proposition 2 provides the link between the HPR and the dividend strip returns.

Proposition 2. *The holding period return is*

$$\begin{aligned}\log R(t,T) = &\int_t^T\int_s^\infty d\log P(s,i)w_s(i)di\\ &+\frac{1}{2}\int_t^T\int_s^\infty \sigma_P(s,i)\left(\sigma_P(s,i) - \int_s^\infty \sigma_P(s,j)w_s(j)dj\right)w_s(i)dids,\end{aligned} \tag{17}$$

where

$$w_t(T) \equiv \frac{P(t,T)}{\int_t^\infty P(t,T)dT} \quad \textit{with} \quad \int_t^\infty w_t(T)dT = 1, \tag{18}$$

$$d\log P(t,T) = d\log D_t - (T-t)de(t,T) + e(t,T)dt, \tag{19}$$

$$\sigma_P(t,T) = \sigma_{D,t} - (T-t)\sigma_e(t,T). \tag{20}$$

Proof. See Appendix A.2.

[11] In practical applications it is common to use the time to maturity, rather than the maturity date, to parametrize dividend strips and equity yields. In terms of time to maturity τ we can define $P_t(\tau)$ and $e_t(\tau)$ as $P_t(\tau) \equiv P(t,t+\tau)$ and $e_t(\tau) \equiv e(t,t+\tau)$. This representation is akin to the Musiela (1993) parametrization used in the interest rate theory.

Equation (17) shows that the HPR can be expressed as the return on a portfolio of dividend strips with weights given by (18) (van Binsbergen and Koijen, 2017) plus a volatility adjustment factor corresponding to the second order term of the Taylor expansion. This adjustment is zero if the volatility of dividend strip returns is a constant function of their maturity. Equation (17) corresponds directly to the approximation of the log return on the portfolio considered in the discrete-time models of Campbell and Viceira (1999) and Campbell, Chan, and Viceira (2003). In our continuous-time model their approximation holds exactly and, unlike in the standard finite asset case, our portfolio consists of a continuum of assets (one dividend strip for each maturity).[12] Equations (19) and (20) provide the expressions for the log dividend strip return $d\log P(t,T)$ and its volatility $\sigma_P(t,T)$ in terms of the equity yield $e(t,T)$ and its volatility $\sigma_e(t,T)$, respectively.

Since the HPR is the return on the value-weighted portfolio of dividend strips, the relationship between the HPR volatility and the equity yield volatility is similar to that between the volatility of a portfolio and the volatility of its constituents. The volatility of a one-period portfolio return depends on the weighted variance-covariance matrix of its constituents. Similarly, the instantaneous HPR volatility at time t will depend on the contemporaneous term structure of equity yield volatilities. For multi-period returns, the autocorrelations and cross-autocorrelations of dividends strips with different maturities come into play. Proposition 3 describes the relation between the variance of HPRs and the variance of dividend strip returns (or equivalently, the variance of equity yields through (20)).

Proposition 3. *The conditional variance of the holding period return is*

$$\mathrm{Var}_t\left(\log R(t,T)\right) = \int_t^T C_{S,t}\left(s,s\right) ds + \int_t^T\!\!\int_t^T C_{S,t}\left(s,u\right) dsdu \tag{21}$$

where

$$C_{S,t}\left(s,s\right) ds = \int_s^\infty\!\!\int_s^\infty \mathrm{E}_t\left[\sigma_P(s,i)\sigma_P(s,j)w_s(i)w_s(j)\right] didjds, \tag{22}$$

$$C_{S,t}\left(s,u\right) dsdu \approx \int_s^\infty\!\!\int_u^\infty \mathrm{Cov}_t\left(d\log P(s,i)w_s(i), d\log P(u,j)w_u(j)\right) didj,\;\; s\neq u. \tag{23}$$

Proof. See Appendix A.3.

The first term of the HPR conditional variance in (21) is the "average short-term variance" term, namely the average snapshot of the term structure of equity risk over horizons in $[t,T]$. Indeed, according to (22) the instantaneous variance of the stock return can be expressed as the (weighted) sum of all the elements of the variance-covariance matrix of the dividend strip returns at a given time t. Therefore, by studying the properties of the dividend strip return volatility (or equivalently of the equity yield volatility through (20)) at a given time as a function of time to maturity, we learn about the ingredients that contribute to the short-term variance of stock returns.

[12] Björk, Kabanov, and Runggaldier (1997) develop a theory that allows to study such measure-valued trading portfolios.

The second term in (21) represents the "average autocovariance" term. From (23) we see that this autocovariance term contains information about the dynamic cross- and auto-correlations between the dividend strip returns at different points in time. Therefore, there is a connection between the equity risk as inferred from the HPR and equity risk as inferred from the dividend strip returns (or equivalently from equity yields through (19)). Importantly, the volatility of the HPR depends on both the average shape of the term structure of dividend strip return volatility as well as on its variation over time.

We next investigate the properties of the term structure of risk of both HPRs and equity yields specific to our model. The first measure of risk is the (annualized) τ-year HPR volatility, which is defined by

$$\mathrm{vol}_{S,t}(\tau) \equiv \sqrt{\frac{1}{\tau}\log\left(\frac{\mathrm{E}_t[\left(S^c_{t+\tau}/S^c_t\right)^2]}{\left(\mathrm{E}_t[S^c_{t+\tau}/S^c_t]\right)^2}\right)}, \tag{24}$$

as in Belo et al. (2015) and Marfè (2015a).[13] Proposition 4 shows that the permanent and transitory components of stock returns have rather different implications for the term structure of HPR volatility. At odds with the empirical evidence provided in Section 2, the model abstracting from the transitory component is only able to produce an upward-sloping term structure of HPR volatility. Accounting for the transitory component helps to overcome this rigidity, and allows the model to potentially generate a term structure of HPR volatility that is empirically supported. The parameter estimation performed in Section 4 shows that Condition (25) is satisfied, and therefore that our model implies a downward-sloping term structure of HPR volatility.

Proposition 4. *Assuming a constant instantaneous return volatility,*[14] *the following holds.*

1. *When stock returns depend only on the permanent component ($S^c_t = Y_t$), the long-term volatility of HPR is* greater *than the short-term volatility, i.e.,*

$$\lim_{\tau\to\infty}\mathrm{vol}_{S,t}(\tau) > \lim_{\tau\to 0^+}\mathrm{vol}_{S,t}(\tau).$$

Moreover, the term structure of HPR volatility is monotonically increasing *for $\tau > 0$, i.e.,*

$$\frac{\partial}{\partial\tau}\mathrm{vol}_{S,t}(\tau) \geq 0.$$

[13] This measure uses a first order approximation of the variance term in

$$\mathrm{vol}_{S,t}(\tau) \equiv \sqrt{\frac{1}{\tau}\mathrm{Var}_t\left(\log\left(S^c_{t+\tau}/S^c_t\right)\right)}.$$

Indeed, the delta method gives $\mathrm{Var}\left[f(X)\right] \approx \left(f'\left(\mathrm{E}[X]\right)\right)^2\mathrm{Var[X]}$. In the log case, $\mathrm{Var}\left[\log(X)\right] \approx \frac{1}{(\mathrm{E}[X])^2}\mathrm{Var[X]} = \frac{\mathrm{E}[X^2]}{(\mathrm{E}[X])^2} - 1$. Finally, using $\log(1+x) \approx x$ we obtain (24).

[14] Due to the lack of closed-form solutions in other cases, Proposition 4 covers only the term structure of HPR volatilities for $v_t = \overline{v}$. As we will see in the subsequent sections, with stochastic v_t the term structure of HPR volatility is, if anything, even more upward sloping when $S^c_t = Y_t$. Intuitively, stochastic volatility increases the uncertainty of stock returns in the long run because shocks accumulate over time (Bansal and Yaron, 2004).

2. *When stock returns depend on both the permanent and the transitory component ($S_t^c = Y_t Z_t$), the long-term volatility of HPR is* smaller *than the short-term volatility, i.e.*

$$\lim_{\tau\to\infty} \mathrm{vol}_{S,t}(\tau) < \lim_{\tau\to 0^+} \mathrm{vol}_{S,t}(\tau),$$

if the volatility of the transitory component is large enough,

$$\sigma_z^2 > \frac{\sigma_x^2}{\kappa_x^2}. \tag{25}$$

Proof. See Appendix A.4.

The second measure of risk is the τ-year equity yield volatility, which is defined by

$$\mathrm{vol}_{e,t}(\tau) \equiv \sqrt{\frac{1}{dt}\mathrm{Var}_t\left(de(t,t+\tau)\right)} = \left|\sigma_e(t,t+\tau)\right|.$$

To understand how the permanent and transitory components influence the term structure of equity yield volatility, we make the simplifying assumption that the impact of the drift of the permanent component $\widehat{x}_t$ on stock returns in (5) and capital gains in (13) is similar (i.e., $a_x \to 1$). Importantly, this assumption is perfectly consistent with our model estimation in Section 4, and is therefore motivated empirically. Proposition 5 shows that, in the absence of a transitory component, the term structure of equity yield volatility is flat, which is inconsistent with the empirical findings of van Binsbergen et al. (2013) and van Binsbergen and Koijen (2017). In contrast, the model considering the transitory component implies (under a mild condition) a downward-sloping term structure of equity yield volatility. That is, modeling the transitory component helps replicate the empirically observed shape of the term structures of equity yield volatility and HPR volatility.

Proposition 5. *Assuming a constant instantaneous return volatility and $a_x \to 1$, the following holds.*

1. *In the model with the permanent component only ($S_t^c = Y_t$), the equity yield volatility is zero at any horizon,*

$$\mathrm{vol}_{e,t}(\tau) = 0 \ \ \forall \tau.$$

2. *In the model with both the permanent and the transitory components ($S_t^c = Y_t Z_t$), the long-term equity yield volatility is* smaller *than the short-term equity yield volatility*

$$\lim_{\tau\to\infty} \mathrm{vol}_{e,t}(\tau) < \lim_{\tau\to 0^+} \mathrm{vol}_{e,t}(\tau),$$

if

$$(a_z-1)\eta_z\kappa_z\Big[(m-m_s)\big((m-m_s)\sqrt{\overline{v}}+\eta_x\big)$$
$$+\big((m-m_s)\sqrt{\overline{v}}+(m_s-r_f-\overline{v}+a_z(-m+r_f+\overline{v}))\eta_z\big)\kappa_z\Big]\neq 0, \tag{26}$$

where η_x and η_z are constants defined in Appendix A.5. Note that Condition (26) is mild because it is a non-equality. When the parameters of the model are estimated empirically, the probability of this non-equality being satisfied is equal to one. Therefore, this condition should be perceived as being always satisfied.

Proof. See Appendix A.5.

From Propositions 4 and 5, we derive Corollary 1. Corollary 1 shows that a downward-sloping term structure of HPR volatility implies a downward-sloping term structure of equity yield volatility. However, the reverse does not necessarily hold. Since there is empirical evidence that the term structure of HPR volatility is downward-sloping (see Section 2 and Campbell and Viceira (2005) for instance), the model predicts that the term structure of equity yield volatility is also downward-sloping. Thus, our theoretical framework contributes to the recent debate on whether or not the term structure of equity risk is downward-sloping (Boguth, Carlson, Fisher, and Simutin, 2012; van Binsbergen and Koijen, 2016; Schulz, 2016).

Corollary 1. *Under the assumptions made in Propositions 4 and 5, a downward-sloping term structure of HPR volatility implies a downward-sloping term structure of equity yield volatility.*

3.5 Portfolio Choice Problem

We next solve the portfolio choice problem of an investor who models returns using permanent and transitory components. The objective of the investor is to choose the fraction of her wealth to be invested in the stock π_t—*the portfolio weight*—that maximizes her expected utility of terminal wealth

$$\sup_{(\pi_s)_{s\in[t,T]}} E_t\left[e^{-\delta T}\frac{W_T{}^{1-\gamma}}{1-\gamma}\right],$$

where δ is the subjective discount rate, γ is the investor's relative risk aversion, and W is the investor's wealth that needs to satisfy the dynamic budget constraint

$$dW_t = r_f W_t dt + \pi_t W_t\left(m+\widehat{x}_t-\kappa_z\widehat{z}_t-r_f\right)dt+\pi_t W_t\sqrt{v_t}d\widehat{B}_t.$$

When the stock return variance is assumed to be constant ($v_t\equiv\bar{v}$), the investor's optimization problem depends only on the investor's wealth W_t and the two filters $\widehat{x}_t$ and $\widehat{z}_t$. The reason is that the posterior variance-covariance matrix Γ_t features a steady state,[15] which we assume has been attained.

Proposition 6 provides the value function and the optimal portfolio under constant volatility.

Proposition 6. *In the model with constant return volatility ($v_t \equiv \overline{v}$), the following hold.*

1. *The investor's optimal value function J takes the form*

$$J(t, W_t, \theta_t) = e^{-\delta t} \frac{W_t^{1-\gamma}}{1-\gamma} F(T-t, \theta_t)^{\gamma}, \tag{27}$$

where $\theta_t = [\widehat{x}_t, \widehat{z}_t]^\top$ and the function $F : \mathbb{R}^+ \times \mathbb{R}^2 \to \mathbb{R}$ is given by

$$F(\tau, \theta) \equiv \exp\left(\frac{1}{2}\theta^\top A(\tau)\theta + \theta^\top B(\tau) + C(\tau)\right),$$

with A, B, C solving the system of matrix ODEs provided in Appendix A.6.

2. *The optimal trading strategy of the investor is*

$$\pi_t^\star = \frac{m + \widehat{x}_t - \kappa_z \widehat{z}_t - r_f}{\gamma \overline{v}} + \pi_t^{\text{hdg}} \tag{28}$$

with the hedging demand given by

$$\pi_t^{\text{hdg}} = \frac{\overline{\gamma}_x - \kappa_z \overline{\gamma}_{zx}}{\overline{v}} \frac{F_x}{F} + \frac{\overline{\gamma}_{zx} - \kappa_z \overline{\gamma}_z + \sigma_z^2}{\overline{v}} \frac{F_z}{F},$$

where F and its derivatives are evaluated at $(T-t, \widehat{x}_t, \widehat{z}_t)$, and $\overline{\gamma}_x, \overline{\gamma}_z, \overline{\gamma}_{zx}$ are elements of the steady-state posterior variance covariance matrix $\overline{\Gamma}$ provided in Appendix A.6.

Proof. See Appendix A.6.

The optimal portfolio in (28) has two components: the myopic demand and the hedging demand. As $t \to T$, the investor's value function given in (27) approaches the terminal utility of wealth, which is independent of the state vector $[\widehat{x}_t, \widehat{z}_t]^\top$. Therefore, the partial derivatives F_x and F_z, and consequently the hedging demand, approach zero as the horizon shrinks. This means that investors with short horizons will choose their stock allocation according to the myopic first term in (28). The hedging demand provides a long-horizon investor with an intertemporal hedge against changes in the two state variables $\widehat{x}_t$ and $\widehat{z}_t$.

We now relax the assumption of constant volatility. In particular, we assume that stock return variance $v_t \in \mathcal{F}_t^S$ follows a square-root process (Heston, 1993):

$$dv_t = \kappa_v(\bar{v} - v_t)dt + \sigma_v \sqrt{v_t} dB_t^v, \tag{29}$$

[15] The steady-state value of Γ_t is obtained by setting the right-hand side of Equations (9), (10), and (11) to zero. At this value, the investor is no longer able to improve upon the accuracy of the estimates.

where κ_v, $\bar{v}$, and σ_v are known constants, and where the shocks to variance are perfectly correlated with shocks to returns, $d\langle B^v, \widehat{B}\rangle_t = dt$. The sign of σ_v determines the sign of the correlation between stock returns and their variance. As shown in Section 4, our estimation yields a negative parameter σ_v and therefore a negative correlation between returns and their variance (see Christoffersen, Jacobs, and Mimouni, 2010; Andersen et al., 2015, among others).[16]

In the model with stochastic volatility, a steady-state for the posterior variance-covariance matrix no longer exists. Therefore, the elements of Γ_t in (8) become state variables. Proposition 7 provides the value function and the optimal portfolio under stochastic volatility.

Proposition 7. *Assuming that the stock return variance is given by (29), the following hold.*

1. *The investor's optimal value function J can be approximated by* [17]

$$J(t, W_t, \theta_t) \approx e^{-\delta t}\frac{W_t^{1-\gamma}}{1-\gamma}F(T-t, \theta_t)^{\gamma}, \tag{30}$$

where $\theta_t = [\widehat{x}_t, \widehat{z}_t, v_t, \gamma_{x,t}, \gamma_{z,t}, \gamma_{zx,t}]^\top$ and the function $F : \mathbb{R}^+ \times \mathbb{R}^6 \to \mathbb{R}$ is given by

$$F(\tau, \theta) \equiv \exp\left(\frac{1}{2}\theta^\top A(\tau)\theta + \theta^\top B(\tau) + C(\tau)\right),$$

with A, B, C solving the system of matrix ODEs defined in Appendix A.7.

2. *The optimal portfolio weight implied by the function J in (30) satisfies*

$$\pi_t^\star = \frac{m + \widehat{x}_t - \kappa_z\widehat{z}_t - r_f}{\gamma v_t} + \pi_t^{\text{hdg}}$$

with the hedging demand given by

$$\pi_t^{\text{hdg}} = -\sigma_v\frac{F_v}{F} + \frac{\gamma_{x,t} - \kappa_z\gamma_{zx,t}}{v_t}\frac{F_x}{F} + \frac{\gamma_{zx,t} - \kappa_z\gamma_{z,t} + \sigma_z^2}{v_t}\frac{F_z}{F},$$

where F and its derivatives are to be evaluated at $(T-t, \widehat{x}_t, \widehat{z}_t, v_t, \gamma_{x,t}, \gamma_{z,t}, \gamma_{zx,t})$.

Proof. See Appendix A.7.

When compared to the constant volatility results in Proposition 6, stochastic volatility implies an additional term in the hedging part of the optimal portfolio.

[16] It is worth mentioning that the perfect correlation between the stock return, dS_t^c/S_t^c, and its variance, dv_t, on the observation filtration implies the same perfect correlation on the full filtration. That is, the return variance does indeed not, as assumed initially, provide more information on x_t and z_t than dS_t^c/S_t^c, and was therefore rightfully omitted from the learning problem in Section 3.2.

[17] The approximation method is borrowed from Andrei and Hasler (2015) who, in the spirit of Campbell and Shiller (1988) and Benzoni, Collin-Dufresne, and Goldstein (2011) approximations, linearize the partial differential equation satisfied by the function F. Evidence of the accuracy of the approximation is provided in Andrei, Carlin, and Hasler (2015).

4 Parameter Estimation and Implied Dynamics

In this section, we first describe the parameter estimation. Then, we discuss the relation between the state variables and the dividend yield, which is a forward-looking measure of equity compensation. We document that most of the dividend yield variation is captured by the transitory component of stock returns. Therefore, models that account for (resp. ignore) the transitory component can (resp. cannot) replicate the observed relation between the dividend yield and future returns as well as between the dividend yield and future dividend growth.

4.1 Parameters Driving Stock Returns

We consider and compare five different models: *1)* the unconstrained PT-SV model that features a permanent component with stochastic drift, a transitory component, and stochastic return volatility, *2)* the PT-CV model that features a permanent component with stochastic drift, a transitory component, and constant return volatility, *3)* the P-SV model that features a permanent component with stochastic drift and stochastic return volatility, *4)* the P-CV model that features a permanent component with stochastic drift and constant return volatility, and *5)* the CM-CV model that features a constant mean return and constant return volatility. The last four models are constrained versions of the unconstrained model and are obtained by setting either a constant mean return ($\sigma_x = \kappa_x = \sigma_z = \kappa_z = 0$), or a constant return variance ($v_t = \overline{v}$), or by ignoring the transitory component ($\sigma_z = \kappa_z = 0$).

The parameter vector, $\Theta \equiv (m, \sigma_z, \kappa_z, \sigma_x, \kappa_x, \overline{v}, \sigma_v, \kappa_v)$, driving the stock return dynamics in (5)–(7), (9)–(11), and (29) is estimated by Maximum Likelihood. Each of the five models described above is fitted to monthly S&P 500 nominal returns from 02/1871 to 02/2016. Data are from Robert Shiller's website. Refer to Appendix B for a complete description of the estimation.

Table 2 provides the estimates and standard errors of the parameter vector Θ. Irrespective of whether return variance is assumed to be constant or stochastic, the transitory component $\widehat{z}_t$ features a relatively high volatility of about 15% and strong persistence (see the right panel of Figure 3). Indeed, its half-life is about 4.5 and 9 years (i.e., business-cycle frequencies) when return variance is assumed to be constant and stochastic, respectively. The significance of the mean-reversion speed of the transitory component increases markedly when stochastic return variance is considered. The Akaike information criteria suggest that adding a transitory component increases the goodness of fit more so when return variance is stochastic than when it is constant. In contrast to $\widehat{z}_t$, the stochastic drift $\widehat{x}_t$ mean-reverts at a high speed (see the left panel of Figure 3). Its half-life is about 3 months and 8 months when return variance is constant and stochastic, respectively. Note that the volatility parameter of the return variance σ_v is negative, implying a negative correlation between returns and their variance (i.e., a leverage effect as in Christie, 1982).

The Akaike information criteria show that accounting for stochastic expected returns increases the goodness of fit to an important degree. Stochastic return variance increases the goodness of fit further and magnifies the benefits of considering a transitory component. When return variance is

Table 2: Parameters Driving Stock Returns.

This table reports the values of the parameters driving stock returns for the different models. P, PT, CV, SV, and CM stand for permanent component with stochastic drift, permanent component with stochastic drift and transitory component, constant volatility, stochastic volatility, and constant mean, respectively. The Akaike information criterion is defined as follows: $AIC = 2k - 2\mathcal{L}$, where k is the number of parameters of the model and $\mathcal{L}$ is the corresponding log-likelihood function. Standard errors are reported in brackets and statistical significance at the 10%, 5%, and 1% levels is labeled with *, **, and ***, respectively. Significance for κ_x, κ_z, κ_v, and $\overline{v}$ are based on one-sided tests because these parameters are positive by definition. Significance for the other parameters are based on two-sided tests. Parameters are estimated using monthly S&P 500 returns from 02/1871 to 02/2016. The last three rows report our choice of preference parameters and the risk-free rate.

Parameter	**Symbol**	**CM-CV**	**P-CV**	**PT-CV**	**P-SV**	**PT-SV**
Long-term mean of	m	0.095***	0.094***	0.095***	0.091***	0.084***
expected return		(0.012)	(0.015)	(0.009)	(0.017)	(0.008)
Volatility of	σ_z			0.148***		0.171***
transitory component				(0.018)		(0.019)
Mean-reversion speed of	κ_z			0.149*		0.075***
transitory component				(0.112)		(0.022)
Volatility of	σ_x		3.168***	3.061***	1.016***	1.046***
the permanent component			(0.523)	(0.317)	(0.041)	(0.024)
Mean-reversion speed of	κ_x		30.066***	28.237***	7.214***	7.419***
the permanent component			(7.443)	(5.033)	(0.571)	(0.304)
Long-term mean of	$\overline{v}$	0.020***	0.018***	0.018***	0.017***	0.017***
return variance		(2.6×10^{-4})	(2.7×10^{-4})	(3.1×10^{-4})	(6.4×10^{-4})	(5.2×10^{-4})
Volatility of	σ_v				−0.042***	−0.042***
return variance					(0.002)	(0.003)
Mean-reversion speed of	κ_v				0.795***	0.895***
return variance					(0.076)	(0.094)
Akaike criterion	AIC	−6207.6	−6355.6	−6353.9	−6482.0	−6492.0
Relative risk aversion	γ	5	5	5	5	5
Subjective discount rate	δ	0.01	0.01	0.01	0.01	0.01
Risk-free rate	r^f	0.05	0.05	0.05	0.05	0.05

constant, however, a transitory component yields over-fitting. Indeed, the likelihood function does not increase sufficiently to offset the two additional parameters' penalty.

To better understand the dynamics of the stochastic drift $\widehat{x}_t$ and the transitory component $\widehat{z}_t$, we regress the historical S&P 500 log dividend yield $\widetilde{y}_t$ (i.e., a forward-looking measure of equity compensation) on these two variables. Table 3 shows that, while $\widehat{x}_t$ and $\widehat{z}_t$ are significant in the univariate regressions, the stochastic drift $\widehat{x}_t$ only explains a small fraction of the dividend yield variation (between 1% and 5%). In contrast, the transitory component $\widehat{z}_t$ has a considerably larger explanatory power (between 37% and 65%). This result is further confirmed by multivariate regressions.

This evidence suggests that the transitory component of returns captures to a large extent the dynamics of the equity risk premium. A negative transitory shock today (low $\widehat{z}_t$) implies a larger future equity compensation (high $\widetilde{y}_t$), whereas the link between the stochastic drift and future

Figure 3: Time Series of the Stochastic Drift and Transitory Component of S&P 500 Returns.
This figure plots the stochastic drift $\widehat{x}_t$ and the transitory component $\widehat{z}_t$ of S&P 500 returns for the PT-CV model (i.e. model featuring a permanent component with stochastic drift, a transitory component, and constant volatility). Data are at the monthly frequency from 02/1871 to 02/2016.

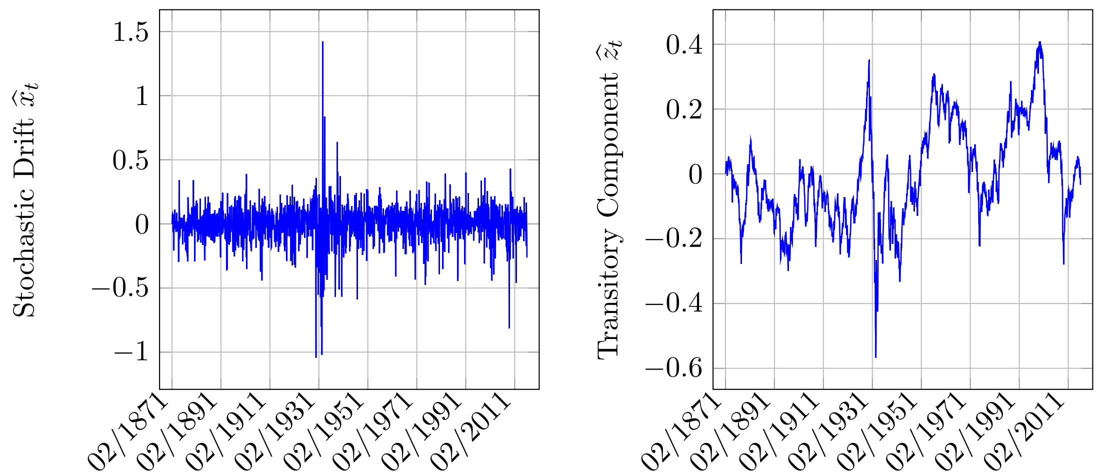


Table 3: Stochastic Drift, Transitory Component, and S&P 500 Dividend Yield.
This table reports the estimates from the regression of the S&P 500 log dividend yield $\widetilde{y}$ on $\widehat{x}$ and $\widehat{z}$:

$$\widetilde{y}_t = a + b_x \widehat{x}_t + b_z \widehat{z}_t + \epsilon_t.$$

Newey-West t-statistics are reported in parentheses and 10%, 5%, and 1% significance levels are denoted with *, **, and ***, respectively. Data are at the monthly frequency from 02/1871 to 02/2016.

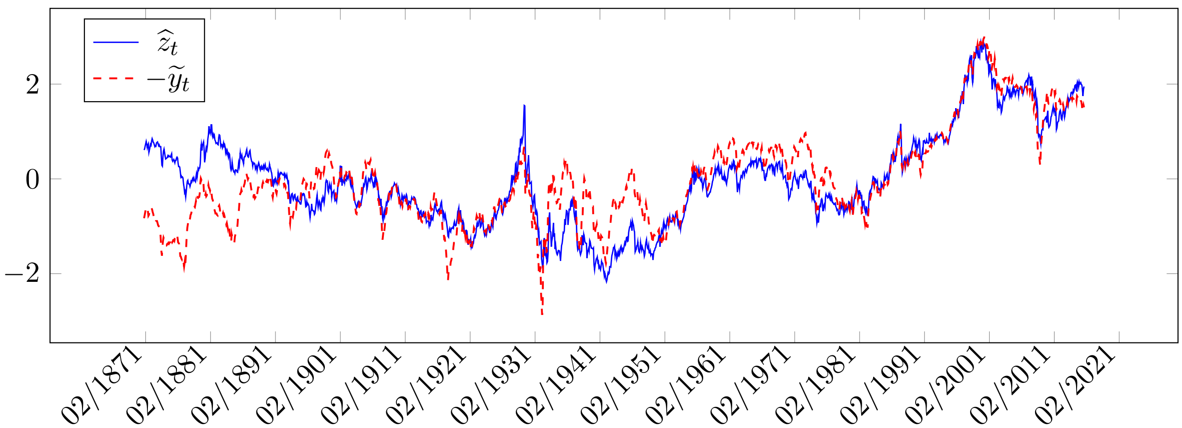
	Constant Volatility			Stochastic Volatility		
	(1)	(2)	(3)	(4)	(5)	(6)
$\widehat{x}$	-0.353***		-0.040	-0.830***		-0.221***
t-stat	(-4.16)		(-0.60)	(-5.57)		(-2.72)
$\widehat{z}$		-1.710***	-1.704***		-0.549***	-0.540***
t-stat		(-11.56)	(-11.27)		(-20.56)	(-19.53)
adj-R^2	0.01	0.37	0.37	0.05	0.65	0.66

equity compensation is much weaker. This is not surprising in light of the results in Figure 2 and Table 1. Dividend yield variations are indeed closely related to short-horizon variance ratios, which capture the importance of transitory risk relative to permanent risk.

To show the importance of modeling the transitory component $\widehat{z}_t$, we plot in Figure 4 the time series of $\widehat{z}_t$ and minus the log dividend yield. Figure 4 shows that $\widehat{z}_t$ matches the cyclical fluctuations

Figure 4: Times Series of the Transitory Component and S&P 500 Dividend Yield.

Time series of the S&P 500 (standardized) minus log dividend yield (i.e., the log price-dividend ratio) and the (standardized) transitory component $\widehat{z}$ for the PT-SV model (i.e., model featuring a permanent component with stochastic drift, a transitory component, and stochastic volatility). Data are at the monthly frequency from 02/1871 to 02/2016.



of $\widetilde{y}_t$ in the 20th century. That is, a model of stock returns that disregards the transitory component fails to capture the time-variation of expected returns (i.e., the dividend yield dynamics) and, hence, cannot accurately describe risk-adjusted returns.

4.2 Parameters Driving Dividends

The parameter vector, $\Theta_d \equiv (m_s, a_x, a_z)$, driving the dividend yield and the dividend growth rate in (14) and (15), respectively, is estimated by the method of moments. That is, Θ_d satisfies

$$\Theta_d = \arg\min \Lambda(\Theta_d)^\top \Lambda(\Theta_d),$$

where the moment vector $\Lambda(\Theta_d)$ consists of the relative error between the actual and model-implied *1)* mean dividend yield, *2)* mean dividend growth, *3)* dividend yield volatility, and *4)* dividend growth volatility.

Table 4 provides the estimates and standard errors of Θ_d as well as the corresponding actual and model-implied moments. While all models accurately match the actual mean dividend yield and mean dividend growth rate, models ignoring the transitory component of stock returns imply an almost constant dividend yield and therefore largely overvalue the volatility of dividend growth. In contrast, the two models that take into account the transitory component match each of the four moments relatively well. Since the parameter a_x is very close to 1 and the parameter a_z is smaller than 1, the model-implied dividend yield is almost exclusively driven by the transitory component and the latter relation is negative. This is perfectly in line with our previous findings that the actual dividend yield is strongly inversely related to the transitory component $\widehat{z}_t$ and negligibly driven by the stochastic drift $\widehat{x}_t$ (see Table 3).

Furthermore, models accounting for the transitory component generate a high correlation between the model-implied and actual dividend yield.[18] This implies that models accounting for the transitory component can replicate the observed relation between the dividend yield and future returns, while models ignoring it cannot, as discussed in Section 4.3 below.

Table 4: Parameters Driving Dividends and Corresponding Moments.
This table reports the values of the parameters driving dividend growth and the dividend yield for the different models. P, PT, CV, SV, and CM stand for permanent component with stochastic drift, permanent component with stochastic drift and transitory component, constant volatility, stochastic volatility, and constant mean, respectively. Standard errors are reported in brackets and statistical significance at the 10%, 5%, and 1% levels is labeled with *, **, and ***, respectively. Parameters are estimated using monthly S&P 500 dividend growth rates and dividend yields from 02/1871 to 02/2016.

Parameter	Symbol	CM-CV	P-CV	PT-CV	P-SV	PT-SV
Long-term expected	m_s	0.047***	0.046***	0.046***	0.042***	0.041***
stock price growth		(6.6×10^{-4})	(7×10^{-4})	(0.001)	(7×10^{-4})	(0.001)
Stock price growth	a_x		1.002***	0.999***	1.009***	0.999***
sensitivity to $\widehat{x}_t$			(0.003)	(0.004)	(0.003)	(0.033)
Stock price growth	a_z			0.465***		0.758***
sensitivity to $\widehat{z}_t$				(0.014)		(0.018)

Moments	Data	CM-CV	P-CV	PT-CV	P-SV	PT-SV
Mean dividend yield	0.044	0.047	0.048	0.050	0.049	0.050
Mean dividend growth	0.035	0.038	0.038	0.034	0.037	0.034
Dividend yield volatility	0.017	0.000	0.000	0.012	0.001	0.012
Dividend growth volatility	0.039	0.134	0.127	0.052	0.106	0.085

4.3 Predictive Regressions

While the dividend yield is a well-known predictor of stock returns (Cochrane, 2008; van Binsbergen and Koijen, 2010), we intend to verify to what extent the model-implied dividend yield replicates the forecasting power of the actual dividend yield.

Using the definition of the dividend yield in (14) and the estimates from Sections 4.1 and 4.2, we obtain the historical model-implied path of the log dividend yield, $\widetilde{y}_t$, as follows:

$$\widetilde{y}_t = \log\left(m - m_s + (1 - a_x)\widehat{x}_t + (a_z - 1)\kappa_z \widehat{z}_t\right).$$

We regress cumulative log S&P 500 returns on either the S&P 500 log dividend yield or the model-implied dividend yield. Monthly data are aggregated at yearly frequency and the forecasting

[18] The correlation is 0.62 and 0.83 in the PT-CV and PT-SV models, respectively, whereas it is only 0.09 and 0.13 in the P-CV and P-SV models (see Table 13 in Appendix D).

horizon ranges from 1 to 10 years. Panel (a) of Table 5 reports the estimates from the regressions performed on the 1872-2015 sample.

Table 5: Return and Dividend Growth Predictability by Dividend Yield.
Panels (a) and (b) report the estimates obtained by regressing the cumulative log S&P 500 return, r, and the cumulative log dividend growth rate, g, respectively, over several horizons on the log dividend yield, $\widetilde{y}$:

$$\sum_{i=1}^{n} r_{t+i} = a_r + b_r \widetilde{y}_t + \epsilon_{r,t}, \qquad \sum_{i=1}^{n} g_{t+i} = a_g + b_g \widetilde{y}_t + \epsilon_{g,t},$$

where $n = \{1, 2, 3, 5, 7, 10\}$ years. The actual log dividend yield and actual log dividend growth rate are those of the S&P 500 ("Data" panels), while the model-implied counterparts ("Model" panels) satisfy $\widetilde{y}_t = \log(m - m_s + (1 - a_x)\widehat{x}_t + (a_z - 1)\kappa_z \widehat{z}_t)$ and $D_t = P_t exp(\widetilde{y}_t)$ (see Section 3.3). Newey-West t-statistics are reported in parentheses and 10%, 5%, and 1% significance levels are denoted with *, **, and ***, respectively. Monthly data are aggregated at yearly frequency over the 1872-2015 sample. P, PT, CV, and SV, stand for permanent component with stochastic drift, permanent component with stochastic drift and transitory component, constant volatility, and stochastic volatility, respectively.

Panel (a), full sample 1872–2015: $\sum_{i=1}^{n} r_{t+i}$	1	2	3	5	7	10
Data						
$\widetilde{y}$	0.067**	0.115**	0.164**	0.281***	0.340***	0.380***
t-stat	(2.12)	(2.15)	(2.36)	(3.33)	(3.49)	(2.74)
adj-R^2	0.02	0.03	0.05	0.09	0.12	0.09
P-CV Model ($\kappa_z = 0$)						
$\widetilde{y}$	4.507	27.559**	15.884	42.830***	21.890	11.694
t-stat	(0.52)	(2.00)	(1.60)	(3.20)	(1.10)	(0.60)
adj-R^2	-0.01	0.02	0.00	0.03	0.00	-0.01
P-SV Model ($\kappa_z = 0$)						
$\widetilde{y}$	1.111	4.544*	3.063*	8.436***	4.555	2.533
t-stat	(0.69)	(1.97)	(1.88)	(3.15)	(1.27)	(0.74)
adj-R^2	-0.00	0.02	0.00	0.04	0.00	-0.00
PT-CV Model						
$\widetilde{y}$	0.106	0.181	0.237	0.274	0.250	0.410
t-stat	(1.63)	(1.44)	(1.41)	(1.48)	(1.27)	(1.62)
adj-R^2	0.02	0.03	0.04	0.04	0.02	0.05
PT-SV Model						
$\widetilde{y}$	0.129**	0.245***	0.347***	0.488***	0.590***	0.883***
t-stat	(2.56)	(2.76)	(3.10)	(3.61)	(3.53)	(3.87)
adj-R^2	0.03	0.05	0.09	0.11	0.14	0.20

Panel (b), full sample 1872–2015: $\sum_{i=1}^{n} g_{t+i}$	1	2	3	5	7	10
Data						
$\widetilde{y}$	-0.071**	-0.099**	-0.111**	-0.111**	-0.145**	-0.228**
t-stat	(-2.21)	(-2.14)	(-2.09)	(-2.00)	(-2.49)	(-2.38)
adj-R^2	0.06	0.04	0.03	0.02	0.03	0.06
P-CV Model ($\kappa_z = 0$)						
$\widetilde{y}$	-2.760	21.971	12.095	45.472***	17.246	7.677
t-stat	(-0.28)	(1.63)	(1.26)	(3.13)	(0.95)	(0.37)
adj-R^2	-0.01	0.01	-0.00	0.03	-0.00	-0.01
P-SV Model ($\kappa_z = 0$)						
$\widetilde{y}$	-0.492	2.763	1.376	6.497***	2.944	1.146
t-stat	(-0.31)	(1.13)	(0.80)	(2.73)	(0.93)	(0.35)
adj-R^2	-0.01	0.00	-0.01	0.02	-0.00	-0.01
PT-CV Model						
$\widetilde{y}$	-0.084***	-0.159***	-0.229***	-0.353***	-0.469***	-0.541***
t-stat	(-4.34)	(-3.90)	(-3.77)	(-4.42)	(-5.21)	(-4.15)
adj-R^2	0.13	0.19	0.26	0.31	0.34	0.27
PT-SV Model						
$\widetilde{y}$	0.036	0.066	0.085	0.106	0.113	0.204
t-stat	(1.27)	(1.17)	(1.10)	(1.06)	(0.94)	(1.14)
adj-R^2	0.00	0.01	0.01	0.01	0.01	0.03

The estimates from actual data show that return predictability is limited (adjusted R^2 varies from 2% to 12%) but the positive relation between future returns and the log dividend yield is significant at every horizon. The models ignoring the transitory component (P-CV and P-SV) are unable to capture the pattern of predictability observed in actual data. Indeed, slopes are unrealistically large and statistically insignificant, while the explanatory power is nil. Instead, the models accounting for the transitory component (PT-CV and PT-SV) imply slopes and adjusted R^2 of magnitudes comparable to those obtained with actual data. While slopes are insignificant in the PT-CV model featuring constant return volatility, they are highly significant in the PT-SV model featuring stochastic return volatility. Consistent with the actual predictability pattern, the PT-SV model implies that the slope and the adjusted R^2 increase with the horizon. Moreover, the adjusted R^2 varies from 3% to 20% which is larger than what is obtained with actual data. Thus, the PT-SV model-implied log dividend yield conveys more precise information about the conditional mean of returns than does the actual log dividend yield. Overall, accounting for the transitory component $\widehat{z}_t$ is important in order to replicate the return predictability properties observed in actual data.

We also investigate the predictability of dividend growth by regressing the cumulative log dividend growth rate on the log dividend yield (Cochrane, 2008; van Binsbergen and Koijen, 2010) using either actual data or the model-implied counterparts. Regression results are reported in Panel (b) of Table 5. Actual data show that dividend growth predictability is limited (adjusted R^2 varies from 2% to 6%) but the negative relation is significant at every horizon. The models ignoring the transitory component (P-CV and P-SV) generate a positive and in some cases significant intertemporal relation, which is inconsistent with actual data. The models accounting for the transitory component (PT-CV and PT-SV), instead, give rise to a more realistic intertemporal relation between the dividend yield and future dividend growth. In particular, the PT-CV model generates a negative and statistically significant relation.[19]

5 Results

In this section, we first illustrate that accounting for the transitory component of stock returns enables us to replicate the observed downward-sloping term structures of HPR volatility and equity yield volatility. Second, we discuss the impact of the transitory component on the optimal portfolio. Third, we compare different investment strategies and show that considering the transitory component implies surges in portfolio returns.

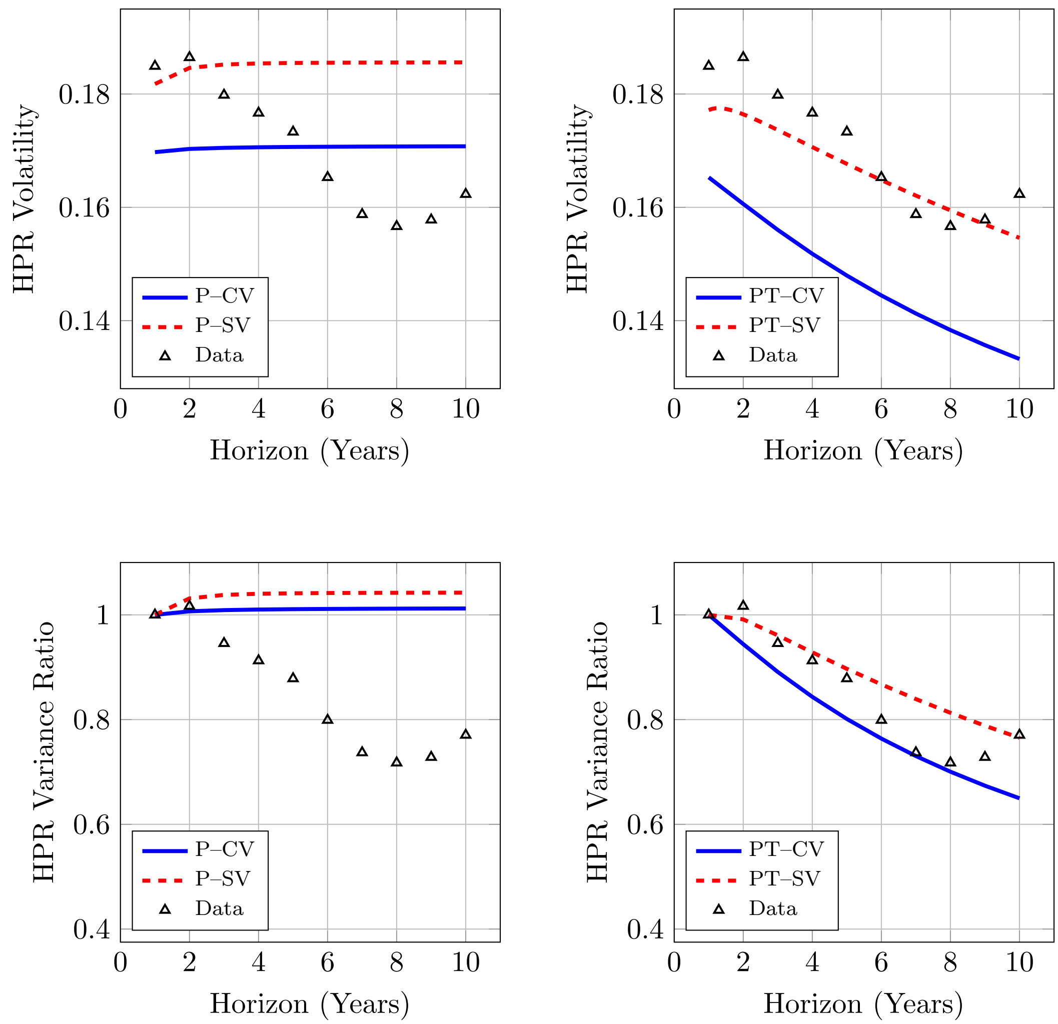
5.1 Model-Implied Timing of Equity Risk

We investigate the model-implied timing of risk by focusing on term structures of HPR volatility and equity yield volatility, as well as on the corresponding variance ratios.

We use the 1-year horizon as a reference and compute the HPR variance ratio, $\mathrm{VR}_{S,t}(\tau)$, and

[19] In Appendix C, we document the predictive power of other variables that are external to our model, such as labour share and financial leverage (see Belo et al. (2015) and Marfè (2016)). While dividend growth is predicted by both ratios in actual data, the model-implied dividend growth is predicted by financial leverage only.

Figure 5: Holding Period Stock Return Volatility and Variance Ratio.
The left panels show the holding period stock return volatility and variance ratio for horizons ranging from 1 to 10 years in the model with a permanent component with stochastic drift (P). The right panels depict the holding period stock return volatility and variance ratio in the model with both a permanent component with stochastic drift and a transitory component (PT). CV stands for constant volatility model and SV stands for stochastic volatility model. State variables are set to their long-term means and the parameter values are reported in Table 2. Data markers display the volatilities of nominal S&P 500 returns from the full sample (1872–2015).

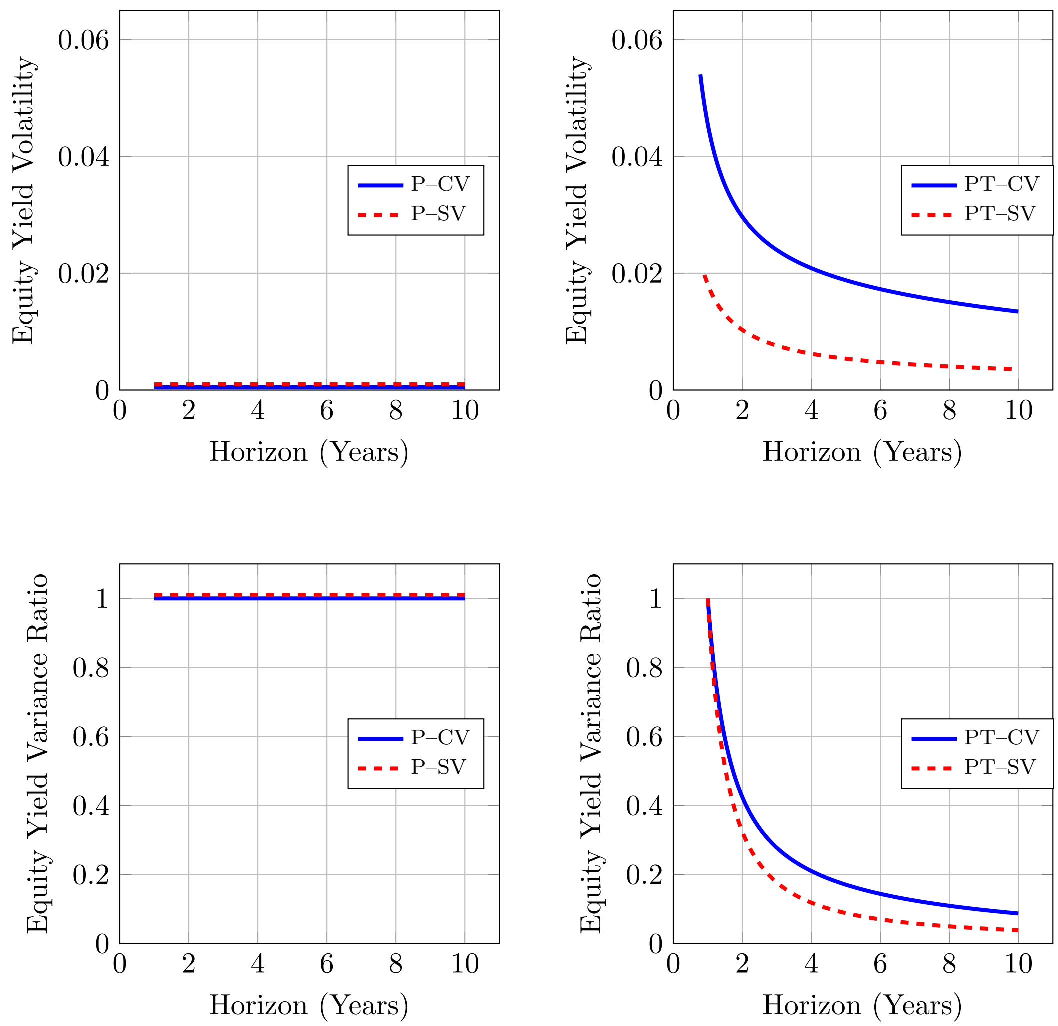


the equity yield variance ratio, $\mathrm{VR}_{P,t}(\tau)$, at the τ-year horizon as

$$\mathrm{VR}_{S,t}(\tau) = \frac{\mathrm{vol}_{S,t}(\tau)^2}{\mathrm{vol}_{S,t}(1)^2}, \qquad \mathrm{VR}_{e,t}(\tau) = \frac{\mathrm{vol}_{e,t}(\tau)^2}{\mathrm{vol}_{e,t}(1)^2}.$$

Figure 5 shows that adding the transitory component helps replicate the observed downward-sloping pattern of HPR volatilities. In line with empirical evidence, the model with both permanent and transitory components (and stochastic volatility) produces a 1-year HPR volatility of about 18% with a decline to about 15% at the 10-year horizon. In contrast, models that ignore the

Figure 6: Equity Yield Volatility and Variance Ratio.
The left panels show the equity yield volatility and variance ratio for horizons ranging from 1 to 10 years in the models with only the permanent component with stochastic drift (P). The right panels depict the equity yield volatility and variance ratio in the model with both permanent component with stochastic drift and transitory component (PT). CV stands for constant volatility model and SV stands for stochastic volatility model. State variables are set to their long-term means and the parameter values are reported in Table 2.

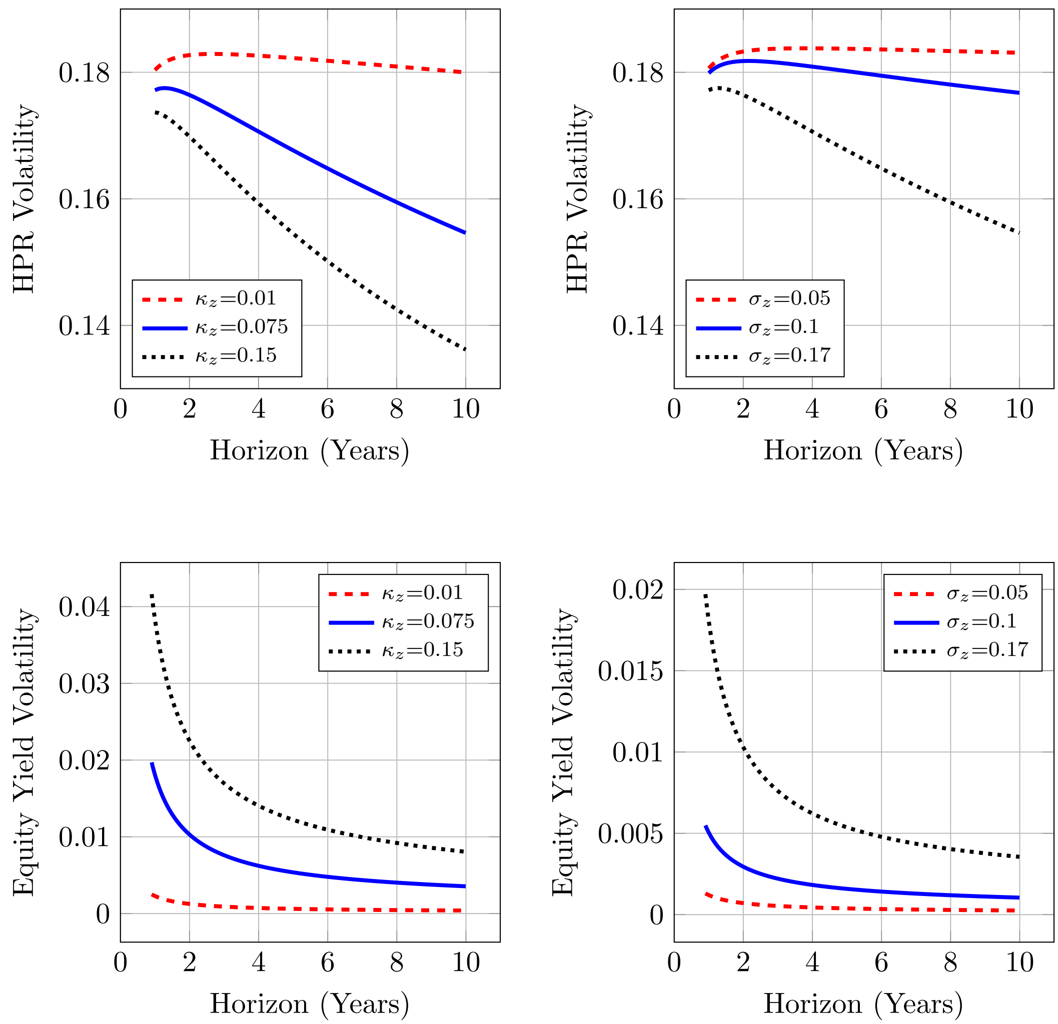


transitory component imply slightly upward-sloping holding period stock return volatilities, and therefore do not accurately replicate the observed timing of risk. Similarly, models with both a permanent component and a transitory component produce downward-sloping variance ratios. The magnitude of model-implied variance ratios for HPR is in line with their empirical counterparts. For example, a 10-year variance ratio in the model with permanent and transitory components is around 0.7, closely matching the value observed in the data. The variance ratios in the model with only a permanent component, on the contrary, are all around one.

Figure 6 shows the model-implied equity yield volatility and variance ratios. As previously shown in Proposition 5, ignoring the transitory component implies flat term structures of equity

Figure 7: Effect of the Mean-Reversion Speed and the Volatility of the Transitory Component on the Timing of Risk.

The left panels show how the volatility of holding period stock returns (top) and equity yields (bottom) react to changes in the mean-reversion speed of the transitory component κ_z. The right panels show how the volatility of holding period stock returns (top) and equity yields (bottom) react to changes in the volatility of the transitory component σ_z. This figure considers the stochastic volatility model. State variables are set to their long-term means and the parameter values are reported in Table 2.



yield volatility and variance ratio. However, accounting for the transitory component allows the model to generate strongly downward-sloping term structures, with both constant and stochastic return volatility. That is, only the models accounting for the transitory component replicate the empirical findings of van Binsbergen et al. (2013) and van Binsbergen and Koijen (2017).

Figure 7 illustrates the sensitivity of the timing of risk to changes in the parameters driving the transitory component: the mean-reversion speed κ_z and the volatility σ_z. The left panels show that when the transitory component mean-reverts faster, risk concentrates in the short term, which implies steeper term structures of risk. The reason is that κ_z governs the speed at which transitory shocks die out. Hence the larger κ_z, the higher the risk in the short run relative to the long run. The right panels show that a more volatile transitory component also implies steeper term structures of risk. The reason is that increasing σ_z increases short-term risk but has no influence on long-term risk. Finally, it is worth noting that the overall level of the HPR volatilities is lower for larger values of κ_z and σ_z, which might appear surprising. The reason is twofold. First, the 0-year HPR volatility is equal to $\sqrt{\bar{v}}$, which is independent of κ_z and σ_z. Second, we have just shown that larger values of κ_z and σ_z imply more negatively-sloped term structures of HPR volatility.[20]

5.2 Optimal Portfolio and Hedging Demands

We next investigate the properties of the investor's portfolio. Table 6 provides the decomposition of the investor's portfolio for different coefficients of relative risk aversion (5, 7, and 10) and investment horizons (1 month, 1 year, and 5 years). The optimal proportion of wealth allocated to the stock, the myopic stock allocation, and the different components of hedging demands are reported using the long-term mean values of the state variables.

Let us first analyze the signs of the portfolio components. The myopic demand is positive as long as the expected excess return on the stock is positive. The signs of the hedging demands depend on the correlation between stock returns and the state variables, as well as on the sensitivity of the marginal utility to changes in the state variables. To hedge transitory shocks, investors tend to increase their risky investment, while to hedge the stochastic drift, investors decrease their risky investment. The opposite signs of the hedging demands are due to the opposite effects the two shocks have on expected returns. Since the transitory component is positively correlated with realized returns and negatively impacts expected returns, the stock is a good hedging instrument against a deterioration in investment opportunities; therefore this hedging demand is positive. On the other hand, positive shocks to the permanent component increase expected returns and coincide with positive realized returns. This makes the stock a bad hedging instrument and therefore implies a negative hedging demand. Similarly, the negative correlation between volatility and realized returns gives rise to a negative hedging demand.

The magnitude of the hedging demands is smaller than that of the myopic stock allocation, but is not negligible. For example, in the constant return volatility model with a 5-year horizon, the hedge of transitory shocks amounts to around 30% of the myopic demand.

The magnitude of the hedging demands increases with the investment horizon. This effect is particularly pronounced for the transitory shock hedging demand and for the volatility hedging demand. Indeed, the magnitude of these hedging demands increases tenfold as the investment horizon increases from 1 month to 1 year. The impact of the horizon on the magnitude of the stochastic drift hedging demand, however, is weaker. In the constant return volatility model, the strong positive relation between the horizon and the transitory shock hedging demand dominates

[20]In the case of equity yields, the impact of transitory shocks is not only influenced by κ_z but also by the loading a_z. The high value of the latter amplifies the sensitivity of equity yields to temporary shocks, and therefore generates a steeper term structure of equity yield volatility.

Table 6: Decomposition of Optimal Stock Allocation.
This table presents the optimal stock allocation, the myopic allocation, and the hedging demands for the two cases: the model with constant return volatility and the model with stochastic return volatility. We consider investors with coefficients of relative risk aversion γ of 5, 7, and 10, and investment horizons $T-t$ of 1 month, 1 year, and 5 years. State variables are set to their long-term means and the parameter values are reported in Table 2.

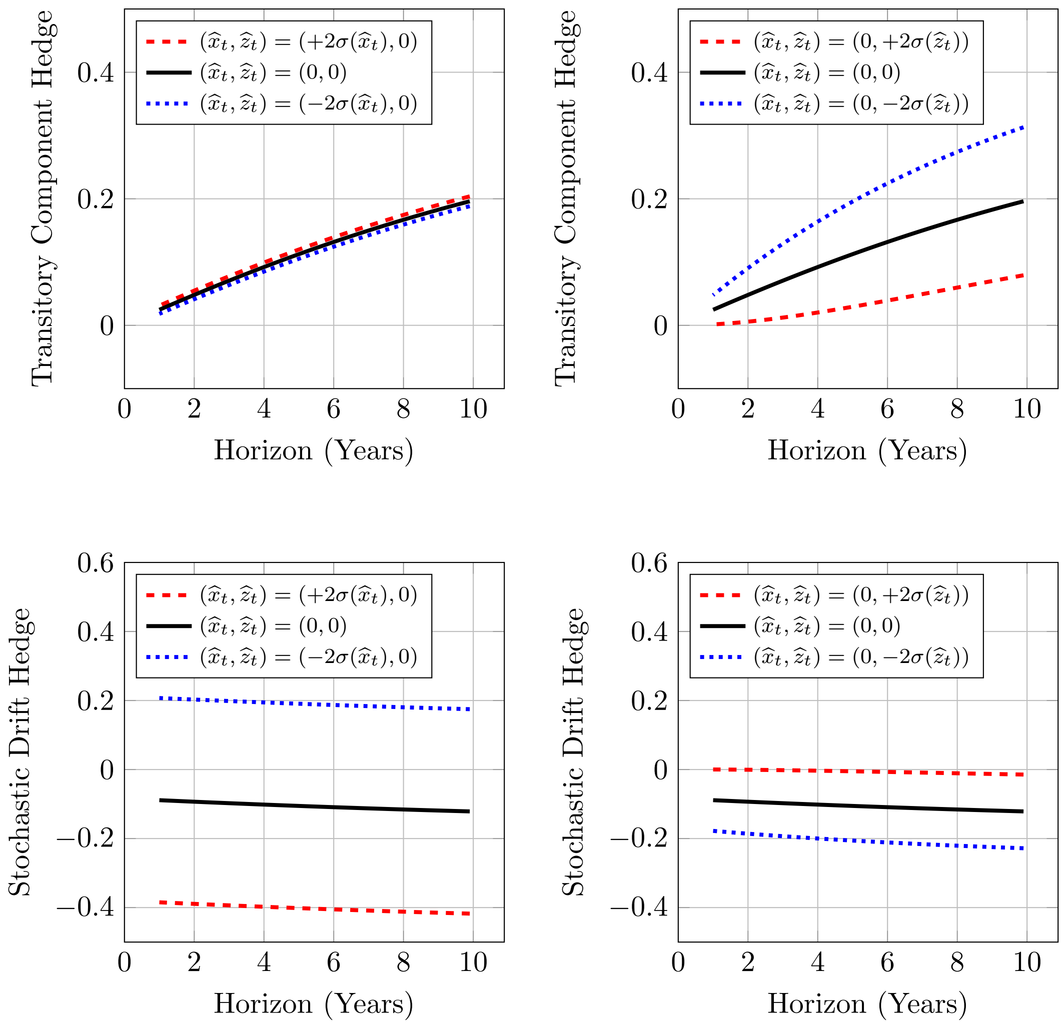
	Constant volatility model			Stochastic volatility model		
γ	1m	1y	5y	1m	1y	5y
	Mean optimal allocation to stocks					
5	0.4048	0.4220	0.5104	0.3475	0.2822	0.2788
7	0.2853	0.2981	0.3645	0.2455	0.1976	0.1948
10	0.1978	0.2070	0.2551	0.1705	0.1364	0.1342
	Myopic stock allocation					
5	0.5017	0.5017	0.5017	0.4121	0.4121	0.4121
7	0.3583	0.3583	0.3583	0.2944	0.2944	0.2944
10	0.2508	0.2508	0.2508	0.2061	0.2061	0.2061
	Stock allocation to hedge the stochastic drift $\widehat{x}_t$					
5	-0.0998	-0.1103	-0.1309	-0.0609	-0.0870	-0.0844
7	-0.0752	-0.0831	-0.0996	-0.0460	-0.0644	-0.0623
10	-0.0546	-0.0604	-0.0729	-0.0335	-0.0463	-0.0446
	Stock allocation to hedge the transitory shock $\widehat{z}_t$					
5	0.0030	0.0306	0.1397	0.0014	0.0100	0.0387
7	0.0022	0.0229	0.1058	0.0011	0.0074	0.0284
10	0.0016	0.0166	0.0772	0.0008	0.0053	0.0202
	Stock allocation to hedge the return variance v_t					
5				-0.0052	-0.0529	-0.0877
7				-0.0039	-0.0397	-0.0657
10				-0.0028	-0.0287	-0.0475

the weak negative relation between the horizon and the stochastic drift hedging demand, resulting in an overall positive horizon effect. That is, investors with a longer horizon tend to invest more in the stock. In the stochastic return volatility model, the strong negative relation between the horizon and the volatility hedging demand dominates. Thus, investors with longer horizons choose a less risky share.

We further investigate the signs, magnitudes, and long-horizon patterns of the hedging demands as the state variables deviate from their long-term means. Figure 8 illustrates the effect of varying the stochastic drift of the permanent component (left panels) and the transitory component (right panels) in the constant return volatility model. The shape of the hedging demands remain unchanged when varying the aforementioned state variables. Indeed, the transitory component hedging demand remains positive and increases with the horizon. The stochastic drift hedging

Figure 8: Effect of the Transitory Component and the Drift of the Permanent Component on Hedging Demands.

The left panels show the hedging demands as a function of the investment horizon for different values of the drift of the permanent component $\widehat{x}_t$. The curves correspond to $\widehat{x}_t = 0, \pm 2\,\sigma\left(\widehat{x}_t\right)$, where $\sigma\left(\widehat{x}_t\right)$ is the long-term standard deviation of the stochastic drift. The right panels show the hedging demands as a function of the investment horizon for different values of the transitory component $\widehat{z}_t$. The curves correspond to $\widehat{z}_t = 0, \pm 2\,\sigma\left(\widehat{z}_t\right)$, where $\sigma\left(\widehat{z}_t\right)$ is the long-term standard deviation of the transitory component. If not stated otherwise, state variables are set to their long-term means and the parameter values are reported in Table 2. The numbers reported are based on the constant return volatility model.



demand depends weakly on the horizon and remains negative.[21]

[21] An exception occurs when $\widehat{x}_t$ is two standard deviations below its long-term mean. The expected excess return becomes negative, the investor shorts the stock, and therefore the sign of the hedging demand is reversed. This effect is temporary and quickly disappears because $\widehat{x}_t$ reverts back at a high speed towards its long-term mean.

Table 7: Unconditional Out-of-Sample Portfolio Moments.
Mean, volatility, and Sharpe ratio are in annualized terms. The investment horizon is $T = 1$, risk aversion is $\gamma = 5$, and initial wealth is $W_0 = \$1$. P, PT, CM, CV, and SV stand for permanent component with stochastic drift, permanent component with stochastic drift and transitory component, constant mean, constant volatility, and stochastic volatility, respectively. Statistics are computed using monthly S&P 500 data from 02/1900 to 02/2016.

	Mean	Volatility	Sharpe Ratio	Skewness	Kurtosis	Terminal Wealth
CM-CV	1.19%	4.24%	0.28	-1.63	18.73	1,162
P-CV	24.38%	39.25%	0.62	7.58	101.88	8.0×10^{11}
PT-CV	23.28%	43.82%	0.53	9.53	153.15	8.5×10^{10}
P-SV	18.54%	31.05%	0.60	4.96	98.38	1.2×10^{9}
PT-SV	21.78%	35.92%	0.61	5.63	98.73	1.4×10^{10}

5.3 Out-of-Sample Portfolio Performance

To investigate the out-of-sample performance of the different investment strategies, we assume that the investor re-estimates the parameters of the models every year.[22] More precisely, we assume that the investor starts by using data from 02/1871 to 01/1900 to estimate the first set of parameters. Assuming an investment horizon of one year, the investor trades and updates her beliefs about $\widehat{x}_t$ and $\widehat{z}_t$ from 02/1900 to 01/1901 by fixing the set of estimated parameters. The investor then uses data from 02/1871 to 01/1901 to estimate a second set of parameters that allows her to trade and update her beliefs from 02/1901 to 01/1902. This process is undertaken until reaching the terminal date 02/2016, and therefore provides a time series of out-of-sample monthly portfolio returns from 02/1900 to 02/2016.[23]

Since the investor re-estimates the parameters of the model every year, parameters are constant over twelve consecutive months and then change. To understand how the shape of the term structure of equity risk varies with economic conditions, we compute the median mean-reversion speed and volatility of the transitory component during NBER recessions and expansions. For the model with constant return volatility (PT-CV model), the median mean-reversion speed is 0.29 in recessions and 0.19 in expansions, whereas the median volatility is 0.18 both in recessions and expansions. For the model with stochastic return volatility (PT-SV model), the median mean-reversion speed is 0.22 in recessions and 0.17 in expansions, whereas the median volatility is 0.18 both in recessions and expansions. Since an increase in either the mean-reversion speed or the volatility of the transitory component implies a more negatively sloped term structure of risk (Section 5.1), the term structure of risk is particularly downward-sloping in recessions. This result is consistent with the findings of van Binsbergen et al. (2013) and Aït-Sahalia et al. (2015), who show that the slope of the term structure of equity yield volatility is pro-cyclical.

Table 7 provides the portfolio return moments of the different strategies, and Table 8 reports the decomposition of the portfolio return volatility and kurtosis into their good and bad components.

[22] In Appendix E we replicate the analysis of the current section *in-sample*, and show that the results are robust.

[23] As in Johannes et al. (2014), we prevent the occurrence of a few extreme portfolio positions typically obtained during periods of very low return volatility by bounding the portfolio weight at -7 and $+7$.

Table 8: Unconditional Out-of-Sample Portfolio Return Volatility and Kurtosis Decomposition.
Volatility is in annualized terms. The investment horizon is $T = 1$ and risk aversion is $\gamma = 5$. P, PT, CM, CV, and SV stand for permanent component with stochastic drift, permanent component with stochastic drift and transitory component, constant mean, constant volatility, and stochastic volatility, respectively. Statistics are computed using monthly S&P 500 data from 02/1900 to 02/2016.

	Good Volatility	Bad Volatility	Good Kurtosis	Bad Kurtosis
CM-CV	2.64%	3.32%	2.34	16.39
P-CV	35.63%	16.47%	100.54	1.33
PT-CV	40.35%	17.10%	151.94	1.21
P-SV	26.01%	16.96%	86.52	11.85
PT-SV	30.62%	18.77%	92.06	6.67

Following Segal et al. (2015), annualized good and bad volatilities, GV and BV, are measured as follows:

$$GV = \sqrt{\frac{\Delta}{N}\sum_{i=1}^{N}\mathbf{1}_{\left\{r^p_{i\Delta}\geq\bar{r}^p\right\}}\left(r^p_{i\Delta}-\bar{r}^p\right)^2}, \qquad BV = \sqrt{\frac{\Delta}{N}\sum_{i=1}^{N}\mathbf{1}_{\left\{r^p_{i\Delta}<\bar{r}^p\right\}}\left(r^p_{i\Delta}-\bar{r}^p\right)^2},$$

where r^p is the portfolio return and $\bar{r}^p$ is the sample average of portfolio returns. Similarly, good and bad kurtosis, GK and BK are computed as follows:

$$GK = \frac{\sqrt{\frac{1}{N}\sum_{i=1}^{N}\mathbf{1}_{\left\{r^p_{i\Delta}\geq\bar{r}^p\right\}}\left(r^p_{i\Delta}-\bar{r}^p\right)^4}}{\mathrm{Var}(r^p)^2}, \qquad BK = \frac{\sqrt{\frac{1}{N}\sum_{i=1}^{N}\mathbf{1}_{\left\{r^p_{i\Delta}<\bar{r}^p\right\}}\left(r^p_{i\Delta}-\bar{r}^p\right)^4}}{\mathrm{Var}(r^p)^2},$$

where $\mathrm{Var}(r^p)$ is the sample variance of portfolio returns.

Tables 7 and 8 show that all learning-based strategies beat the constant mean and volatility strategy (CM-CV) in terms of Sharpe ratio, good relative to bad volatility, skewness, and good relative to bad kurtosis. That is, learning significantly improves portfolio performance in all dimensions. This result is particularly interesting given previous evidence that the out-of-sample predictive power and portfolio performance of a constant mean and volatility model are particularly hard to beat (Welch and Goyal, 2008; Johannes et al., 2014).

Adding a transitory component to the model increases skewness, good kurtosis, and good volatility, decreases bad kurtosis, and has a weak impact on bad volatility irrespective of whether return volatility is constant or stochastic. The Sharpe ratio, however, decreases when return volatility is constant and is unaffected when it is stochastic. This shows that the benefits of adding a transitory component to the model are particularly important when return volatility is stochastic.

Surprisingly, stochastic volatility in the model with both permanent and transitory components (PT) has out-of-sample implications that are quite different from the in-sample implications. Indeed, while stochastic volatility tends to increase the Sharpe ratio and decrease both skewness and kurtosis out-of-sample, it decreases the Sharpe ratio and increases both skewness and kur-

Table 9: Out-of-Sample Portfolio Moments in NBER Recession and Expansion.
Mean, volatility, and Sharpe ratio are in annualized terms. The investment horizon is $T = 1$ and risk aversion is $\gamma = 5$. P, PT, CM, CV, and SV stand for permanent component with stochastic drift, permanent component with stochastic drift and transitory component, constant mean, constant volatility, and stochastic volatility, respectively. Statistics are computed using monthly S&P 500 data from 02/1900 to 02/2016.

	Mean		Volatility		Sharpe Ratio		Skewness		Kurtosis	
	Exp.	Rec.	Exp.	Rec.	Exp.	Rec.	Exp.	Rec.	Exp.	Rec.
CM-CV	2.61%	-2.99%	3.35%	6.00%	0.78	-0.50	-0.35	-1.68	5.77	15.28
P-CV	16.72%	46.87%	25.44%	64.14%	0.66	0.73	6.18	5.46	108.37	46.41
PT-CV	14.46%	49.17%	27.23%	72.92%	0.53	0.67	6.88	6.85	122.85	69.16
P-SV	16.45%	24.70%	23.81%	46.16%	0.69	0.54	5.63	3.60	99.26	58.45
PT-SV	19.23%	29.29%	29.23%	50.69%	0.66	0.58	4.39	5.35	56.92	79.07

tosis in-sample. This shows that in-sample portfolio performance analyses can be misleading and might therefore be completely uninformative regarding the performance of actual real-time portfolio strategies.

We have provided evidence that the observed slope of the term structure of HPR variance ratios is negative unconditionally, and steepens as economic conditions deteriorate (Section 2). In addition, the model-implied term structure of HPR variance ratios is potentially downward-sloping if the transitory component of returns is considered, whereas it is upward-sloping if the transitory component is ignored (Section 5.1). Therefore, the portfolio benefits of considering the transitory component should be larger in bad times than in good times. To investigate this matter, we report in Tables 9 and 10 the portfolio moments of each strategy conditional on being in NBER recession or expansion. As expected, the unconditional portfolio benefits obtained by considering a transitory component mostly come from the significantly better performance observed in NBER recession. Indeed, the relative increase in skewness, good kurtosis, and good volatility obtained by considering the transitory component is larger in recession than in expansion. In addition, the transitory component decreases bad volatility and bad kurtosis in recession but not necessarily in expansion. When return volatility is stochastic, the transitory component increases the Sharpe ratio in recession more than it decreases the Sharpe ratio in expansion, lending further support to the notion that the benefits of the transitory component concentrate in bad times.

To further quantify the relative performance of the different strategies without directly relying on portfolio return moments, we follow Johannes et al. (2014) and define the certainty equivalent return, $CER_{m,t}$, of model m at time t, as follows:

$$J(t, X_{m,t}; m) \equiv e^{-\delta t} \frac{\left(W_t e^{(T-t)CER_{m,t}}\right)^{1-\gamma}}{1-\gamma}, \qquad m \in \{\text{CM-CV}, \text{P-CV}, \text{PT-CV}, \text{P-SV}, \text{PT-SV}\},$$

where $J(.,.;m)$ and $X_{m,t}$ are the value function and the vector of state variables associated with model m, respectively. That is, the certainty equivalent return is defined as the yield of a fictitious riskless asset that makes the investor indifferent between implementing her optimal risky strategy

Table 10: Out-of-Sample Portfolio Return Volatility and Kurtosis Decomposition in NBER Recession and Expansion.

Volatility is in annualized terms. The investment horizon is $T = 1$ and risk aversion is $\gamma = 5$. P, PT, CM, CV, and SV stand for permanent component with stochastic drift, permanent component with stochastic drift and transitory component, constant mean, constant volatility, and stochastic volatility, respectively. Statistics are computed using monthly S&P 500 data from 02/1900 to 02/2016.

	Good Volatility		Bad Volatility		Good Kurtosis		Bad Kurtosis	
	Exp.	Rec.	Exp.	Rec.	Exp.	Rec.	Exp.	Rec.
CM-CV	2.28%	3.61%	2.46%	4.79%	2.11	2.04	3.66	13.24
P-CV	21.68%	59.06%	13.32%	25.02%	101.32	46.15	7.05	0.26
PT-CV	23.43%	68.08%	13.87%	25.13%	114.96	68.96	7.89	0.20
P-SV	20.18%	38.11%	12.63%	26.04%	94.58	49.65	4.68	8.81
PT-SV	24.60%	43.60%	15.78%	25.85%	54.46	73.08	2.46	5.99

Table 11: Out-of-Sample Ratio of Certainty Equivalent Returns.

CER, Unc, Exp., and Rec. stand for certainty equivalent return, unconditional, expansion, and recession, respectively. The investment horizon is $T = 1$ and risk aversion is $\gamma = 5$. P, PT, CV, and SV stand for permanent component with stochastic drift, permanent component with stochastic drift and transitory component, constant volatility, and stochastic volatility, respectively. Statistics are computed using monthly S&P 500 data from 02/1900 to 02/2016. The unconditional mean (median) CERs in the two baseline cases are 15.60% (15.19%) in the P-CV model and 13.24% (12.16%) in the P-SV model.

	PT-CV/P-CV			PT-SV/P-SV		
	Unc.	Exp.	Rec	Unc.	Exp.	Rec
Ratio of Mean CER	1.0214	1.0142	1.0414	1.5898	1.6679	1.3849
Ratio of Median CER	0.9839	0.9787	1.0091	1.2072	1.2064	1.2119

and investing her entire wealth in this riskless asset.

Table 11 provides the ratio of certainty equivalent returns obtained by considering and then ignoring the transitory component of stock returns. The transitory component helps to increase the certainty equivalent return, and especially so when stock return volatility is stochastic. Indeed, the mean and median certainty equivalent return increase by about 60% and 20%, respectively, when volatility is stochastic. When return volatility is constant, however, the mean and median certainty equivalent returns increase and decrease by 2%, respectively. Since return volatility is indeed time-varying (Engle, 1982; Bollerslev, 1986), ignoring the transitory component of returns, and therefore mismodeling the observed timing of equity risk, is costly to investors. Furthermore, Table 11 shows that the benefits of considering the transitory component tend to be concentrated in recessions. The reason is that the term structure of equity risk is strongly downward-sloping in recessions, and the aim of the transitory component is precisely to model this downward-sloping pattern.

6 Conclusion

This paper provides evidence that properly modeling the shape of the term structure of equity risk significantly increases investors' portfolio performance. We assume that equity returns are driven by a permanent component with stochastic drift and volatility, and a transitory component. In contrast to existing models considering mean-reverting expected returns and stochastic return volatility, we show that our model is flexible enough to generate either flat, upward-sloping, or downward-sloping term structures of equity risk. The model replicates the observed dynamics of the dividend yield and its ability to predict future returns and future dividend growth rates. Importantly, the model mimics the observed downward-sloping term structure of equity risk. Both in-sample and out-of-sample, trading strategies that account for the observed shape of the term structure of equity risk significantly outperform those that do not. Indeed, properly modeling the timing of equity risk implies surges in portfolio returns, which significantly increases the strategy's certainty equivalent return. Finally, we show that out-performance is concentrated in recessions, periods in which the term structure of equity risk is the most downward-sloping.

References

Aït-Sahalia, Y., M. Karaman, and L. Mancini (2015). The term structure of variance swaps and risk premia. Unpublished manuscript.

Alvarez, F. and U. J. Jermann (2005, November). Using Asset Prices to Measure the Persistence of the Marginal Utility of Wealth. *Econometrica 73*(6), 1977–2016.

Andersen, T. G., O. Bondarenko, and M. T. Gonzalez-Perez (2015). Exploring return dynamics via corridor implied volatility. *The Review of Financial Studies 28*(10), 2902.

Andersen, T. G., N. Fusari, and V. Todorov (2015). The risk premia embedded in index options. *Journal of Financial Economics 117*(3), 558 – 584.

Andrei, D., B. Carlin, and M. Hasler (2015). Asset pricing with structural uncertainty and structural disagreement. *Working Paper*.

Andrei, D. and M. Hasler (2015). Investor attention and stock market volatility. *Review of Financial Studies 28*(1), 33–72.

Babaoglu, K., P. Christoffersen, S. L. Heston, and K. Jacobs (2016). Option valuation with volatility components, fat tails, and non-monotonic pricing kernels. *Working Paper*.

Bansal, R., D. Kiku, and A. Yaron (2010, September). Long run risks, the macroeconomy, and asset prices. *American Economic Review 100*(2), 542–46.

Bansal, R. and A. Yaron (2004). Risks for the long run: A potential resolution of asset pricing puzzles. *Journal of Finance 59*(4), 1481–1509.

Belo, F., P. Collin-Dufresne, and R. S. Goldstein (2015). Dividend dynamics and the term structure of dividend strips. *Journal of Finance 70*(3), 1115–1160.

Benzoni, L., P. Collin-Dufresne, and R. S. Goldstein (2011). Explaining asset pricing puzzles associated with the 1987 market crash. *Journal of Financial Economics 101*(3), 552 – 573.

Björk, T., Y. Kabanov, and W. Runggaldier (1997). Bond market structure in the presence of marked point processes. *Mathematical Finance 7*(2), 211–239.

Boguth, O., M. Carlson, A. J. Fisher, and M. Simutin (2012). Leverage and the limits of arbitrage pricing: Implications for dividend strips and the term structure of equity risk premia. *Working*.

Bollerslev, T. (1986, April). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics 31*(3), 307–327.

Campbell, J. Y., Y. L. Chan, and L. M. Viceira (2003). A multivariate model of strategic asset allocation. *Journal of Financial Economics 67*(1), 41–80.

Campbell, J. Y. and J. H. Cochrane (1999). By force of habit: A consumption-based explanation of aggregate stock market behavior. *Journal of Political Economy 107*, 205–251.

Campbell, J. Y. and R. J. Shiller (1988). The dividend-price ratio and expectations of future dividends and discount factors. *Review of Financial Studies 1*(3), 195–228.

Campbell, J. Y. and L. Viceira (2005). The Term Structure of the Risk-Return Tradeoff. *Financial Analysts Journal 61*, 34–44.

Campbell, J. Y. and L. M. Viceira (1999, May). Consumption and portfolio decisions when expected returns are time varying. *The Quarterly Journal of Economics 114*(2), 433–495.

Chacko, G. and L. M. Viceira (2005). Dynamic consumption and portfolio choice with stochastic volatility in incomplete markets. *Review of Financial Studies 18*(4), 1369–1402.

Christie, A. A. (1982, December). The stochastic behavior of common stock variances : Value, leverage and interest rate effects. *Journal of Financial Economics 10*(4), 407–432.

Christoffersen, P., K. Jacobs, and K. Mimouni (2010). Volatility dynamics for the s&p500: Evidence from realized volatility, daily returns, and option prices. *Review of Financial Studies 23*(8), 3141–3189.

Cochrane, J. H. (2008). The dog that did not bark: A defense of return predictability. *Review of Financial Studies 21*(4), 1533–1575.

Croce, M. M., M. Lettau, and S. C. Ludvigson (2015). Investor information, long-run risk, and the term structure of equity. *Review of Financial Studies 28*(3), 706–742.

Dechow, P., R. Sloan, and M. Soliman (2004). Implied equity duration: A new measure of equity risk. *Review of Accounting Studies 9*, 197–228.

Detemple, J. B. (1986, June). Asset pricing in a production economy with incomplete information. *Journal of Finance 41*(2), 383–91.

Duffie, D. (2001). *Dynamic Asset Pricing*. Princeton University Press.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica 50*(4), pp. 987–1007.

Fama, E. F. and K. R. French (1988). Permanent and temporary components of stock prices. *Journal of Political Economy 96*(2), 246–273.

Gabaix, X. (2012). Variable rare disasters: An exactly solved framework for ten puzzles in macro-finance. *Quarterly Journal of Economics 127*(2), 645–700.

Greenwald, D. L., M. Lettau, and S. C. Ludvigson (2014). Origins of stock market fluctuations. Unpublished manuscript, National Bureau of Economic Research.

Hansen, L. P. and J. A. Scheinkman (2009, 01). Long-Term Risk: An Operator Approach. *Econometrica 77*(1), 177–234.

Hasler, M. and R. Marfè (2016). Disaster recovery and the term structure of dividend strips. *Journal of Financial Economics 122*(1), 116 – 134.

Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies 6*(2), 327–343.

Johannes, M., A. Korteweg, and N. Polson (2014). Sequential learning, predictability, and optimal portfolio returns. *The Journal of Finance 69*(2), 611–644.

Jondeau, E. and M. Rockinger (2016). Forecasting financial returns with a structural macroeconomic model. Unpublished manuscript.

Kim, T. S. and E. Omberg (1996). Dynamic nonmyopic portfolio behavior. *Review of Financial Studies 9*(1), 141–61.

Liptser, R. S. and A. N. Shiryaev (2001). *Statistics of Random Processes II.* Springer Verlag, New York.

Marfè, R. (2015a). Corporate fraction and the equilibrium term structure of equity risk. *Review of Finance*.

Marfè, R. (2015b). Labor rigidity and the dynamics of the value premium. Unpublished manuscript.

Marfè, R. (2016). Income insurance and the equilibrium term structure of equity. *Forthcoming in the Journal of Finance*.

Markowitz, H. (1952). Portfolio selection. *The Journal of Finance 7*(1), 77–91.

Menzly, L., T. Santos, and P. Veronesi (2004, February). Understanding predictability. *Journal of Political Economy 112*(1), 1–47.

Moskowitz, T. J., Y. H. Ooi, and L. H. Pedersen (2012). Time series momentum. *Journal of Financial Economics 104*(2), 228–250.

Musiela, M. (1993). Stochastic pdes and term structure models. *Journees Internationales des Finance*.

Pástor, u. and R. F. Stambaugh (2012). Are stocks really less volatile in the long run? *The Journal of Finance 67*(2), 431–478.

Schulz, F. (2016, October). On the timing and pricing of dividends: Comment. *American Economic Review 106*(10), 3185–3223.

Segal, G., I. Shaliastovich, and A. Yaron (2015). Good and bad uncertainty: Macroeconomic and financial market implications. *Journal of Financial Economics 117*(2), 369 – 397.

van Binsbergen, J., M. Brandt, and R. Koijen (2012, June). On the timing and pricing of dividends. *American Economic Review 102*(4), 1596–1618.

van Binsbergen, J., W. Hueskes, R. Koijen, and E. Vrugt (2013). Equity yields. *Journal of Financial Economics 110*(3), 503 – 519.

van Binsbergen, J. H. and R. Koijen (2017). The term structure of returns: Facts and theory. *Journal of Financial Economics 124*(1), 121.

van Binsbergen, J. H. and R. S. J. Koijen (2010). Predictive regressions: A present-value approach. *The Journal of Finance 65*(4), 1439–1471.

van Binsbergen, J. H. and R. S. J. Koijen (2016, October). On the timing and pricing of dividends: Reply. *American Economic Review 106*(10), 3224–37.

Wachter, J. A. (2002). Portfolio and consumption decisions under mean-reverting returns: An exact solution for complete markets. *The Journal of Financial and Quantitative Analysis 37*(1), pp. 63–91.

Wachter, J. A. (2013). Can time-varying risk of rare disasters explain aggregate stock market volatility? *Journal of Finance 68*(3), 987–1035.

Weber, M. (2016). The term structure of equity returns: Risk or mispricing? Unpublished manuscript.

Welch, I. and A. Goyal (2008). A comprehensive look at the empirical performance of equity premium prediction. *Review of Financial Studies 21*(4), 1455–1508.

Xia, Y. (2001). Learning about predictability: The effects of parameter uncertainty on dynamic asset allocation. *The Journal of Finance 56*(1), pp. 205–246.

Zariphopoulou, T. (2001). A solution approach to valuation with unhedgeable risks. *Finance and Stochastics 5*(1), 61–82.

Appendix

A Proofs

A.1 Proof of Proposition 1

See Liptser and Shiryaev (2001).

A.2 Proof of Proposition 2

Let the dynamics of the price of the dividend strip with maturity i and the corresponding equity yield be given by

$$\frac{dP(t,i)}{P(t,i)} = \mu_P(t,i)dt + \sigma_P(t,i)d\widehat{B}_t,$$
$$de(t,i) = \mu_e(t,i)dt + \sigma_e(t,i)d\widehat{B}_t.$$

From the definition of the dividend strip in (16), we can interpret S_t^c as the value process for the self-financing portfolio that consists of a continuum of dividend strips:

$$S_t^c = n_t S_t = \int_t^\infty n_t P(t,i)di,$$

where

$$n_t \equiv \exp\left(\int_0^t \frac{D_u}{S_u}du\right)$$

is the number of dividend strips of each maturity held during the period $[t, t+dt]$. The corresponding relative weight $w_t(i)$ is given by

$$w_t(i) \equiv \frac{P(t,i)}{\int_t^\infty P(t,i)di}, \quad \text{with} \quad \int_t^\infty w_t(i)di = 1.$$

$w_t(i)$ tells us the relative proportion of the total portfolio value invested at time t in the dividend strip that pays $D_i di$ in the interval $[i, i+di]$. As in van Binsbergen and Koijen (2017), the stock return is the instantaneous return on the value-weighted portfolio of strips

$$\frac{dS_t^c}{S_t^c} = \int_t^\infty \frac{dP(t,i)}{P(t,i)} w_t(i)di.$$

Applying Ito's lemma yields

$$\begin{aligned}\frac{dS_t^c}{S_t^c} &= \int_t^\infty \left(d\log P(t,i) + \frac{1}{2}\sigma_P^2(t,i)dt\right) w_t(i)di \\ &= \int_t^\infty d\log P(t,i) w_t(i)di + \frac{1}{2}\int_t^\infty \sigma_P^2(t,i)w_t(i)didt. \end{aligned} \tag{31}$$

Moreover, we have

$$d\log S_t^c = \frac{dS_t^c}{S_t^c} - \frac{1}{2}\left(\frac{dS_t^c}{S_t^c}\right)^2 = \frac{dS_t^c}{S_t^c} - \frac{1}{2}v_t dt, \tag{32}$$

where

$$v_t dt = \left(\frac{dS_t^c}{S_t^c}\right)^2 = \int_t^\infty \int_t^\infty \sigma_P(t,i)\sigma_P(t,j)w_t(i)w_t(j)didjdt. \tag{33}$$

Putting (31), (32), and (33) together provides the relation between the instantaneous stock return and dividend strip returns

$$d\log S_t^c = \int_t^\infty d\log P(t,i)w_t(i)di + \frac{1}{2}\int_t^\infty \left(\sigma_P^2(t,i) - \int_t^\infty \sigma_P(t,i)\sigma_P(t,j)w_t(j)dj\right)w_t(i)didt. \tag{34}$$

The HPR is by definition given by

$$\log R(t,T) = \int_t^T d\log S_s^c.$$

Using (34) allows us to write

$$\begin{aligned}\log R(t,T) = &\int_t^T \int_s^\infty d\log P(s,i)w_t(i)di \\ &+ \frac{1}{2}\int_t^T \int_s^\infty \left(\sigma_P^2(s,i) - \int_s^\infty \sigma_P(s,i)\sigma_P(s,j)w_s(j)dj\right) w_s(i)dids.\end{aligned}$$

Moreover, from the definitions of the dividend strip price and the equity yield, we know that

$$\begin{aligned} d\log P(t,T) &= d\log D_t + e(t,T)dt - de(t,T)(T-t) \\ \sigma_P(t,T) &= \sigma_D - \sigma_e(t,T)(T-t). \end{aligned}$$

□

A.3 Proof of Proposition 3

The conditional variance of the holding period return is

$$\mathrm{Var}_t\left[\log R(t,T)\right] = \mathrm{Cov}_t\left(\int_t^T d\log S_s^c, \int_t^T d\log S_u^c\right) = \int_t^T\int_t^T \mathrm{Cov}_t\left(d\log S_s^c, d\log S_u^c\right). \tag{35}$$

For $s = u$, the integrand in (35) is the instantaneous variance of log returns. Using (34) and Fubini's theorem, we can write it as follows:

$$\begin{aligned}&\mathrm{Cov}_t\left(d\log S_s^c, d\log S_s^c\right) = \mathrm{Var}_t\left(d\log S_s^c\right) = E_t\left[\left(d\log S_s^c - E_t\left[d\log S_s^c\right]\right)^2\right] \\ &= E_t\left[\left(\int_s^\infty \sigma_P(s,i)w_s(i)di\right)^2\right] ds = E_t\left[\int_s^\infty\int_s^\infty \sigma_P(s,i)\sigma_P(s,j)w_s(i)w_s(j)didj\right] ds \\ &= \int_s^\infty\int_s^\infty E_t\left[\sigma_P(s,i)\sigma_P(s,j)w_s(i)w_s(j)\right] didjds.\end{aligned}$$

For $s \neq u$, we approximate (34) by abstracting from the second-order volatility adjustment term. We obtain

$$\begin{aligned}\mathrm{Cov}_t\left(d\log S^c_s, d\log S^c_u\right) &\approx \mathrm{Cov}_t\left(\int_s^\infty d\log P(s,i) w_s(i) di, \int_u^\infty d\log P(u,j) w_u(j) dj\right)\\ &= \int_s^\infty\int_u^\infty \mathrm{Cov}_t\left(d\log P(s,i) w_s(i), d\log P(u,j) w_u(j)\right) di dj.\end{aligned}$$

□

A.4 Proof of Proposition 4

To prove the statements in the first part of the proposition we start by computing the HPR variance at a τ-year horizon in the model with only the permanent component. For this, we use the fact that the moment generating function of the logarithm of the stock price is given by

$$\widehat{MGF}(t,\tau,u) = E_t\left[\left(S^c_{t+\tau}\right)^u\right] = e^{u\log(S^c_t) + \widehat{h}_0(\tau;u) + \widehat{h}_1(\tau;u)\widehat{x}_t}, \tag{36}$$

where $\widehat{h}_0$ and $\widehat{h}_1$ solve

$$\begin{aligned}\widehat{h}_1'(\tau) &= u - \kappa_x \widehat{h}_1(\tau),\\ \widehat{h}_0'(\tau) &= \frac{1}{2}\overline{v}u^2 + u\left(m - \frac{\overline{v}}{2}\right) + u\overline{\gamma}_x\widehat{h}_1(\tau) + \frac{1}{2}\frac{\overline{\gamma}_x^2}{\overline{v}}\widehat{h}_1^2(\tau)\end{aligned}$$

with initial conditions $\widehat{h}_0(0) = \widehat{h}_1(0) = 0$. Solving these two ODEs and using (36) together with the definition of $\mathrm{vol}_{S,t}(\tau)$ in (24) yields

$$\mathrm{vol}_{S,t}(\tau)^2 = \frac{(\overline{\gamma}_x + \overline{v}\kappa_x)^2}{\overline{v}\kappa_x^2} - \frac{1}{\tau}\frac{e^{-2\kappa_x\tau}(1-e^{\kappa_x\tau})((1-3e^{\kappa_x\tau})\overline{\gamma}_x - 4e^{\kappa_x\tau}\overline{v}\kappa_x)\overline{\gamma}_x}{2\overline{v}\kappa_x^3}. \tag{37}$$

Taking the limits of (37) and using the fact that the steady-state posterior variance $\overline{\gamma}_x$ is a (non-negative) solution to

$$\sigma_x^2 - 2\kappa_x\overline{\gamma}_x - \frac{1}{\overline{v}}\overline{\gamma}_x^2 = 0$$

we obtain

$$\begin{aligned}\lim_{\tau\to 0^+}\mathrm{vol}_{S,t}(\tau)^2 &= \overline{v},\\ \lim_{\tau\to\infty}\mathrm{vol}_{S,t}(\tau)^2 &= \overline{v} + \frac{\sigma_x^2}{\kappa_x^2}.\end{aligned}$$

The derivative of (37) with respect to the horizon is

$$\partial_\tau \mathrm{vol}_{S,t}(\tau)^2 = \frac{\overline{\gamma}_x^2 e^{-\kappa_x\tau}}{2\kappa_x^3\tau^2\overline{v}}\left(\left(3 + 4\kappa_x\frac{\overline{v}}{\overline{\gamma}_x}\right)e^{\kappa_x\tau} + (1+2\kappa_x\tau)e^{-\kappa_x\tau} - 4\left(1+\kappa_x\tau\right)\left(1+\kappa_x\frac{\overline{v}}{\overline{\gamma}_x}\right)\right). \tag{38}$$

With parameters κ_x, $\overline{\gamma}_x$, and $\overline{v}$ being nonnegative, in order to show that the term structure of HPR volatility is nondecreasing for $\tau > 0$ it is enough to show that the expression in the parentheses in

(38) is nonnegative, i.e.

$$\left(3 + 4\kappa_x \frac{\overline{v}}{\overline{\gamma}_x}\right) e^{\kappa_x \tau} + (1 + 2\kappa_x \tau) e^{-\kappa_x \tau} - 4\left(1 + \kappa_x \tau\right)\left(1 + \kappa_x \frac{\overline{v}}{\overline{\gamma}_x}\right) \geq 0.$$

The left hand side of the above inequality can be rearranged into a sum of two parts

$$e^{-\kappa_x \tau}\left[1 + 2\kappa_x \tau + 3e^{2\kappa_x \tau} - 4e^{\kappa_x \tau}\left(1 + \kappa_x \tau\right)\right] + \left[4e^{\kappa_x \tau}\kappa_x \frac{\overline{v}}{\overline{\gamma}_x} - 4\left(1 + \kappa_x \tau\right)\kappa_x \frac{\overline{v}}{\overline{\gamma}_x}\right]. \tag{39}$$

The second bracket is positive because $e^{\kappa_x \tau} > 1 + \kappa_x \tau$ and it is enough to prove that the first bracket of (39) is nonnegative:

$$1 + 2\kappa_x \tau + 3e^{2\kappa_x \tau} - 4e^{\kappa_x \tau}\left(1 + \kappa_x \tau\right) \geq 0.$$

Rearranging the left-hand side of the above inequality gives

$$\begin{aligned}
&1 + 2\kappa_x \tau + 3e^{2\kappa_x \tau} - 4e^{\kappa_x \tau}\left(1 + \kappa_x \tau\right) \\
&= \left(e^{\kappa_x \tau} - (1 + \kappa_x \tau)\right)^2 + 2e^{2\kappa_x \tau} - 2e^{\kappa_x \tau}(1 + \kappa_x \tau) - (\kappa_x \tau)^2 \\
&> \left(e^{\kappa_x \tau} - (1 + \kappa_x \tau)\right)^2 + 2e^{\kappa_x \tau}(1 + \kappa_x \tau + \frac{(\kappa_x \tau)^2}{2}) - 2e^{\kappa_x \tau}(1 + \kappa_x \tau) - (\kappa_x \tau)^2 \\
&= \left(e^{\kappa_x \tau} - (1 + \kappa_x \tau)\right)^2 + (\kappa_x \tau)^2 (e^{\kappa_x \tau} - 1) \\
&\geq 0.
\end{aligned}$$

This proves that the term structure of HPR volatility in the P-CV model is monotonically increasing.

To prove the statements in the second part of the proposition, we need the moment generating function of the logarithm of the stock price in the PT-CV model.

$$\widehat{MGF}(t, \tau, u) = E_t\left[\left(S^c_{t+\tau}\right)^u\right] = e^{u \log(S^c_t) + \widehat{g}_0(\tau;u) + \widehat{g}_1(\tau;u)\widehat{x}_t + \widehat{g}_2(\tau;u)\widehat{x}_t}, \tag{40}$$

where $\widehat{g}_0$, $\widehat{g}_1$ and $\widehat{g}_2$ solve

$$\begin{aligned}
\widehat{g}_1'(\tau) &= u - \kappa_x \widehat{g}_1(\tau), \\
\widehat{g}_2'(\tau) &= -\kappa_z\left(u + \widehat{g}_2(\tau)\right), \\
\widehat{g}_0'(\tau) &= \frac{1}{2}\overline{v}u^2 + u(m - \frac{\overline{v}}{2}) + u\left((\overline{\gamma}_{zx} - \overline{\gamma}_z \kappa_z + \sigma_z^2)\widehat{g}_2(\tau) + (\overline{\gamma}_x - \overline{\gamma}_z x \kappa_z)\widehat{g}_1(\tau)\right) \\
&\quad + \frac{1}{2\overline{v}}\left((\overline{\gamma}_{zx} - \overline{\gamma}_z \kappa_z + \sigma_z^2)\widehat{g}_2(\tau) + (\overline{\gamma}_x - \overline{\gamma}_z x \kappa_z)\widehat{g}_1(\tau)\right)^2,
\end{aligned}$$

with initial conditions $\widehat{g}_0(0) = \widehat{g}_1(0) = \widehat{g}_2(0) = 0$. Solving these two ODEs and using (40) together with the definition of $\text{vol}_{S,t}(\tau)$ gives us the expression for the HPR volatility in the PT-CV model. Using the obtained expression and the fact that the elements of the steady-state posterior variance-covariance matrix solve

$$\sigma_x^2 - 2\kappa_x \overline{\gamma}_x - \frac{1}{\overline{v}} \left(\overline{\gamma}_x - \kappa_z \overline{\gamma}_{zx}\right)^2 = 0,$$
$$\sigma_z^2 - 2\kappa_z \overline{\gamma}_z - \frac{1}{\overline{v}} \left(\sigma_z^2 - \kappa_z \overline{\gamma}_z + \overline{\gamma}_{zx}\right)^2 = 0,$$
$$-\overline{\gamma}_{zx}(\kappa_x + \kappa_z) - \frac{1}{\overline{v}} \left(\overline{\gamma}_x - \kappa_z \overline{\gamma}_{zx}\right)\left(\sigma_z^2 - \kappa_z \overline{\gamma}_z + \overline{\gamma}_{zx}\right) = 0,$$

we can derive the limiting expressions for $\mathrm{vol}_{S,t}(\tau)^2$:

$$\lim_{\tau \to 0^+} \mathrm{vol}_{S,t}(\tau)^2 = \overline{v},$$
$$\lim_{\tau \to \infty} \mathrm{vol}_{S,t}(\tau)^2 = \overline{v} + \frac{\sigma_x^2}{\kappa_x^2} - \sigma_z^2.$$

We see that the long-term variance (volatility) of HPRs is smaller than the short-term variance (volatility) if

$$\frac{\sigma_x^2}{\kappa_x^2} - \sigma_z^2 < 0.$$

□

A.5 Proof of Proposition 5

To compute the price of the dividend strip we recall that

$$P(t,T) = E_t\left[\frac{\xi_T}{\xi_t} D_T\right] = \frac{1}{\xi_t} E_t\left[e^{\log \xi_T + \log D_T}\right].$$

For the conditional expectation

$$H(t, \xi_t, D_t, \theta_t; T) \equiv E_t\left[e^{\log \xi_T + \log D_T}\right] \tag{41}$$

we guess an (approximate) exponential affine solution of the form

$$H(t, \xi_t, D_t, \theta_t; T) = e^{\log \xi_t + \log D_t + a(T-t) + \theta_t^\top b(T-t)}, \tag{42}$$

where $\theta_t = [\widehat{x}_t, \widehat{z}_t, v_t, \gamma_{x,t}, \gamma_{z,t}, \gamma_{zx,t}]^\top$ is the vector of state variables. Given the dynamics of the state-price density and the dividend process in (12) and (15), as well as the dynamics of the state variables provided in (6)–(10) and (29), it follows from the Feynman-Kac Theorem that H, as given in (41), needs to satisfy the following PDE

$$
\begin{aligned}
&H_t - r_f \xi H_\xi + \mu_D D H_D - \kappa_x x H_x - \kappa_z z H_z + \kappa_v (\bar{v} - v) H_v \\
&\quad + \left(\sigma_x^2 - 2\kappa_x \gamma_x - \frac{1}{v} \left(\gamma_x - \gamma_{zx} \kappa_z \right)^2 \right) H_{\gamma_x} + \left(\sigma_z^2 - 2\kappa_z \gamma_z - \frac{1}{v} \left(\gamma_{zx} - \gamma_z \kappa_z + \sigma_z^2 \right)^2 \right) H_{\gamma_z} \\
&\quad + \left(- \left(\kappa_x + \kappa_z \right) \gamma_{zx} - \frac{1}{v} \left(\gamma_x - \gamma_{zx} \kappa_z \right) \left(\gamma_{zx} - \gamma_z \kappa_z + \sigma_z^2 \right) \right) H_{\gamma_{zx}} \\
&\quad + \frac{1}{2v} \left(m + x - \kappa_z z - r_f \right)^2 \xi^2 H_{\xi\xi} + \frac{1}{2} \sigma_D^2 D^2 H_{DD} + \frac{1}{2v} (\gamma_x - \kappa_z \gamma_{zx})^2 H_{xx} \\
&\quad + \frac{1}{2v} (\gamma_{zx} - \kappa_z \gamma_z + \sigma_z^2)^2 H_{zz} + \frac{1}{2} \sigma_v^2 v H_{vv} - \frac{1}{\sqrt{v}} (m + x - \kappa_z z - r_f) \sigma_D H_{\xi D} + \sigma_D \sigma_v \sqrt{v} H_{Dv} \\
&\quad - \frac{1}{v} (m + x - \kappa_z z - r_f)(\gamma_x - \kappa_z \gamma_{zx}) H_{\xi x} - \frac{1}{v} (m + x - \kappa_z z - r_f)(\gamma_{zx} - \kappa_z \gamma_z + \sigma_z^2) H_{\xi z} \\
&\quad - (m + x - \kappa_z z - r_f) \sigma_v H_{\xi v} + \frac{1}{\sqrt{v}} \sigma_D (\gamma_x - \kappa_z \gamma_{zx}) H_{Dx} + \frac{1}{\sqrt{v}} \sigma_D (\gamma_{zx} - \kappa_z \gamma_z + \sigma_z^2) H_{Dz} \\
&\quad + \frac{1}{v} \left(\gamma_{zx} - \kappa_z \gamma_z + \sigma_z^2 \right) \left(\gamma_x - \kappa_z \gamma_{zx} \right) H_{zx} + (\gamma_x - \kappa_z \gamma_{zx}) \sigma_v H_{xv} + (\gamma_{zx} - \kappa_z \gamma_z + \sigma_z^2) \sigma_v H_{zv} = 0.
\end{aligned}
$$

Plugging the guess (42) (and the resulting partial derivatives from the guess solution) into the PDE and performing a first-order Taylor expansion around the long-term means of the state variables $\theta_t = [\widehat{x}_t, \widehat{z}_t, v_t, \gamma_{x,t}, \gamma_{z,t}, \gamma_{zx,t}]^\top$ allows us to obtain a system of ODEs that determines the function $a(\cdot)$ and the vector valued function $b(\cdot)$, with the first-order conditions $a(0) = b_i(0) = 0$ for $i = 1, ..., 6$. Finally, equity yields are given by

$$
e(t, T) = -\frac{1}{T - t} \left(a(T - t) + \theta_t^\top b(T - t) \right). \tag{43}
$$

The diffusion of the equity yield is recovered from the expression for $e(t, T)$ in (43) by applying Ito's lemma:

$$
\sigma_e(t, T) = -\frac{1}{T - t} \left[\frac{1}{\sqrt{v_t}} \left(\gamma_{x,t} - \kappa_z \gamma_{zx,t} \right), \frac{1}{\sqrt{v_t}} \left(\gamma_{zx,t} - \kappa_z \gamma_{z,t} + \sigma_z^2 \right), \sigma_v \sqrt{v_t}, 0, 0, 0 \right] b(T - t).
$$

In the P-CV model ($S_t^c = Y_t$ and constant instantaneous volatility, $v_t = \overline{v}$), the volatility of equity yields is given by

$$
\mathrm{vol}_{e,t}(\tau) = \frac{1}{\tau} \left| \frac{\overline{\gamma}_x}{\sqrt{\overline{v}}} b(\tau) \right|,
$$

where $\overline{\gamma}_x$ solves $\dfrac{d\gamma_{x,t}}{dt} = 0$. The function $b_1(\tau)$ is given by

$$
b(\tau) = -\frac{c_2}{c_1} \left(\exp(-\tau c_1 / c_3) - 1 \right),
$$

with

$$\begin{aligned}
c_1 &= (a_x-1)^2\sqrt{\overline{v}}\frac{\overline{\gamma}_x^2}{\overline{v}} + \left(\frac{\overline{\gamma}_x}{\sqrt{\overline{v}}} + \sqrt{\overline{v}}\kappa_x\right)(m-m_s)^2,\\
c_2 &= (a_x-1)\left((m-m_s)^2\sqrt{\overline{v}} - (a_x-1)(r_f+\overline{v})\eta_x + m(a_x\eta_x+\sqrt{\overline{v}}\kappa_x) - m_s(\eta_x+\sqrt{\overline{v}}\kappa_x)\right),\\
c_3 &= (m-m_s)^2\sqrt{\overline{v}}.
\end{aligned}$$

For $a_x \to 1$, we have $b(\tau) = 0$ and $\text{vol}_{e,t}(\tau) = 0$.

In the PT-CV model ($S_t^c = Y_t Z_t$ and constant instantaneous volatility, $v_t = \overline{v}$), the volatility of equity yields is given by

$$\text{vol}_{e,t}(\tau) = \frac{1}{\tau}\left|\eta_x b_1(\tau) + \eta_z b_2(\tau)\right|,$$

where

$$\eta_x \equiv \frac{\overline{\gamma}_x - \kappa_z\overline{\gamma}_{zx}}{\sqrt{\overline{v}}}, \quad \eta_z \equiv \frac{\overline{\gamma}_{zx} - \kappa_z\overline{\gamma}_z + \sigma_z^2}{\sqrt{\overline{v}}} \tag{44}$$

and the steady-state posterior variance-covariance matrix

$$\overline{\Gamma} \equiv \begin{pmatrix} \overline{\gamma}_z & \overline{\gamma}_{zx} \\ \overline{\gamma}_{zx} & \overline{\gamma}_x \end{pmatrix}$$

is obtained by solving the following system of equations $\frac{d\gamma_{z,t}}{dt} = 0$, $\frac{d\gamma_{x,t}}{dt} = 0$, and $\frac{d\gamma_{zx,t}}{dt} = 0$. Functions $b_1(\tau)$ and $b_2(\tau)$ solve the system of ODEs

$$\begin{aligned}
A_1 b_1'(\tau) + B_1 b_2(\tau) + C_1 b_1(\tau) + D_1 = 0,\\
A_2 b_2'(\tau) + B_2 b_1(\tau) + C_2 b_2(\tau) + D_2 = 0,
\end{aligned}$$

where

$$\begin{aligned}
A_1 &= -(m-m_s)^2\sqrt{\overline{v}}, \quad A_2 = -A_1,\\
B_1 &= \eta_z\left(-(m-m_s)^2 + (a_x-1)\sqrt{\overline{v}}(-(a_x-1)\eta_x + (a_z-1)\eta_z\kappa_z)\right),\\
B_2 &= -\eta_x\kappa_z\left((m-m_s)^2 + (a_z-1)\sqrt{\overline{v}}((a_x-1)\eta_x - (a_z-1)\eta_z\kappa_z)\right),\\
C_1 &= -\left((m-m_s)^2(\eta_x+\sqrt{\overline{v}}\kappa_x) + (a_x-1)\sqrt{\overline{v}}\eta_x((a_x-1)\eta_x - (a_z-1)\eta_z\kappa_z)\right),\\
C_2 &= \kappa_z\left((m-m_s)^2(\sqrt{\overline{v}}-\eta_z) + (a_z-1)\sqrt{\overline{v}}\eta_z(-(a_x-1)\eta_x + (a_z-1)\eta_z\kappa_z)\right),\\
D_1 &= (a_x-1)\left((m-m_s)^2\sqrt{\overline{v}} - (a_x-1)(r_f+\overline{v})\eta_x - m_s(\eta_x+\sqrt{\overline{v}}\kappa_x) + m(a_x\eta_x+\sqrt{\overline{v}}\kappa_x)\right)\\
&\quad + (a_z-1)(m_s - r_f - \overline{v} + a_x(-m+r_f+\overline{v}))\eta_z\kappa_z,\\
D_2 &= \kappa_z\Big((a_z-1)(m-m_s)^2\sqrt{\overline{v}} - (a_x-1)m_s\eta_x - (a_z-1)m_s(\sqrt{\overline{v}}-\eta_z)\kappa_z\\
&\quad + (a_z-1)(r_f+\overline{v})(-(a_x-1)\eta_x + (a_z-1)\eta_z\kappa_z) + m((a_x-1)a_z\eta_x + (a_z-1)(\sqrt{\overline{v}}-a_z\eta_z)\kappa_z)\Big).
\end{aligned}$$

For $a_x \to 1$, we have

$$\begin{aligned}\lim_{\tau\to 0^+} \mathrm{vol}_{e,t}(\tau) &= \frac{1}{A_2}\left|D_1\eta_x - D_2\eta_z\right| \\ &= \frac{1}{(m-m_s)^2\sqrt{\bar{v}}}\Big|(a_z-1)\eta_z\kappa_z\Big((m-m_s)\big((m-m_s)\sqrt{\bar{v}}+\eta_x\big) \\ &+ \big((m-m_s)\sqrt{\bar{v}} + (m_s - r_f - \bar{v} + a_z(-m+r_f+\bar{v}))\eta_z\big)\kappa_z\Big)\Big|.\end{aligned}$$

The short-term volatility of the equity yield is strictly greater than zero as long as the term within the absolute value is non-zero. Moreover, the long-term equity yield volatility satisfies[24]

$$\lim_{\tau\to\infty} \mathrm{vol}_{e,t}(\tau) = 0.$$

We see that

$$\lim_{\tau\to\infty} \mathrm{vol}_{e,t}(\tau) \le \lim_{\tau\to 0^+} \mathrm{vol}_{e,t}(\tau),$$

with strict inequality if

$$\begin{aligned}&(a_z-1)\eta_z\kappa_z\Big((m-m_s)\big((m-m_s)\sqrt{\bar{v}}+\eta_x\big) \\ &+ \big((m-m_s)\sqrt{\bar{v}} + (m_s - r_f - \bar{v} + a_z(-m+r_f+\bar{v}))\eta_z\big)\kappa_z\Big) \neq 0.\end{aligned}$$

□

A.6 Proof of Proposition 6

The optimal value function is of the form $J(t, W, x, z)$, and the HJB equation is as follows:

$$\begin{aligned}&J_t - \kappa_z z J_z + \frac{1}{2}\eta_z^2 J_{zz} - \kappa_x x J_x + \frac{1}{2}\eta_x^2 J_{xx} + \eta_z\eta_x J_{zx} + \sup_{\pi}\Big\{r_f W J_W \\ &+ \pi W\left(m + x - \kappa_z z - r_f\right) J_W + \frac{1}{2}\pi^2 W^2 \bar{v} J_{WW} + \pi W\sqrt{\bar{v}}\eta_x J_{Wx} + \pi W\sqrt{\bar{v}}\eta_z J_{Wz}\Big\} = 0, \quad (45)\end{aligned}$$

where η_x and η_z are as defined in (44).

Solving the first-order condition yields the following optimal portfolio weight

$$\pi^\star = -\frac{J_W}{W J_{WW}}\frac{m + x - \kappa_z z - r_f}{\bar{v}} - \frac{J_{Wx}}{W J_{WW}}\frac{\eta_x}{\sqrt{\bar{v}}} - \frac{J_{Wz}}{W J_{WW}}\frac{\eta_z}{\sqrt{\bar{v}}}. \quad (46)$$

Inspired by results in Kim and Omberg (1996) and Zariphopoulou (2001), we make the conjec-

[24]Note that the following parameter restrictions need to hold for the long-term moments of equity yields to be determinate

$$\begin{aligned}A_2 &> 0, \\ C_2 - C_1 &> 0, \\ (C_2 - C_1) - \sqrt{(C_1+C_2)^2 - 4B_1B_2} &> 0.\end{aligned}$$

ture

$$J(t,W,x,z) = e^{-\delta t}\frac{W^{1-\gamma}}{1-\gamma}F(t,x,z)^{\gamma}. \tag{47}$$

Plugging this conjecture back into the HJB equation (45) implies that the function $F(t,x,z)$ solves the linear second-order PDE

$$\begin{aligned} &F_t + \left[-\kappa_x x + \frac{(1-\gamma)(m+x-\kappa_z z - r_f)\eta_x}{\gamma\sqrt{\bar{v}}}\right]F_x + \left[-\kappa_z z + \frac{(1-\gamma)(m+x-\kappa_z z-r_f)\eta_z}{\gamma\sqrt{\bar{v}}}\right]F_z \\ &+\frac{1}{2}\eta_x^2 F_{xx} + \frac{1}{2}\eta_z^2 F_{zz} + \eta_x\eta_z F_{xz} + \frac{1-\gamma}{\gamma}\left[r_f - \frac{\delta}{1-\gamma} + \frac{(m+x-\kappa_z z - r_f)^2}{2\gamma\bar{v}}\right]F = 0. \end{aligned} \tag{48}$$

The solution to PDE (48) is given by

$$F(\tau,\theta) = \exp\left(\frac{1}{2}\theta^{\top}A(\tau)\theta + \theta^{\top}B(\tau) + C(\tau)\right), \tag{49}$$

where $\theta = [x,z]^{\top}$, $A(\tau)$ is a symmetric 2×2 matrix-valued function, $B(\tau)$ is a 2×1 matrix-valued function and $C(\tau)\in\mathbb{R}$ with A, B, and C solving the system of matrix Riccati differential equations

$$\begin{aligned} \dot{A}(\tau) &= \frac{1}{2}A(\tau)q_1^{\top}q_1A(\tau) + (q_2+q_3q_1)A(\tau) + A(\tau)(q_2 + (q_3q_1)^{\top}) + q_4 \\ \dot{B}(\tau) &= \frac{1}{2}A(\tau)q_1^{\top}q_1B(\tau) + (q_2+q_3q_1)B(\tau) + A(\tau)q_1^{\top}q_5 + q_6 \\ \dot{C}(\tau) &= \frac{1}{4}B^{\top}(\tau)q_1^{\top}q_1B(\tau) + \frac{1}{2}q_7^{\top}A(\tau)q_7 + B(\tau)^{\top}q_1^{\top}q_5 + q_8 \end{aligned}$$

with coefficients

$$\begin{aligned} q_1 &= \begin{pmatrix} \eta_x & \eta_z \\ \eta_x & \eta_z \end{pmatrix}, \quad q_2 = \begin{pmatrix} -\kappa_x & 0 \\ 0 & -\kappa_z \end{pmatrix}, \\ q_3 &= \frac{1-\gamma}{\gamma}\frac{1}{\sqrt{\bar{v}}}\begin{pmatrix} 1 & 0 \\ 0 & -\kappa_z \end{pmatrix}, \quad q_4 = \frac{1-\gamma}{\gamma}\frac{1}{\gamma\bar{v}}\begin{pmatrix} 1 & -\kappa_z \\ -\kappa_z & \kappa_z^2 \end{pmatrix}, \\ q_5 &= \frac{1-\gamma}{\gamma}\frac{(m-r_f)}{\sqrt{\bar{v}}}\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad q_6 = \frac{1-\gamma}{\gamma}\frac{(m-r_f)}{\gamma\bar{v}}\begin{pmatrix} 1 \\ -\kappa_z \end{pmatrix}, \\ q_7 &= \begin{pmatrix} \eta_x \\ \eta_z \end{pmatrix}, \quad q_8 = \frac{1-\gamma}{\gamma}\left[r_f - \frac{\delta}{1-\gamma} + \frac{(m-r_f)^2}{2\gamma\bar{v}}\right] \end{aligned}$$

The boundary conditions are $A_{ij}(0) = B_i(0) = C(0) = 0$, where A_{ij} and B_i denote the components of A and B. Substituting Equation (49) in Equation (47) and plugging the result in Equation (46) yields the optimal portfolio weight.

□

A.7 Proof of Proposition 7

For the problem with stochastic return volatility v_t, the optimal value function is of the form $J(t, W, x, z, v, \gamma_x, \gamma_z, \gamma_{zx})$, and the HJB equation is as follows:

$$
\begin{aligned}
&J_t - \kappa_x x J_x - \kappa_z z J_z + \kappa_v(\bar{v} - v) J_v \\
&+ \left(\sigma_x^2 - 2\kappa_x\gamma_x - \frac{1}{v}\left(\gamma_x - \gamma_{zx}\kappa_z\right)^2\right) J_{\gamma_x} + \left(\sigma_z^2 - 2\kappa_z\gamma_z - \frac{1}{v}\left(\gamma_{zx} - \gamma_z\kappa_z + \sigma_z^2\right)^2\right) J_{\gamma_z} \\
&+ \left(-\left(\kappa_x + \kappa_z\right)\gamma_{zx} - \frac{1}{v}\left(\gamma_x - \gamma_{zx}\kappa_z\right)\left(\gamma_{zx} - \gamma_z\kappa_z + \sigma_z^2\right)\right) J_{\gamma_{zx}} \\
&+ \frac{1}{2v}(\gamma_{zx} - \kappa_z\gamma_z + \sigma_z^2)^2 J_{zz} + \frac{1}{2v}(\gamma_x - \kappa_z\gamma_{zx})^2 J_{xx} + \frac{1}{2}\sigma_v^2 v J_{vv} \\
&+ \frac{1}{v}\left(\gamma_{zx} - \kappa_z\gamma_z + \sigma_z^2\right)\left(\gamma_x - \kappa_z\gamma_{zx}\right) J_{zx} + (\gamma_x - \kappa_z\gamma_{zx})\sigma_v J_{xv} + (\gamma_{zx} - \kappa_z\gamma_z + \sigma_z^2)\sigma_v J_{zv} \\
&+ r_f W J_W + \sup_{\pi}\Big\{\pi W\left(m + x - \kappa_z z - r_f\right) J_W + \frac{1}{2}\pi^2 W^2 v J_{WW} + \pi W(\gamma_x - \kappa_z\gamma_{zx}) J_{Wx} \\
&+ \pi W(\gamma_{zx} - \kappa_z\gamma_z + \sigma_z^2) J_{Wz} + \pi W v\sigma_v J_{Wv}\Big\} = 0.
\end{aligned}
\tag{50}
$$

Solving the first-order condition yields the following optimal portfolio weight

$$
\begin{aligned}
\pi^\star = &- \frac{J_W}{W J_{WW}} \frac{(m + x - \kappa_z z - r_f)}{v} - \frac{J_{Wx}}{W J_{WW}} \frac{\gamma_x - \kappa_z\gamma_{zx}}{v} \\
&- \frac{J_{Wz}}{W J_{WW}} \frac{\gamma_{zx} - \kappa_z\gamma_z + \sigma_z^2}{v} - \frac{J_{Wv}}{W J_{WW}}\sigma_v.
\end{aligned}
\tag{51}
$$

After plugging Equation (51) into the HJB equation (50), and guessing that

$$
J(t, W, x, z, v_t, \gamma_x, \gamma_z, \gamma_{zx}) = e^{-\delta t}\frac{W^{1-\gamma}}{1-\gamma} F(T - t, x, z, v, \gamma_x, \gamma_z, \gamma_{zx})^\gamma,
$$

we obtain a nonlinear PDE for the F function.

We look for an approximate analytic solution to this PDE of the form

$$
F(\tau, \theta) = \exp\left(\frac{1}{2}\theta^\top A(\tau)\theta + \theta^\top B(\tau) + C(\tau)\right),
\tag{52}
$$

where $\theta = [x, z, v, \gamma_x, \gamma_z, \gamma_{zx}]^\top$, $A(\tau)$ is a symmetric 6×6 matrix-valued function, $B(\tau)$ is a 6×1 matrix-valued function, and $C(\tau) \in \mathbb{R}$. Plugging (52) in the PDE and performing a second order Taylor expansion around the long-run means of the state variables allows us to use the variable separation method to obtain a coupled system of ODEs for coefficient functions A, B, and C.

□

B Maximum Likelihood Estimation

Based on the dynamics provided in (5)-(7), (9)-(11), and (29), the monthly log-return r_t, the transitory component $\widehat{z}_t$, the stochastic drift $\widehat{x}_t$, the return variance v_t, and the posterior variance-covariances γ_z, γ_x, and γ_{zx} are discretized as follows:

$$
\begin{aligned}
r_{t+\Delta} &= (m - v_t/2 + \widehat{x}_t - \kappa_z \widehat{z}_t)\,\Delta + \sqrt{v_t \Delta}\epsilon_{t+\Delta}, \\
\widehat{z}_{t+\Delta} &= e^{-\kappa_z \Delta}\widehat{z}_t + \left(\gamma_{zx} - \gamma_z \kappa_z + \sigma_z^2\right) \sqrt{\frac{1 - e^{-2\kappa_z \Delta}}{2\kappa_z v_t}}\epsilon_{t+\Delta}, \\
\widehat{x}_{t+\Delta} &= e^{-\kappa_x \Delta}\widehat{x}_t + (\gamma_x - \gamma_{zx}\kappa_z) \sqrt{\frac{1 - e^{-2\kappa_x \Delta}}{2\kappa_x v_t}}\epsilon_{t+\Delta}, \\
v_{t+\Delta} &= e^{-\kappa_v \Delta} v_t + \overline{v}\left(1 - e^{-\kappa_v \Delta}\right) + \operatorname{sign}(\sigma_v)\sqrt{l_t}\epsilon_{t+\Delta}, \\
\gamma_{z,t+\Delta} &= \gamma_{z,t} + \left[-\frac{\left(\gamma_{zx,t} - \gamma_{z,t}\kappa_z + \sigma_z^2\right)^2}{v_t} - 2\gamma_{z,t}\kappa_z + \sigma_z^2\right]\Delta, \\
\gamma_{x,t+\Delta} &= \gamma_{x,t} + \left[-\frac{\left(\gamma_{x,t} - \gamma_{zx,t}\kappa_z\right)^2}{v_t} - 2\gamma_{x,t}\kappa_x + \sigma_x^2\right]\Delta, \\
\gamma_{zx,t+\Delta} &= \gamma_{zx,t} + \left[-\frac{\left(\gamma_{x,t} - \gamma_{zx,t}\kappa_z\right)\left(\gamma_{zx,t} - \gamma_{z,t}\kappa_z + \sigma_z^2\right)}{v_t} - (\kappa_x + \kappa_z)\,\gamma_{zx,t}\right]\Delta,
\end{aligned}
$$

where $\Delta = 1/12 = 1$ month, $\epsilon_t \sim \mathrm{N}(0,1)$ follows a normal distribution with mean 0 and variance 1, and $l_t \equiv \frac{\sigma_v^2}{\kappa_v} v_t (e^{-\kappa_v \Delta} - e^{-2\kappa_v \Delta}) + \frac{\sigma_v^2}{2\kappa_v}\overline{v}(1 - e^{-\kappa_v \Delta})^2$. Conditional on knowing the parameters and the initial values $\widehat{z}_0$, $\widehat{x}_0$, v_0, $\gamma_{z,0}$, $\gamma_{x,0}$, and $\gamma_{zx,0}$, observing monthly returns r_t allows us to sequentially back out ϵ_t and therefore $\widehat{z}_t$, $\widehat{x}_t$, v_t, $\gamma_{z,t}$, $\gamma_{x,t}$, and $\gamma_{zx,t}$.

The 8-dimensional vector of parameters is $\Theta \equiv (m, \sigma_z, \kappa_z, \sigma_x, \kappa_x, \overline{v}, \sigma_v, \kappa_v)$ and the log-likelihood function $\mathcal{L}(\Theta; r_\Delta, r_{2\Delta}, \ldots, r_{N\Delta})$ satisfies

$$
\begin{aligned}
\mathcal{L}\left(\Theta; r_\Delta, r_{2\Delta}, \ldots, r_{N\Delta}\right) = \sum_{i=1}^{N} & \log\left(\frac{1}{\sqrt{2\pi v_t \Delta}}\right) \\
& - \frac{1}{2}\frac{\left(r_{i\Delta} - \left(m - \frac{1}{2}v_t + \widehat{x}_{(i-1)\Delta} - \kappa_z \widehat{z}_{(i-1)\Delta}\right)\Delta\right)^2}{v_t \Delta},
\end{aligned} \tag{53}
$$

where N is the number of observations. The vector of parameters Θ is chosen to maximize the log-likelihood function provided in Equation (53).

C Predictive Regressions

Table 12 reports the estimates obtained by regressing cumulative dividend growth rates on either the labour share or the financial leverage from the nonfinancial corporate sector in the post-war sample (1946-2014). Actual dividend growth rates are predicted by both the labour share and the financial leverage with positive and significant slopes increasing with the forecasting horizon (Belo et al., 2015; Marfè, 2016). Consistent with actual data, model-implied dividend growth rates are positively and significantly predicted by the financial leverage, with slopes increasing with the forecasting horizon. The explanatory power, however, is larger than in actual data. The slope of the relation between the model-implied dividend growth rate and the labour share is positive and increases with the forecasting horizon, in line with actual data. This positive relation, however, is insignificant when using model-implied data, whereas it is strongly significant when using actual data. Although *(i)* the labour share and the financial leverage are variables that are external to our model and *(ii)* our estimation methodology is not set to capture any of these dividend growth predictability patterns, model-implied dividend growth rates preserve to some extent the predictability observed in actual data.

Table 12: Dividend Growth Predictability by External Variables.
This table reports the estimates obtained by regressing the cumulative log dividend growth rate, g, over several horizons on either the labour share, LS, or the financial leverage, FL:

$$\sum_{i=1}^{n} g_{t+i} = a + b\,x_t + \epsilon_t,$$

where $n = \{1,2,3,5,7,10\}$ years and $x = \{LS, FL\}$. The actual log dividend growth rate is that of the S&P 500, while the model-implied counterpart satisfies $\tilde{y}_t = \log(m - m_s + (1-a_x)\widehat{x}_t + (a_z - 1)\kappa_z \widehat{z}_t)$ and $D_t = P_t \exp(\tilde{y}_t)$ (see Section 3.3). The labour share (labour compensations over value added) and the financial leverage (debt over equity) are from the nonfinancial corporate sector (Flows of Funds). Newey-West t-statistics are reported in parentheses and 10%, 5%, and 1% significance levels are denoted with *, **, and ***, respectively. Monthly data are aggregated at yearly frequency over the 1946-2014 sample. P, PT, CV, and SV, stand for permanent component with stochastic drift, permanent component with stochastic drift and transitory component, constant volatility, and stochastic volatility, respectively.

Predictability by labor share: $\sum_{i=1}^{n} g_{t+i}$	1	2	3	5	7	10
Data						
LS	-0.428	-0.264	1.224	4.598***	4.885***	5.004*
t-stat	(-1.03)	(-0.31)	(0.93)	(4.50)	(3.05)	(1.81)
adj-R^2	0.00	-0.01	0.01	0.20	0.20	0.12
P-CV Model ($k_z = 0$)						
LS	-0.402	0.142	0.836	3.597	4.869	4.581
t-stat	(-0.41)	(0.08)	(0.35)	(1.17)	(1.10)	(0.55)
adj-R^2	-0.01	-0.02	-0.01	0.01	0.01	-0.01
P-SV Model ($k_z = 0$)						
LS	-0.374	0.112	0.791	3.356	4.722	4.230
t-stat	(-0.41)	(0.06)	(0.34)	(1.07)	(1.09)	(0.52)
adj-R^2	-0.01	-0.02	-0.01	0.00	0.01	-0.01
PT-CV Model						
LS	0.448	0.939	1.196	0.821	-0.526	-2.634
t-stat	(1.28)	(1.29)	(0.99)	(0.34)	(-0.19)	(-0.87)
adj-R^2	0.02	0.03	0.02	-0.01	-0.02	0.00
PT-SV Model						
LS	-0.004	0.238	0.273	0.425	0.623	3.007
t-stat	(-0.01)	(0.48)	(0.37)	(0.34)	(0.36)	(0.74)
adj-R^2	-0.02	-0.01	-0.01	-0.01	-0.01	0.00

Predictability by financial leverage: $\sum_{i=1}^{n} g_{t+i}$	1	2	3	5	7	10
Data						
FL	0.202	0.607***	0.912***	0.985***	0.663**	0.693**
t-stat	(1.60)	(3.58)	(4.33)	(4.35)	(2.28)	(2.14)
adj-R^2	0.02	0.10	0.16	0.17	0.07	0.07
P-CV Model ($k_z = 0$)						
FL	1.208***	2.115***	2.511***	3.886***	4.877***	6.953***
t-stat	(4.60)	(4.47)	(3.94)	(5.71)	(7.01)	(8.06)
adj-R^2	0.19	0.31	0.34	0.46	0.53	0.71
P-SV Model ($k_z = 0$)						
FL	1.097***	2.013***	2.429***	3.732***	4.705***	6.683***
t-stat	(4.55)	(4.52)	(4.02)	(5.61)	(6.94)	(8.11)
adj-R^2	0.18	0.31	0.34	0.46	0.53	0.71
PT-CV Model						
FL	-0.076	-0.152	-0.244	-0.135	0.092	0.800
t-stat	(-0.52)	(-0.58)	(-0.72)	(-0.30)	(0.18)	(1.30)
adj-R^2	-0.01	-0.00	0.01	-0.01	-0.02	0.04
PT-SV Model						
FL	0.525***	0.873***	0.952***	1.487***	2.022***	2.927***
t-stat	(4.80)	(4.47)	(4.38)	(5.37)	(4.72)	(5.97)
adj-R^2	0.17	0.28	0.30	0.40	0.47	0.65

D Actual and Model-Implied Dividend Yield

Table 13 reports the correlation between the actual S&P 500 log dividend yield and the model-implied log dividend yield, between the actual log dividend yield and the state variables $\widehat{x}$ and $\widehat{z}$, and between the model-implied log dividend yield and the state variables $\widehat{x}$ and $\widehat{z}$. Correlations are computed over the full sample (1872-2015) as well as over the 1872-1945 and 1946-2015 sub-samples.

Table 13: Actual and Model-Implied Dividend Yields

This table reports the correlations and p-values between the actual S&P 500 log dividend yield and the model-implied log dividend yield, between the actual log dividend yield and the state variables $\widehat{x}$ and $\widehat{z}$, and between the model-implied log dividend yield and the state variables $\widehat{x}$ and $\widehat{z}$. The model-implied log dividend yield satisfies $\tilde{y}_t = \log(m - m_s + (1 - a_x)\widehat{x}_t + (a_z - 1)\kappa_z \widehat{z}_t)$ (see Section 3.3). Correlations are computed over the full sample (1872-2015) as well as over the 1872-1945 and 1946-2015 sub-samples. P, PT, CV, and SV, stand for permanent component with stochastic drift, permanent component with stochastic drift and transitory component, constant volatility, and stochastic volatility, respectively.

	1872–2015				1872–1945				1946–2015			
	P-CV	P-SV	PT-CV	PT-SV	P-CV	P-SV	PT-CV	PT-SV	P-CV	P-SV	PT-CV	PT-SV
corr($\tilde{y}^{\text{actual}}, \tilde{y}^{\text{model}}$)	0.09	0.13	0.62	0.83	0.23	0.32	0.43	0.29	0.01	0.02	0.42	0.95
p-value	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.80)	(0.65)	(0.00)	(0.00)
corr($\tilde{y}^{\text{actual}}, \widehat{x}$)	-0.09	-0.13	-0.12	-0.23	-0.23	-0.32	-0.24	-0.33	-0.01	-0.02	-0.02	-0.17
p-value	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.80)	(0.63)	(0.49)	(0.00)
corr($\tilde{y}^{\text{actual}}, \widehat{z}$)	0.00	0.00	-0.61	-0.81	0.00	0.00	-0.45	-0.31	0.00	0.00	-0.38	-0.96
p-value	(1.00)	(1.00)	(0.00)	(0.00)	(1.00)	(1.00)	(0.00)	(0.00)	(1.00)	(1.00)	(0.00)	(0.00)
corr($\tilde{y}^{\text{model}}, \widehat{x}$)	-1.00	-1.00	-0.15	-0.20	-1.00	-1.00	-0.15	-0.18	-1.00	-1.00	-0.15	-0.19
p-value	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
corr($\tilde{y}^{\text{model}}, \widehat{z}$)	0.00	0.00	-0.59	-0.98	0.00	0.00	-0.98	-1.00	0.00	0.00	-0.99	-0.99
p-value	(1.00)	(1.00)	(0.00)	(0.00)	(1.00)	(1.00)	(0.00)	(0.00)	(1.00)	(1.00)	(0.00)	(0.00)

E In-sample Portfolio Performance

In this section, we compare the in-sample performance of different investment strategies. Specifically, we use the parameters estimated over the entire sample (see Table 2) to implement strategies that invest in the S&P 500 and in a riskless asset with constant return. In-sample, strategies that consider the transitory component of stock returns significantly outperform those that ignore it.

Consider an investor with an investment horizon equal to T and an initial wealth equal to \$1. The investor starts implementing her strategy at time $t = 0$ and rebalances every month. Upon reaching the horizon $t = T$, the investor reiterates the process until reaching the terminal date of the sample. Since data are at the monthly frequency, monthly portfolio excess returns are provided over the entire sample.

Table 14 shows that considering a stochastic drift (P-CV) increases the mean, volatility, Sharpe ratio, skewness, and kurtosis of portfolio returns compared to the model with constant mean and constant volatility (CM-CV). Because the increase in the mean is large, terminal wealth is several orders of magnitude larger when expected stock returns are assumed to be stochastic.

In contrast, considering stochastic return variance (P-SV) decreases the mean and Sharpe ratio of portfolio returns but increases their volatility, skewness, and kurtosis (see Table 14). Although both volatility and kurtosis increase, it is important to note that it is in fact good volatility and good kurtosis that increase, while bad volatility and bad kurtosis remain at levels observed under constant stock return variance (see Table 15). That is, when considering stochastic stock return variance, the investor enjoys the increase in good return volatility, larger return spikes because of higher skewness, and a larger number of return spikes because of higher good kurtosis.

When the investor accounts for the presence of the transitory component (PT-CV, PT-SV), the mean, volatility, skewness, and kurtosis of her portfolio returns increase (see Table 14). Since increases in mean and volatility are of similar magnitudes, the Sharpe ratio remains the same. Importantly, the increase in volatility does not hurt the investor because the increase in good volatility is larger than that in bad volatility (see Table 15). Moreover, the increase in kurtosis is purely beneficial to the investor because good kurtosis increases and bad kurtosis decreases. In other words, the investor enjoys higher returns, larger return spikes, a larger number of return spikes, a lower number of extreme negative returns, and is not hurt by the increase in volatility when accounting for the transitory component of stock returns.

These results show that accounting for the transitory component of stock returns has no significant impact on the left tail of the distribution of portfolio returns and mainly impacts its right tail through higher good volatility, skewness, and good kurtosis. Properly modeling the behaviour of short-term equity returns implies surges in portfolio returns, which are captured by measures of performance such as good volatility, skewness, and good kurtosis. This shows that modeling the transitory component of stock returns is beneficial to the investor, and even more so when return variance is considered to be stochastic. Furthermore, it is worth noting that these benefits are not captured by the most common measure of portfolio performance, namely the Sharpe ratio.

Tables 14 and 15 show that the portfolio performance benefits of considering the transitory component are robust to changes in both investment horizon and risk aversion. An increase in the investment horizon has a weak impact on the mean, volatility, and Sharpe ratio, but it increases skewness and good kurtosis across all strategies. An increase in risk aversion has no impact on the Sharpe ratio because it decreases the mean and volatility by the same percentage across all strategies. Furthermore, an increase in risk aversion impacts skewness and kurtosis only when return volatility is stochastic. In this case, an increase in risk aversion increases skewness and good kurtosis, while it decreases bad kurtosis.

Table 16 confirms our previous statement that modeling the transitory component of equity returns is particularly beneficial to the investor when return volatility is stochastic. Indeed, the certainty equivalent return obtained by considering the transitory component is about 1% and 10% larger than that obtained by ignoring it when return volatility is constant and stochastic, respectively. Since the fraction of wealth invested in the stock increases with the horizon and decreases with risk aversion for all investment strategies (see Table 6), all strategies converge to riskless strategies when the investment horizon decreases and when risk aversion increases. For this reason, the ratio of certainty equivalent returns increases with the horizon and decreases with risk aversion.

Table 14: In-Sample Portfolio Moments vs. Investment Horizon and Risk Aversion. Mean, volatility, and Sharpe ratio are in annualized terms. In columns 1 to 3, the investment horizon is set to 1 year. In columns 4 to 6, risk aversion is $\gamma = 5$ and the investment horizon is set to 1 month, 1 year, and 5 years, respectively. P, PT, CM, CV, and SV stand for permanent component with stochastic drift, permanent component with stochastic drift and transitory component, constant mean, constant volatility, and stochastic volatility, respectively. Statistics are computed using monthly S&P 500 data from 02/1871 to 02/2016.

	Risk Aversion			Horizon		
	5	7	10	1m	1y	5y
			Mean			
CM-CV	2.11%	1.51%	1.05%	2.11%	2.11%	2.11%
P-CV	21.93%	15.57%	10.85%	21.95%	21.93%	21.93%
PT-CV	22.31%	15.83%	11.03%	22.29%	22.31%	22.40%
P-SV	17.87%	12.67%	8.82%	18.49%	17.87%	17.70%
PT-SV	19.34%	13.71%	9.55%	19.94%	19.34%	19.35%
			Volatility			
CM-CV	6.52%	4.65%	3.26%	6.52%	6.52%	6.52%
P-CV	34.81%	24.72%	17.23%	34.83%	34.81%	34.81%
PT-CV	36.66%	26.03%	18.14%	36.62%	36.66%	37.01%
P-SV	42.59%	30.35%	21.20%	43.10%	42.59%	42.65%
PT-SV	46.41%	33.08%	23.12%	46.81%	46.41%	46.61%
			Sharpe Ratio			
CM-CV	0.32	0.32	0.32	0.32	0.32	0.32
P-CV	0.63	0.63	0.63	0.63	0.63	0.63
PT-CV	0.61	0.61	0.61	0.61	0.61	0.61
P-SV	0.42	0.42	0.42	0.43	0.42	0.42
PT-SV	0.42	0.41	0.41	0.43	0.42	0.42
			Skewness			
CM-CV	0.55	0.55	0.55	0.55	0.55	0.55
P-CV	9.10	9.10	9.10	9.10	9.10	9.10
PT-CV	9.73	9.73	9.72	9.72	9.73	9.87
P-SV	25.68	25.86	25.99	24.85	25.68	25.84
PT-SV	26.84	27.03	27.16	26.02	26.84	26.71
			Kurtosis			
CM-CV	20.71	20.71	20.71	20.71	20.71	20.71
P-CV	124.24	124.20	124.16	124.33	124.24	124.25
PT-CV	141.47	141.32	141.21	141.20	141.47	146.44
P-SV	881.43	890.23	896.66	841.50	881.43	889.25
PT-SV	942.14	951.18	957.76	901.42	942.14	935.33
			Terminal Wealth			
CM-CV	2.2×10^4	1.0×10^4	5.9×10^3	2.2×10^4	2.2×10^4	2.2×10^4
P-CV	1.7×10^{14}	2.8×10^{11}	1.5×10^9	1.7×10^{14}	1.7×10^{14}	1.7×10^{14}
PT-CV	1.8×10^{14}	3.2×10^{11}	1.7×10^9	1.8×10^{14}	1.8×10^{14}	1.9×10^{14}
P-SV	4.8×10^{11}	3.9×10^9	7.2×10^7	8.7×10^{11}	4.8×10^{11}	3.9×10^{11}
PT-SV	2.0×10^{12}	1.1×10^{10}	1.6×10^8	3.5×10^{12}	2.0×10^{12}	1.9×10^{12}

Table 15: In-Sample Portfolio Return Volatility and Kurtosis Decomposition vs. Investment Horizon and Risk Aversion.

Volatility is in annualized terms. In columns 1 to 3, the investment horizon is set to 1 year. In columns 4 to 6, risk aversion is $\gamma = 5$ and the investment horizon is set to 1 month, 1 year, and 5 years, respectively. P, PT, CM, CV, and SV stand for permanent component with stochastic drift, permanent component with stochastic drift and transitory component, constant mean, constant volatility, and stochastic volatility, respectively. Statistics are computed using monthly S&P 500 data from 02/1871 to 02/2016.

	Risk Aversion			Horizon		
	5	7	10	1m	1y	5y
			Good Volatility			
CM-CV	4.50%	3.22%	2.25%	4.50%	4.50%	4.50%
P-CV	32.37%	22.99%	16.02%	32.39%	32.37%	32.37%
PT-CV	34.23%	24.30%	16.93%	34.19%	34.23%	34.54%
P-SV	40.64%	28.97%	20.25%	40.98%	40.64%	40.72%
PT-SV	44.40%	31.67%	22.15%	44.64%	44.40%	44.57%
			Bad Volatility			
CM-CV	4.71%	3.37%	2.36%	4.71%	4.71%	4.71%
P-CV	12.80%	9.09%	6.33%	12.82%	12.80%	12.80%
PT-CV	13.13%	9.32%	6.50%	13.12%	13.13%	13.30%
P-SV	12.76%	9.03%	6.28%	13.36%	12.76%	12.69%
PT-SV	13.50%	9.55%	6.64%	14.09%	13.50%	13.64%
			Good Kurtosis			
CM-CV	16.09	16.09	16.09	16.09	16.09	16.09
P-CV	123.75	123.70	123.67	123.84	123.75	123.76
PT-CV	141.00	140.85	140.74	140.73	141.00	145.96
P-SV	880.82	889.64	896.09	840.77	880.82	888.66
PT-SV	941.67	950.73	957.32	900.89	941.67	934.86
			Bad Kurtosis			
CM-CV	4.62	4.62	4.62	4.62	4.62	4.62
P-CV	0.50	0.49	0.49	0.50	0.50	0.49
PT-CV	0.47	0.47	0.47	0.47	0.47	0.49
P-SV	0.61	0.59	0.57	0.73	0.61	0.60
PT-SV	0.47	0.45	0.44	0.54	0.47	0.47

Table 16: In-Sample Ratio of Certainty Equivalent Returns.
CER stands for certainty equivalent return. In columns 1 to 3, the investment horizon is set to 1 year. In columns 4 to 6, risk aversion is $\gamma = 5$ and the investment horizon is set to 1 month, 1 year, and 5 years, respectively. P, PT, CV, and SV stand for permanent component with stochastic drift, permanent component with stochastic drift and transitory component, constant volatility, and stochastic volatility, respectively. Statistics are computed using monthly S&P 500 data from 02/1871 to 02/2016.

	Risk Aversion			Horizon		
	5	7	10	1m	1y	5y
			Ratio of Mean CER			
PT-CV/P-CV	1.0127	1.0110	1.0092	1.0126	1.0127	1.0195
PT-SV/P-SV	1.1037	1.0880	1.0717	1.0950	1.1037	1.1170
			Ratio of Median CER			
PT-CV/P-CV	1.0096	1.0084	1.0072	1.0030	1.0096	1.0170
PT-SV/P-SV	1.0927	1.0771	1.0618	1.0770	1.0927	1.0939