

Can Reinvestment Risk Explain the Dividend and Bond Term Structures?*

Andrei S. Gonçalves[†]

This Version: January, 2018

[Link to Most Recent Version](#)

Abstract

Contradicting leading asset pricing models, recent evidence indicates the term structure of dividend discount rates is downward sloping at long maturities despite the typical upward sloping bond yield curve. This paper empirically shows that reinvestment risk explains both the dividend and bond term structures. Intuitively, dividend claims hedge reinvestment risk because dividend present values rise as expected returns decline. This hedge is more effective for longer-term claims because they are more sensitive to discount rate variation, resulting in a downward sloping dividend term structure. For bonds, as expected equity returns decline, nominal interest rates rise, and bond prices fall. Consequently, bonds are exposed to reinvestment risk, and this exposure increases with duration, giving rise to an upward sloping bond term structure. This reinvestment risk channel has direct implications for the long run risks framework.

JEL Classification: E32; E43; G11; G12.

Keywords: Dividend Term Structure; Bond Term Structure; Dividend Risk Premia; Reinvestment Risk; Intertemporal CAPM.

*I thank Christian Mueller-Glissmann at Goldman Sachs International for providing me with dividend futures data. I am deeply indebted to René Stulz, Lu Zhang, and Kewei Hou for their invaluable advice and guidance. I am also grateful for the very helpful comments from Zahi Ben-David, Michael Brennan (discussant), Justin Birru, Oliver Boguth (discussant), John Campbell, Andrew Chen, Stéphane Chrétien (discussant), Spencer Couts, Robert Dittmar (discussant), Harry DeAngelo, José Faias (discussant), Anastassia Fedyk (discussant), Shan Ge, Steffen Hitzemann, Sehoon Kim, Ralph Koijen, Alan Kwan, Xiaoji Lin, Tyler Muir, Andrea Rossi, Mike Schwert, Andrea Tamoni (discussant), and Mike Weisbach as well as seminar participants at The Ohio State University, the 2017 FRA, EFA, NFA, and FMA meetings, the 2017 USC Marshall Ph.D. Conference in Finance, the 2017 Trans-Atlantic Doctoral Conference, and the 2017 LubraFin Conference. All remaining errors are my own. First draft: December 2016.

[†]Finance Ph.D. Candidate, the Ohio State University. 810 Fisher Hall, 2100 Neil Avenue, Columbus, OH 43210. E-mail: goncalves.11@fisher.osu.edu.

Introduction

Discounting future cash flows is at the heart of investment decisions. But, how do discount rates vary with cash flow maturity? For risky cash flows, like dividends, recent evidence indicates the term structure of discount rates decreases beyond the first few years (Giglio, Maggiori, and Stroebl (2015); Binsbergen and Koijen (2017); Weber (2017)). The relatively low discount rates of long-term dividends contrast with the fact that interest rates tend to increase in bond duration and with leading asset pricing models as they predict a flat or upward sloping dividend term structure (Binsbergen, Brandt, and Koijen (2012)). The opposite term structure behavior of bonds and dividends is a paradox and the fact that prominent models fail to capture such a fundamental property of financial markets is unsettling.

As Campbell (2018) points out, “A recent literature has asked how to reconcile a downward-sloping risky term structure with asset pricing theory...[but]...it is not easy (p. 291).” This paper sheds light on this issue by studying the risk differences between short- and long-term bonds and dividend claims. In particular, I empirically demonstrate that reinvestment risk is a potential explanation for the *downward* sloping dividend term structure at long maturities and the *upward* sloping bond term structure.

I define reinvestment risk as exposure to declines in market reinvestment rates. I model this risk in the context of the Intertemporal Capital Asset Pricing Model (ICAPM) of Campbell (1993). The model features market returns as well as interest rate and equity premium shocks as risk factors. While negative market returns decrease current wealth, negative interest rate and equity premium shocks reduce bond and equity reinvestment rates respectively. In the ICAPM, declines in current wealth and in reinvestment rates decrease lifetime utility, and hence are relevant sources of risk.

The new insight is that the dividend and bond term structures have opposite slopes because long-term dividend claims hedge reinvestment risk while long-term bonds are exposed to such risk. Intuitively, *decreases* in reinvestment rates are associated with *increases* in dividend present values through lower discount rates. This hedge is valuable and more relevant for longer-term claims given their higher sensitivity to discount rate variation. Therefore, investors have higher demand for longer-term dividend claims, which translates into lower risk premia for these assets. In the case of bonds, prices decrease when equity reinvestment rates decline because nominal interest rates are negatively correlated with the equity premium (Fama and Schwert (1977); Campbell (1987); Ferson

(1989); Shanken (1990); Brennan (1997)). Consequently, bonds are exposed to equity reinvestment risk despite hedging against interest rate declines. Given their higher duration, longer-term bonds are more exposed to equity reinvestment risk, and thus command higher risk premia.

I empirically test this reinvestment risk mechanism. For the bond term structure, I use returns on six portfolios containing bonds with maturities up to 1, 2, 3, 4, 5, and 10 years (the data is available since 1952). In the case of the dividend term structure, the biggest challenge is that dividend claims only start trading in the 21st century, and thus tests of risk-based mechanisms are likely to be contaminated by the great recession. I overcome this challenge by developing a novel methodology to estimate risk exposures on dividend claims and applying it to aggregate U.S. dividends accruing in one to ten years. While the previous paragraph holds expected dividends fixed to simplify the intuition, my tests fully incorporate cash flow shocks. I calculate dividend returns (and risk exposures) based on dividend present values, which I obtain by discounting expected dividends by expected equity returns every month from 1952 to 2015. I rely on predictive regressions over alternative horizons to estimate expectations for dividends and equity returns, and all predictive variables used are commonly applied in the asset pricing literature.

The first major empirical finding is that longer-term dividend claims have lower risk premia within the ICAPM despite their higher market risk, and this result is a consequence of their ability to hedge reinvestment risk (Figures 1a and 1c). Specifically, I find that market betas increase in dividend maturity, while equity premium betas strongly decrease. The market beta evidence reveals that the dividend term structure is upward sloping under the CAPM and is consistent with the previous literature (Ang and Liu (2004); Brennan and Xia (2006); Binsbergen, Brandt, and Koijen (2012)). In contrast, the equity premium beta pattern represents a novel empirical result and is consistent with long-term dividend claims hedging reinvestment risk. To obtain risk premia, I combine betas with risk prices implied by the ICAPM of Campbell (1993) when relative risk aversion is calibrated to match the equity premium and the bond term structure. The ICAPM-based risk premia are lower for longer-term dividend claims, which indicates that the reinvestment risk channel dominates the market risk effect and produces a downward sloping dividend term structure.

The second main result is that longer-term bonds are more exposed to reinvestment risk, resulting in an upward sloping bond term structure (Figures 1b and 1c). In particular, long term bonds hedge against interest rate declines, but are also highly exposed to equity premium shocks. The later effect dominates so that the bond term structure of reinvestment risk is upward sloping. Market risk

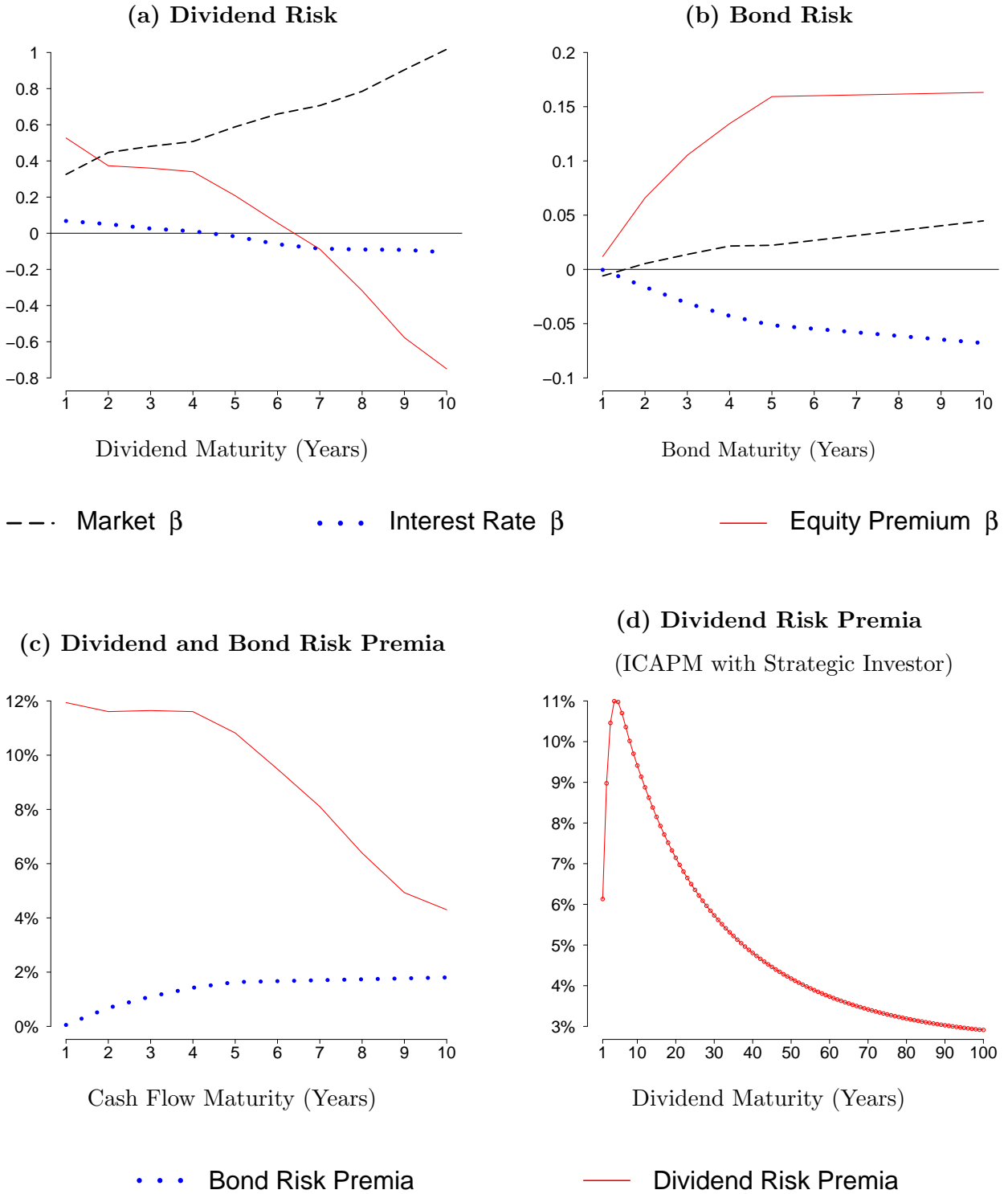


Figure 1

The Dividend and Bond Term Structures

The graphs report betas and risk premia for dividend claims and bond portfolios. Panels (a) and (b) focus on betas relative to the three ICAPM risk factors (market return, interest rate, and equity premium shocks) while panels (c) and (d) concentrate on risk premia. All betas are in market beta units (i.e., covariance normalized by market variance) and all risk premia are constructed by combining betas with risk prices. Risk prices for panel (c) are based on the ICAPM with a passive investor (Campbell (1993)) and a relative risk aversion (=6.3) calibrated to match the equity premium and the bond term structure. Risk prices for panel (d) are based on the ICAPM with a strategic investor (see Campbell, Chan, and Viceira (2003)). All necessary parameters are estimated over my main sample period (1952 to 2016). Empirical details are available in subsections 3.1 and 4.1.

also increases in bond maturity, but the upward sloping bond term structure is mostly driven by reinvestment risk as the standard CAPM cannot produce a realistic bond-term structure.

The reinvestment risk explanation for the dividend and bond term structures is consistent with economic theory and several stylized empirical facts. For instance, I formalize the calibration of relative risk aversion as a Generalized Method of Moments (GMM) estimation and find an implied relative risk aversion of 6.3, which is comparable to recent estimates (Campbell et al. (2017)) and below the upper bound Mehra and Prescott (1985) argue to be reasonable for asset pricing models. Moreover, I expand the estimation to other testing assets and find that the ICAPM quantitatively captures the term structures of Treasury and corporate bonds and reasonably fits the equity, value, and credit premiums (all simultaneously).

My approach of studying the dividend term structure based on dividend present values builds on the traditional methodology of decomposing stock return variation into cash flow and discount rate news (Campbell and Shiller (1989); Campbell (1991); Chen, Da, and Zhao (2013)). My innovation is to rewrite Campbell and Shiller (1989)'s valuation identity in a way that allows me to decompose equity returns into their underlying dividend returns over long sample periods. This alternative way to estimate dividend returns is an important contribution because it allows future research to test alternative models and explore new properties of the dividend term structure. Empirical testing has been limited because data on derivative contracts used to calculate dividend returns are not available for a sufficiently long sample period.

However, dividend returns obtained from present values are not traded asset returns. Consequently, the dividend term structure documented in Figures 1a and 1c is based on the contribution of each dividend to the equity premium as opposed to the more common method of averaging excess returns over time. In Section 1, I provide a detailed explanation of this point and prove that the proposed methodology provides a valid decomposition of the equity premium into its underlying dividend risk premia.

I complement Figures 1a and 1c with a test based on a short-term dividend strategy that uses a proprietary database of quoted prices of dividend futures provided to me by Goldman Sachs. I find that a portfolio of short-term dividend futures has lower market β , higher equity premium β , and higher ICAPM-implied risk premium than a long-duration portfolio of dividend futures. All of these results are consistent with the ones obtained using dividend returns based on present values.

The results in Figures 1a and 1c are striking in that they suggest reinvestment risk produces both

a downward sloping dividend term structure and an upward sloping bond term structure. However, they are likely to overstate the effect of reinvestment risk and consequently the slope of the dividend term structure. The reason is that the ICAPM of [Campbell \(1993\)](#) considers a passive investor who does not change her portfolio allocation in response to changes in reinvestment rates (i.e., the investor is not strategic). If the marginal investor is strategic, then she can endogenously decrease her exposure to reinvestment risk, which makes it less of a concern.

To explore this issue, I consider a fully structural ICAPM in which the marginal investor responds optimally to changes in reinvestment rates (see [Campbell, Chan, and Viceira \(2003\)](#)). I derive expressions for the risk exposures and risk premia of dividend claims within this framework and study the dividend term structure by applying the derived expressions to the data. This approach allows me to study the ICAPM pricing of dividend claims with arbitrary maturities (it is not based on dividend present values). This benefit, of course, comes at the expense of requiring more economic/econometric assumptions than the previous tests. The results are, however, interesting and empirically credible.

Figure 1d displays the dividend term structure of risk premia for the ICAPM with a strategic investor. For the first few maturities, the market risk effect dominates, inducing an upward sloping dividend term structure. However, the reinvestment risk effect dominates for dividend claims with longer maturities so that long-term dividend claims have relatively low risk premia. Overall, the ICAPM produces a hump-shaped dividend term structure with particularly low discount rates for very long-term dividends. These results are striking in that they match the otherwise puzzling pattern uncovered in the literature. For instance, [Binsbergen and Koijen \(2017\)](#) report (for S&P 500 dividend claims) average returns that increase over the first few years, but decrease over the long-term as dividend claims with maturities from three to five years outperform the S&P 500, which is a long-duration portfolio of dividend claims.

The term structures in Figure 1 are robust to extensive variation in my empirical design. For instance, results are qualitatively similar when the sample goes as far back as 1928 or when it excludes the great depression/recession. Moreover, results are robust to excluding any of the predictive variables used to estimate expected returns and dividend growth. I also study robustness to several measurement, econometric, and economic decisions and find that more than 300 specifications yield results that are qualitatively similar to the ones obtained in my main analysis.

The reinvestment risk channel I propose to explain the dividend and bond term structures has direct implications to the long-run risks framework ([Bansal and Yaron \(2004\)](#); [Bansal, Kiku, and](#)

Yaron (2012)). I show (through simulations) that the typical calibration of the long-run risks model does not account for the dividend term structure because it fails to produce the downward sloping term structure of reinvestment risk uncovered in my empirical analysis. However, the long-run risks framework correctly prices reinvestment risk and the relationship between this framework and the ICAPM implies that there must exist an alternative specification for the consumption growth process that captures both the bond and dividend term structures. Such specification might not reflect the consumption growth of a representative consumer (see Malloy, Moskowitz, and Vissing-Jørgensen (2009)), but it can still retain the main aspects of the model.

In summary, this paper has two key implications. The first is that investors should discount longer-term equity cash flows at lower rates if they price reinvestment risk as typical agents in intertemporal asset pricing theory do. The second is that reinvestment risk provides an empirically credible explanation for the coexistence of a *downward* sloping dividend term structure at long maturities and an *upward* sloping bond term structure in financial markets.

Research on how discount rates vary with cash flow maturity can be traced back to a literature focused on the importance of the discount rate term structure to the valuation of projects and firms (Gordon (1962); Brennan (1973); Bogue and Roll (1974); Fama (1977); Myers and Turnbull (1977)). My paper contributes to this early work by providing evidence suggesting that long-term equity cash flows are good hedges for reinvestment risk, and thus should require relatively low discount rates in present value applications.

Several studies find that long-term risky cash flows (in particular dividend claims) have relatively low discount rates so that the term structure of risky cash flows is downward sloping at long maturities (see Binsbergen and Koijen (2017) for a review of this literature).¹ This result represents a puzzle

¹This literature builds on Brennan (1998)'s insight that variation in dividend prices provides an important signal for changes in the fundamental value of cash flows. Binsbergen, Brandt, and Koijen (2012) use option prices and the put-call parity to show that claims on S&P500 short-term dividends have higher average returns than the index itself, which is a long duration portfolio of dividend claims. Binsbergen and Koijen (2017) studies dividend futures contracts and find similar results while Binsbergen et al. (2013) emphasizes that, despite being downward sloping on average, the dividend term structure is pro-cyclical from 2002 to 2012. Cejnek and Randl (2017) and Li and Wang (2017) further explore the predictability power of dividend prices and Cejnek and Randl (2016) relate the dividend term structure to an options-based downside risk factor. Manley and Mueller-Glissmann (2008) and Wilkens and Wimschulte (2010) provide valuable institutional details about the market for dividend derivatives. Some papers argue that the dividend term structure is no longer downward sloping when we consider microstructure noise (Boguth et al. (2012)), dividend taxation (Schulz (2016)), or dealer funding costs (Song (2016)). However, these criticisms only apply to the evidence based on option prices (not to dividend futures contracts), and thus cannot explain the low discount rates on long-term dividend claims. I provide further details about all three points and some counterarguments to them in Appendix B.3. Evidence of relatively low discount rates for long-term cash flows is also found in real estate prices (Giglio, Maggiori, and Stroebele (2015); Giglio et al. (2015)) and in portfolios sorted on equity cash flow duration (Dechow, Sloan, and Soliman (2004); Weber (2017)).

because it contradicts the flat or upward sloping dividend term structure implied by some of the most prominent asset pricing models.² My contribution to this literature is to *empirically* demonstrate the plausibility of reinvestment risk as a simple *explanation* for the downward sloping dividend term structure and the upward sloping bond term structure. My paper is unique within this literature in that it focuses on an empirical test of a theoretical mechanism as opposed to providing evidence within a simulated economy.

The reinvestment risk channel is substantially different from the three classes of models that have been proposed to capture the dividend term structure. The first class assumes an exogenous stochastic discount factor and uses it to price dividends (and sometimes bonds) through non-arbitrage pricing restrictions (Lettau and Wachter (2007, 2011); Ang and Ulrich (2012); Kragt, de Jong, and Driessen (2016)). The second class modifies preferences (Andries, Eisenbach, and Schmalz (2015); Eisenbach and Schmalz (2016)) or beliefs (Croce, Lettau, and Ludvigson (2014)) in order to consider investors who worry more about short-term risks. The third class modifies the economic growth process in a way that induces shorter-term dividends to be riskier.³ The mechanism in my paper differs substantially from these as reinvestment risk pricing does not require investors to worry more about the short-term, and the term structure of reinvestment risk is driven by discount rate movements, not cash flow risk. In simple words, it is not about the risk that short-term dividends have, but rather about the hedge benefit they do not have.

The results in this paper contribute to the well established ICAPM and bond term structure literatures by demonstrating that a simple ICAPM specification simultaneously produces a downward sloping dividend term structure and an upward sloping bond term structure.⁴ One paper that is close to mine in these literatures is Brennan and Xia (2006). They calibrate a version of Merton (1973)'s ICAPM based on Treasury bond yields and inflation data and use it to study the valuation of

²Standard models generating a flat or upward sloping dividend term structure include (i) the Sharpe (1964) and Lintner (1965) CAPM; (ii) the habit formation model in Campbell and Cochrane (1999); (iii) the long-run risk model in Bansal and Yaron (2004) and Bansal, Kiku, and Yaron (2012) and (iv) the disaster risk framework in Barro (2006), Gabaix (2012) and Wachter (2013).

³Mechanisms include stationary leverage policy (Belo, Collin-Dufresne, and Goldstein (2015)), variation in risk prices added to a latent time-varying leverage in dividend growth (Doh and Wu (2016)), dividends that are cointegrated with consumption (Marfè (2015)) or that recover after disasters (Hasler and Marfè (2016)), older vintages of capital that are more exposed to productivity shocks (Ai et al. (2017)), and labor rigidity in the form of sticky wages (Favilukis and Lin (2016)) or income insurance (Marfè (2017)).

⁴The bond term structure literature is summarized by Piazzesi (2010) and Duffee (2013). The ICAPM was first proposed by Merton (1973) and a closely related discrete version was developed by Campbell (1993), which is the framework I rely on in this paper. Given the long history of the ICAPM, many papers empirically tests and/or further develop the model. Some examples are Campbell (1996), Ferson and Harvey (1999), Campbell, Chan, and Viceira (2003), Brennan, Wang, and Xia (2004), Campbell and Vuolteenaho (2004), Petkova (2006), Campbell, Polk, and Vuolteenaho (2009), Bali and Engle (2010) and Campbell et al. (2017).

cash flows with different horizons. My paper differs from theirs in many aspects, with one being particularly important. While they estimate a negative risk price for interest rate shocks, I find that relative risk aversion above one restricts risk prices on equity premium and interest rate shocks to be positive. Consequently, my paper captures the logic of reinvestment risk and fundamentally differs from theirs in terms of how investors discount cash flows with different horizons.

The documented evidence also relates to studies on the risk and return characteristics of portfolios with different cash flow duration (Dechow, Sloan, and Soliman (2004); Lettau and Wachter (2007, 2011); Hansen, Heaton, and Li (2008); Da (2009); Weber (2017)). This literature is connected to the value premium (Fama and French (1992, 1993, 1996)) and finds that longer-duration stocks have lower expected returns. One natural question is whether the differences between short- and long-duration portfolios are driven by the dividend term structure. My results provide guidance to answer this question because they imply that term structure differences in duration sorted portfolios are, at least in part, driven by heterogeneity in reinvestment risk exposure.

The rest of the paper is organized as follows. Section 1 provides a theoretical motivation for the methods I use to study the dividend term structure. Section 2 details the ICAPM mechanism that induces reinvestment risk to be priced. In turn, Sections 3 and 4 report results based on the ICAPM with a passive and strategic investor respectively. Section 5 explores the relationship between the reinvestment risk channel I study and the long-run risks framework. Finally, Section 6 concludes by discussing some implications of my results and avenues for future research. The Appendix contains derivations and empirical/econometric details as well as supplementary empirical results.

1 The Equity Premium Term Structure Decomposition

This section motivates my empirical approach to study the dividend term structure. The punchline is that the equity premium can be decomposed into its underlying dividend risk premia. Moreover, this decomposition can be performed using either dividend prices (calculated by discounting dividends using the Stochastic Discount Factor) or dividend present values (calculated by discounting dividends using equity discount rates). The method based on dividend prices is often used in the literature (and in my Section 4), while the present value approach is novel and the focus of Section 3. The main advantage of this method is that it allows me to study the dividend term structure over a long sample period without a fully structural model.

For the rest of this paper, flow variables are stated in real terms (unless otherwise noted) and technical derivations can be found in Appendix A.

1.1 Two Alternative Equity Premium Decompositions

This subsection outlines two alternative equity premium decompositions, which I label “the price method” and “the present value method”. The present value method is novel and allows me to study the dividend term structure without imposing strong economic restrictions. In contrast, the price method requires stronger assumptions to be empirically analyzed, but is more comparable to the literature and allows for a more in depth economic understanding of the dividend term structure.

To understand the price method, note that, in the absence of arbitrage opportunities, there exists a Stochastic Discount Factor (SDF), M_t , such that the value of an equity index, P_t , is given by the discounted value of its dividends $\{D_{t+h}\}_{h=1}^{\infty}$:

$$P_t = \sum_{h=1}^{\infty} \mathbb{E}_t [D_{t+h} \cdot M_{t \rightarrow t+h}] \equiv \sum_{h=1}^{\infty} P_t^{(h)} \quad (1)$$

where the second equality defines the price of a dividend accruing in h years (also called h -year dividend strip).

Manipulating equation 1, we have that returns on the equity index, $R_{e,t}$, are equivalent to returns on a portfolio of dividend strips (with weights $w_t^{(h)} = P_t^{(h)} / P_t$):

$$R_{e,t} = \sum_{h=1}^{\infty} w_{t-1}^{(h)} \cdot R_{ds,t}^{(h)} \quad (2)$$

and this equation allows us to decompose the equity premium into a term structure of dividend risk premia.

An alternative equity premium decomposition can be derived based on the valuation identity explored in Campbell and Shiller (1988, 1989). Start from the definition of a gross equity index return and isolate price to get $P_t = (P_{t+1} + D_{t+1}) / R_{e,t+1}$, which can be iterated forward to yield:

$$P_t = \sum_{h=1}^{\infty} \mathbb{E}_t \left[D_{t+h} \cdot \left(\prod_{j=1}^h R_{e,t+j} \right)^{-1} \right] \equiv \sum_{h=1}^{\infty} PV_t^{(h)} \quad (3)$$

The present value of a dividend accruing in h years, $PV^{(h)}$, represents its contribution to the current equity price. I refer to these as dividend PVs to differentiate them from the dividend strip terminology reserved for $P^{(h)}$. While $P^{(h)}$ is obtained by discounting dividends using the SDF,

$PV^{(h)}$ discounts using equity discount rates. Appendix A.2 explains the similarities and distinctions between $P^{(h)}$ and $PV^{(h)}$, but it is clear from equations 1 and 3 that the two objects provide alternative decompositions of equity prices.

Manipulating equation 1, we can derive the analogue of equation 2 for dividend PV returns. However, it is empirically easier to work with a log-linear version of this equation, which I detail in the next subsection.

1.2 Dividend Present Values and the Log-Linear Valuation Identity

This subsection derives a log-linear identity linking the equity premium to the term structure of dividend risk premia.

Assuming log returns and dividend growth are conditionally homoskedastic and normally distributed, equation 3 implies that dividend PVs depend only on current dividends, expected dividend growth and expected equity returns (ignoring constants):

$$\ln(PV_t^{(h)}) = \ln(D_t) + \mathbb{E}_t \left[\sum_{j=1}^h \Delta d_{t+j} - r_{e,t+j} \right] \quad (4)$$

where Δd is the log dividend growth and r_e is the log equity return.

I combine dividend PV returns ($R_{pv,t}^{(h)} = PV_t^{(h-1)} / PV_{t-1}^{(h)}$) with equation 4 along with a log-linear stock return approximation (Campbell and Shiller (1989); Campbell (1991)) to show that (log) stock returns can also be viewed as returns on a portfolio of dividend claims:

$$r_{e,t} - \mathbb{E}_{t-1}[r_{e,t}] \simeq \sum_{h=1}^{\infty} w_h \cdot \left(r_{pv,t}^{(h)} - \mathbb{E}_{t-1}[r_{pv,t}^{(h)}] \right) \quad (5)$$

where $r_{pv}^{(h)}$ are dividend PV log returns, $w_h = \rho^{h-1} - \rho^h$ are weights that decrease in horizon and satisfy $\sum_{h=1}^{\infty} w_h = 1$, and ρ is Campbell and Shiller (1989)'s log-linearization constant.⁵

Finally, letting the log SDF be given by $m = \log(M)$, we have that equation 5 implies:

$$\beta_{e,t} \simeq \sum_{h=1}^{\infty} w_h \cdot \beta_{pv,t}^{(h)} \quad (6)$$

$$\mathbb{E}_t[R_{e,t+1} - R_{f,t+1}] \simeq \sum_{h=1}^{\infty} w_h \cdot \beta_{pv,t}^{(h)} \cdot \lambda_t \quad (7)$$

⁵In my empirical analysis, I use $\rho = \exp(\overline{pd}) / [\exp(\overline{pd}) + 1]$, which is approximately 0.97 in my main sample. Figure A.1 reports the dividend PV return weights, $w_h = \rho^{h-1} - \rho^h$, for $\rho = 0.97$.

where R_f is the gross risk free rate, $\beta_{pv,t}^{(h)} = Cov_t(r_{pv,t+1}^{(h)} - r_{f,t+1}, -m_{t+1})/Var_t(m_{t+1})$ captures the respective dividend claim risk exposure, and $\beta_{pv,t}^{(h)} \cdot \lambda_t$ represents its risk premium.

Equations 6 and 7 provide a way to study the dividend term structure. Equation 6 shows that the exposure of the equity index to the log SDF is a weighted average of its dividend β s. This equation is valid with other risk factors (whether priced or not), and thus can be applied without imposing economic assumptions. Similarly, equation 7 demonstrates that the equity premium is a weighted average of its dividend risk premia. Since risk premia depend on dividend risk exposures and on risk prices, this equation can be used to study the dividend term structure, but it requires an explicit model for the SDF (Section 2 details the SDF used in my empirical analysis).

Note that dividend PV returns are not traded asset returns, and thus $\mathbb{E}_t[R_{pv,t+1}^{(h)} - R_{f,t+1}] \neq \beta_{pv,t}^{(h)} \cdot \lambda_t$. In fact, $\mathbb{E}_t[R_{pv,t+1}^{(h)}] = \mathbb{E}_t[R_{e,t+1}]$ holds for all h by construction. This result means that average excess returns on dividend PVs are irrelevant for studying the dividend term structure. Instead, $\beta_{pv,t}^{(h)} \cdot \lambda_t$ is the object of interest since it represents the respective dividend's contribution to the equity premium.

1.3 Sources of Variation in Dividend PV Returns

To study the dividend term structure, it is necessary to understand dividend β s, which depend on the sources of variation in dividend PV returns. This subsection demonstrates that dividend PV returns depend on shocks to dividend growth as well as on news about future (i) dividend growth; (ii) equity premium; and (iii) interest rates.

To detail the sources of variation in dividend PV returns, I decompose the dividend PVs in equation 4 as (ignoring constants):

$$\ln(PV_t^{(h)}) = \ln(D_t) + \underbrace{\mathbb{E}_t \left[\sum_{j=1}^h \Delta d_{t+j} \right]}_{g_t^{(h)}} - \underbrace{\mathbb{E}_t \left[\sum_{j=1}^h r_{e,t+j} - r_{f,t+j} \right]}_{ep_t^{(h)}} - \underbrace{\mathbb{E}_t \left[\sum_{j=1}^h r_{f,t+j} \right]}_{ir_t^{(h)}} \quad (8)$$

and combine this equation with the definition of dividend PV returns to obtain:

$$\begin{aligned} r_{pv,t}^{(h)} - \mathbb{E}_{t-1}[r_{pv,t}^{(h)}] &= \Delta d_t - \mathbb{E}_{t-1}[\Delta d_t] \\ &+ \left(g_t^{(h-1)} - \mathbb{E}_{t-1}[g_t^{(h-1)}] \right) - \left(ep_t^{(h-1)} - \mathbb{E}_{t-1}[ep_t^{(h-1)}] \right) - \left(ir_t^{(h-1)} - \mathbb{E}_{t-1}[ir_t^{(h-1)}] \right) \end{aligned}$$

or in more compact notation (with \sim representing shocks to contemporaneous variables and N news

about future information):

$$\tilde{r}_{pv,t}^{(h)} = \tilde{\Delta}d_t + N_{g,t}^{(h-1)} - N_{ep,t}^{(h-1)} - N_{ir,t}^{(h-1)} \quad (9)$$

Equation 9 shows that all dividend claims are exposed to an identical dividend growth shock ($\tilde{\Delta}d$), but they are subject to different news about future (i) dividend growth ($N_g^{(h)}$); (ii) equity premium ($N_{ep}^{(h)}$); and (iii) interest rates ($N_{ir}^{(h)}$). For instance, the annual return on a 1-year dividend claim is only subject to dividend growth shocks. In contrast, the annual return on a 10-year dividend claim is exposed not only to dividend growth shocks, but also to news about the remaining nine years dividend growth, equity premium, and interest rate. As a consequence, the heterogeneity in dividend PV returns with different maturities originates from the risk heterogeneity in these three components.

A key advantage of studying the dividend term structure using dividend PV returns is that they can be constructed based on proxies for dividend growth, equity premium, and interest rate news. This is substantially different from dividend strip returns, which require either observations of $P^{(h)}$ in financial markets or a fully specified SDF to recover $P^{(h)}$.

Section 3 studies the dividend term structure based on dividend PV returns and provides empirical results that rely on little economic structure. The key intuition that long-term dividend claims are good hedges for reinvestment risk is clear from this analysis. There are, however, important nuances in the dividend term structure. For instance, the term structure seems upward sloping over the short term in the U.S. despite being downward sloping over the long-horizon (see [Binsbergen and Koijen \(2017\)](#)). Section 4 presents results based on $P^{(h)}$ constructed from a fully structural ICAPM and shows that these aspects are also captured by the model.

The next section details the ICAPM structure I rely on throughout the rest of this paper to capture the reinvestment risk logic.

2 An ICAPM with Reinvestment Risk

In this section, I show that a passive investor with Epstein-Zin preferences dislikes reinvestment risk and prices it accordingly. I provide only the key steps, but the model is an ICAPM, with all derivations detailed in [Campbell \(1993\)](#) and an important empirical test performed by [Campbell and Vuolteenaho \(2004\)](#). The model here is also the starting point for the fully structural ICAPM framework in Section 4, which features a strategic investor who responds optimally to changes in

reinvestment rates.

2.1 The SDF from the Perspective of a Passive Investor

An investor with Epstein-Zin recursive preferences (Epstein and Zin (1989); Weil (1989)) who has his wealth invested in the market portfolio (i.e., is a passive investor) has a log SDF, $m = \ln(M)$, with shocks given by:

$$\tilde{m}_t = - (1 - \theta) \cdot \tilde{r}_{m,t} - \frac{\theta}{ies} \cdot \tilde{\Delta}c_t$$

where \tilde{x} represents a shock to a generic variable x , Δc is the log consumption growth, r_m is the log return on the market portfolio, and $\theta = (1 - rra) / (1 - 1/ies)$, with rra representing relative risk aversion and ies representing intertemporal elasticity of substitution.

Letting W represent the wealth and C the consumption of the relevant investor, the log return on the market portfolio can be written as $r_{m,t} = -\ln((W_{t-1} - C_{t-1})/C_{t-1}) + \Delta c_t - \ln(C_t/W_t)$. Hence, we have $\tilde{\Delta}c_t = \tilde{r}_{m,t} + \tilde{l}\tilde{n}(C_t/W_t)$ and can write \tilde{m}_t as:

$$\tilde{m}_t = - rra \cdot \tilde{r}_{m,t} - \frac{\theta}{ies} \cdot \tilde{l}\tilde{n}(C_t/W_t) \quad (10)$$

Assuming log returns and the log SDF are (conditionally joint) homoskedastic and normally distributed and using the pricing equation for the market portfolio, $\mathbb{E}_{t-1}[M_t \cdot R_{m,t}] = 1$, and a log-linear approximation to the budget constraint, $W_t = R_{m,t} \cdot (W_{t-1} - C_{t-1})$, it is possible to write shocks to the log consumption-wealth ratio as:

$$\begin{aligned} \tilde{l}\tilde{n}\left(\frac{C_t}{W_t}\right) &\approx (1 - ies) \cdot (\mathbb{E}_t - \mathbb{E}_{t-1}) \left[\sum_{h=1}^{\infty} \rho^h \cdot r_{m,t+h} \right] \\ &\equiv (1 - ies) \cdot N_{\mathbb{E},t} \end{aligned} \quad (11)$$

where $N_{\mathbb{E},r}$ represents news to expected returns or reinvestment rates (discount rate news in the terminology used in Campbell (1993) and Campbell and Vuolteenaho (2004)).

Equation 11 says that the consumption-wealth ratio responds to reinvestment rates with the exact behavior depending on whether the income or the substitution effect dominates.⁶ Equation 11 can be substituted into equation 10 to capture the reinvestment risk logic. If we assume the market portfolio is the equity market portfolio, then we have $N_{\mathbb{E},r} = N_{ep} + N_{ir}$ and:

⁶If $ies < 1$, the income effect dominates and an increase in reinvestment rates induces the investor to save less (consume more as a fraction of wealth). Conversely, the substitution effect dominates if $ies > 1$ and an increase in reinvestment rates induces the investor to save more (consume less as a fraction of wealth).

$$\tilde{m}_t = - rra \cdot \tilde{r}_{m,t} - (rra - 1) \cdot N_{ep,t} - (rra - 1) \cdot N_{ir,t} \quad (12)$$

where

$$N_{ep,t} = (\mathbb{E}_t - \mathbb{E}_{t-1}) \left[\sum_{h=1}^{\infty} \rho^h \cdot (r_{e,t+h} - r_{f,t+h}) \right] \quad \text{and} \quad N_{ir,t} = (\mathbb{E}_t - \mathbb{E}_{t-1}) \left[\sum_{h=1}^{\infty} \rho^h \cdot r_{f,t+h} \right]$$

Equation 12 indicates that both market risk (through \tilde{r}_m) and reinvestment risk (through N_{ep} and N_{ir}) are priced. Moreover, it specifies the respective risk prices as functions of a single parameter (relative risk aversion). Intuitively, while negative market returns decrease current wealth, negative interest rate (N_{ir}) and equity premium (N_{ep}) news reduce reinvestment rates on the market portfolio. In the ICAPM, declines in current wealth and in reinvestment rates both decrease lifetime utility. Therefore, \tilde{r}_m , N_{ep} , and N_{ir} are relevant sources of risk with all three having positive risk price (as long as $rra > 1$).

The $rra > 1$ condition for reinvestment risk to have a positive risk price is a consequence of two offsetting effects. An asset that performs well when reinvestment rates increase is desirable because it allows the investor to profit from the higher reinvestment rate, but undesirable because it increases the investor's exposure to reinvestment risk. When $rra > 1$, the later effect dominates and reinvestment risk has a positive risk price.

This ICAPM demonstrates why reinvestment risk is priced from the perspective of a long-term investor who holds the market portfolio. However, as in [Campbell and Vuolteenaho \(2004\)](#), the model is silent on why the investor would not change her portfolio allocation in response to changes in expected returns. Thus, the results in [Section 3](#), which rely on this model, are better understood as providing a microeconomic foundation for why a passive long-term investor would not perceive short-term dividend claims or long-term bonds as anomalous investment opportunities despite their relatively high risk premia.

[Section 4](#) imposes further economic assumptions to study the term structure based on dividend strip prices, $P^{(h)}$. In doing so, it considers an ICAPM in which the marginal investor responds optimally to changes in reinvestment rates. I defer the details to [Section 4](#).

2.2 From ICAPM to Risk Premia

From [Equation 12](#), we have:

$$\mathbb{E}[r_j - r_f] + \frac{1}{2}(\sigma_j^2 - \sigma_f^2) = rra \cdot Cov(\tilde{r}_{j,t} - \tilde{r}_{f,t}, \tilde{r}_{p,t}) + (rra - 1) \cdot Cov(\tilde{r}_{j,t} - \tilde{r}_{f,t}, N_{ep,t} + N_{ir,t})$$

or, in β -pricing notation:

$$\mathbb{E}[r_j - r_f] + \frac{1}{2}(\sigma_j^2 - \sigma_f^2) = \beta_{j,m} \cdot \lambda_m + \beta_{j,ep} \cdot \lambda_{ep} + \beta_{j,ir} \cdot \lambda_{ir} \quad (13)$$

where $\sigma_j^2 = Var(\tilde{r}_{j,t})$ is the respective return variance, $\mathbb{E}[r_j - r_f] + \frac{1}{2}(\sigma_j^2 - \sigma_f^2) \approx \mathbb{E}[R_j - R_f]$ represents risk premia, $\lambda_m = \sigma_m^2 \cdot rra$ and $\lambda_{ep} = \lambda_{ir} = \sigma_m^2 \cdot (rra - 1)$ are average risk prices, and $\beta_{j,x} = Cov(r_{j,t} - r_{f,t}, x_t) / \sigma_m^2$ captures risk quantities.⁷

Under the ICAPM, the risk premium of any given asset depends on compensation for exposing investors to market risk ($\beta_{j,m} \cdot \lambda_m$) and reinvestment risk, with the later operating through equity premium risk ($\beta_{j,ep} \cdot \lambda_{ep}$) and interest rate risk ($\beta_{j,ir} \cdot \lambda_{ir}$).

The focus of this paper is on the dividend and bond term structures. In the next section, I first examine the risk term structures of dividend claims and bond portfolios by estimating the term structure of β s relative to all three ICAPM risk factors and then calibrate rra to study the dividend and bond term structures through equation 13.

3 ICAPM Term Structures: Passive Investor

In this section, I study the dividend and bond term structures from the perspective of a passive long-term investor such as the one featured in the ICAPM of the previous section. The model imposes theoretical constraints on risk factors and risk prices, but is not fully structural in the sense that the marginal investor is not allowed to respond to shifts in reinvestment rates. Section 4 provides results using a fully structural ICAPM. The main advantage of relying on a passive investor is that this allows me to study the dividend term structure without imposing strong economic/econometric assumptions to derive empirical results. The key disadvantage of imposing less structure is that I cannot derive dividend strip prices using this framework and rely instead on the present value approach (described in Section 1) to study the dividend term structure. Subsection 3.6 complements

⁷When reporting average risk premium for any given asset j , I use the expression $\hat{\mathbb{E}}[r_j - r_f] + \frac{1}{2}(\hat{\sigma}_j^2 - \hat{\sigma}_f^2)$ as opposed to $\hat{\mathbb{E}}[R_j - R_f]$ (even though I tend to use the later notation for simplicity). Since β s are estimated using log returns, I consider the risk premia based on log returns to be more directly linked to the ICAPM implied risk premia (right hand side of equation 13). However, the two risk premia expressions are identical in the ICAPM (because of the conditional Normality assumption) and tend to be relatively close to each other in the data.

the present value approach by studying the ICAPM pricing implications to a short-term dividend strategy based on dividend futures, which are directly traded in financial markets.

3.1 Empirical Design

This subsection details the empirical design used for the results in Section 3. It explains how I estimate the risk exposures of dividends and bonds and briefly describes the sample construction and variables used in my empirical analysis. In the interest of clarity, I defer to Appendix D.1 most of the discussion on the results from various robustness tests. In short, the term structure results obtained using my primary empirical approach are robust to extensive variation to the empirical design (I report key summary results from 345 specifications).

a) Estimating Dividend and Bond Risk Exposures

To properly measure risk exposures of dividend claims, it is important to reasonably capture shocks to the expectation of dividend growth (Δd), equity excess returns ($r_e - r_f$) and interest rates (r_f). I also need unexpected returns for bond portfolios (and other assets). I take the simplest approach of using annual predictive regressions to capture all necessary shocks:⁸

$$\Delta d_{t+h} = b_g^{(h)'} z_t + \epsilon_{g,t+h}^{(h)} \quad \forall h = 1, 2, \dots, 10 \quad (14a)$$

$$r_{e,t+h} - r_{f,t+h} = b_{ep}^{(h)'} z_t + \epsilon_{ep,t+h}^{(h)} \quad \forall h = 1, 2, \dots, 10 \quad (14b)$$

$$r_{f,t+h} = b_{ir}^{(h)'} z_t + \epsilon_{ir,t+h}^{(h)} \quad \forall h = 1, 2, \dots, 10 \quad (14c)$$

$$r_{j,t+1} - r_{f,t+1} = b_j' z_t + \epsilon_{j,t+1} \quad (14d)$$

where z_t is a $k \times 1$ vector of predictive variables (with a constant as its first element) and r_j is the return on portfolio j .

It is possible to use this system to write all ICAPM risk factors and dividend PV unexpected returns as scaled changes in the state variables: $A_1' z_t - A_0' z_{t-1}$, with A_0 and A_1 representing (vector) functions of the parameters in the system (all expressions are derived in Appendix C.1). This result

⁸The system in 14a to 14d implicitly assumes $b_g^{(h)} = b_{ep}^{(h)} = b_{ir}^{(h)} = 0$ for all h higher than ten. This has no direct effect on the relative comparison between the first ten year dividend claims explored in my empirical analysis, but it can affect the risk factors used. This restriction imposes, for example, that information known as of December 2000 does not help predicting excess returns from December of 2010 to December of 2011. This is a reasonable approach if real dividend growth, excess returns, and real interest rates are stationary since the predictive coefficients of a stationary time series have to converge to zero as h increases. In Appendix D.1, I find consistent results after exploring two alternative approaches: (i) allow for predictability up to 15 years and (ii) use a vector autoregressive system after 10 years to account for predictability even at very long horizons.

suggests that all covariances (and hence β s) of dividend claims relative to risk factors can be recovered using the parameters in equations 14a, 14b, and 14c coupled with the variance-covariance matrix $VCov([z_t \ z_{t-1}])$. This is in fact the case, and Appendix C.1 also details all expressions linking covariances to the relevant parameters. For the risk exposures of bond portfolios (and other assets), we have $Cov(\tilde{r}_{j,t} - \tilde{r}_{f,t}, N_{ep,t}) = Cov(\epsilon_{j,t}, N_{ep,t})$, with analogous expressions for other risk factors. Therefore, all relevant risk exposures can be recovered from the system in 14a to 14d.

I estimate the system in two steps. The first step uses Generalized Method of Moments (GMM) with Ordinary Least Squares (OLS) orthogonality conditions while imposing consistency of unconditional averages across equations. The second step adjusts the GMM estimates to account for the small sample bias in predictive regressions with persistent predictors (see [Stambaugh \(1999\)](#) for the bias and [Amihud and Hurvich \(2004\)](#) for the correction).

For flow variables, such as returns and dividend growth, I use monthly observations of annual flows to estimate the system, which means that my observations overlap for eleven months.⁹ The use of GMM allows me to estimate asymptotic standard errors accounting for the overlapping observations as well as for the cross dependence among equations. Standard errors of dividend PV risk exposures and other parameters derived from the system are obtained through the delta method. Appendix C provides a more detailed explanation about the system implications, GMM estimation, bias correction, and standard errors.

When reporting results for dividend claims, I calculate returns with average maturity of h years to be more comparable to the dividend strips reported in the literature. For instance, my reported 1-year dividend claim is an average of $\tilde{r}_{pv,t}^{(1)}$ and $\tilde{r}_{pv,t}^{(2)}$ so that annual returns have an average duration of one year.

My approach of relying only on predictive regressions to derive implications for the risk term structure of dividend claims is analogous to the “local projection method” for impulse response functions advocated by [Jordà \(2005\)](#). The main advantage relative to the standard approach of using a Vector Autoregressive System (VAR) is that the local projection method is much more robust to misspecification. Another important advantage in the context of my analysis is that predictive regressions allow me to directly match both short- and long-run dynamics of returns and dividend growth, while

⁹The rationale for this approach is based on two considerations. First, I estimate predictive regressions up to ten years, and thus annual variables reduce the number of predictive regressions from 120 to 10 while keeping monthly observations to retain part of the power of monthly frequency estimation. Second, annual dividend growth does not suffer from the seasonality issues that are important at monthly or quarterly frequencies.

the VAR captures short-term dynamics and infer the long-term ones based on the structure assumed. Since the goal is to infer term structure patterns, properly capturing the differences between short- and long-term dynamics is crucial. The disadvantage of using predictive regressions is that this approach limits the dividend maturities I am able to study.

Section 5 explores the dividend term structure within a fully structural ICAPM that relies on a VAR structure, so the results there are complementary to the ones in this section. The key result that long-term dividend claims have low discount rates because they hedge reinvestment risk is valid under both analyses. However, Section 5 considers the SDF of an investor who strategically responds to shifts in reinvestment rates and so there are important economic (and quantitative) differences between the two analyses, which I detail in Section 5.

b) Sample Construction

Here, I provide only a brief description of the sample period and variables used in my empirical analysis, but a more detailed description of the measurement and data sources for all variables is provided in Appendix B.

My main sample goes from 1/1952 to 12/2016. The starting date is restricted by the availability of the bond portfolios used to study the bond term structure (starting in 1952 also makes my results directly comparable to papers in the bond term structure literature).¹⁰ Nevertheless, my robustness analysis also reports qualitatively similar results when dividend PV returns and risk factors are obtained based on alternative sample periods, including a sample that starts in 1/1928.

For the bond term structure, I use returns on six CRSP portfolios containing bonds with maturities up to 1, 2, 3, 4, 5, and 10 years. The same portfolios are used by [Binsbergen and Koijen \(2017\)](#) to study the bond term structure.

For the dividend term structure (and risk factors), I estimate the system in equations 14a to 14c. The estimation requires measuring interest rates (r_f), market returns ($r_m = r_e$), and dividend growth (Δd). The annual interest rate over a given year is equal to the one year Treasury log yield as of the beginning of the period. Market returns and dividend growth are based on a value-weighted portfolio containing all common stocks available in the CRSP dataset and their measurement accounts for delistings. Dividend growth and all returns are deflated using the CPI index.

¹⁰The availability of bond portfolios coincides with the Fed-Treasury Accord of 1951 that restored independence to the Fed, affecting monetary policy and, consequently, dramatically changing the time-series behavior of nominal interest rates.

The econometric system also relies on seven predictive variables embedded into z_t (all measured in natural log units): dividend growth (Δd), dividend yield (dy), equity payout yield ($epoy$), one year Treasury yield (ty), term spread (TS), credit spread (CS) and value spread (VS).¹¹ All of these variables have been explored in the literature as important predictors of dividend growth, equity returns, and/or interest rates.¹² My robustness analysis (Appendix D.1) excludes predictive variables one at a time to demonstrate the robustness of the term structure patterns uncovered to the predictive variables used.

3.2 Correlations among Risk Factors, Predictive Variables, and Returns

This subsection focuses on Table 1, which reports correlations among shocks to risk factors (r_m , N_{ep} , and N_{ir}), predictive variables (Δd , dy , $epoy$, ty , TS , CS , and VS), and returns on bond portfolios and dividend claims ($r_b^{(h)}$ and $r_{pv}^{(h)}$). The correlation matrix is estimated based on my main sample (1952-2016) and is split into three panels to simplify exposition. These correlations are important because they are intrinsically connected to the risk term structures of dividends and bonds. Shocks to risk factors and dividend PV returns are based on equations A7a to A8e, with empirical estimates of the link between dividend PV returns and z_t reported in Table A.1.¹³

Panel A reports correlations between shocks to risk factors and predictive variables and its most important observation is that the correlation between equity premium news and shocks to the one year Treasury yield is -0.58. This negative correlation between the equity premium and nominal interest

¹¹Dividend yield is the log of aggregate dividends over a normalized index price. Equity payout yield is the log of (one plus) aggregate net equity payout over market equity and only relies on CRSP data. The term spread is the difference between the 10-year and 1-year log Treasury yields. The credit spread is the difference between Moody’s corporate BAA and AAA log yields. The value spread is the difference between the log book-to-market ratios of the value and growth portfolios in Fama and French (1993, 1996) HML factor. The value spread monthly observations are adjusted to account for within year movements in market equity (see the Internet Appendix of Campbell and Vuolteenaho (2004)). My use of the HML portfolios (as opposed to focusing on small stocks as in Campbell and Vuolteenaho (2004)) assures the value spread behavior is not dominated by small stocks, but results are robust to using the small value spread (see Appendix D.1).

¹²All predictive variables are designed to capture either cash flow growth or yields in financial markets. Chen and Zhao (2009) use lagged dividend growth as a dividend growth predictor. Several papers use the dividend yield as a predictor for both dividend growth and stock returns, with theoretical justification provided by the valuation identity (Campbell and Shiller (1989)). Modifications to this valuation identity can be used to motivate many additional valuation ratios as predictors for all three relevant variables. I follow Boudoukh et al. (2007) and Larrain and Yogo (2008) and use the equity payout yield since it does not suffer from structural breaks that dividend yield does (see Lettau and Nieuwerburgh (2007) and Boudoukh et al. (2007) for evidence). The treasury yield (Fama and Schwert (1977); Fama (1981)), the term spread (Campbell (1987); Fama and French (1989)), and the credit spread (Keim and Stambaugh (1986)) are classical interest rate and equity return predictors. Finally, Campbell and Vuolteenaho (2004), Campbell, Polk, and Vuolteenaho (2009), and Campbell et al. (2017) rely on the value spread as an important predictor of stock returns.

¹³Shocks to predictive variables, z_t , are not included in my econometric system (are not required in my methodology to construct risk factors or dividend PV returns). I report statistics related to z_t shocks in Tables 1 and A.1. For that, I use the residuals from a regression of z_t onto z_{t-1} .

rates is well documented in the literature (Fama and Schwert (1977); Campbell (1987); Ferson (1989); Shanken (1990); Brennan (1997)) and is ultimately responsible for the fact that bonds are exposed to reinvestment risk despite dividend claims hedging such risk. Different theoretical mechanisms have been proposed for the negative correlation between interest rates and the equity premium and they typically operate through expected inflation (Fama (1981); Geske and Roll (1983); Stulz (1986)), but the empirical fact is more important than its theoretical foundation for the purpose of this paper.

Panel B reports correlations between returns on bond portfolios and dividend PVs. The key observation is that the term-structure dimension is an important driver of the return correlation structure of dividend claims and bond portfolios. For instance, the 5- and 10-year dividend PV returns have a 55% correlation, but the 1- and 10-year dividend PV returns have a correlation of only 22%. Similarly, the portfolio composed by bonds with maturity up to 10 years has a 94% return correlation with the one composed by up to 5-year bonds and a return correlation of only 16% with the one composed by up to 1-year bonds.

Finally, Panel C reports correlations between returns and shocks to predictive variables and risk factors. The main result is that the correlations between returns and risk factors are largely responsible for the beta patterns described in the introduction and further explored in the next subsection. In particular, returns on longer-term dividend claims have higher correlation with market returns, but lower correlation with equity premium and interest rate news. Similarly, returns on longer-term bonds have higher correlation with market returns and more negative correlation with interest rate news. The one exception is that the correlations between bond portfolio returns and equity premium news follow a u-shape, which means that bond volatilities are crucial for longer-term bonds to have higher exposure to equity-premium news.

This subsection is useful in understanding how bond portfolios and dividend strips correlate with risk factors. However, what ultimately determines risk premium is the exposure to market and reinvestment risk (i.e., β_s) and such exposures also depend on volatilities. The volatilities of risk factors, $\sigma(r_m) = 15\%$, $\sigma(N_{ep}) = 20\%$, and $\sigma(N_{ir}) = 5\%$, indicate that reinvestment risk is an important source of risk and operate mostly through equity premium shocks. Next subsection studies risk exposures more directly by focusing on β_s .

Table 1
Correlations: Returns and Shocks to Predictive Variables and Risk Factors

The table reports correlations among shocks to risk factors, predictive variables, and returns on bonds and dividend claims. r_m represents equity market returns and N_{ep} and N_{ir} capture equity premium and interest rate news. The predictive variables are the log dividend growth (Δd), dividend yield (dy), equity payout yield ($epoy$), one year Treasury yield (ty), term spread (TS), credit spread (CS), and value spread (VS). $r_{pv}^{(h)}$ ($r_b^{(h)}$) represents the h -year dividend PV (bond portfolio) returns. The correlation matrix is estimated based on my main sample (1952-2016) and is split into three panels to simplify exposition. Panel A reports correlations between shocks to predictive variables and risk factors while Panel B focuses on correlations between returns and Panel C on correlations between returns and shocks to predictive variables and risk factors. Details about the construction of risk factors, predictive variables, and returns can be found in subsection 3.1.

PANEL A: Shocks to Predictive Variables and Risk Factors											PANEL B: Bond and Dividend Strip Returns						
	r_m	N_{ep}	N_{ir}	Δd	dy	$epoy$	ty	TS	CS	VS		$r_b^{(1)}$	$r_b^{(5)}$	$r_b^{(10)}$	$r_{pv}^{(1)}$	$r_{pv}^{(5)}$	$r_{pv}^{(10)}$
r_m	1										$r_b^{(1)}$	1					
N_{ep}	-0.44	1									$r_b^{(5)}$	0.22	1				
N_{ir}	-0.34	-0.07	1								$r_b^{(10)}$	0.16	0.97	1			
Δd	0.15	0.13	0.37	1							$r_{pv}^{(1)}$	0.02	0.05	0.00	1		
dy	-0.94	0.48	0.47	0.20	1						$r_{pv}^{(5)}$	-0.05	0.12	0.11	0.87	1	
$epoy$	-0.50	0.62	-0.04	0.23	0.57	1					$r_{pv}^{(10)}$	-0.20	0.02	0.08	0.22	0.55	1
ty	-0.12	-0.58	0.58	0.05	0.13	-0.06	1										
TS	-0.02	0.65	-0.22	-0.10	-0.02	0.13	-0.79	1									
CS	-0.48	0.59	-0.10	-0.07	0.45	0.36	-0.23	0.25	1								
VS	0.13	-0.26	-0.26	-0.11	-0.17	-0.04	-0.13	0.07	0.22	1							

PANEL C: Returns and Shocks to Predictive Variables and Risk Factors																	
	$r_b^{(1)}$	$r_b^{(2)}$	$r_b^{(3)}$	$r_b^{(4)}$	$r_b^{(5)}$	$r_b^{(10)}$	$r_{pv}^{(1)}$	$r_{pv}^{(2)}$	$r_{pv}^{(3)}$	$r_{pv}^{(4)}$	$r_{pv}^{(5)}$	$r_{pv}^{(6)}$	$r_{pv}^{(7)}$	$r_{pv}^{(8)}$	$r_{pv}^{(9)}$	$r_{pv}^{(10)}$	
r_m	-0.19	0.06	0.09	0.10	0.09	0.15	0.33	0.47	0.52	0.54	0.64	0.75	0.81	0.89	0.95	0.98	
N_{ep}	0.28	0.54	0.51	0.49	0.49	0.40	0.41	0.30	0.29	0.27	0.17	0.05	-0.08	-0.27	-0.45	-0.54	
N_{ir}	-0.04	-0.54	-0.61	-0.63	-0.64	-0.67	0.21	0.16	0.08	0.04	-0.06	-0.21	-0.30	-0.31	-0.29	-0.31	
Δd	-0.06	-0.07	-0.08	-0.06	-0.06	-0.02	0.57	0.57	0.51	0.52	0.55	0.43	0.31	0.24	0.16	0.10	
dy	0.16	-0.08	-0.12	-0.13	-0.11	-0.15	-0.13	-0.26	-0.33	-0.35	-0.44	-0.59	-0.69	-0.79	-0.88	-0.93	
$epoy$	0.13	0.05	0.04	0.05	0.05	0.01	0.32	0.20	0.14	0.15	0.09	0.01	-0.14	-0.37	-0.52	-0.56	
ty	-0.20	-0.94	-0.96	-0.94	-0.93	-0.89	-0.20	-0.17	-0.22	-0.24	-0.23	-0.20	-0.19	-0.14	-0.05	-0.02	
TS	0.24	0.72	0.66	0.60	0.57	0.45	0.39	0.33	0.34	0.32	0.28	0.20	0.13	0.02	-0.09	-0.13	
CS	0.22	0.22	0.21	0.20	0.23	0.19	-0.09	-0.19	-0.12	-0.05	-0.09	-0.14	-0.10	-0.13	-0.28	-0.45	
VS	-0.03	0.11	0.12	0.11	0.11	0.13	-0.10	-0.12	-0.07	0.04	0.10	0.11	0.17	0.22	0.21	0.16	

3.3 The Dividend and Bond Risk Term Structures

This subsection explores how β s vary with dividend and bond maturity. The most important message is that exposure to equity premium shocks can potentially generate opposite risk premia term structures for bonds and dividends while market and interest rate exposures cannot.

Figure 2 displays the risk term structures of dividends and bonds based on the three ICAPM risk factors considered (market returns, equity premium news, and interest rate news). All β s are in market β units (i.e., covariance normalized by market variance), and hence are directly comparable. I perform statistical tests for $H_0 : \beta^{(h)} = \beta_{Base}$ with filled black dots indicating significance at 10% level, blue crosses at 5% level, and red stars at 1% level. The base assets are the respective low risk-premium assets (10-year dividend claim and the 1-year bond portfolio).

Market β s increase in dividend and bond maturity (Figures 2a and 2b). In particular, the 10-year dividend claim has more than two times the market β of the 1-year dividend claim, with the increasing market β pattern being statistically significant for several dividend claims and indicating that the CAPM generates a counterfactual upward sloping dividend term structure. In the case of bonds, market β s increase in maturity, but not by much ($\beta^{(10)} - \beta^{(1)} = 0.051$). Considering that the typical risk price in the CAPM is below 10%, the CAPM generates a bond term structure that is too flat relative to what is observed in the data (spread between 10- and 1-year bond portfolios is below 0.5% under the CAPM while in the data it is close to 1.5%).

In stark contrast, the equity premium β s decrease in dividend maturity despite increasing in bond maturity, with long-term dividend claims hedging equity premium shocks and long-term bonds being exposed to such shocks (Figures 2c and 2d). The risk differences across dividend maturities are sizable, with a β difference between the 1- and 10-year dividend claims that is close to 1.3 market β units (and statistically significant at 5% level). The differences are also large for bonds, with a β spread between the 1- and 10-year bond portfolios higher than 0.15 market β units, much larger than the respective market β spread (0.051).

Finally, interest rate β s tend to decrease in dividend and bond maturity. However, the differences are smaller than the ones observed for equity premium β s and are statistically insignificant.

Overall, the largest risk exposure differences are observed when considering equity premium shocks. Moreover, the term structure of equity premium β s goes in the right direction in explaining both the *downward sloping* dividend term structure and the *upward sloping* bond term structure.

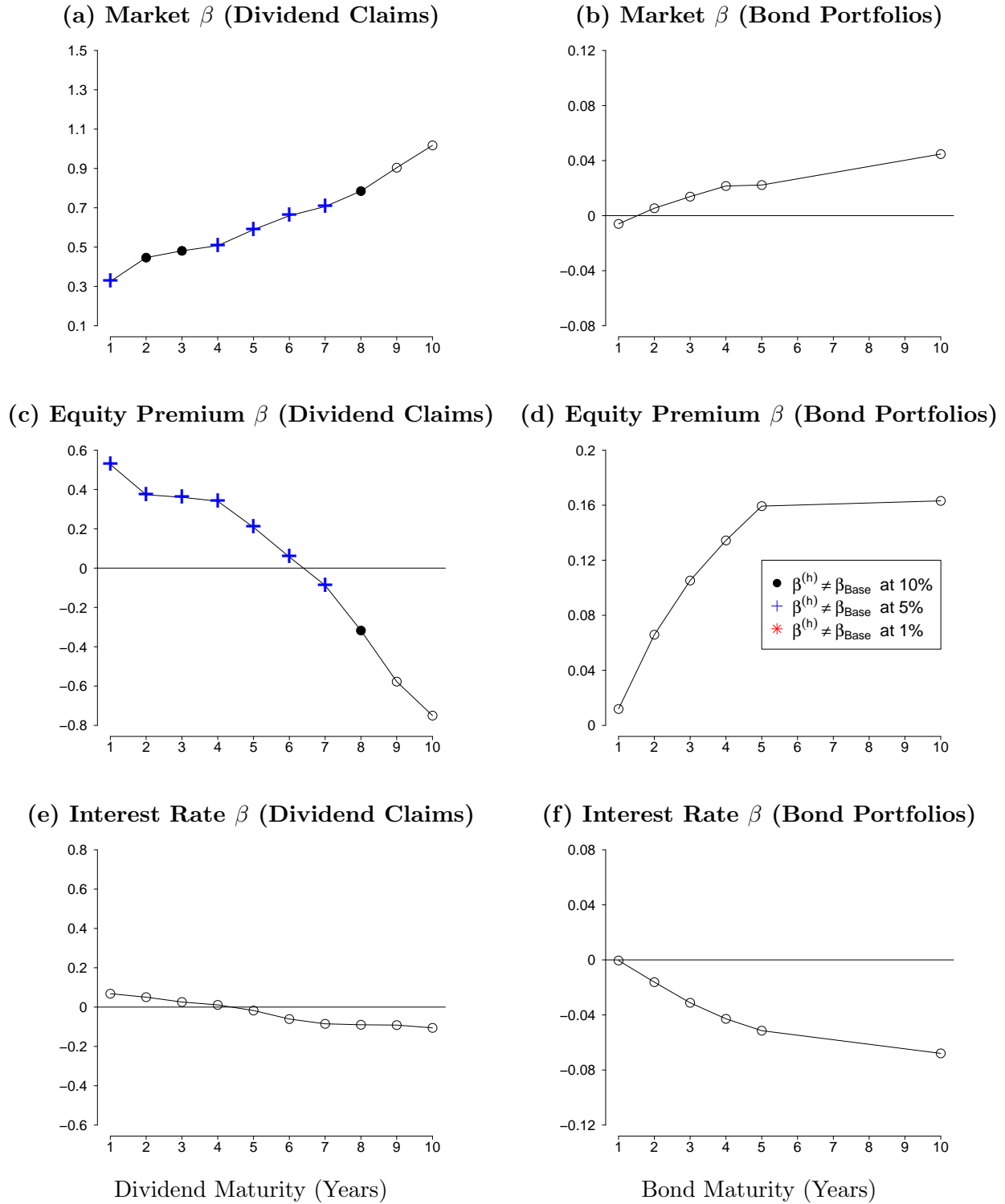


Figure 2
The Dividend and Bond Risk Term Structures

The graphs report β s relative to unexpected market returns (\tilde{r}_m), equity premium news (N_{ep}), and interest rate news (N_{ir}) based on the system in equations 14a to 14d estimated over my main sample period (1952-2016). All β s are transformed to market β units (i.e., covariance normalized by market variance) and the ranges in the y axes are the same for different β s within the same asset class so that β s are directly comparable. I perform statistical tests for $H_0 : \beta^{(h)} = \beta_{Base}$ with the legend detailing the significance levels. The base assets are the respective low risk-premium assets (10-year dividend claim and 1-year bond portfolio). Empirical details can be found in subsection 3.1.

It is natural to wonder what drives the equity premium β pattern. Note that if dividend growth and interest rates were unpredictable, we would have $\ln(PV_t^{(h)}) = \ln(D_t) - h \cdot \bar{r}_{e,t \rightarrow t+h}$, with the last term being the average expected equity return from t to $t+h$. In this case, a permanent *decrease* in the equity premium would coincide with an *increase* in dividend present values due to their inverse relation with expected returns, producing a hedge against reinvestment rate declines. The multiplication by h amplifies this hedge effect for longer-term dividends, which generates the downward sloping term structure of equity premium β s observed in the previous section, with long-term dividend claims having large negative equity premium β s.

In reality, there are other important effects. For instance, decreases in expected returns are not permanent. Since expected returns are stationary, the decrease in $\bar{r}_{e,t \rightarrow t+h}$ is stronger for lower h (a mean reversion effect), which counteracts the duration mechanism of the previous paragraph. Moreover, expected dividend growth and interest rates also vary and correlate with the equity premium.

Figure 3a demonstrates that the duration mechanism is the dominant effect empirically. Each line represents the term structure of equity premium β s when only one source of heterogeneity in dividend strip returns (in equation 9) is considered. Specifically, the blue dotted line keeps only dividend growth news ($N_g^{(h)}$), the red solid line uses only equity premium news ($N_{ep}^{(h)}$), and the black dashed line considers only interest rate news ($N_{ir}^{(h)}$).

The downward sloping term structure of equity premium β s is entirely driven by the duration effect from equity premium news. Interest rate news have a neutral effect and dividend growth news create an *increasing* pattern in equity premium β s, effectively dampening the duration mechanism discussed.

It is interesting to see from Figure 3b that the market risk term structure is also mostly driven by equity premium shocks. This indicates that variation in discount rates is fundamentally important in understanding risk exposures in financial markets.

In summary, this section demonstrates that reinvestment risk (mostly through equity premium shocks) is a relevant source of risk and can potentially generate opposite term structure slopes for dividends and bonds. However, whether this is the case depends on the risk prices associated with the three risk factors considered. Next section relies on the ICAPM of Section 2 to calibrate risk prices and explores the implications of reinvestment risk to the dividend and bond term structures.

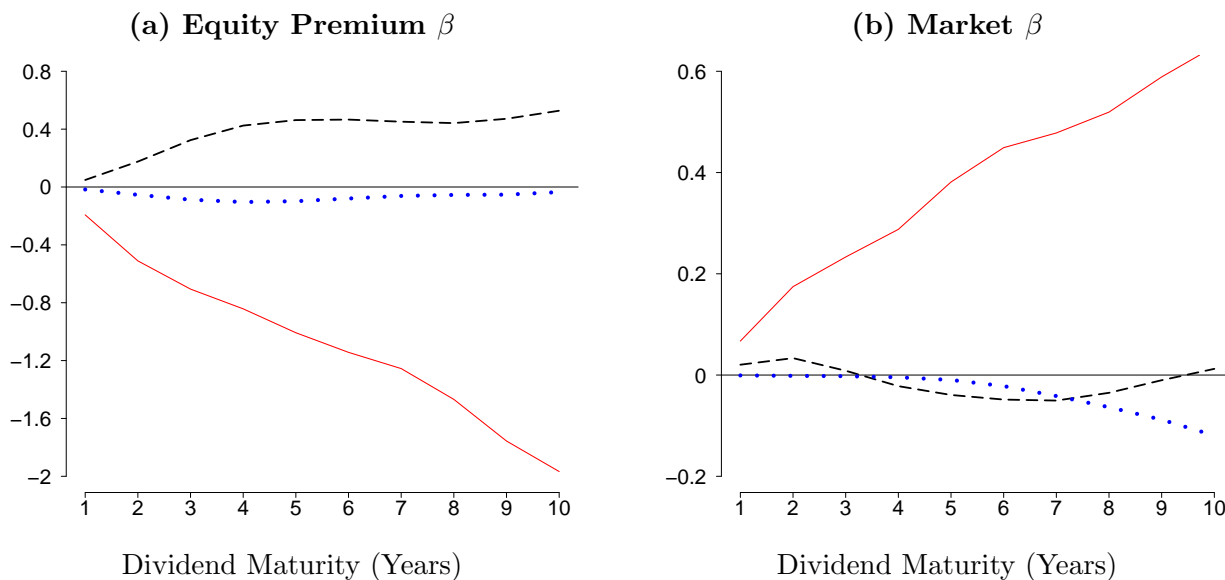


Figure 3

Decomposing the Dividend Term Structures of Equity Premium and Market β s

The graphs report components of equity premium and market β s based on the system in equations 14a to 14c estimated over my main sample period (1952-2016). All β s are in market β units (i.e., covariance normalized by market variance). Each line keeps only one source of heterogeneity in dividend PV returns (equation 9). Specifically, the blue dotted line keeps only dividend growth news ($N_g^{(h)}$), the red solid line uses only equity premium news ($N_{ep}^{(h)}$), and the black dashed line maintains only interest rate news ($N_{ir}^{(h)}$). Empirical details can be found in subsection 3.1.

3.4 The Risk Premia Term Structures of Dividends and Bonds

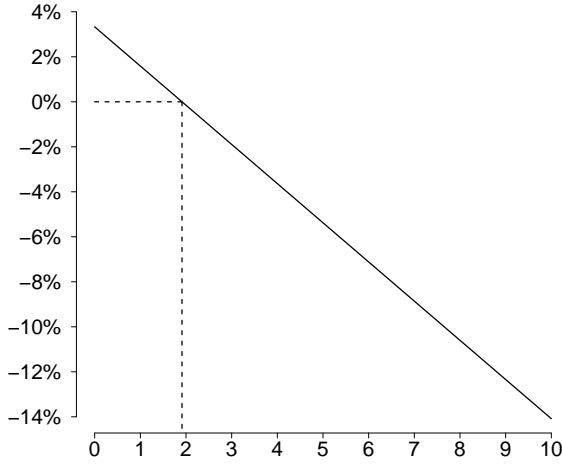
In this subsection, I use the β -pricing equation 13 to study the dividend and bond term structures relying only on a reasonable level of relative risk aversion. The main finding is that the ICAPM simultaneously generates a downward sloping dividend term structure and an upward sloping bond term structure due to reinvestment risk being priced.

a) General Relative Risk Aversion

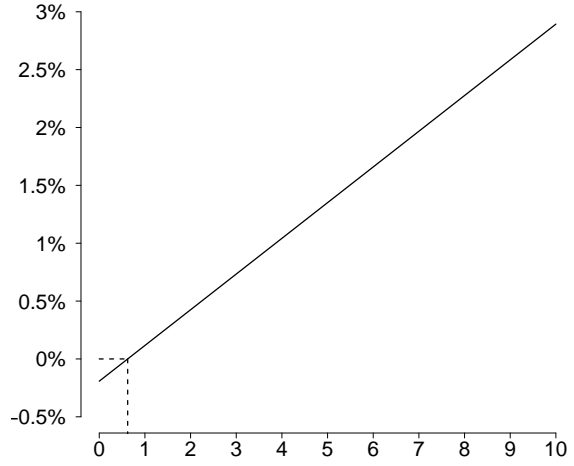
Given a level of relative risk aversion (rra), we can use the β -pricing equation 13 to obtain dividend and bond term structures within the ICAPM. However, there is a more general way to understand the ICAPM implications to the dividend and bond term structures. Using equation 13 and the dividend risk premium approximation, $\mathbb{E} \left[r_{pv}^{(h)} - r_f \right] + \frac{1}{2} \left(\sigma_{pv,h}^2 - \sigma_f^2 \right) \approx \mathbb{E} \left[R_{pv}^{(h)} - R_f \right]$, we can write the dividend term premium as a linear function of the relative risk aversion (and similarly for the bond term premium):

(a) Dividend Term Premium

$$\mathbb{E} \left[R_{pv}^{(10)} - R_{pv}^{(1)} \right] = 3.3 - 1.7 \cdot rra$$

Relative Risk Aversion (rra)**(b) Bond Term Premium**

$$\mathbb{E} \left[R_b^{(10)} - R_b^{(1)} \right] = -0.2 + 0.3 \cdot rra$$

Relative Risk Aversion (rra)**Figure 4****Dividend and Bond Term Premia as a Functions of Relative Risk Aversion**

The graphs report the dividend and bond term premia ($\mathbb{E} [R^{(10)} - R^{(1)}]$) as linear functions of relative risk aversion (rra) based on the term premium β -pricing equation 15. The intercepts and slopes are based on the β s estimated over my main sample period (1952-2016) and studied in subsection 3.3. Empirical details can be found in subsection 3.1.

$$\mathbb{E} \left[R_{pv}^{(10)} - R_{pv}^{(1)} \right] = \underbrace{- (\Delta\beta_{ep} + \Delta\beta_{ir}) \cdot \sigma_m^2}_{\text{Intercept}} + \underbrace{(\Delta\beta_m + \Delta\beta_{ep} + \Delta\beta_{ir}) \cdot \sigma_m^2}_{\text{Slope}} \cdot rra \quad (15)$$

where $\Delta\beta_m$, $\Delta\beta_{ep}$, and $\Delta\beta_{ir}$ represent the respective β differences between the 10- and 1-year dividend claims.

Figure 4 displays this linear function for both dividend claims and bond portfolios using parameter estimates for the intercepts and slopes based on the β s estimated in subsection 3.3.

Two key messages originate from Figure 4. First, as relative risk aversion increases, the slope of the dividend term structure (dividend term premium) decreases and the slope of the bond term structure (bond term premium) increases. Second, even moderately high relative risk aversion ($rra > 5$) implies a large economic difference between short- and long-term dividend risk premia ($\mathbb{E} [R_{pv}^{(10)} - R_{pv}^{(1)}] > 5\%$) and a reasonable bond term premium ($\mathbb{E} [R_b^{(10)} - R_b^{(1)}] > 1\%$).

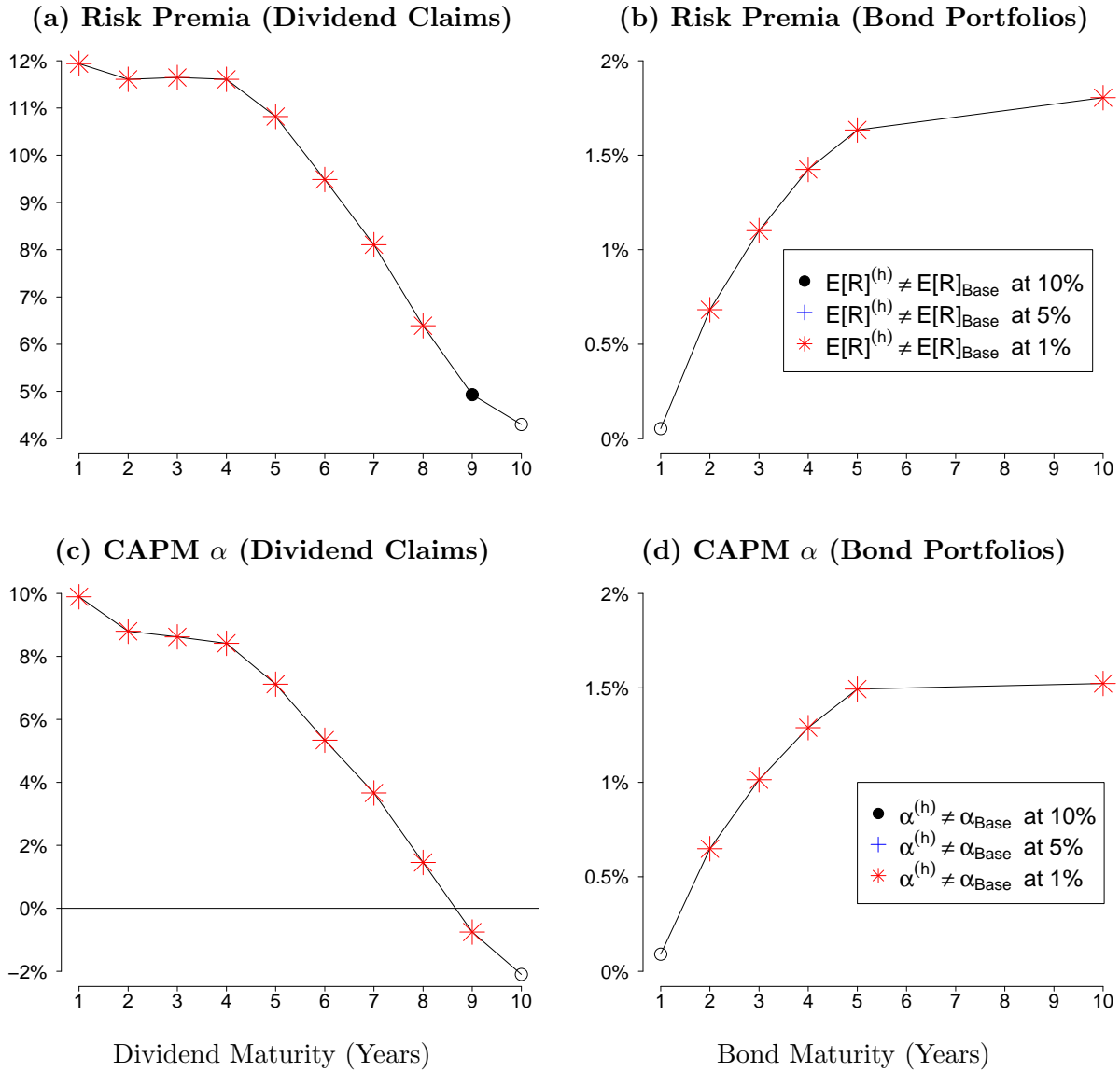


Figure 5

The Dividend and Bond Term Structures of Risk Premia and CAPM α s

The graphs report the (ICAPM-implied) dividend and bond term structures of risk premia and CAPM pricing errors (α s). All risk premia are based on equation 13 with risk aversion calibrated to $r_{ra} = 6.3$ and β s estimated from the system in equations 14a to 14d (studied in subsection 3.3). Empirical details can be found in subsection 3.1.

b) Calibrated Relative Risk Aversion

While the results in Figure 4 are general enough to not depend on a specific level of rra , we can get a full picture of the dividend and bond term structures by calibrating rra to a reasonable level. I use the value $rra = 6.3$, which seems arbitrary, but is actually the outcome of a GMM estimation designed to fit (as well as possible) the equity premium and the bond term structure. All details of the GMM estimation as well as an analysis of the ICAPM fit to equity and bond portfolios is provided in the next subsection. It is worth mentioning that 6.3 is a reasonable value because (i) it is comparable to the 7.2 value estimated in [Campbell et al. \(2017\)](#) (using an ICAPM related to the one in this paper) and (ii) is below the $rra = 10$ upper bound [Mehra and Prescott \(1985\)](#) argue to be reasonable for asset pricing models.

Figures 5a and 5b display the dividend and bond term structures obtained with $rra = 6.3$. The dividend term structure is strongly downward sloping, with the spread between the 1- and 10-year dividend claims being higher than 7%. In contrast, the bond term structure is upward sloping, with a 1.75% spread between the risk premiums of the shortest- and longest-term bond portfolios (close to the 1.4% average return differential between the two portfolios).

Despite the strong economic significance of the dividend and bond term structure slopes, there is substantial uncertainty in estimating risk factors and dividend PV returns. Consequently, statistical differences between risk premia are only confirmed when taking all shocks as given (analogous to the approach used in [Campbell and Vuolteenaho \(2004\)](#)). I start by considering all uncertainty associated with estimating risk-premia and find that standard errors are large even relative to the economically sizable premia, and thus most risk premia differences are not statistically significant. To understand the economic implication of this result, I then recalculate standard errors taking as given the estimates for risk factors and unexpected returns on dividend claims and bond portfolios. I perform statistical tests for $H_0 : \mathbb{E} [R^{(10)}] = \mathbb{E} [R_{base}]$ with the base assets being the respective low risk-premium assets (10-year dividend claim and 1-year bond portfolio). The results (in Figures 5a and 5b) demonstrate that there are strong statistical differences between the risk premia of short- and long-term dividend claims and bond portfolios after we take shocks as given.

Figures 5c and 5d show similar results for CAPM pricing errors (αs). For dividend claims, market βs go in the wrong direction in explaining the dividend term structure, and thus the term structure of CAPM αs is even more pronounced than the one based on risk premia. In the case of bonds, market

β s go in the right direction, but their slope is small so that there is still substantial differences in CAPM α s across the shortest- and longest-term bond portfolios (close to 1.5% α differential).

Many empirical decisions can affect the results in subsections 3.3 and 3.4. As such, Appendix D.1 provides an extensive robustness analysis to demonstrate that the key results remain the same under alternative empirical choices (I report key summary results from 345 specifications). The overall robust conclusion is that an ICAPM in which reinvestment risk is priced produces both a downward sloping dividend term structure and an upward sloping bond term structure.

3.5 Estimating Relative Risk Aversion

The results in the previous subsection are striking. They imply that a standard ICAPM simultaneously produces a *downward* sloping dividend term structure and an *upward* sloping bond term structure when risk aversion is calibrated to $rra = 6.3$. The rra calibration process is actually based on a GMM estimation, which I detail in this subsection. I also provide a brief analysis of the ICAPM fit to equity and bond portfolios. The main result is that the ICAPM not only matches the dividend and bond term structures, but also reasonably captures the equity, value, and credit premiums.

a) Estimation Procedure

Estimating ICAPM risk prices requires a set of testing assets. I use excess returns (relative to the risk-free asset) on equity and bond portfolios: the market portfolio, the three (size controlled) book-to-market sorted portfolios in Fama and French (1993) (i.e., HML portfolios), the Barclays' mid-term and long-term AAA, AA, A, and BAA corporate bond portfolios and the six CRSP Treasury bond portfolios with maturities up to 1, 2, 3, 4, 5 and 10 years (the same ones used for the bond term structure in previous subsections). Data sources and further details are provided in Appendix B.

I estimate the ICAPM by augmenting the system in equations 14a to 14d with the β -pricing equation 13 for the testing assets. I also add moment equations for the covariances between testing assets and risk factors to assure that standard errors jointly account for estimation uncertainty from all parameters. After adding equation 13 to the GMM procedure, the system becomes over-identified (more moment conditions than parameters). I estimate parameters based on their respective (just identified) moment equations, except for the risk prices, which are estimated by solving $\hat{\lambda} = \underset{\lambda}{argmin} (\alpha' \alpha)$ with $\alpha_j = \hat{\mathbb{E}} [r_j - r_f] + \frac{1}{2} (\hat{\sigma}_j^2 - \hat{\sigma}_f^2) - \lambda' \cdot \hat{\beta}_j$. The entire estimation procedure is specified in terms

of a unified GMM and details are provided in Appendix C.¹⁴

Risk prices, $\lambda = [\lambda_m \ \lambda_{ep} \ \lambda_{ir}]$, are always estimated based on equation 13 and the ICAPM of Campbell (1993) implies $\lambda = \sigma_m^2 \cdot [rra \ (rra - 1) \ (rra - 1)]$, which leaves only one parameter (rra) to fully determine all ICAPM risk prices. I keep this as my main specification as it is used in all other sections. However, measurement error and misspecifications can potentially break the validity of the risk price restrictions. As such, I also estimate an “Unconstrained ICAPM” in which the restrictions on λ_{ir} and λ_{ep} are relaxed and only $\lambda_{ir} \geq 0$ and $\lambda_{ep} \geq 0$ are imposed (this model has three parameters: rra , λ_{ir} , and λ_{ep}).

b) Estimation Results

Table 2 reports relative risk aversion (rra) and risk prices for the market (λ_m), equity premium (λ_{ep}), and interest rate (λ_{ir}) risk factors estimated over my main sample period (1952-2016). Columns (1a) and (1b) estimate a CAPM while other columns report results from the two alternative ICAPM specifications. I focus my discussion on the ICAPM in columns (2a) and (2b), which imposes the risk price restrictions, $\lambda_m = \sigma_m^2 \cdot rra$ and $\lambda_{ep} = \lambda_{ir} = \sigma_m^2 \cdot (rra - 1)$. However, columns (3a) and (3b) also present results in which risk prices are not restricted (the Unconstrained ICAPM). In parenthesis and square brackets, I report t-statistics and p-values accounting for all estimation uncertainty while in curly brackets I report these statistics conditioning on observed shocks (as in Campbell and Vuolteenaho (2004)).

The table shows that the estimated rra for the ICAPM is 6.3 (the calibrated value) when I match the equity premium and the bond term structure and 8.0 when I also add corporate bond portfolios as well as HML portfolios. Therefore, the $rra = 6.3$ calibrated is conservative in the sense that term structure results would be even more pronounced with $rra = 8.0$ (see Figure 4).

The ICAPM-specific risk prices (λ_{ep} and λ_{ir}) are economically important, especially λ_{ep} . In particular, λ_{ep} is 12.2% (16.0%) when I match the equity premium and Treasury bond portfolios (all testing assets), which implies a 12.2% (16.0%) annual risk premium for an asset with $\beta_{ep} = 1$ (which

¹⁴Since the data on corporate bond portfolios start in 1973, the estimator is technically the “long GMM estimator” described in Lynch and Wachter (2013), which allows for samples of unequal length. It is identical to regular GMM, except that each moment equation uses all data available to its estimation and the standard errors are adjusted accordingly. This is not the most efficient estimator described in Lynch and Wachter (2013), but is the simplest one and it suffices for the purpose of this paper. Even though the testing assets are observed over samples of unequal length, I still use an identity weighting matrix. This approach partially compensates for the lower estimation uncertainty in bond moment conditions relative to stocks. In the robustness analysis (Appendix D.1), I demonstrate that dividend and bond term structure patterns are similar when using other weighting matrices to estimate rra .

Table 2
ICAPM Risk Prices

The table reports relative risk aversion (rra) and risk prices for the market (λ_m), equity premium (λ_{ep}), and interest rate (λ_{ir}) risk factors estimated over my main sample period (1952-2016). Columns (1a) and (1b) estimate a CAPM while other columns focus on the two alternative ICAPM specifications. The ICAPM in Columns (2a) and (2b) report results based on the model in Section 2, which imposes the risk price restrictions, $\lambda_m = \sigma_m^2 \cdot rra$ and $\lambda_{ep} = \lambda_{ir} = \sigma_m^2 \cdot (rra - 1)$. In contrast, columns (3a) and (3b) present results for the Unconstrained ICAPM, which only imposes $\lambda_m = \sigma_m^2 \cdot rra$, $\lambda_{ep} \geq 0$ and $\lambda_{ir} \geq 0$. The testing assets are excess returns (relative to the risk-free asset) on equity and bond portfolios: the market portfolio, the three (size controlled) book-to-market sorted portfolios in Fama and French (1993) (i.e., HML portfolios), the Barclays' mid-term and long-term AAA, AA, A, and BAA corporate bond portfolios, and the six CRSP Treasury bond portfolios with maturities up to 1, 2, 3, 4, 5, and 10 years. The (t-statistics) are provided below risk price estimates and [p-values] for Walt tests with null $\lambda_m = \lambda_{ep} = \lambda_{ir} = 0$ and $\lambda_{ep} = \lambda_{ir} = 0$ are reported at the bottom of the table. I also provide {t-statistics} and {p-values} for the same tests, but conditioning on observations of shocks and risk factors. Details about the construction of risk factors and model estimation can be found in subsections 3.1 and 3.5.

	Equity Premium + Treasury Bonds			All Testing Assets		
	CAPM (1a)	ICAPM (2a)	Unrest. ICAPM (3a)	CAPM (1b)	ICAPM (2b)	Unrest. ICAPM (3b)
rra	2.9 (2.76) {2.76}	6.3 (1.68) {2.70}	4.7 (1.58) {2.43}	3.7 (2.97) {2.97}	8.0 (1.40) {2.78}	8.8 (1.67) {3.58}
λ_m	6.7% (3.84) {3.85}	14.5% (1.58) {3.10}	10.7% (1.61) {2.41}	8.4% (4.49) {4.50}	18.3% (1.33) {3.39}	20.2% (1.63) {3.62}
λ_{ep}		12.2% (1.35) {2.59}	6.2% (1.05) {1.34}		16.0% (1.18) {2.94}	17.0% (1.74) {2.94}
λ_{ir}		12.2% (1.35) {2.59}	2.1% (0.09) {0.13}		16.0% (1.18) {2.94}	25.2% (0.70) {1.14}
$H_0 : \lambda_m = \lambda_{ep} = \lambda_{ir} = 0$	[0.0%] {0.0%}	[0.0%] {0.0%}	[0.0%] {0.0%}	[0.0%] {0.0%}	[0.0%] {0.0%}	[5.4%] {0.0%}
$H_0 : \lambda_{ep} = \lambda_{ir} = 0$		[17.8%] {0.9%}	[98.8%] {89.6%}		[23.9%] {0.3%}	[5.6%] {0.0%}

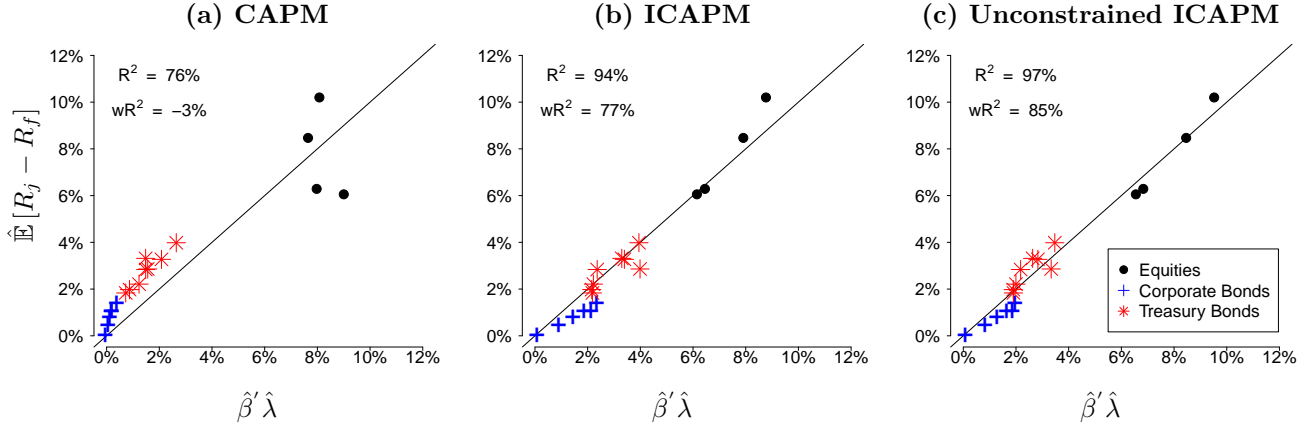


Figure 6
Predicted vs Realized Risk Premia for Testing Assets

The graphs display predicted (based on the right hand side of equation 13) and realized (based on the left hand side of equation 13) risk premia for the testing assets. The testing assets are excess returns on equity and bond portfolios: the market portfolio, the three (size controlled) book-to-market sorted portfolios in Fama and French (1993) (i.e., HML portfolios), the Barclays' mid-term and long-term AAA, AA, A, and BAA corporate bond portfolios and the six CRSP Treasury bond portfolios with maturities up to 1, 2, 3, 4, 5 and 10 years. $R^2 = 1 - \sum \alpha_j^2 / \sum (\hat{\mathbb{E}}[R_j - R_f] - \bar{\mathbb{E}}[R_j - R_f])^2$ captures the fraction of average return variability explained by the respective model and $wR^2 = 1 - \sum (\alpha_j - \bar{\alpha}_{class})^2 / \sum (\hat{\mathbb{E}}[R_j - R_f] - \bar{\mathbb{E}}_{class}[R_j - R_f])^2$ captures the same quantity, but for average return variability within each asset class (equities, corporate bonds, and Treasury bonds). All quantities are estimated from their respective available observations within my main sample period (1952-2016). Empirical details can be found in subsections 3.1 and 3.5.

is less than the β_{ep} difference between the 1- and 10-year dividend claims).

Nevertheless, ICAPM-specific risk prices only display strong statistical significance when we take shocks as given (curly brackets). In this case, all prices are strongly significant and the hypothesis that the ICAPM does not add to the CAPM ($\lambda_{ep} = \lambda_{ir} = 0$) can be rejected. This indicates that the relatively weak statistical results for ICAPM risk prices when all uncertainty is taken into account is not due to a weak relationship between β s and average returns, but instead due to the fact that there is substantial uncertainty in estimating N_{ep} and N_{ir} .

The general conclusion is that the ICAPM risk factors (as estimated) matter economically and statistically when pricing equity and bond portfolios. However, the estimated N_{ep} and N_{ir} are noisy proxies for the true equity premium and interest rate news and this source of estimation uncertainty has a non-trivial impact on the statistical power of ICAPM tests.

c) Pricing Errors

The ICAPM not only matches the dividend and bond term structures, but also reasonably captures the equity, value, and credit premia.

Figure 6 displays predicted (based on the right hand side of equation 13) and realized (based on the left hand side of equation 13) risk premia for the testing assets over my main sample period (1952-2016). A perfect model has all points aligned in the 45 degree line (ignoring sampling error).

From Figure 6a, the CAPM correctly predicts that the risk premia increase from Treasury bonds to corporate bonds and then to equities. As a result, it captures a reasonable fraction of the risk premia variation across asset classes ($R^2 = 76\%$). However, its performance substantially deteriorates within each asset class. It does not capture the Treasury bond term structure, fails to match the risk premium level on corporate bonds, and completely misses the value premium. As a result, its within asset class explanatory power is even negative ($wR^2 = -3\%$), indicating that the average risk premium for each asset class better predicts realized risk premia than the CAPM.

The ICAPM substantially improves upon the CAPM despite also having only one parameter to capture cross-sectional variation in risk premia (Figure 6b). The R^2 increases dramatically (from 76% to 94%) and most of the improvement originates from better capturing within asset class variation (wR^2 increases from -3% to 77%).

Despite the substantial improvement, the ICAPM produces a bond term structure that is slightly stronger than in the data and a value premium that is slightly weaker than in the data. Increasing the ICAPM flexibility by dropping the risk price restrictions improves the model performance by better capturing the bond term structure and value premium (Figure 6c), but the improvement is marginal (R^2 goes from 94% to 97% and wR^2 from 77% to 85%) since the original fit was already good.

Figure A.2 decomposes risk premia for the testing assets into the effect of each ICAPM risk factor to shed light on the mechanism behind the good fit to equity and bond portfolios. However, since this paper's main objective is not to demonstrate the ICAPM ability to capture multiple asset pricing effects, but rather to explore its implications for the bond and dividend term structures, I do not further study the ICAPM pricing ability. Campbell et al. (2017) provide extensive evidence that the ICAPM can price equity portfolios sorted on several dimensions.

3.6 ICAPM Pricing of a Short-Term Dividend Strategy

All previous subsections focus on dividend PV returns. In this subsection, I show that the ICAPM also reasonable captures the average return of a short-term dividend strategy designed to capture the spread between S&P 500 dividend strips and the index itself while simultaneously hedging interest rate changes. The mechanism is the same as in previous subsections: the S&P 500 is a better hedge for equity premium shocks than short-term dividend claims.

a) The Short-Term Dividend Strategy

The short-term dividend strategy in this subsection is based on dividend futures. A S&P 500 dividend future is a standardized contract where, at maturity, the buyer pays the futures price, which is determined today, and the seller pays the S&P 500 dollar amount of dividends during a certain calendar year. For instance, the 2015 S&P 500 dividend futures contract was quoted at \$41.65 in December 1st, 2014. On the third Friday of December 2015, the buyer of the futures contract would pay \$41.65, and the seller would pay the cash dividend amount on the S&P 500 index that was paid out between the third Friday in December of 2014 and the third Friday in December of 2015. The contract is settled based on the sum of all dividends paid throughout the year, and there is no reinvestment of the dividends.¹⁵

[Binsbergen and Kojien \(2017\)](#) demonstrate that (up to a first order approximation):

$$R_{sp500,t} - R_{b,t} \approx \sum_{h=1}^{\infty} w_{df,t-1}^{(h)} \cdot R_{df,t}^{(h)} \quad (16)$$

where R_{sp500} is the gross return on the S&P 500, R_b is the gross return on a bond portfolio with duration that matches the S&P 500, $R_{df}^{(h)}$ is the net return on a h -year dividend futures contract, and $w_{df}^{(h)}$ are portfolio weights.

Similar to [Binsbergen and Kojien \(2017\)](#), I construct the excess return series (representing a short-term dividend strategy) $xR^{(st-mkt)} = R_{df}^{(st)} - [R_{sp500} - R_b^{(10)}]$, with $R_{df}^{(st)}$ having a duration of approximately 1.5 years. This is a long-short strategy that (i) longs a short-term dividend future (a position that buys the short-term dividend claim and forgoes the short-term bond return) and (ii) shorts a long-duration portfolio of dividend futures (a position that buys the S&P 500 index and

¹⁵[Manley and Mueller-Glissmann \(2008\)](#); [Wilkins and Wimschulte \(2010\)](#) provide valuable institutional details about the market for dividend derivatives and [Binsbergen et al. \(2013\)](#); [Cejnek and Randl \(2016, 2017\)](#); [Binsbergen and Kojien \(2017\)](#) explore different aspects of the dividend term structure using dividend futures data.

forgoes the long-term bond return). This strategy provides exposure to short-term dividend strips while attempting to hedge shocks to equity markets and changes in the term structure of interest rates.

Equation 16 implies we can understand $xR^{(st-mkt)}$ as providing information on the term-structure of dividend futures as long as we treat $R_b^{(10)}$ as a proxy for R_b . This is a conservative assumption since the S&P 500 duration is likely higher than ten years (so that R_b is likely to provide higher risk premium than $R_b^{(10)}$).

To construct $xR^{(st-mkt)}$, I obtain R_{sp500} from Datastream, $R_b^{(10)}$ from CRSP (same measurement as in previous subsections), and $R_{df,t}^{(st)}$ from three data sources. First, I use a database of daily quoted prices for S&P 500 dividend futures (from 02-Jan-2006 to 14-Oct-2016) provided to me by Goldman Sachs, who is an important player in the dividend futures market. This is a proprietary database, which Goldman Sachs uses firm-wide both as a pricing source and to mark the internal trading books to the market. Second, I use Datastream to update the dividend futures database until 31-Dec-2016 (dividend futures are now exchange traded, so the recent data is available in Datastream). Finally, to obtain a longer sample period, I use the data from [Binsbergen, Brandt, and Kojien \(2012\)](#) to create $R_{df,t}^{(st)}$ from February-1996 to December-2005. This data relies on S&P 500 option contracts and the put-call parity to recover S&P 500 dividend strip prices. The $R_{df,t}^{(st)}$ from this period is not exactly a portfolio of dividend futures, but instead a long-short portfolio with similar exposures as a portfolio of dividend futures and an average duration close to 1.5 years. I still refer to $R_{df,t}^{(st)}$ as a dividend futures position.

From these three data sources, I calculate overlapping annual returns on the short-term dividend futures position, $R_{df,t}^{(st)}$, to match the ICAPM risk factors. As the ICAPM pricing equation 13 is based on log excess returns, the final return series for the short-term dividend strategy is given by $xr^{(st-mkt)} = \ln(R_f + xR^{(st-mkt)}) - \ln(R_f)$. Since $xr_t^{(st-mkt)}$ is based on annual returns observed monthly, it goes from January 1997 to December 2016. Further details on the construction of $xR^{(st-mkt)}$ are provided in [Appendix B.3](#).

b) ICAPM pricing for the Short-Term Dividend Strategy

Table 3 summarizes the risk and risk premium properties of the short-term dividend strategy. All statistics use monthly observations of annual returns and risk factors over the period from January 1997 to December 2016, with risk factors constructed using my main sample period (1952-2016).

Table 3
ICAPM Pricing of Short-Term Dividend Strategy (1997-2016)

The table reports risk exposures and risk premium properties of $r^{(st-mkt)}$, a short-term dividend strategy (in the spirit of [Binsbergen and Koijen \(2017\)](#)) that (i) longs a short-term dividend future (a position that buys the short-term dividend claim and forgo the short-term bond return) and (ii) shorts a long-duration portfolio of dividend futures (a position that buys the S&P 500 index and forgo the long-term bond return). All statistics use monthly observations of annual returns and risk factors over the period from January 1997 to December 2016, with risk factors constructed using my main sample period (1952-2016). I report β s relative to unexpected market returns (β_m), equity premium news (β_{ep}), and interest rate news (β_{ir}); the realized risk premium estimated based on gross returns ($\mathbb{E}[R]$) as well as on log returns ($\mathbb{E}[r] + \frac{\sigma}{2}$), the ICAPM-implied risk premium ($\beta' \lambda$), and the ICAPM pricing error ($\alpha = \mathbb{E}[r] + \frac{\sigma}{2} - \beta' \lambda$). All β s are transformed to market β units (i.e., covariance normalized by market variance) and risk premia and pricing errors are calculated under both $rra = 6.3$ and $rra = 8.0$, which are the values reported in [Table 2](#). The (t-statistics) are reported below each statistic and account for all estimation uncertainty and {t-statistics} take shocks as given. Details about the construction of risk factors and model estimation can be found in subsections [3.1](#) and [3.5](#).

	Risk Exposures			Realized Premium		ICAPM($rra = 6.3$)		ICAPM($rra = 8.0$)	
	β_m	β_{ep}	β_{ir}	$\mathbb{E}[R]$	$\mathbb{E}[r] + \frac{\sigma}{2}$	$\beta' \lambda$	α	$\beta' \lambda$	α
$\mathbf{x}r^{(st-mkt)}$	-1.07	1.47	-0.01	3.0%	2.3%	2.4%	-0.1%	3.9%	-1.7%
$\mathbf{(t}_{stat})$	(-6.25)	(1.41)	(-0.08)	(0.55)	(0.50)		(-0.01)		(-0.18)
$\mathbf{\{t}_{stat}\}}$	{-6.68}	{2.51}	{-0.11}				{-0.02}		{-0.18}

The β columns indicate that the key risk patterns reported in subsection [3.3](#) hold for the short-term dividend strategy. In particular, the short-term dividend futures position has a lower market β and higher equity premium β than the long-duration portfolio of dividend futures (the $R_{sp500} - R_b^{(10)}$ position), with the equity premium β spread being more pronounced. Both β spreads are statistically significant (when we take shocks as given) and the spread magnitudes are economically important. The spread in interest rate β s is lower in magnitude and very weak statistically, and so it is not a key determinant of the risk term structure of dividend futures.

The $\mathbb{E}[R]$ and $\mathbb{E}[r] + \frac{\sigma}{2}$ columns indicate that the short-term dividend futures position pays a premium relative to the long-duration portfolio of dividend futures, which is consistent with the literature (see [Binsbergen, Brandt, and Koijen \(2012\)](#); [Binsbergen and Koijen \(2017\)](#)). This short-term dividend premium is not statistically significant for the U.S. market (as in the original papers), but it is economically important. [Binsbergen and Koijen \(2017\)](#) show that the premium is statistically significant when data from other countries (in Europe and Japan) are pooled together with data on U.S. dividend futures. I focus on the U.S. because I construct ICAPM risk factors to capture reinvestment risk in the U.S. market.

Combining the risk exposures with ICAPM risk prices, I calculate the ICAPM-implied risk premium for the short-term dividend strategy and found results consistent with the evidence previously reported (in subsection 3.4). In particular, the short-term dividend futures position has a higher risk premium than the long-duration portfolio of dividend futures. This is true whether I use the $rra = 6.3$ calibration in column (3a) of Table 2 or the $rra = 8.0$ calibration in columns (3b). Moreover, pricing errors (α s) are relatively small and zero from a statistical perspective.

The main conclusion is that, while not perfect, the ICAPM with a passive investor (detailed in Section 2) captures the essence of the dividend term structure reported in Binsbergen, Brandt, and Kojen (2012); Binsbergen and Kojen (2017). This result complements the empirical findings in the previous subsections, which are based on dividend PV returns. Despite the two approaches to estimate ICAPM risk exposures being different, they both lead to similar conclusions: long-term dividend claims have relatively low risk premia because they are good hedges for reinvestment risk.¹⁶

4 ICAPM Term Structures: Strategic Investor

The ICAPM results from Section 3 are striking in that they suggest an ICAPM produces both a downward sloping dividend term structure and an upward sloping bond term structure. However, the ICAPM used considers a passive investor and is silent on why the investor would not change her portfolio allocation in response to changes in expected returns. Moreover, the results are based on a present value decomposition of the equity premium as opposed to the common approach of studying dividend strip returns (subsection 3.6 is the exception as it relies on dividend futures data). This section addresses these issues. It considers a fully structural ICAPM in which the marginal investor responds optimally to changes in reinvestment rates. Moreover, I study the dividend term structure based on dividend strip β s and risk premia derived from the model.

4.1 The ICAPM with a Strategic Investor

This subsection details the SDF of an investor who responds optimally to changes in reinvestment rates (i.e., a strategic investor) and derives risk premia expressions for dividend strips and bond

¹⁶Dividend futures data can also be used to demonstrate that dividend PV returns correlate in expected ways with returns obtained directly from financial markets. Figure A.3 makes this point by demonstrating that (i) short-term dividend PV excess returns correlate strongly with returns on the short-term dividend futures position while long-term dividend PV excess returns do not; and (ii) the opposite pattern is true for the correlation between dividend PV excess returns and excess returns on the S&P 500 itself (which is a long-duration dividend strip portfolio).

portfolios. For that, I rely on the ICAPM framework in [Campbell, Chan, and Viceira \(2003\)](#) (which is a multivariate extension of [Campbell and Viceira \(1999\)](#)). I also explain the calibration of the model to the data. I omit proofs and technical derivations as they can be found in the internet Appendix of [Campbell, Chan, and Viceira \(2003\)](#). The exceptions are the novel analytical expressions for dividend strip β s and risk premia (I provide derivations for these expressions in Appendix [A.3](#)).

a) The SDF from the Perspective of a Strategic Investor

Analogously to Section [2](#), the log SDF of a strategic investor with Epstein-Zin recursive preferences is given by (ignoring constants):

$$m_t = - rra \cdot r_{p,t} - \frac{\theta}{ies} \cdot \ln(C_t/W_t) \quad (17)$$

where $r_p = \ln(R_p)$ are log returns on the investor's wealth portfolio (not necessarily the aggregate wealth portfolio), with $R_{p,t} = R_{f,t} + \gamma'_{t-1}(R_t - R_{f,t})$ depending on γ_t (the endogenous portfolio weights).

An investor who responds to changes in reinvestment rates endogenously creates time-varying risk in the SDF as the risk level of her wealth portfolio varies as a function of her allocation to different assets. Consequently, the log-linear approximation to the consumption-wealth ratio given in equation [10](#) is no longer valid. In this case, we instead have the following log-linear approximation:

$$\begin{aligned} \tilde{\ln}\left(\frac{C_t}{W_t}\right) &\approx (1 - ies)(\mathbb{E}_t - \mathbb{E}_{t-1}) \left[\sum_{h=1}^{\infty} \rho^h \cdot r_{p,t+h} \right] + \frac{1}{2} \frac{ies}{\theta} (\mathbb{E}_t - \mathbb{E}_{t-1}) \left[\sum_{h=1}^{\infty} \rho^h \cdot Var_{t+h-1}(m_{t+h} + r_{p,t+h}) \right] \\ &\equiv (1 - ies) \cdot N_{\mathbb{E},t} + \frac{ies}{\theta} \cdot N_V \end{aligned} \quad (18)$$

We can get log SDF shocks from equation [17](#) and substitute the new consumption-wealth ratio approximation to get (see [Bansal et al. \(2014\)](#); [Campbell et al. \(2017\)](#)):

$$\tilde{m}_t = - rra \cdot \tilde{r}_{p,t} - (rra - 1) \cdot N_{\mathbb{E},t} + N_{V,t} \quad (19)$$

Equation [19](#) provides an expression for the SDF of a strategic investor that allows us to think about reinvestment risk. Unfortunately, the risk factors depend on the endogenous portfolio allocation, γ_t . As such, I cannot link the risk factors to r_e and r_f without taking a stand on γ_t .

I make structural assumptions such that γ_t and $cw_t = \ln(C_t/W_t)$ can be directly solved for as functions of a set of state variables. This is the approach in [Campbell, Chan, and Viceira \(2003\)](#).

I assume the strategic investor can allocate capital to the (nominal) risk-free rate to receive the (unknown and stochastic) real log return $r_{f,t}$. She also has access to a set of risky assets with excess log returns denoted by xr_t . Moreover, z_t represents a vector with other variables that capture variation in expected returns.

I define $s_t = [r_{f,t} \quad xr_t \quad z_t]$ to be the state vector which summarizes the current information set and I assume s_t evolves according to a first-order vector autoregression:

$$s_{t+1} = \Phi_0 + \Phi_1 \cdot s_t + u_{t+1} \quad (20)$$

where $u_t \stackrel{i.i.d}{\sim} N(0, \Sigma)$.

Log returns on the investor's wealth portfolio can be approximated as:

$$r_{p,t+1} \approx r_{f,t} + \gamma'_t xr_{t+1} + \frac{1}{2} \cdot \gamma'_t \cdot [\text{diag}(\Sigma_{xr}) - \Sigma_{xr} \cdot \gamma_t] \quad (21)$$

The given structure implies the following (approximate) consumption and portfolio allocation policies for an investor with Epstein-Zin recursive preferences:

$$\gamma_t = \Gamma_0 + \Gamma_1 \cdot s_t \quad (22)$$

$$cw_t = C_0 + C_1' \cdot s_t + s_t' C_2 s_t \quad (23)$$

with policy parameters implicitly depending on rra , ies , ρ , and the VAR parameters (except that γ_t does not depend on ies given ρ).

Using these policy functions, we can recover the log SDF by substituting equations 20 to 23 into equation 17.

I solve for the policy parameters (Γ s and C s) following the procedure outlined in [Campbell, Chan, and Viceira \(2003\)](#). The idea is to start from a guess of ρ and solve for all Γ s and C s recursively until ρ is consistent with the policy parameters (the steps that lead to the policy parameters for a given value of ρ are available in [Campbell, Chan, and Viceira \(2003\)](#)'s internet Appendix). After solving for Γ s and C s, I recover \tilde{r}_p , $N_{\mathbb{E}r}$, and N_V to explore the reinvestment risk implications to the dividend and bond term structures.

b) Risk Premia on Dividend Strips and Bond Portfolios

The ICAPM-based risk premium of any asset j is given by:

$$\mathbb{E}_t [xr_{j,t+1}] + \frac{1}{2} (\sigma_{j,t}^2 - \sigma_{f,t}^2) = rra \cdot Cov_t(xr_{j,t+1}, r_{p,t+1}) + \frac{\theta}{ies} \cdot Cov_t(xr_{j,t+1}, cw_{t+1}) \quad (24a)$$

$$\begin{aligned} &= rra \cdot Cov_t(xr_{j,t+1}, r_{p,t+1}) \\ &+ (rra - 1) \cdot Cov_t(xr_{j,t+1}, N_{\mathbb{E}r,t+1}) \\ &+ Cov_t(xr_{j,t+1}, N_{V,t+1}) \end{aligned} \quad (24b)$$

Assuming $\mathbb{E}_t [xr_{j,t+1}]$ evolves as in the VAR (equation 20), we can use the policy rules previously derived to solve for the risk premium of asset j (whether the asset is included in r_p or not). I use this approach to study the bond term structure based on the bond portfolios in Section 3.

For dividend strips, I assume Δd_t is part of s_t , define $pd_t^{(h)} \equiv \ln(P_t^{(h)}/D_t)$, and use the recursive relation $pd_t^{(h)} = pd_{t+1}^{(h-1)} + \Delta d_{t+1} - r_{ds,t+1}^{(h)}$ (with boundary condition $pd_t^{(0)} = 0$) together with the policy functions (cw_t and γ_t) to solve for dividend risk and risk premia as linear functions of the state variables (details in Appendix A.3):¹⁷

$$Var_t \left(r_{ds,t+1}^{(h)} \right) = \sigma_{ds,h}^2 \quad (25a)$$

$$Cov_t \left(xr_{ds,t+1}^{(h)}, r_{p,t+1} \right) = \pi_{0,p}^{(h)} + \pi_{1,p}^{(h)'} s_t \quad (25b)$$

$$Cov_t \left(xr_{ds,t+1}^{(h)}, cw_{t+1} \right) = \pi_{0,cw}^{(h)} + \pi_{1,cw}^{(h)'} s_t \quad (25c)$$

$$\mathbb{E}_t \left[xr_{ds,t+1}^{(h)} \right] + \frac{1}{2} \cdot (\sigma_{ds,h}^2 - \sigma_f^2) = \pi_{0,\mathbb{E}r}^{(h)} + \pi_{1,\mathbb{E}r}^{(h)'} s_t \quad (25d)$$

I explore the unconditional term structures (by taking unconditional averages of the expressions provided) as well as time variation in the dividend term structure, which is difficult to do without a fully structural framework such as the one in this section.

c) Model Calibration

To calibrate the model to the data, I need to define which variables proxy for $r_{f,t}$, xr_t , and z_t . To be consistent with the previous section, I use the same $r_{f,t}$ and a z_t with the same predictive variables: dividend growth (Δd), dividend yield (dy), equity payout yield ($epoy$), one year Treasury yield (ty),

¹⁷As in Section 3, empirical results are based on equal weighted portfolios of consecutive maturities. Since I focus on annual returns, this procedure assures the average duration of a portfolio indexed by h is equal to h years.

term spread (TS), credit spread (CS), and value spread (VS). Analogously to the previous section, I impose VAR restrictions such that only z_t predicts future information. For xr_t , I assume the strategic investor has access to the risk-free asset as well as the equity market index ($xr_e = r_e - r_f$) and the short-term dividend strategy ($xr^{(st-lt)}$) from subsection 3.6. The inclusion of $xr^{(st-lt)}$ assures the ICAPM results are consistent with the pricing of the short-term dividend strategy. However, the dividend term structure is similar when the investor only has access to r_f and r_e (compare Figure 7 to panel (j) in Figure A.7).

I calibrate the VAR to capture the annual dynamics of z_t and the long-term predictability in r_f and xr_e . The calibration details are provided in Appendix C.4, but the idea is to make sure the volatility of long-term discount rates (which determines $N_{\mathbb{E}r}$ volatility) matches the volatility observed in the data. Focusing on the short-term dynamics of r_f and xr_e substantially understates long-term discount rate volatility, which has non-trivial consequences for the importance of reinvestment risk relative to market risk (see Appendix D.2). This result is largely consistent with Giglio and Kelly (2017), who find that VAR term structure models calibrated to match the price dynamics of short maturity claims produce too little volatility for long maturity claims. In my case, calibration to the dynamics of short-term discount rates produces too little volatility for long-term discount rates.

In terms of preference parameters, I calibrate the intertemporal elasticity of substitution to $ies = 0.5$ and follow Campbell, Chan, and Viceira (2003) in setting the time discount factor to $\delta = 0.92$ at the annual frequency. I then consider a grid for rra from 2 to 15 (with 0.2 intervals) and evaluate the fit to the six bond portfolios in Section 3 (in terms to mean squared pricing errors). I study the *dividend* term structure with $rra = 10.2$ as this value gives the best fit to the *bond* term structure.¹⁸

4.2 Unconditional Term Structures

This subsection explores the unconditional dividend term structure in the ICAPM with a strategic investor. It also discusses the ICAPM fit to bond portfolios to assure the model captures both the dividend and bond term structures.

Panel (a) of Figure 7 displays dividend risk premia originated from three sources of risk: market

¹⁸Campbell, Chan, and Viceira (2003) argues that ies has only second order effects on the SDF within the ICAPM. Consistent with this claim, $ies = 1.5$ produces a similar (but with slightly smaller slope) dividend term structure (compare Figure 7 to panel (g) in Figure A.7). I also evaluate the dividend term structure with $rra = 6.3$, which is the value used in most of Section 3. Again, I find similar qualitative results for the dividend term structure, but with a somewhat smaller slope (compare Figure 7 to panel (h) in Figure A.7).

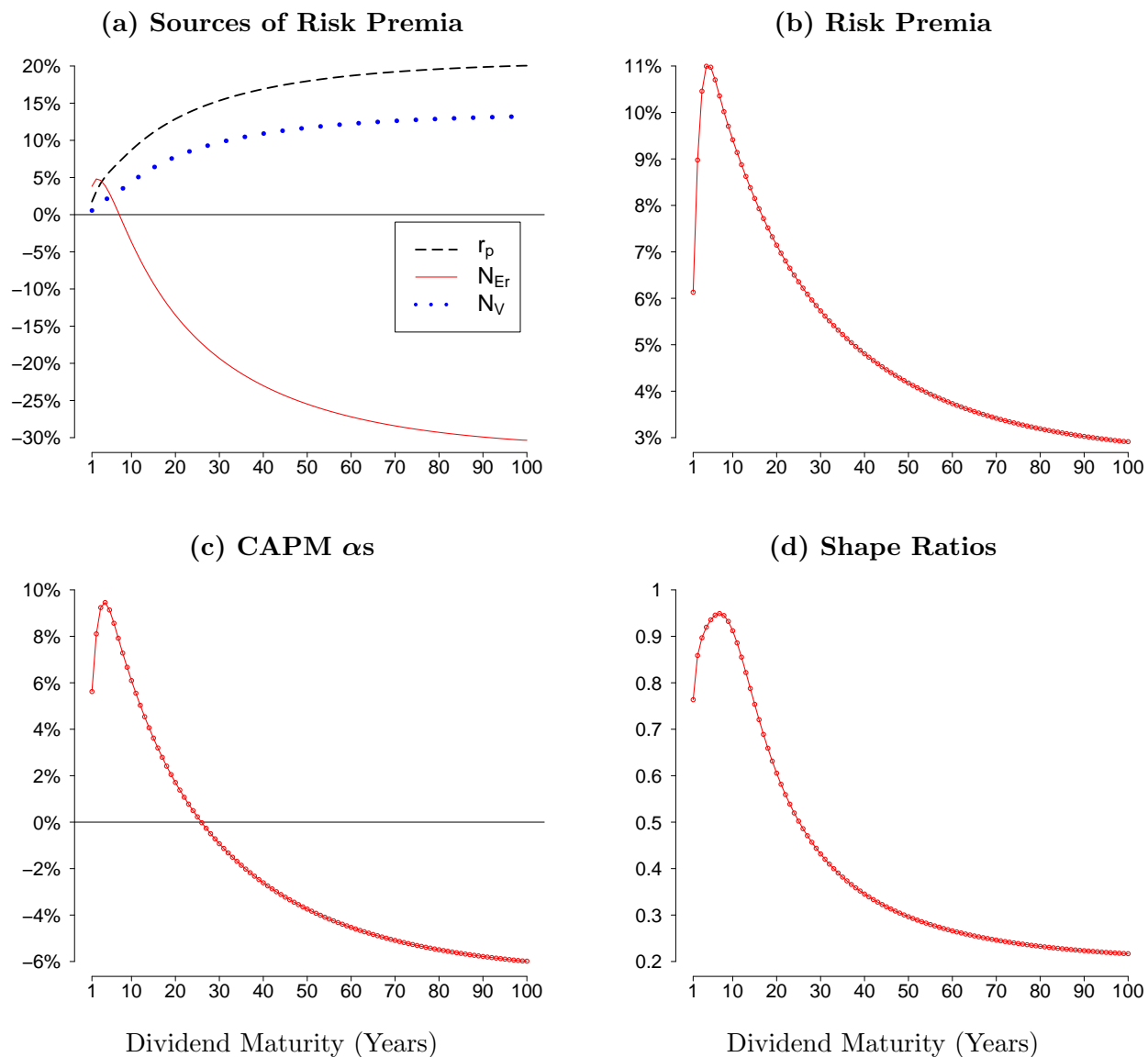


Figure 7
Dividend Term Structures in the ICAPM with a Strategic Investor

The graphs report the dividend term structures of (i) risk premia, (ii) CAPM α s, and (iii) Sharpe ratios within the ICAPM with an investor who responds to changes in reinvestment rates (i.e., a strategic investor). The Risk premia term structure can be decomposed into the effect of market risk (r_p), reinvestment risk (N_{Er}), and volatility risk (N_V) and panel (a) provides such decomposition (based on equation 24b). All results are obtained based on my main sample period (1952-2016) and details for the ICAPM and calibration can be found in subsection 4.1.

risk from variation in the investor’s portfolio returns (r_p), reinvestment risk from variation in long-run expected returns (N_{Er}), and volatility risk from changes in the long-run volatility of r_p and cw (N_V).¹⁹

First, as in the previous section, longer-term dividend strips have higher market risk and are better hedges for reinvestment risk, with the second effect dominating the first. Second, longer-term dividend strips are more exposed to volatility shocks. This effect happens because variation in expected returns are linked to risk variation in this ICAPM framework (since there is no variation in risk aversion). Consequently, volatility increases coincide with higher discount rates, which lowers dividend strip prices and exposes them to volatility shocks. Duration amplifies this effect, making longer-term dividend strips more exposed to volatility shocks.²⁰

Panel (b) combines the risk premia from the three risk sources to produce the dividend term structure of risk premia. For the first few maturities, the market and volatility risk effects dominate, inducing an upward sloping dividend term structure ($\mathbb{E}[R_{ds}^{(1)}] = 6.1\%$ and $\mathbb{E}[R_{ds}^{(5)}] = 11.6\%$). However, the reinvestment risk effect dominates for dividend strips with longer maturities so that long-term dividend strips have very low risk premia (they are roughly 3% at maturities close to 100 years). Overall, the ICAPM produces a hump-shaped dividend term structure with particularly low discount rates for very long-term dividends.

The dividend risk premia produced by the ICAPM are striking in that they match the otherwise puzzling pattern uncovered in the literature. For instance, [Binsbergen and Koijen \(2017\)](#) reports (for S&P 500 dividend strips) average returns that increase over the first few years, but decrease over the long-term as the dividend strips with maturities from three to five years outperform the S&P 500, which is a long-duration portfolio of dividend strips.

However, the dividend term structure seems much less steep than the one produced by the ICAPM in Section 3. The reason is simple. If the marginal investor can respond to shifts in reinvestment rates by changing her portfolio allocation (as opposed to being a passive), then reinvestment risk becomes

¹⁹Note that N_V can be written as:

$$N_V = \frac{1}{2} (\mathbb{E}_t - \mathbb{E}_{t-1}) \left[\sum_{h=1}^{\infty} \rho^h \cdot Var_{t+h-1} \left((rra - 1) \cdot r_{p,t+h} + \frac{\theta}{ies} \cdot cw_{t+h} \right) \right]$$

²⁰This effect is present because volatility changes are induced by shifts in portfolio allocation, which are directly connected to changes in discount rates. It is unclear whether a model with exogenous volatility shocks would produce the same term structure of volatility risk. The reason is that volatility shocks tend to be less persistent than discount rate shocks. As such, it is possible that short-term cash flows are more affected by transitory volatility shocks. Exploring the term structure of volatility risk is an interesting avenue for future work. The framework developed by [Bansal et al. \(2014\)](#) and [Campbell et al. \(2017\)](#) seems a natural starting point for that.

less of a concern. Of course, it remains economically important and still overturn the market risk effect, but only over long maturities so that the term structure is hump-shaped and much less steep than in Section 3. Nevertheless, the reinvestment risk mechanism is the same in both sections and this is an important point because the two sections rely on quite different opportunity sets for the ICAPM long-term investor.

Panels (c) and (d) demonstrate that similar hump-shaped term structures are observed for CAPM α s and Sharpe ratios. In the case of CAPM α s, the term structure differences are even more pronounced, with the 4-year dividend strip displaying $\alpha = 9.5\%$ while the 100-year dividend strip has $\alpha = -6.0\%$. For Sharpe ratios, the differences are also large, with short-term dividend strips having Sharpe ratios between 0.75 and 0.95 while very long-term dividend strips have Sharpe ratios below 0.3. Both term structures (Sharpe ratios and CAPM α s) are consistent with the patterns reported in [Binsbergen and Kojen \(2017\)](#) (they do not report CAPM α s, but these can be inferred from the reported average returns and CAPM β s).

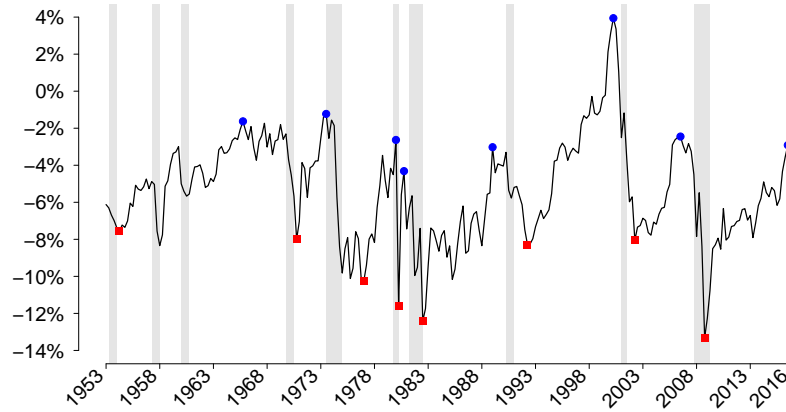
Figure A.5 displays ICAPM-predicted risk premia against realized risk premia for the same testings assets used in subsection 3.5. The ICAPM almost perfectly matches the equity premium and bond term structure and also reasonably captures risk premia on corporate bond portfolios. However, it produces a value premium that is weaker than in the data (roughly half the value premium is explained by the ICAPM) with all HML portfolios having ICAPM-predicted risk premia lower than in the data. This indicates that the reinvestment risk channel (at least as tested here) is only a partial explanation for the value premium if the marginal investor responds to shifts in reinvestment rates by changing her portfolio allocation.

Overall, the ICAPM with a strategic investor produces the same reinvestment risk mechanism of the previous section. That is, long-term dividend claims have relatively low discount rates because they are good hedges for reinvestment risk. At the same time, the model nicely captures the upward sloping bond term structure. The next subsection explores time variation in dividend risk premia.

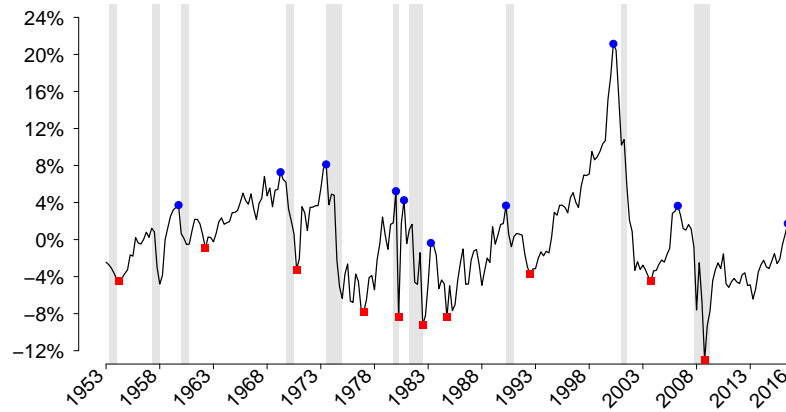
4.3 Time Variation in the Dividend Term Structure

In this subsection, I explore ICAPM-implied time variation in the dividend term structure. I argue that the buy and hold dividend term structure slope decreases during recessions and increases in expansions (which is consistent with [Binsbergen et al. \(2013\)](#)) and that this pattern is driven by variation in the relative risk of short- and long-term *holding periods* as opposed to the relative risk

(a) Term Spread in Buy and Hold $\mathbb{E}[r]$ for Dividend Claims



(b) Term Spread in Buy and Hold $\mathbb{E}[r]$ for the Aggregate Equity Market



(c) Term Spread in 1-Year $\mathbb{E}[R]$ for Dividend Claims

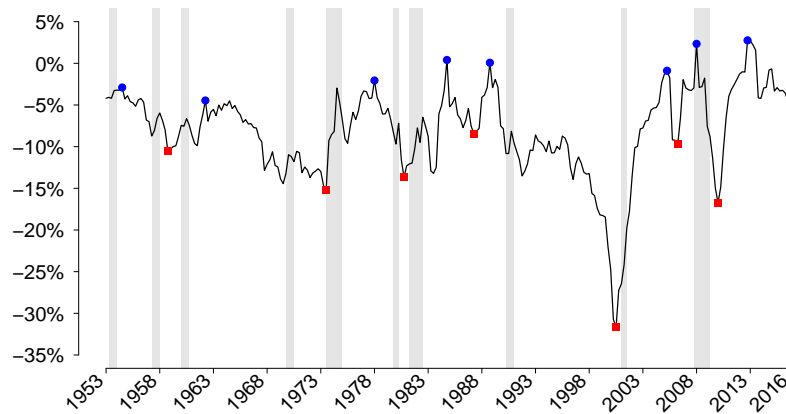


Figure 8
Expected Return Term Spreads

The graphs report time variation in the term spread (100-year minus 5-year) of (i) buy and hold expected returns on dividend strips ($\mathbb{E}_t[\bar{r}_{ds,t \rightarrow t+100}^{(100)}] - \mathbb{E}_t[\bar{r}_{ds,t \rightarrow t+5}^{(5)}]$); (ii) buy and hold expected returns on the equity market ($\mathbb{E}_t[\bar{r}_{e,t \rightarrow t+100}] - \mathbb{E}_t[\bar{r}_{e,t \rightarrow t+5}]$); and (iii) 1-year expected returns on dividend strips ($\mathbb{E}[\bar{R}_{ds,t+1}^{(100)}] - \mathbb{E}[\bar{R}_{ds,t+1}^{(5)}]$). The term structure points (5 and 100 years) were selected to match the highest and lowest points in the unconditional dividend term structure. Shaded regions represent recessionary periods (as defined by the National Bureau of Economic Research - NBER). Blue circles (red squares) represent local peaks (troughs) of the respective time series (details about the algorithm used to identify peaks/troughs are available upon request). All results are obtained based on my main sample period (1952-2016) and details for the ICAPM and calibration can be found in subsection 4.1.

of short- and long-term *dividends*, which is a novel result.

Panel (a) of Figure 8 displays (from January 1953 to December 2016) the term spread in buy and hold $\mathbb{E}[r]$ of dividend strips. I consider the spread between (ICAPM-implied) 100-year expected returns, $\mathbb{E}[\bar{r}_{ds,t \rightarrow t+100}^{(100)}]$, relative to five-year expected returns, $\mathbb{E}[\bar{r}_{ds,t \rightarrow t+5}^{(5)}]$. Analogous spreads are constructed for the equity market ($\mathbb{E}[\bar{r}_{e,t \rightarrow t+100}] - \mathbb{E}[\bar{r}_{e,t \rightarrow t+5}]$) in panel (b) and for 1-year expected returns on dividend strips ($\mathbb{E}[\bar{r}_{ds,t+1}^{(100)}] - \mathbb{E}[\bar{r}_{ds,t+1}^{(5)}]$) in panel (c). The term structure points were selected to match the highest and lowest points in the unconditional dividend term structure. The gray shaded area identifies recessionary periods, the blue circles represent local peaks, and the red squares refer to local troughs.

The buy and hold dividend term spread (Panel (a)) tends to decrease during recessions and increase in expansions (i.e., it is procyclical), reaching its local peaks close to the beginning of recessions and local troughs close to recession ends (a pattern consistent with [Binsbergen et al. \(2013\)](#)). While this spread is based both on cash flow maturity and holding period, the spread in panel (b) is entirely driven by holding period and the spread in panel (c) is only based on cash flow maturity. Having all three spreads allows me to ask whether the variation in the buy and hold dividend term spread is driven by holding period or cash flow maturity.

Panel (b) demonstrates that the term spread in buy and hold $\mathbb{E}[r]$ for the aggregate equity market is also procyclical, with variation that is remarkably similar to the one in panel (a). In contrast, the term spread in (one-year holding period) dividend strip expected returns (Panel (c)) increases in recessions and decreases in expansions (if anything, it is countercyclical). Therefore, the buy and hold dividend term spread is entirely driven by variation in the relative risk of short- and long-term *holding periods* (as in panel (b)) as opposed to the relative risk of short- and long-term *dividends* (as in panel (c)).

The mechanism is quite simple. The equity premium increases in recessions and is expected to revert to its mean over the long-term. Consequently, the equity term spread decreases over recessionary periods. The opposite happens during expansions. The term spread in buy and hold expected returns of dividend strips follows the same pattern because the level of dividend risk premia strongly correlates with the equity premium.

In summary, variation in the buy and hold dividend term structure slope is entirely driven by time variation in level of risk premia as opposed to time variation in the term structure slope of risk premia.

5 Implications to the Long-Run Risks Framework

The results from the previous two sections suggest that reinvestment risk, as captured by the ICAPM, produces both a downward sloping dividend term structure at long maturities and an upward sloping bond term structure. However, the results are somewhat puzzling since the ICAPM structure is very similar to the one in the long-run risks model (Bansal and Yaron (2004); Bansal, Kiku, and Yaron (2012)), which produces an upward sloping dividend term structure (see Binsbergen, Brandt, and Koijen (2012)).

This section clarifies the differences between the two models and explores the implications of reinvestment risk to the long-run risks model. I show (through simulations) that the typical calibration of the long-run risks model does not account for the dividend term structure because it fails to produce the downward sloping term structure of reinvestment risk uncovered in Sections 3 and 4. However, the long-run risks framework correctly prices reinvestment risk and the relationship between this framework and the ICAPM implies that there must exist an alternative specification for the consumption growth process that captures both the bond and dividend term structures. Such specification might not reflect the consumption growth of a representative consumer (see Malloy, Moskowitz, and Vissing-Jørgensen (2009)), but it can still retain the main aspects of the model. As such, I view the long-run risks *framework* as potentially consistent with the dividend and bond term structures.

Both the ICAPM of Sections 3 and 4 and the long-run risks model rely on an investor with Epstein-Zin recursive preferences. However, the two models differ in the type of assumptions they make about the economy. The long-run risks model takes a stand on who the marginal investor is and on the dynamics of her consumption process. From those assumptions and the structure of an endowment economy, the model derives an SDF, which provides direct implications to asset prices. In contrast, the ICAPM takes a stand on the assets available to the marginal investor (not on who she is) and the dynamics of those assets (variation in investment opportunities). From those assumptions and the structure of a portfolio decision problem, the model derives an SDF, which also provides asset pricing implications. Despite the differences, these models are two sides of the same coin. In particular, for any given ICAPM, there exists a long-run risks model that produces the exact same results (see Campbell et al. (2017) for a detailed explanation of this point).²¹ Therefore, it is worth

²¹The consumption process dynamics in the long-run risks model that recover the ICAPM results can be very different

trying to understand why the typical calibration of the long-run risks model does not account for the dividend term structure despite the fact that the ICAPM does.

Table 4 explores this issue. Panel A provides results based on a simulation of the long-run risks model using the calibration in [Bansal and Yaron \(2004\)](#) while Panel B focuses on the calibration in [Bansal, Kiku, and Yaron \(2012\)](#). In both cases, I rely on the solution method outlined in [Binsbergen, Brandt, and Koijen \(2012\)](#) except that I solve for the price-consumption ratio (and recover the wealth-consumption ratio) analytically as in [Bansal, Kiku, and Yaron \(2012\)](#).²² I simulate 101,000 years of monthly data, burn the first 1,000 years to assure the simulation is not influenced by the initial conditions, and report results based on the following 100,000 years. All statistics are calculated from monthly data and risk-premia information is multiplied by twelve so that the magnitudes are roughly comparable to the previous sections. My discussion focuses on Panel A and later details how the basic results differ for the calibration in Panel B.

The first row in Panel A clearly demonstrates that risk premia increases in dividend maturity in the long-run risks model. Since the model is based on an Epstein-Zin investor, we can decompose the dividend risk premia into the components associated with the three risk factors I explore in Section 4: market risk (\tilde{r}_m), reinvestment risk ($N_{\mathbb{E}r}$), and volatility risk (N_V).²³ I do so in the subsequent three rows. Market and volatility risks have small upward sloping term structure effects. As in the previous sections, reinvestment risk has a strong term structure effect. However, unlike the previous sections, reinvestment risk induces an upward sloping dividend term structure. Instead of providing a hedge against reinvestment risk shocks, dividend strips are exposed to reinvestment risk, with longer-term claims more exposed to this risk given their higher duration.

To understand how the long-run risks model produces reinvestment risk betas that increase in dividend maturity, the next four rows explore the properties of discount rate and cash flow news in the model. Row $Cor(N_{\mathbb{E}r}, N_{\mathbb{E}r}^{(h)})$ shows that when reinvestment rates decline, $N_{\mathbb{E}r} < 0$, so do dividend

from the typical dynamics researchers have in mind when referring to long-run risks. Nevertheless, such dynamics exist. Here, I am generically referring to the long-run risks model as any model with a fully specified consumption process which is assumed to originate from the optimal consumption decisions of an Epstein-Zin agent.

²²I find results similar to the ones in this section after solving the long-run risks model using a non-linear algorithm to approximate the wealth-consumption ratio as a function of the two state variables. As such, my results are not induced by the analytical log-linear solution I use.

²³Note that, for any given asset, the sum of the risk premium components from the risk factors is not equal to the average risk premium over the simulation. The difference comes from the fact that the components are based on unconditional covariances while the model-implied pricing equation is based on the unconditional expectation of the conditional covariance, which requires solving for the conditional covariances as functions of the state variables. I focus on unconditional covariances as it is much simpler and is typically what researchers use to evaluate models empirically. However, any difference in the two approaches is minor and the basic intuition for why the long-run risks model fail should remain the same.

Table 4
Reinvestment Risk in the Long-run Risks Model

The table reports (long run risks) model-implied statistics for dividend strips and the equity portfolio. Panel A provides results based on a simulation of the long-run risks model using the calibration in [Bansal and Yaron \(2004\)](#) while Panel B focuses on the calibration in [Bansal, Kiku, and Yaron \(2012\)](#). In both cases, I rely on the solution method outlined in [Bansal, Brandt, and Koijen \(2012\)](#) except that I solve for the price-consumption ratio (and recover the wealth-consumption ratio) analytically as in [Bansal, Kiku, and Yaron \(2012\)](#). I simulate 101,000 years of monthly data, burn the first 1,000 years to assure the simulation is not influenced by the initial conditions, and report results based on the following 100,000 years. All statistics are calculated from monthly data and risk-premia information is multiplied by twelve so that the magnitudes are roughly comparable to the previous sections.

PANEL A: Calibration in Bansal and Yaron (2004)											
Maturity (in Years) =	1	3	5	7	9	10	20	30	40	50	Market
$\mathbb{E}[R_{ds}^{(h)} - R_f]$	1.3%	3.4%	4.7%	5.4%	5.9%	6.1%	6.7%	6.8%	6.8%	6.8%	5.3%
$rra \cdot Cov(\tilde{r}^{(h)}, \tilde{r}_m)$	0.5%	1.2%	1.6%	1.9%	2.1%	2.1%	2.4%	2.4%	2.4%	2.4%	1.9%
$(rra - 1) \cdot Cov(\tilde{r}^{(h)}, N_{Er})$	0.9%	2.1%	2.9%	3.4%	3.6%	3.7%	4.0%	4.0%	4.0%	4.0%	3.2%
$-1 \cdot Cov(\tilde{r}^{(h)}, N_V)$	0.0%	0.0%	0.0%	0.1%	0.1%	0.2%	0.3%	0.3%	0.3%	0.3%	0.2%
$Cor(N_V, N_{Er}^{(h)})$	-0.12	0.02	0.16	0.28	0.38	0.42	0.60	0.63	0.64	0.64	0.45
$Cor(N_{Er}, N_{Er}^{(h)})$	0.97	1.00	1.00	0.98	0.96	0.95	0.86	0.84	0.83	0.83	0.94
$Cor(N_{Er}, N_g^{(h)})$	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
$\sigma(N_g^{(h)})/\sigma(N_{Er}^{(h)})$	3.6	4.2	4.3	4.2	4.1	4.1	3.6	3.5	3.5	3.4	3.8
$Cor(\tilde{r}^{(h)}, N_{Er}^{(h)})$	0.22	0.49	0.60	0.62	0.61	0.60	0.49	0.45	0.44	0.44	0.54
PANEL B: Calibration in Bansal, Kiku, and Yaron (2012)											
Maturity (in Years) =	1	3	5	7	9	10	20	30	40	50	Market
$\mathbb{E}[R_{ds}^{(h)} - R_f]$	2.5%	3.5%	4.2%	4.8%	5.2%	5.4%	7.0%	8.3%	9.5%	10.5%	6.1%
$rra \cdot Cov(\tilde{r}^{(h)}, \tilde{r}_m)$	2.0%	2.4%	2.7%	2.8%	3.0%	3.1%	3.7%	4.1%	4.6%	5.0%	3.2%
$(rra - 1) \cdot Cov(\tilde{r}^{(h)}, N_{Er})$	0.4%	1.0%	1.2%	1.3%	1.3%	1.3%	0.9%	0.4%	0.0%	-0.3%	0.9%
$-1 \cdot Cov(\tilde{r}^{(h)}, N_V)$	0.0%	0.1%	0.3%	0.6%	0.9%	1.0%	2.5%	3.8%	5.0%	6.0%	1.5%
$Cor(N_V, N_{Er}^{(h)})$	-0.03	0.11	0.26	0.41	0.54	0.59	0.87	0.94	0.96	0.97	0.69
$Cor(N_{Er}, N_{Er}^{(h)})$	0.77	0.85	0.92	0.97	1.00	1.00	0.93	0.85	0.81	0.78	0.92
$Cor(N_{Er}, N_g^{(h)})$	0.79	0.79	0.79	0.79	0.79	0.79	0.79	0.79	0.79	0.79	0.79
$\sigma(N_g^{(h)})/\sigma(N_{Er}^{(h)})$	3.1	3.6	3.5	3.4	3.1	3.0	1.9	1.3	1.0	0.9	2.1
$Cor(\tilde{r}^{(h)}, N_{Er}^{(h)})$	0.10	0.24	0.29	0.30	0.27	0.25	0.02	-0.14	-0.25	-0.33	0.07

discount rates, $N_{\mathbb{E}r}^{(h)}$, inducing higher dividend strip prices. This is the mechanism responsible for generating a downward sloping term structure of reinvestment risk in the ICAPM. Row $Cor(N_{\mathbb{E}r}, N_g^{(h)})$ shows that dividend growth shocks counteract the discount rate channel because lower reinvestment rates are strongly associated with lower expected dividend growth (0.99 correlation for all maturities), inducing a price decrease when reinvestment rates decline. Row $\sigma(N_g^{(h)})/\sigma(N_{\mathbb{E}r}^{(h)})$ shows that dividend growth news are substantially more volatile than discount rate news in the model, which means that the dividend growth channel dominates so that dividend strip prices decrease when reinvestment rates decline: dividend strips are exposed to reinvestment risk. This dividend growth mechanism is so strong that dividend strip prices (and equity prices) tend to decrease when dividend discount rates decline (see row $Cov(r_{ds}^{(h)}, N_{\mathbb{E}r}^{(h)})$), which is highly counterfactual.²⁴

The reason for the strong correlation between reinvestment rates and dividend growth news is straightforward. The long-run risks model implies $N_{cg} = ies \cdot N_{\mathbb{E}r} - \frac{ies-1}{rra-1} \cdot N_V$ (see equation 12 in [Bansal et al. \(2014\)](#)). As long as volatility shocks are small relative to consumption growth shocks, $Cor(N_{\mathbb{E}r}, N_{cg})$ is high. This is the case in [Bansal and Yaron \(2004\)](#)'s calibration of the model, which produces $Cor(N_{\mathbb{E}r}, N_{cg}) = 0.99$. Since expected dividend growth and expected consumption growth are perfectly correlated and driven by a univariate process, we also have $Cor(N_{\mathbb{E}r}, N_g^{(h)}) = 0.99$ for any maturity h .

The calibration in [Bansal, Kiku, and Yaron \(2012\)](#) (Panel B) implies stronger long-run volatility shocks (because of the higher autocorrelation in the volatility process), which alleviates some of the issues in [Bansal and Yaron \(2004\)](#)'s calibration. For instance, $Cor(N_{\mathbb{E}r}, N_{cg}) = 0.79$ in this calibration (in contrast to the 0.99 values of the previous calibration). Nevertheless, the strong positive correlation between discount rate and cash flow news coupled with the much higher volatility of cash flow news continues to induce a counterfactual term structure of reinvestment risk (increasing for the first ten years and decreases after that), which still shuts down the strong reinvestment risk mechanism outlined in the previous sections.

Another issue with the model is that, in both calibrations, volatility shocks represent the only source of variation in expected returns. As such, they are (strongly) positively correlated with dividend discount rates (see row $Cov(N_V, N_{\mathbb{E}r}^{(h)})$). This channel induces an upward sloping term

²⁴The literature studying the relative importance of discount rate and cash flow news to equity returns is extensive and generally finds a negative correlation between equity returns and discount rate news (e.g., [Campbell \(1991\)](#); [Campbell and Ammer \(1993\)](#); [Chen, Da, and Zhao \(2013\)](#)). In my empirical analysis, I find $Cor(\tilde{r}_e, N_{ep} + N_{ir}) = -0.52$, which is largely consistent with the literature, but inconsistent with the implications from the long-run risks model.

structure of volatility risk (the price of longer-term dividend strips decrease by more when volatility increases). In [Bansal, Kiku, and Yaron \(2012\)](#)'s calibration, this channel is very strong because volatility shocks are highly persistent, inducing large variability in N_V . Since there is no reinvestment risk channel counteracting this effect, the model produces a strongly upward sloping dividend term structure.

It is important to point out that the results produced by the long-run risks model would be counterfactual even if the dividend term structure were upward sloping. The reason is that the term structure of reinvestment risk is strongly downward sloping in the data, which is not the case in either calibration of the long-run risks model. Hence, the mechanism the model uses to generate an upward sloping dividend term structure is counterfactual.

Table 4 makes it clear that the typical calibration of the long-run risks model fails to capture the dividend term structure due to the model-implied properties of cash flow and discount rate news. However, this does not mean the framework is wrong. The results in Section 4 coupled with the close relation between the ICAPM and the long-run risks model necessarily imply that there exists an alternative specification for the consumption growth process that captures both the bond and dividend term structures.

In any calibration of the long-run risks model, the price of reinvestment risk is consistent with the ICAPM. The issue is only on risk quantities. Two adjustments to the model might go a long way in capturing the dividend term structure. First, weakening the link between expected dividend growth and expected consumption growth and increasing the volatility of discount rates (relative to cash flow growth) should help the model to produce an empirically credible term structure of reinvestment risk. Second, weakening the link between volatility and discount rate shocks through the introduction of time-varying risk premia (see [Melino and Yang \(2003\)](#)) or slightly decreasing the persistence of volatility shocks would help to make sure the reinvestment risk channel dominates the volatility channel, as it happens in the ICAPM. Exploring these possibilities is an interesting avenue for future research, but it is outside the scope of this paper.

6 Conclusion

In this paper, I empirically study the dividend and bond term structures from the perspective of intertemporal asset pricing theory, which features investors who care not only about their current

wealth, but also about their wealth reinvestment rate. The key new finding is that reinvestment risk provides a simple explanation for the *downward* sloping dividend term structure (at long maturities) recently documented in the literature while simultaneously producing an *upward* sloping bond term structure, consistent with historical data. The underlying reason is that long-term dividend claims hedge reinvestment risk while long-term bonds are exposed to such risk.

The findings in this paper have important implications beyond providing an empirically credible explanation for the dividend and bond term structures. First, the results suggest that equity cash flows accruing further in the future should require lower discount rates in present value applications because they hedge reinvestment risk. As such, this paper uncovers a relative advantage of longer-term investments projects with higher cash flow duration. This result is consequential because many real and financial investment opportunities differ dramatically in the timing of their expected cash flows. Hence, ignoring term structure differences in discount rates can lead to severe capital misallocation.

Second, the higher intertemporal hedging value of longer-term dividend claims over several decades indicates that the dividend term structure recently documented in the literature is likely to be an equilibrium outcome as opposed to an anomaly in the recent years. This result is important because the dividend term structure has been used as an out-of-sample test for asset pricing models, which could be misleading if the term structure facts were specific to the short period of data available on dividend contracts (Cochrane (2017)).

Third, my results demonstrate that longer-term dividend claims have intertemporal characteristics that make them more attractive to long-term investors. In particular, their prices tend to go up when investors' portfolios are expected to perform poorly in the future. Thus, portfolio managers interested in boosting returns by investing in short-term dividend strategies (such as the new dividend-focused ETFs) should consider how affected their investors are by decreases in reinvestment rates. This consideration is particularly important for pension funds since a significant fraction of retirees' wealth is invested for the long-run, which means that declines in reinvestment rates induce a substantial reduction in lifetime utility.

However, this paper also raises important new questions. First, are the premia for short-duration stocks (Dechow, Sloan, and Soliman (2004); Lettau and Wachter (2007, 2011); Hansen, Heaton, and Li (2008); Da (2009); Weber (2017)) and the low discount rates of long-term real estate assets (Giglio, Maggiori, and Stroebel (2015); Giglio et al. (2015)) consistent with the reinvestment risk

explanation for the dividend term structure? Second, can we use the methodology provided in this paper to empirically test alternative models and explore new properties of the dividend term structure? Third, how should firms incorporate reinvestment risk when making capital budgeting decisions? These are examples of potential avenues for future research related to the term structure of discount rates.

References

- Ai, Hengjie, Mariano M. Croce, Anthony M. Diercks, and Kai Li, 2017, News shocks and production-based term structure of equity returns, Working Paper.
- Amihud, Yakov, and Clifford M. Hurvich, 2004, Predictive regressions: A reduced-bias estimation method, *Journal of Financial and Quantitative Analysis* 39, 813–841.
- Andries, Marianne, Thomas M. Eisenbach, and Martin C. Schmalz, 2015, Asset pricing with horizon-dependent risk aversion, Staff Reports 703, Federal Reserve Bank of New York.
- Ang, Andrew, and Jun Liu, 2004, How to discount cashflows with time-varying expected returns?, *Journal of Finance* 59, 2745–2783.
- Ang, Andrew, and Maxim Ulrich, 2012, Nominal bonds, real bonds, and equity, Working Paper.
- Bali, Turan G., and Robert F. Engle, 2010, The intertemporal capital asset pricing model with dynamic conditional correlations, *Journal of Monetary Economics* 57, 377–390.
- Bansal, Ravi, Dana Kiku, Ivan Shaliastovich, and Amir Yaron, 2014, Volatility, the macroeconomy, and asset prices, *Journal of Finance* 69, 2471–2511.
- Bansal, Ravi, Dana Kiku, and Amir Yaron, 2012, An empirical evaluation of the long-run risks model for asset prices, *Critical Finance Review* 1, 183–221.
- Bansal, Ravi, and Amir Yaron, 2004, Risks for the long run: A potential resolution of asset pricing puzzles, *Journal of Finance* 59, 1481–1509.
- Barro, Robert J., 2006, Rare disasters and asset markets in the twentieth century, *Quarterly Journal of Economics* 121, 823–866.

- Belo, Frederico, Pierre Collin-Dufresne, and Robert S. Goldstein, 2015, Dividend dynamics and the term structure of dividend strips, *Journal of Finance* 70, 1115–1160.
- Bianchi, Daniele, and Andrea Tamoni, 2016, The dynamics of expected returns: Evidence from multi-scale time series modeling, Working Paper.
- Binsbergen, Jules H. van, Michael W. Brandt, and Ralph S. J. Koijen, 2012, On the timing and pricing of dividends, *American Economic Review* 102, 1596–1618.
- Binsbergen, Jules H. van, Wouter Hueskes, Ralph Koijen, and Evert Vrugt, 2013, Equity yields, *Journal of Financial Economics* 110, 503–519.
- Binsbergen, Jules H. van, and Ralph S. J. Koijen, 2010, Predictive regressions: A present-value approach, *Journal of Finance* 65.
- Binsbergen, Jules H. van, and Ralph S. J. Koijen, 2017, The term structure of returns: Facts and theory, *Journal of Financial Economics* 124, 1–21.
- Binsbergen, Jules H. van, and Ralph S.J. Koijen, 2016, On the timing and pricing of dividends: Reply, Working Paper.
- Bogue, Marcus C., and Richard Roll, 1974, Capital budgeting of risky projects with "imperfect" markets for physical capital, *Journal of Finance* 29, 601–613.
- Boguth, Oliver, Murray Carlson, Adlai Fisher, and Mikhail Simutin, 2011, Dividend strips and the term structure of equity risk premia: A case study of limits to arbitrage, Working Paper.
- Boguth, Oliver, Murray Carlson, Adlai Fisher, and Mikhail Simutin, 2012, Leverage and the limits of arbitrage pricing: Implications for dividend strips and the term structure of equity risk premia, Working Paper.
- Boudoukh, Jacob, Roni Michaely, Matthew Richardson, and Michael R. Roberts, 2007, On the importance of measuring payout yield: Implications for empirical asset pricing, *Journal of Finance* 62, 877–915.
- Brennan, Michael J., 1973, An approach to the valuation of uncertain income streams, *Journal of Finance* 28, 661–674.

- Brennan, Michael J., 1997, The term structure of discount rates, *Financial Management* 26, 81–90.
- Brennan, Michael J., 1998, Stripping the S&P 500 index, *Financial Analysts Journal* 54, 12–22.
- Brennan, Michael J., Ashley W. Wang, and Yihong Xia, 2004, Estimation and test of a simple model of intertemporal capital asset pricing, *Journal of Finance* 59, 1743–1775.
- Brennan, Michael J., and Yihong Xia, 2006, Risk and valuation under an intertemporal capital asset pricing model, *Journal of Business* 79, 1–36.
- Campbell, John Y., 1987, Stock returns and the term structure, *Journal of Financial Economics* 18, 373–399.
- Campbell, John Y., 1991, A variance decomposition for stock returns, *The Economic Journal* 101, 157–179.
- Campbell, John Y., 1993, Intertemporal asset pricing without consumption data, *American Economic Review* 83, 487–512.
- Campbell, John Y., 1996, Understanding risk and return, *Journal of Political Economy* 104, 298–345.
- Campbell, John Y., 2018, *Financial Decisions and Markets: A Course in Asset Pricing*, chapter 9.4, The Term Structure of Risky Assets, 290–295 (Princeton University Press).
- Campbell, John Y., and John Ammer, 1993, What moves the stock and bond markets? a variance decomposition for long-term asset returns, *Journal of Finance* 48, 3–37.
- Campbell, John Y., Yeung Lewis Chan, and Luis M. Viceira, 2003, A multivariate model of strategic asset allocation, *Journal of Financial Economics* 67, 41–80.
- Campbell, John Y., and John H. Cochrane, 1999, By force of habit: A consumption-based explanation of aggregate stock market behavior, *Journal of Political Economy* 107, 205–251.
- Campbell, John Y., Stefano Giglio, Christopher Polk, and Robert Turley, 2017, An intertemporal CAPM with stochastic volatility, *Journal of Financial Economics* Forthcoming.
- Campbell, John Y., Christopher Polk, and Tuomo Vuolteenaho, 2009, Growth or glamour? fundamentals and systematic risk in stock returns, *Review of Financial Studies* 23, 305–344.

- Campbell, John Y., and Robert J. Shiller, 1988, Stock prices, earnings, and expected dividends, *Journal of Finance* 43, 661–676.
- Campbell, John Y., and Robert J. Shiller, 1989, The dividend-price ratio and expectations of future dividends and discount factors, *Review of Financial Studies* 1, 195–228.
- Campbell, John Y., and Luis M. Viceira, 1999, Consumption and portfolio decisions when expected returns are time varying, *The Quarterly Journal of Economics* 114, 433–495.
- Campbell, John Y., and Tuomo Vuolteenaho, 2004, Bad beta, good beta, *American Economic Review* 94, 1249–1275.
- Cejnek, Georg, and Otto Randl, 2016, Risk and return of short-duration equity investments, *Journal of Empirical Finance* 36, 181–198.
- Cejnek, Georg, and Otto Randl, 2017, Dividend risk premia, Working Paper.
- Chen, Long, 2009, On the reversal of return and dividend growth predictability: A tale of two periods, *Journal of Financial Economics* 92, 128–151.
- Chen, Long, Zhi Da, and Xinlei Zhao, 2013, What drives stock price movements?, *Review of Financial Studies* 26, 841–876.
- Chen, Long, and Xinlei Zhao, 2009, Return decomposition, *Review of Financial Studies* 22, 5213–5249.
- Cochrane, John H., 2017, Macro-finance, *Review of Finance* Forthcoming.
- Croce, Mariano M., Martin Lettau, and Sydney C. Ludvigson, 2014, Investor information, long-run risk, and the term structure of equity, *Review of Financial Studies* 28, 706–742.
- Da, Zhi, 2009, Cash flow, consumption risk, and the cross-section of stock returns, *Journal of Finance* 64, 923–956.
- Dechow, Patricia M., Richard G. Sloan, and Mark T. Soliman, 2004, Implied equity duration: A new measure of equity risk, *Review of Accounting Studies* 9, 197–228.
- Doh, Taeyoung, and Shu Wu, 2016, The equilibrium term structure of equity and interest rates, Technical Report RWP 16-11, The Federal Reserve Bank of Kansas City.

- Duffee, Gregory R., 2013, Bond pricing and the macroeconomy, in George M. Constantinides, Milton Harris, and René M. Stulz, eds., *Handbook of the Economics of Finance*, volume 2, first edition, chapter 13, 907–967 (Elsevier Science).
- Eisenbach, Thomas M., and Martin C. Schmalz, 2016, Anxiety in the face of risk, *Journal of Financial Economics* 121, 414–426.
- Epstein, Larry G., and Stanley E. Zin, 1989, Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework, *Econometrica* 57, 937–969.
- Fama, Eugene F., 1977, Risk-adjusted discount rates and capital budgeting under uncertainty, *Journal of Financial Economics* 5, 3–24.
- Fama, Eugene F., 1981, Stock returns, real activity, inflation, and money, *The American Economic Review* 71, 545–565.
- Fama, Eugene F., and Kenneth R. French, 1989, Business conditions and expected returns on stocks and bonds, *Journal of Financial Economics* 25, 23–49.
- Fama, Eugene F., and Kenneth R. French, 1992, The cross-section of expected stock returns, *Journal of Finance* 47, 427–465.
- Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3–56.
- Fama, Eugene F., and Kenneth R. French, 1996, Multifactor explanations of asset pricing anomalies, *Journal of Finance* 55, 55–84.
- Fama, Eugene F., and G. William Schwert, 1977, Asset returns and inflation, *Journal of Financial Economics* 5, 115–146.
- Favilukis, Jack, and Xiaoji Lin, 2016, Wage rigidity: A quantitative solution to several asset pricing puzzles, *Review of Financial Studies* 29, 148–192.
- Ferson, Wayne E., 1989, Changes in expected security returns, risk, and the level of interest rates, *Journal of Finance* 44, 1191–1217.
- Ferson, Wayne E., and Campbell R. Harvey, 1999, Conditioning variables and the cross section of stock returns, *Journal of Finance* 54, 1325–1360.

- Gabaix, Xavier, 2012, Variable rare disasters: an exactly solved framework for ten puzzles in macro-finance, *Quarterly Journal of Economics* 127, 645–700.
- Geske, Robert, and Richard Roll, 1983, The fiscal and monetary linkage between stock returns and inflation, *Journal of Finance* 38, 1–33.
- Giglio, Stefano, and Bryan Kelly, 2017, Excess volatility: Beyond discount rates, *Quarterly Journal of Economics* forthcoming.
- Giglio, Stefano, Matteo Maggiori, and Johannes Stroebel, 2015, Very long-run discount rates, *Quarterly Journal of Economics* 130, 1–53.
- Giglio, Stefano, Matteo Maggiori, Johannes Stroebel, and Andreas Weber, 2015, Climate change and long-run discount rates: Evidence from real estate, Working Paper.
- Golez, Benjamin, 2014, Expected returns and dividend growth rates implied by derivative markets, *Review of Financial Studies* 27, 790–822.
- Gordon, Myron J., 1962, The savings investment and valuation of a corporation, *The Review of Economics and Statistics* 44, 37–51.
- Gorton, Gary, and Andrew Metrick, 2012, Securitized banking and the run on repo, *Journal of Financial Economics* 104, 425–451.
- Hansen, Lars Peter, John C. Heaton, and Nan Li, 2008, Consumption strikes back? measuring long-run risk, *Journal of Political Economy* 116, 260–302.
- Hansen, Lars Peter, and Kenneth J. Singleton, 1982, Generalized instrumental variables estimation of nonlinear rational expectations models, *Econometrica* 50, 1269–1286.
- Hasler, Michael, and Roberto Marfè, 2016, Disaster recovery and the term structure of dividend strips, *Journal of Financial Economics* 122, 116–134.
- Jordà, Òscar, 2005, Estimation and inference of impulse responses by local projections, *American Economic Review* 95, 161–182.
- Keim, Donald B., and Robert F. Stambaugh, 1986, Predicting returns in the stock and bond markets, *Journal of Financial Economics* 17, 357–390.

- Koijen, Ralph S.J., and Stijn Van Nieuwerburgh, 2011, Predictability of returns and cash flows, *Annual Review of Financial Economics* 3, 467–491.
- Kragt, Jac., Frank de Jong, and Joost Driessen, 2016, The dividend term structure, Netspar Discussion Paper 11/2014-055.
- Larrain, Borja, and Motohiro Yogo, 2008, Does firm value move too much to be justified by subsequent changes in cash flow?, *Journal of Financial Economics* 87.
- Lettau, Martin, and Stijn Van Nieuwerburgh, 2007, Reconciling the return predictability evidence, *Review of Financial Studies* 41, 1607–1652.
- Lettau, Martin, and Jessica A. Wachter, 2007, Why is long-horizon equity less risky? a duration-based explanation of the value premium, *Journal of Finance* 62, 55–92.
- Lettau, Martin, and Jessica A. Wachter, 2011, The term structures of equity and interest rates, *Journal of Financial Economics* 101, 90–113.
- Li, Ye, and Chen Wang, 2017, Rediscover predictability: Information from the relative prices of long-term and short-term dividends, Working Paper.
- Lintner, John, 1965, The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets, *Review of Economics and Statistics* 47, 13–37.
- Lynch, Anthony W., and Jessica A. Wachter, 2013, Using samples of unequal length in generalized method of moments estimation, *Journal of Financial and Quantitative Analysis* 48, 277–307.
- Malloy, Christopher J., Tobias J. Moskowitz, and Annette Vissing-Jørgensen, 2009, Long-run stockholder consumption risk and asset returns, *Journal of Finance* 64, 2427–2479.
- Manley, Richard, and Christian Mueller-Glissmann, 2008, The market for dividends and related investment strategies, *Financial Analysts Journal* 64, 17–29.
- Marfè, Roberto, 2015, Corporate fraction and the equilibrium term structure of equity risk, *Review of Finance* 20, 855–905.
- Marfè, Roberto, 2017, Income insurance and the equilibrium term-structure of equity, *Journal of Finance* Forthcoming.

- Mehra, Rajnish, and Edward C. Prescott, 1985, The equity premium: A puzzle, *Journal of Monetary Economics* 15, 145–161.
- Melino, Angelo, and Alan X. Yang, 2003, State-dependent preferences can explain the equity premium puzzle, *Review of Economic Dynamics* 6, 806–830.
- Merton, Robert C., 1973, An intertemporal capital asset pricing model, *Econometrica* 41, 867–887.
- Myers, Stewart C., and Stuart M. Turnbull, 1977, Capital budgeting and the capital asset pricing model: Good news and bad news, *Journal of Finance* 32, 321–333.
- Newey, Whitney K., and Kenneth D. West, 1987, A simple, positive-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* 55, 703–708.
- Petkova, Ralitsa, 2006, Do the fama-french factors proxy for innovations in predictive variables?, *Journal of Finance* 61, 581–612.
- Piazzesi, Monika, 2010, Affine term structure models, in Yacine Ait-Sahalia, and Lars Peter Hansen, eds., *Handbook of Financial Econometrics*, volume 1, chapter 12, 691–766 (Elsevier).
- Pope, Alun Lloyd, 1990, Biases of estimators in multivariate non-gaussian autoregressions, *Journal of Time Series Analysis* 11, 249–258.
- Rebonato, Riccardo, and Peter Jäckel, 1999, The most general methodology for creating a valid correlation matrix for risk management and option pricing purposes, *Journal of Risk* 2, 17–27.
- Sabbatucci, Riccardo, 2015, Are dividends and stock returns predictable? new evidence using M&A cash flows, Working Paper.
- Schulz, Florian, 2016, On the timing and pricing of dividends: Comment, *American Economic Review* 106, 3185–3223.
- Shanken, Jay, 1990, Intertemporal asset pricing: An empirical investigation, *Journal of Econometrics* 45, 99–120.
- Sharpe, William F., 1964, Capital asset prices: A theory of market equilibrium under conditions of risk, *Journal of Finance* 19, 425–442.
- Shumway, Tyler, 1997, The delisting bias in crsp data, *Journal of Finance* 52, 327–340.

- Sialm, Clemens, 2009, Tax changes and asset pricing, *American Economic Review* 99, 1356–1383.
- Song, Yang, 2016, Dealer funding costs: Implications for the term structure of dividend risk premia, Working Paper.
- Stambaugh, Robert F., 1999, Predictive regressions, *Journal of Financial Economics* 54, 375–421.
- Stulz, René M., 1986, Asset pricing and expected inflation, *Journal of Finance* 41, 209–223.
- Wachter, Jessica A., 2013, Can time-varying risk of rare disasters explain aggregate stock market volatility?, *Journal of Finance* 68, 987–1035.
- Weber, Michael, 2017, Cash flow duration and the term structure of equity returns, *Journal of Financial Economics* Forthcoming.
- Weil, Philippe, 1989, The equity premium puzzle and the risk-free rate puzzle, *Journal of Monetary Economics* 24, 401–421.
- Wilkins, Sascha, and Jens Wimschulte, 2010, The pricing of dividend futures in the european market: A first empirical analysis, *Journal of Derivatives & Hedge Funds* 16, 136–143.

APPENDIX

Appendix A Technical Derivations

A.1 Present Value Term Structure Decomposition

This subsection derives the term structure decompositions of equity returns, risk, and risk premium in equations 5, 6, and 7:

$$\tilde{r}_{e,t} \simeq \sum_{h=1}^{\infty} w_h \cdot \tilde{r}_{pv,t}^{(h)} \quad (5)$$

$$\beta_{e,t} \simeq \sum_{h=1}^{\infty} w_h \cdot \beta_{pv,t}^{(h)} \quad (6)$$

$$\mathbb{E}_t [R_{e,t+1} - R_{f,t+1}] \simeq \sum_{h=1}^{\infty} w_h \cdot \beta_{pv,t}^{(h)} \cdot \lambda_t \quad (7)$$

I start from a log-linear approximation to unexpected stock returns ([Campbell and Shiller \(1989\)](#); [Campbell \(1991\)](#)):

$$\tilde{r}_{e,t} \simeq \tilde{\Delta}d_t + N_{g,t} - N_{ep,t} - N_{ir,t} \quad (A.1)$$

where

$$N_{g,t} = (\mathbb{E}_t - \mathbb{E}_{t-1}) \left[\sum_{h=1}^{\infty} \rho^h \cdot \Delta d_{t+h} \right] \quad \text{are real dividend growth news}$$

$$N_{ep,t} = (\mathbb{E}_t - \mathbb{E}_{t-1}) \left[\sum_{h=1}^{\infty} \rho^h \cdot (r_{e,t+h} - r_{f,t+h}) \right] \quad \text{are equity premium news}$$

$$N_{ir,t} = (\mathbb{E}_t - \mathbb{E}_{t-1}) \left[\sum_{h=1}^{\infty} \rho^h \cdot r_{f,t+h} \right] \quad \text{are real interest rate news}$$

Comparing equations A.1 (derived in [Campbell \(1991\)](#)) and 5 (derived in Section 1.2), we have that the $\tilde{\Delta}d_t$ component cancels out in equation 5. Thus, showing that

$$\lim_{H \rightarrow \infty} \left(\sum_{h=1}^H \rho^h \cdot A_h \right) = \lim_{H \rightarrow \infty} \left(\sum_{h=1}^H w_{h+1} \cdot \sum_{j=1}^h A_j \right)$$

holds for a generic variable A is sufficient to prove equation 5.

Start by reorganizing the right hand side of this equation as:

$$\sum_{h=1}^H w_{h+1} \cdot \sum_{j=1}^h A_j = A_1 \cdot \left(\sum_{h=1}^H w_{h+1} \right) + A_2 \cdot \left(\sum_{h=2}^H w_{h+1} \right) + \dots + A_H \cdot w_{H+1}$$

Now, use $w_h = \rho^{h-1} - \rho^h$ and the formula for the finite sum of a geometric progression to get:

$$\sum_{h=1}^H w_{h+1} \cdot \sum_{j=1}^h A_j = \sum_{h=1}^H \left(\frac{\rho^h - \rho^{h+1} + \rho^{H+1} - \rho^H}{1 - \rho} \right) \cdot A_h$$

As H goes to infinity, the $\rho^{H+1} - \rho^H$ term vanishes since $\rho < 1$ and, thus, all we need to show is that $\frac{\rho^h - \rho^{h+1}}{1 - \rho} = \rho^h$, which can be demonstrated by multiplying both sides of this equation by $1 - \rho$. Hence, we have our result and equation 5 holds.

Equation 6 follows directly from subtracting the real interest rate and taking covariance with respect to $-m$ on both sides of equation 5. To derive equation 7, start by noting that the approximation $\mathbb{E}_t [r_{j,t+1}] + 0.5 \cdot \sigma_j^2 \simeq \mathbb{E}_t [R_{j,t+1}] - 1$ and the non-arbitrage condition, $\mathbb{E}_t [\exp \{m_{t+1} + r_{j,t+1}\}] = 1$, imply $\mathbb{E}_t [R_{e,t+1} - R_{f,t+1}] = \beta_{e,t} \cdot \lambda_t$. Then, substituting equation 6 into this expression yields equation 7 and concludes all derivations.

To help visualize the equity return term structure decomposition, Figure A.1 (in Appendix E.2) shows the weights, $w_h = \rho^{h-1} - \rho^h$, for h varying from 1 to 100 years when $\rho = 0.97$ (the value estimated from my sample period and used in Section 3). The weights decrease exponentially as in a weighted moving average process since $w_h = \rho^{h-1} \cdot (1 - \rho)$.

A.2 Relationship Between Dividend PV Returns and Dividend Strip Returns

This subsection derives the relationship between dividend PV returns and dividend strip returns (see Section 1). The key result is that the only difference between the two types of dividend returns is that shocks to future market excess returns have different impact on the two types of dividend returns.

Let $P_t^{(h)}$ represent the h -year dividend strip price. Defining its gross return as $R_{ds,t}^{(h)} = P_t^{(h-1)} / P_{t-1}^{(h)}$, solving prices forward and imposing the boundary condition, $P_t^{(0)} = D_t$, we have:

$$\begin{aligned} \ln \left(P_t^{(h)} \right) &= \ln (D_t) + \mathbb{E}_t \left[\sum_{j=1}^h \Delta d_{t+j} \right] - \mathbb{E}_t \left[\sum_{j=1}^h r_{ds,t+j}^{(h-j)} - r_{t+j}^f \right] - \mathbb{E}_t \left[\sum_{j=1}^h r_{t+j}^f \right] \\ &= \ln (D_t) + \mathbb{E}_t \left[\sum_{j=1}^h \Delta d_{t+j} \right] - \mathbb{E}_t \left[\sum_{j=1}^h \frac{\mathbb{E}_t \left[r_{ds,t+j}^{(h-j)} - r_{t+j}^f \right]}{\mathbb{E}_t \left[r_{e,t+j} - r_{t+j}^f \right]} \cdot \left(r_{e,t+j} - r_{t+j}^f \right) \right] - \mathbb{E}_t \left[\sum_{j=1}^h r_{t+j}^f \right] \\ &= \ln (D_t) + \underbrace{\mathbb{E}_t \left[\sum_{j=1}^h \Delta d_{t+j} \right]}_{g_t^{(h)}} - \underbrace{\mathbb{E}_t \left[\sum_{j=1}^h \psi_t (j, h) \cdot \left(r_{e,t+j} - r_{t+j}^f \right) \right]}_{ep_{ds,t}^{(h)}} - \underbrace{\mathbb{E}_t \left[\sum_{j=1}^h r_{t+j}^f \right]}_{ir_t^{(h)}} \quad (\text{A.2}) \end{aligned}$$

The only distinction between $P_t^{(h)}$ and $PV_t^{(h)}$ (i.e., equations 8 and A.2) is the equity premium component. Equation A.2 can be combined with the definition of gross returns to yield unexpected log real returns:

$$\tilde{r}_{ds,t}^{(h)} = \tilde{\Delta}d_t + N_{g,t}^{(h-1)} - N_{ds,ep,t}^{(h-1)} - N_{ir,t}^{(h-1)} \quad (\text{A.3})$$

where $N_{ds,ep,t}^{(h)} = ep_{ds,t}^{(h)} - \mathbb{E}_{t-1}[ep_{ds,t}^{(h)}]$.

A comparison of equations A.3 and 9 shows that dividend PV returns and dividend strip returns are very similar. The only distinction between them is that equity premium shocks are weighted differently when calculating returns based on dividend prices. With a flat term structure ($\mathbb{E}_t[r_{ds,t+j}^{(h)} - r_{t+j}^f] = \mathbb{E}_t[r_{e,t+j} - r_{t+j}^f]$), we have that returns based on present values are identical to the ones based on prices. With a downward sloping term structure, the relative weights $\psi_t(j, h)$ vary over time and tend to decrease in $h - j$, but their weighted average across j s converges to one as h goes to infinity. This result indicates that, even when the term structure is not flat, the dividend PV returns and dividend strip returns are connected.

A.3 Dividend Strip Prices in the ICAPM with a Strategic Investor

In this subsection, I derive equations 25a to 25d. Let $r_{f,t} = \mathbf{1}'_{rf}z_t$ and $\Delta d = \mathbf{1}'_{\Delta d}z_t$ (we can always include Δd and assign and impose VAR coefficients to assign zero loadings to it if Δd is assumed to have no predictive power).

Start by applying equation 24a to dividend strips:

$$\mathbb{E}_t \left[xr_{ds,t+1}^{(h)} \right] + \frac{1}{2} (\sigma_{ds,h}^2 - \sigma_f^2) = rra \cdot Cov_t \left(xr_{ds,t+1}^{(h)}, r_{p,t+1} \right) + \frac{\theta}{ies} \cdot Cov_t \left(xr_{ds,t+1}^{(h)}, cw_{t+1} \right) \quad (\text{A.4})$$

From the definition of a dividend strip gross return, $R_{ds,t+1}^{(h)} = P_{t+1}^{(h-1)}/P_t^{(h)}$, we have that log excess returns are given by:

$$xr_{ds,t+1}^{(h)} = \Delta d_{t+1} - r_{f,t+1} + pd_{t+1}^{(h-1)} - pd_t^{(h)} \quad (\text{A.5})$$

$$= \pi_{0,pd}^{(h-1)} + \left[\mathbf{1}_{\Delta d} - \mathbf{1}_{rf} + \pi_{1,pd}^{(h-1)} \right]' s_{t+1} - pd_t^{(h)} \quad (\text{A.6})$$

where the second equality follows from the conjecture (which I later verify) that $pd_t^{(h-1)} = \pi_{0,pd}^{(h-1)} + \pi_{1,pd}^{(h-1)'} z_t$.

The risk free variance is $Var_t(r_{f,t+1}) = \sigma_f^2 = \mathbf{1}'_{rf} \Sigma \mathbf{1}_{rf}$. From equation A.6, we have:

$$\begin{aligned}\sigma_{ds,h}^2 &= Var_t \left(r_{ds,t+1}^{(h)} \right) \\ &= \left[\mathbf{1}_{\Delta d} + \pi_{1,pd}^{(h-1)} \right]' \Sigma \left[\mathbf{1}_{\Delta d} + \pi_{1,pd}^{(h-1)} \right]\end{aligned}$$

as well as

$$\begin{aligned}Cov_t \left(xr_{ds,t+1}^{(h)}, r_{p,t+1} \right) &= Cov_t \left(\left[\mathbf{1}_{\Delta d} - \mathbf{1}_{rf} + \pi_{1,pd}^{(h-1)} \right]' s_{t+1}, r_{f,t+1} + \gamma_t' xr_{t+1} \right) \\ &= [\mathbf{1}_{xr} \gamma_t + \mathbf{1}_{rf}]' \Sigma \left[\mathbf{1}_{\Delta d} - \mathbf{1}_{rf} + \pi_{1,pd}^{(h-1)} \right] \\ &= [\mathbf{1}_{xr} \mathbf{\Gamma}_0 + \mathbf{1}_{rf}]' \Sigma \left[\mathbf{1}_{\Delta d} - \mathbf{1}_{rf} + \pi_{1,pd}^{(h-1)} \right] + [\mathbf{1}_{xr} \mathbf{\Gamma}_1 s_t]' \Sigma \left[\mathbf{1}_{\Delta d} - \mathbf{1}_{rf} + \pi_{1,pd}^{(h-1)} \right] \\ &= [\mathbf{1}_{xr} \mathbf{\Gamma}_0 + \mathbf{1}_{rf}]' \Sigma \left[\mathbf{1}_{\Delta d} - \mathbf{1}_{rf} + \pi_{1,pd}^{(h-1)} \right] + \left[\mathbf{1}_{\Delta d} - \mathbf{1}_{rf} + \pi_{1,pd}^{(h-1)} \right]' \Sigma [\mathbf{1}_{xr} \mathbf{\Gamma}_1] s_t \\ &= \pi_{0,p}^{(h)} + \pi_{1,p}^{(h)'} s_t\end{aligned}$$

and

$$\begin{aligned}Cov_t \left(xr_{ds,t+1}^{(h)}, cw_{t+1} \right) &= Cov_t \left(\left[\mathbf{1}_{\Delta d} - \mathbf{1}_{rf} + \pi_{1,pd}^{(h-1)} \right]' z_{t+1}, \mathbf{C}_0 + \mathbf{C}'_1 \cdot s_{t+1} + s'_{t+1} \mathbf{C}_2 s_{t+1} \right) \\ &= Cov_t \left(\left[\mathbf{1}_{\Delta d} - \mathbf{1}_{rf} + \pi_{1,pd}^{(h-1)} \right]' u_{t+1}, \left[\mathbf{C}'_1 + \Phi'_0 \mathbf{C}_2 + \Phi'_0 \mathbf{C}'_2 \right] u_{t+1} \right) \\ &\quad + Cov_t \left(\left[\mathbf{1}_{\Delta d} - \mathbf{1}_{rf} + \pi_{1,pd}^{(h-1)} \right]' u_{t+1}, \left[s'_t \Phi'_1 \mathbf{C}_2 + s'_t \Phi'_1 \mathbf{C}'_2 \right] u_{t+1} \right) \\ &= \left[\mathbf{C}'_1 + \Phi'_0 \left(\mathbf{C}_2 + \mathbf{C}'_2 \right) \right] \Sigma \left[\mathbf{1}_{\Delta d} - \mathbf{1}_{rf} + \pi_{1,pd}^{(h-1)} \right] \\ &\quad + \left[s'_t \Phi'_1 \left(\mathbf{C}_2 + \mathbf{C}'_2 \right) \right] \Sigma \left[\mathbf{1}_{\Delta d} - \mathbf{1}_{rf} + \pi_{1,pd}^{(h-1)} \right] \\ &= \left[\mathbf{C}'_1 + \Phi'_0 \left(\mathbf{C}_2 + \mathbf{C}'_2 \right) \right] \Sigma \left[\mathbf{1}_{\Delta d} - \mathbf{1}_{rf} + \pi_{1,pd}^{(h-1)} \right] \\ &\quad + \left[\mathbf{1}_{\Delta d} - \mathbf{1}_{rf} + \pi_{1,pd}^{(h-1)} \right] \Sigma \left[\left(\mathbf{C}'_2 + \mathbf{C}_2 \right) \Phi_1 \right] s_t \\ &= \pi_{0,cw}^{(h)} + \pi_{1,cw}^{(h)'} s_t\end{aligned}$$

Substituting the linear expressions for σ_f^2 , $\sigma_{ds,h}^2$, $Cov_t\left(xr_{ds,t+1}^{(h)}, r_{p,t+1}\right)$, and $Cov_t\left(xr_{ds,t+1}^{(h)}, cw_{t+1}\right)$ into equation A.4 yields:

$$\begin{aligned} \mathbb{E}_t \left[xr_{ds,t+1}^{(h)} \right] + \frac{1}{2} \cdot (\sigma_{ds,h}^2 - \sigma_f^2) &= rra \cdot \left[\pi_{0,p}^{(h)} + \pi_{1,p}^{(h)'} s_t \right] + \frac{\theta}{ieS} \cdot \left[\pi_{0,cw}^{(h)} + \pi_{1,cw}^{(h)'} s_t \right] \\ &+ \frac{1}{2} \cdot \left(\left[\mathbf{1}_{\Delta d} + \pi_{1,pd}^{(h-1)} \right]' \Sigma \left[\mathbf{1}_{\Delta d} + \pi_{1,pd}^{(h-1)} \right] - \mathbf{1}_{rf}' \Sigma \mathbf{1}_{rf} \right) \\ &= \pi_{0,\mathbb{E}r}^{(h)} + \pi_{1,\mathbb{E}r}^{(h)'} s_t \end{aligned}$$

which concludes the derivation of equations 25a to 25d as long as the conjecture $pd_t^{(h-1)} = \pi_{0,pd}^{(h-1)} + \pi_{1,pd}^{(h-1)'} s_t$ is correct.

To verify the conjecture, note that equation A.5 implies:

$$\begin{aligned} pd_t^{(h)} &= \mathbb{E}_t \left[pd_{t+1}^{(h-1)} \right] + \mathbb{E}_t [\Delta d_{t+1}] - \mathbb{E}_t [r_{f,t+1}] - \mathbb{E}_t \left[xr_{ds,t+1}^{(h)} \right] \\ &= \pi_{0,pd}^{(h-1)} - \pi_{0,\mathbb{E}r}^{(h)} + \left(\mathbf{1}_{\Delta d} - \mathbf{1}_{rf} + \pi_{1,pd}^{(h-1)} - \pi_{1,\mathbb{E}r}^{(h)} \right)' \left[\Phi_0 + \Phi_1 \cdot s_t \right] \\ &= \pi_{0,pd}^{(h)} + \pi_{1,pd}^{(h)'} s_t \end{aligned}$$

which verifies that as long as the conjecture is true for $h - 1$, it must be true for h .

Finally, note that the boundary condition is $pd_t^{(0)} = 0$, which means that the conjecture is true for $h = 0$. Therefore (from the result above) it must be true for all $h > 0$.

Appendix B Measurement and Data Sources

B.1 Main Variables

This subsection details the measurement and data sources I rely on for main analyses (Sections 3 and 4)

a) Risk Free Rate (r^f)

The annual interest rate over a given year is equal to the one year Treasury log yield as of the beginning of the period, which comes from Global Financial Data until 12/1951 and from the Center of Research in Security Prices (CRSP) Fama-Bliss discount bond file after that. For the period before December of 1940, the one year log yield is a linear interpolation of the respective yields at three months and ten years due to lack of data availability.

The one-year Treasury yield (and not the one-month Treasury bill rate) represents the risk-free rate in my analysis because I measure returns (and all flow variables) based on monthly observations of overlapping annual periods. Hence, the annual return on a one-month Treasury bill is not known as of the beginning of each period.

However, in my robustness analysis (subsection D.1), I also explore an alternative risk-free rate proxy based on annual returns on a one-month Treasury bill. The data is from Kenneth French's data library.

b) Equity Market Returns (r_e) and Dividend Growth (Δd)

Market returns and dividend growth are based on a value-weighted portfolio containing all common stocks available in the CRSP dataset and their measurement accounts for delistings. I do not use the CRSP value-weighted index because it includes all issues listed on NYSE, NASDAQ and AMEX with, on average, 5.3% of the market capitalization in the index referring to non common stock issues (see Sabbatucci (2015)). Moreover, accounting for delistings requires a "bottom-up" approach.

I start by adjusting returns for delistings. For each firm I can identify a delisting (delisting code available and different from 100), I adjust the (ex- and cum-dividend) return for the month in which the distribution of proceeds took place by assigning the delisting return to that month. If no delisting return is available, I base the delisting return on Shumway (1997)'s findings and assign to the delisting

month a -30% if the delisting was for cause (delisting code between 400 and 599) and 0% otherwise. I assign a 0% return to all months between delisting and distribution when there is a temporal gap between the two events. Despite being appropriate, accounting for delistings has almost no impact as the return correlation between my final equity portfolio and the CRSP value-weighted index is above 99.9%.

With ex- and cum-dividend returns accounting for delistings, I construct returns and dividends based on a value-weighted equity portfolio. I start by selecting all common shares (share codes 10 and 11) listed on NYSE, NASDAQ or AMEX (exchange code 1, 2, and 3). Then, I calculate value-weighted cum- and ex-dividend monthly returns. Based on the ex-dividend returns, I construct a normalized time series of prices and use $D_t^{monthly} = (R_t^{cum} - R_t^{ex}) \cdot P_{t-1}$ for this portfolio to get a time series of monthly dividends per share. Obtaining dividends from cum- and ex-dividend returns is a standard procedure in the literature (see [Kojien and Nieuwerburgh \(2011\)](#) and references therein).

Finally, annual dividends (D_t) are simply the sum of the monthly dividends over the respective period. I sum the dividend as opposed to reinvesting them into the stock market to avoid introducing properties of returns into dividend growth (see [Chen \(2009\)](#) and [Binsbergen and Kojien \(2010\)](#)). In my main analysis, I construct a monthly series of annual dividend growth from [Campbell and Shiller \(1989\)](#)'s log-linear stock return approximation: $\Delta d_t = r_{e,t} - (k - \rho \cdot dy_t + dy_{t-1})$ in which $dy_t = \ln(D_t/P_t)$ is the dividend yield. I measure dividend growth directly from [Campbell and Shiller \(1989\)](#)'s approximation because many of the ICAPM implications I study (as well as my dividend PV returns) rely on such approximation.

In my robustness analysis I measure annual dividend growth as $\Delta d = \ln(D_t/D_{t-12})$ and find almost identical results. I also explore an alternative dividend measurement that accounts for M&A activity (mergers and acquisitions) because [Sabbatucci \(2015\)](#) shows that M&A's paid in cash have an economically important effect both on dividend growth and on dividend yield.

The adjustment is simply to incorporate an extra step after the delisting adjustment, but before aggregating the ex- and sum-dividend returns. For the same firms/months adjusted for delisting returns, I check if the distribution was from an M&A paid with cash (distribution code between 3000 and 3400). If so, I subtract the distribution over lagged price from the ex-dividend return for that firm, limiting the minimum ex-dividend return to -100%. This procedure assures the dividend calculated from $D_{it} = (R_{it}^{cum} - R_{it}^{ex}) \cdot P_{it-1}$ contains the M&A cash payment.

Accounting for M&A activity has a non-trivial influence on dividend growth and dividend payout

measurement. For instance, I find a correlation of 61.5% between my dividend growth measure that account for M&A activity and the one that does not. Similarly the correlation between dividend yields using the two alternative dividend measurements is 93.2%. The dividend yield adjusted for M&A is also less persistent as the annual autocorrelation is 0.85 in comparison to the 0.90 for the dividend yield without M&A adjustment. All of these observations are consistent with [Sabbatucci \(2015\)](#), who argues that dividends should include cash originated from M&A activity. Despite M&A activity having a significant effect on the dividend growth measurement, my results are similar whether I adjust for M&A or not.

c) Predictive Variables ($z_t = [\Delta d \quad dy \quad epoy \quad ty \quad TS \quad CS \quad VS]$)

Dividend growth (Δd) and dividend yield (dy) have the same measurement as previously detailed in this subsection.

Equity payout yield ($epoy$) is the log of (one plus) aggregate net equity payout over market equity and only relies on CRSP data: $y_{epo} = \log(1 + NPO/ME)$. I first create a time-series of (normalized) market equity (ME) following a procedure similar to what was previously detailed for prices. The only difference is that accumulation is done using aggregate market equity growth as opposed to ex dividend returns (this accounts for equity issuances and repurchases). I then construct a time-series of monthly total dividends (TD) the same way I do it for dividends per share, but using ME as the basis. Finally, to get net payout (NPO), I use a procedure in the spirit of [Larrain and Yogo \(2008\)](#), but at the aggregate level. Specifically, I define the aggregate number of shares as $N = ME/P$ and find monthly net payout using $NPO_t = TD_t - (N_t - N_{t-1}) \cdot P_t$. Annual net payout is simply the sum of the monthly NPO 's over the respective period.

The term spread (TS) is the difference between the ten years and the one year log Treasury yields and the credit spread (CS) is the difference between Moody's corporate BAA and AAA log yields. All bond yields are obtained from Global Financial Data except for the one year Treasury yield (ty), which follows the interest rate construction explained previously (comes from CRSP Fama-Bliss discount bond file when available).

The value spread is the difference between the log book-to-market ratios of the value and growth portfolios in [Fama and French \(1993, 1996\)](#) HML factor (obtained from Kenneth French's data library). The value spread monthly observations are adjusted to account for within year movements in market equity (see the Internet Appendix of [Campbell and Vuolteenaho \(2004\)](#)). My use of the

HML portfolios (as opposed to focusing on small stocks as in [Campbell and Vuolteenaho \(2004\)](#)) assures the value spread behavior is not dominated by small stocks, but results are robust to using the small value spread (see subsection [D.1](#)).

B.2 Testing Assets (Including Bond Portfolios)

This section briefly describes the data sources for testing assets in subsection [3.5](#) and in the robustness analysis (subsection [D.1](#)).

For all returns described here, monthly observations of annual log returns are constructed from overlapping annual gross returns.

The equity market portfolio and risk-free rate follow the same construction as in the previous subsection. The book-to-market portfolios are based on [Fama and French \(1993\)](#)'s 3-by-2 sorts on size and book-to-market. Each book-to-market portfolio equal weights the respective two size sorted portfolios, which assures that the high minus low book-to-market portfolio in my analysis is exactly the HML value factor in [Fama and French \(1993\)](#) (data obtained from Kenneth French's data library). Data on Treasury bond portfolios (1952-2016) and corporate bond portfolios (1973-2016) are obtained from CRSP and Datastream respectively. The same bond portfolios are used by [Binsbergen and Koijen \(2017\)](#) to study the Treasury and corporate bond term structures.

Some specifications in the robustness analysis (subsection [D.1](#)) use alternative portfolios to capture the bond term structure or the value premium. The bond replacement portfolios are also from the CRSP database and represent seven zero-coupon nominal Treasury bond portfolios with maturities of 1, 2, 5, 7, 10, 20 and 30 years. The value premium replacement portfolios are five portfolios (quintiles) sorted on book-to-market only with the data obtained from Kenneth French's data library as well.

B.3 Short-Term Dividend Strategy

Excess returns on the short-term dividend strategy I study in subsection [3.6](#) are given by $xR^{(st-mkt)} = R_{df}^{(st)} - [R_{sp500} - R_b^{(10)}]$. Here, I provide further details on the construction of $R_{df}^{(st)}$. As described in subsection [3.6](#), the data comes from three sources: (i) a database of daily quoted prices for dividend futures (from 02-Jan-2006 to 14-Oct-2016) provided to me by Goldman Sachs; (ii) Datastream (from 15-Oct-2016 to 31-Dec-2016); and (iii) the data explored in [Binsbergen, Brandt, and Koijen \(2012\)](#) obtained directly from Ralph Koijen's data library (from Jan-1996 to Dec-2005).

For the first two datasets, I have quoted prices for dividend futures of consecutive maturities, which

allows me to directly calculate monthly returns on dividend futures. I follow [Binsbergen and Koijen \(2017\)](#) and obtain returns on the one- and two-year dividend futures by linearly interpolating the returns of the respective closest maturities (so each constant maturity dividend future is a portfolio of dividend futures). I get $R_{df}^{(st)}$ as an equal-weighted portfolio of the one- and two-year dividend futures and calculate annual overlapping returns by compounding monthly returns.

For the last dataset, I have a time-series of monthly returns representing a portfolio of dividend strips with duration ranging from 1.3 to 1.9 years.²⁵ From these returns and the returns on the 2-year bond portfolio of Section 3, I calculate annual overlapping returns on a long-short portfolio that buys the short-term dividend portfolio and forgoes the returns on the short-term bond portfolio (similar to a portfolio of dividend futures). This approach recognizes that no dividend future is available in the early period and simply switches (over that period) to a strategy designed to create the same type of exposure a portfolio of short-term dividend futures has.

Note that a fixed definition of $R_{df}^{(st)}$ is not a requirement to construct a short-term dividend strategy. We can consider an investor who builds a short-term dividend strategy in which the exact definition of the dividend position changes over time. That is how my short-term dividend strategy should be interpreted.

However, three potential issues are worth mentioning. Because my early period relies on data directly from [Binsbergen, Brandt, and Koijen \(2012\)](#), it is subject to the same criticisms as the original paper. The literature has elaborated on three issues. I detail them below:

a) Taxes

[Schulz \(2016\)](#) argues that the differential tax treatment of dividends and capital gains implies that it is important to adjust dividend strip returns by taxes and doing so has a significant effect on average dividend strip returns. While the earlier part of my sample (until Dec-2005) is potentially subject to this criticism, the later part is not as it is based on dividend futures, which are not subject to the dividend taxation (all profits are taxed as capital gain). Moreover, [Binsbergen and Koijen \(2016\)](#)

²⁵These monthly returns rely on non-arbitrage in the options market. Specifically, the put-call parity implies:

$$P_{ds}^{(0 \rightarrow h)} = \text{put}_h - \text{call}_h + \text{Stock Price} - \text{Strike Price} \cdot e^{-h \cdot y^{(h)}}$$

where $P_{ds}^{(0 \rightarrow h)}$ is the price of all dividends to be paid until the options expire (h years). This parity relation shows that purchasing the short-term dividend is equivalent to buying a put option, writing a call option, buying the stocks in the index, and borrowing cash. [Binsbergen, Brandt, and Koijen \(2012\)](#) construct dividend strip prices by matching each of the components in this equation (for S&P 500 contracts) within the same minute of the last trading day of each month. I rely on their dividend strategy 1 (as labeled in the original paper).

demonstrate that differential tax rates of dividends versus capital gains estimated from ex-dividend day returns (which [Schulz \(2016\)](#) relies on) are “suspect and imprecisely measured, peaking at over 100% in some periods.” The reason is that abnormal price changes around dividend payment dates are not only due to taxes, something [Schulz \(2016\)](#) recognizes (see [Binsbergen and Koijen \(2016\)](#) for a detailed description of other effects). In fact, a regression of Schultz’s differential tax series on a direct estimate of differential tax series (as reported in [Sialm \(2009\)](#)) produces an R^2 below 1%. Finally, if one relies on the tax rate differential in [Sialm \(2009\)](#), the adjustment is much weaker and it remains true that S&P 500 dividend strips outperform the index significantly in economic terms.

b) Dealer Funding Costs

[Song \(2016\)](#) argues that funding costs to derivative dealers’ shareholders for carrying and hedging inventories can induce failures of the put-call parity when the non-arbitrage relation is evaluated using standard proxies for interest rates (like the LIBOR rate used by [Binsbergen, Brandt, and Koijen \(2012\)](#) to get dividend prices). He provides an estimated adjustment to the put-call parity that takes into account this aspect and shows that dividend strip returns are substantially smaller using this adjustment (from January 2013 to January 2016). As with taxes, this criticism only applies to the earlier part of my sample, which uses dividend strip prices based on the put-call parity. While the mechanism provided in the paper is interesting, it is unclear whether its effect is relevant for my analysis.

There are two important issues to consider. First, the sample period in [Song \(2016\)](#) is economically different from the relevant one in my analysis. [Song \(2016\)](#) assumes repo rates are the appropriate measures of funding costs that include the extra cost for carrying and hedging inventories faced by derivative dealers’ shareholders. However, his empirical analysis goes from January 2013 to January 2016. Even if we take as given the claim that repo rates are the correct discount rates for options from 2013 to 2016, they may not be from 1996 to 2005 (period in which I rely on the put-call parity), before the financial crisis. Moreover, the supply of credit through repurchase agreements was substantially affected by the financial crisis (see [Gorton and Metrick \(2012\)](#)), which makes it unclear whether the spread from 2013 to 2016 is representative of the spread from 1996 to 2005 (the results in [Song \(2016\)](#) depend entirely on the spread between repo rates and LIBOR rates).

Second, the claim that repo rates are better measures of derivative dealers’ funding costs than LIBOR rates does not seem appropriate. On theoretical grounds, the cash position in the put-call-

parity formula is risk free, and, as such, the appropriate discount rate is the risk free government bond yield.²⁶ In empirical terms, investors can create a risk-free bond by buying a put option and writing a covered call on a non-dividend paying security (as pointed out by [Binsbergen and Koijen \(2016\)](#)). Given that options are exchange traded and therefore do not carry any bank-specific credit risk, the interest rate from such position is risk-free and also not directly affected by credit supply. As such, we can use it as a proxy for derivative dealers' funding costs (they can actually fund positions using this approach). [Golez \(2014\)](#) estimates implied interest rates in derivative markets from 1994 to 2008 by combining data on options and futures and finds that the interest rate used in U.S. option markets generally lies on top of, or below (in particular during the financial crisis), the LIBOR rate (see Figure 5 in the working paper version). Hence, if anything, the LIBOR rate is a conservative measure of derivative dealers' funding costs over the relevant sample period for my study (1996 to 2005).

c) Market Microstructure Noise

[Boguth et al. \(2011\)](#) argue that market microstructure noise can explain the empirical properties of dividend strips calculated using the put-call parity. In particular, they show that “negligible pricing frictions in underlying asset markets can become greatly magnified when using no-arbitrage arguments to price derivative claims”. This effect is particularly strong when the replicating portfolio contains partially offsetting positions that induce an amplification mechanism relative to the original market frictions. As with taxes and dealer funding costs, this argument only applies to my early sample (when the put-call parity is used).

To deal with potential market micro structure noise, [Boguth et al. \(2011\)](#) suggest using more robust measures of dividend strip returns. One of the suggestions is to use annual returns (as oppose to monthly) as they show that microstructure frictions are proportionately much smaller for annual returns. I follow their suggestion and use only annual returns for all statistics calculated (including β s and average returns). As such, I see my methodology as relatively immune to the significant market micro structure effects explored in [Boguth et al. \(2011\)](#).

²⁶[Binsbergen, Brandt, and Koijen \(2012\)](#) uses the LIBOR rate instead as it is conservative but [Binsbergen and Koijen \(2016\)](#) point out that the use of LIBOR rates that include the bank credit risk premium as the relevant interest rate in the option pricing literature is similar to the fallacy of using a company-wide (instead of a project-specific) discount rate when computing the net present value of a project.

Appendix C Econometric Details

C.1 Predictive System in Section 3

This subsection details how I construct the dividend risk and risk premia term structures based on dividend PV returns. The relevant equations are 14a, 14b, and 14c:

$$\Delta d_{t+h} = b_g^{(h)'} z_t + \epsilon_{g,t+h}^{(h)} \quad \forall h = 1, 2, \dots, 10 \quad (14a)$$

$$r_{e,t+h} - r_{f,t+h} = b_{ep}^{(h)'} z_t + \epsilon_{ep,t+h}^{(h)} \quad \forall h = 1, 2, \dots, 10 \quad (14b)$$

$$r_{f,t+h} = b_{ir}^{(h)'} z_t + \epsilon_{ir,t+h}^{(h)} \quad \forall h = 1, 2, \dots, 10 \quad (14c)$$

where z_t is a $k \times 1$ vector of predictive variables (with a constant as its first element).

Let $B_{ep}^{(h)} \equiv \sum_{j=1}^h b_{ep}^{(j)}$ be the sum of the equity premium regression coefficients up to horizon h and $B_{ep} \equiv \sum_{h=1}^{\infty} \rho^h \cdot b_{ep}^{(h)}$ be the infinite discounted sum of them (with analogous quantities for interest rates and dividend growth). Then, the ICAPM risk factors (with the market portfolios being the equity index) can be written as:

$$\tilde{r}_{m,t} = \epsilon_{ep,t}^{(1)} + \epsilon_{ir,t}^{(1)} \quad (A7a)$$

$$N_{ep,t} = B_{ep}' z_t - \frac{1}{\rho} \left(B_{ep} - \rho \cdot B_{ep}^{(1)} \right)' z_{t-1} \quad (A7b)$$

$$N_{ir,t} = B_{ir}' z_t - \frac{1}{\rho} \left(B_{ir} - \rho \cdot B_{ir}^{(1)} \right)' z_{t-1} \quad (A7c)$$

Similarly, equation 9 can be used to write dividend PV unexpected returns and its components as:²⁷

²⁷To build intuition for these equations, consider $N_{g,t}^{(1)}$. I do not assume a Vector Autoregressive (VAR) structure, but if I did (with Γ representing the VAR parameters), we could simplify equation A8b so that $N_{g,t}^{(1)}$ would be a linear combination of the VAR shocks (this is a general result):

$$N_{g,t}^{(1)} = \mathbf{1}_{\Delta d} \cdot \left[\Gamma' z_t - (\Gamma + \Gamma^2 - \Gamma)' z_{t-1} \right] = \mathbf{1}_{\Delta d} \cdot \Gamma' \cdot \left(z_t - \Gamma' z_{t-1} \right) = \mathbf{1}_{\Delta d} \cdot \Gamma' \cdot \epsilon_t$$

where $\mathbf{1}_{\Delta d}$ represents an indicator variable selecting Δd .

$$\tilde{\Delta}d_t = \epsilon_{g,t}^{(1)} \quad (\text{A8a})$$

$$N_{g,t}^{(h)} = B_g^{(h)'} z_t - \left(B_g^{(h+1)} - B_g^{(1)} \right)' z_{t-1} \quad (\text{A8b})$$

$$N_{ep,t}^{(h)} = B_{ep}^{(h)'} z_t - \left(B_{ep}^{(h+1)} - B_{ep}^{(1)} \right)' z_{t-1} \quad (\text{A8c})$$

$$N_{ir,t}^{(h)} = B_{ir}^{(h)'} z_t - \left(B_{ir}^{(h+1)} - B_{ir}^{(1)} \right)' z_{t-1} \quad (\text{A8d})$$

$$\tilde{r}_{pv,t}^{(h)} = \tilde{\Delta}d_t + N_{g,t}^{(h-1)} - N_{ep,t}^{(h-1)} - N_{ir,t}^{(h-1)} \quad (\text{A8e})$$

Next, I derived these expressions and also detail a misspecification adjustment I apply to them. The adjustment imposes equation 5 (which is equivalent to imposing the stock return approximation identity of Campbell (1991)), but is additive and fixed across hs , and thus has no term structure effect.

a) Getting Risk Factors

I derive equations A7b and A7c relating $N_{ep}^{(h)}$, and $N_{ir}^{(h)}$ to the system parameters.

From the definition of N_{ir} , we have (derivation is analogous for N_{ep}):

$$\begin{aligned} N_{ir,t} &= \mathbb{E}_t \left[\sum_{h=1}^{\infty} \rho^h \cdot r_{f,t+h} \right] - \mathbb{E}_{t-1} \left[\sum_{h=1}^{\infty} \rho^h \cdot r_{f,t+h} \right] \\ &= \sum_{h=1}^{\infty} \rho^h \cdot b_{ir}^{(h)'} z_t - \frac{1}{\rho} \cdot \sum_{h=1}^{\infty} \rho^{h+1} \cdot b_{ir}^{(h+1)'} z_{t-1} \\ &= \sum_{h=1}^h \rho^h \cdot b_{ir}^{(h)'} z_t - \frac{1}{\rho} \cdot \sum_{h=1}^{\infty} \rho^h \cdot b_{ir}^{(h)'} z_{t-1} - b_{ir}^{(1)'} z_{t-1} \\ &= \underbrace{B_{ir}'}_{A'_{ir,1}} z_t - \frac{1}{\rho} \underbrace{\left(B_{ir} - \rho \cdot B_{ir}^{(1)} \right)'}_{A'_{ir,0}} z_{t-1} \end{aligned}$$

which completes the derivation of equation A7c.

b) Getting Dividend PV Returns

I derive equations [A8b](#), [A8c](#), and [A8d](#) relating $N_g^{(h)}$, $N_{ep}^{(h)}$, and $N_{ir}^{(h)}$ to the system parameters and then substitute them (and equation [A8a](#)) on equation [A8e](#) to get dividend PV returns.

From the definition of $N_g^{(h)}$, we have (derivation is analogous for $N_{ep}^{(h)}$, and $N_{ir}^{(h)}$):

$$\begin{aligned}
N_{g,t}^{(h)} &= \mathbb{E}_t \left[\sum_{j=1}^h \Delta d_{t+j} \right] - \mathbb{E}_{t-1} \left[\sum_{j=1}^h \Delta d_{t+j} \right] \\
&= \sum_{j=1}^h b_g^{(j)'} z_t - \sum_{j=1}^h b_g^{(j+1)'} z_{t-1} \\
&= \sum_{j=1}^h b_g^{(j)'} z_t - \sum_{j=1}^{h+1} b_g^{(j)'} z_{t-1} - b_g^{(1)'} z_{t-1} \\
&= B_g^{(h)'} z_t - \left(B_g^{(h+1)} - B_g^{(1)} \right)' z_{t-1}
\end{aligned}$$

which completes the derivation of equation [A8b](#).

Then, defining $B_{pv}^{(h)} = B_g^{(h)} - B_{ep}^{(h)} - B_{ir}^{(h)}$, we can use equation [A8e](#) to get the expression linking dividend PV (unexpected) returns to the system parameters:

$$\begin{aligned}
\tilde{r}_{pv,t}^{(h)*} &= \epsilon_{g,t}^{(1)} + B_{pv}^{(h)'} z_t - \left(B_{pv}^{(h+1)} - B_{pv}^{(1)} \right)' z_{t-1} \\
&= \underbrace{\left(\mathbf{1}_{\Delta d} + B_{pv}^{(h)} \right)'}_{A_{pv,1}^{(h)*}} z_t - \underbrace{\left(b_g^{(1)} + B_{pv}^{(h+1)} - B_{pv}^{(1)} \right)'}_{A_{pv,0}^{(h)*}} z_{t-1}
\end{aligned}$$

where $\mathbf{1}_{\Delta d}$ is an indicator variable selecting Δd_t in z_t .

While this equation represents dividend strip unexpected returns, it is subject to specification error. For instance, the system does not impose equation [5](#), which is the key equation relating dividend PV returns to equity returns (this is equivalent to not imposing the log-linear approximation in [Campbell \(1991\)](#)).

To minimize the impact of misspecification error, I take a slightly different approach. I calculate dividend PV returns in excess of equity markets (as implied by $\tilde{r}_{pv,t}^{(h)*}$) and later combine this with actual equity returns. First, I get:

$$\tilde{r}_{pv,t}^{(h)*} - \tilde{r}_{e,t}^* = \left[A_{pv,1}^{(h)*'} z_t - A_{pv,0}^{(h)*'} z_{t-1} \right] - \sum_{h=1}^{\infty} w_h \cdot \tilde{r}_{pv,t}^{(h)*} \quad (\text{A.9})$$

and then I calculate:

$$\begin{aligned}\tilde{r}_{pv,t}^{(h)} &= \left(\tilde{r}_{pv,t}^{(h)*} - \tilde{r}_{e,t}^* \right) + \tilde{r}_{e,t} \\ &= \tilde{r}_{pv,t}^{(h)*} + \underbrace{\left(\tilde{r}_{e,t} - \tilde{r}_{e,t}^* \right)}_{\text{Adjustment}}\end{aligned}$$

In words, I calculate dividend PV returns and add a (identical) misspecification adjustment to all maturities so that equation 5 holds for $\tilde{r}_{pv,t}^{(h)}$ and $\tilde{r}_{e,t}$. If there is no misspecification in the original measure (i.e., equation 5 is satisfied for $\tilde{r}_{pv,t}^{(h)*}$), then the adjustment is zero. If there is misspecification, then there is an adjustment that imposes 5. Since the misspecification adjustment is additive and the same for all dividend PV returns, this approach has no impact on any term structure result I present other than a parallel shift to make dividend strip returns comparable to the market returns. For instance, the adjustment cancels out when looking at β s or risk premia differences across two dividend PV returns with different dividend maturities. This adjustment is not performed in Section 4 as the econometric structure I use in that section automatically imposes equation 5.

c) Getting Risk Exposures of Dividend Claims

As demonstrated above, dividend PV unexpected returns and risk factors can both be written as linear combinations of z_t and z_{t-1} . From these linear combinations, getting variances and covariaces is straightforward. For instance:

$$\begin{aligned}\text{Cov} \left(\tilde{r}_{pv,t}^{(h)}, N_{ir,t} \right) &= \text{Cov} \left(A_{pv,1}^{(h)'} z_t - A_{pv,0}^{(h)'} z_{t-1}, A_{ir,1}' z_t - A_{ir,0}' z_{t-1} \right) \\ &= A_{pv,1}^{(h)'} \text{VCov} (z_t, z_t) A_{ir,1} + A_{pv,0}^{(h)'} \text{VCov} (z_t, z_t) A_{ir,0} \\ &\quad - A_{pv,1}^{(h)'} \text{VCov} (z_t, z_{t-1}) A_{ir,0} - A_{pv,0}^{(h)'} \text{VCov} (z_{t-1}, z_t) A_{ir,1}\end{aligned}$$

Since β s are ratios of covariances and variances, they are also functions of the parameters of the system in equations 14a to 14c and the variance-covariance matrixes $\text{VCov} ([z_t \ z_{t-1}])$.

Even though all β s reported in the text are normalized to market β units (covariance over market variance), I still perform all statistical tests using the actual projection β s (i.e., covariance over the respective risk factor variance). Since variances are positive, one is zero if and only if the other also is, but projections have better statistical properties.

C.2 GMM Estimation of the Predictive System of Section 3

The GMM estimation of the predictive system is explained in subsection 3.1. There are, however, a few details that I only briefly mention in the main text. I elaborate on these details in this section.

a) Adjustment for Small Sample Bias

Predictive regressions as the ones in the given system suffer from small sample bias when the predictors are persistent (Stambaugh (1999)). The bias tends to be particularly severe when sum of coefficients are used (which is equivalent to long-run cumulative regressions typically used in the literature). While there is no obvious term structure direction in the bias, my estimation approach corrects for it (results not correcting for the bias are similar and presented in Appendix D.1).

To address the small sample bias issue, I use the method in Amihud and Hurvich (2004). It consists of first estimating the VAR system for z_t with a bias-free method and then adding the estimated residuals as control variables in the other predictive regressions. Amihud and Hurvich (2004) show that this approach is a natural generalization of the “Stambaugh bias correction” (often applied in the literature) to a setting with multiple predictors. To get a bias free estimate of the VAR system for z_t , I use the method in Pope (1990) and treat as independent observations only the total number of year (as opposed to the number of months), which is a conservative adjustment for the fact that I use overlapping annual observations when estimating the VAR.

b) Standard Errors

As pointed out in subsection 3.1, I use monthly frequency observations of annual variables to estimate the system in equations 14a to 14d, which means that my observations overlap for eleven months. I estimate standard errors accounting for that by designing a special approach to estimate the spectral density matrix.

The estimation is done in two steps. First, all moment condition residuals are aligned normally except for moment condition residuals associated with regressions in which $h > 1$ (i.e., residuals of $\hat{\mathbb{E}}[\epsilon_{t+2} \cdot (1 z_t)], \dots, \hat{\mathbb{E}}[\epsilon_{t+10} \cdot (1 z_t)]$). These residuals are aligned with other moment condition residuals based on the calendar month of $\epsilon_{t+2}, \dots, \epsilon_{t+10}$. For example, $\epsilon_{1/2010} \cdot (1 z_{1/2000})$ is aligned with $\epsilon_{1/2010} \cdot (1 z_{1/2001}), \dots, \epsilon_{1/2010} \cdot (1 z_{1/2009})$. Second, using the new alignment, I estimate the spectral moment matrix relying on Newey and West (1987) and accounting for autocorrelation up to

18 months.

In summary, this procedure allows for arbitrary correlation among any two moment condition residuals observed over same calendar period and accounts for up to 6 months of autocorrelation in moment condition residuals beyond the overlapping period.

It is worth pointing out that the asymptotic standard errors I obtain for the GMM are identical to asymptotic standard errors for the bias corrected estimator. The reason is that the bias correction converges to zero at the rate of the sample size, which means that the bias corrected estimator has the same asymptotic distribution as the GMM estimator with no correction.

C.3 *rra* Estimation

After the system in equations 14a to 14d is augmented by the moment conditions in 13, it becomes an overidentified system. As pointed out in subsection 3.5, I still estimate all parameters based on their respective (just identified) moment equations, except for the risk prices, which are estimated by solving $\hat{\lambda} = \underset{\lambda}{argmin} (\alpha' \alpha)$ with $\alpha_j = \hat{\mathbb{E}} [r_j - r_f] + \frac{1}{2} (\hat{\sigma}_j^2 - \hat{\sigma}_f^2) - \lambda' \cdot \hat{\beta}_j$ with hats indicating sample analogue quantities.

The given procedure is specified in terms of a unified GMM. Letting the full vector of moment conditions be given by $g(\theta)$ with θ representing all q parameters to be estimated, we have that any GMM procedure can be stated as $A \cdot g(\hat{\theta}) = 0$ (i.e., q linear combinations of the moment conditions are set to zero). My approach sorts the parameters in θ such that risk prices are at the end of the vector ($\theta = (\theta_0 \ \lambda)$) and specifies A to be block diagonal. The first diagonal block is an identity matrix and refers to all moment conditions used to estimate the parameters in θ_0 . This block simply sets each moment condition to zero.

The second diagonal block refers to the testing asset pricing errors, which represent the moment conditions used to estimate risk prices. If we represent the GMM weighting matrix by Ω , the second diagonal block is given by $\frac{\partial g(\lambda)'}{\partial \lambda} \cdot \Omega$ and defines the linear combinations of pricing errors that need to be set to zero to minimize the GMM objective function. In my main specification, I use an identity matrix, but in the robustness analysis I also explore two alternative diagonal matrices (details available in subsection D.1).

In some of the estimation steps, I have testing assets observed over different periods. For instance, returns on corporate bond portfolios are not available before 1972. The estimation procedure remains the same, but each moment condition is adjusted to depend only on the period in which it can be

calculated. All details for the adjustment can be found in [Lynch and Wachter \(2013\)](#) since I am using their so called “long estimator”.

To get standard errors, I estimate the spectral density matrix in two steps. In the first step, I demean the moment equation residuals (as suggested by [Hansen and Singleton \(1982\)](#)), realign residuals, and use [Newey and West \(1987\)](#) accounting for autocorrelation up to 18 months (as explained in subsection [C.2](#)). Since the testing asset moment equations have different sample lengths, I get all variance/covariance terms in the [Newey and West \(1987\)](#) approach using pairwise complete observations (as opposed to using the method in [Lynch and Wachter \(2013\)](#), which would not be valid given the overlapping observations in my framework).

This estimation approach means that the spectral density matrix estimate might not be positive definite depending on the specification. Since I need at least a positive semidefinite matrix to use the delta method for other parameters, I take a second step. In this step, I apply a transformation to the spectral density matrix that maintains the diagonal (i.e., standard errors) fixed and adjusts the correlations to make the final covariance matrix positive semidefinite. The correlation matrix transformation I use is explained in [Rebonato and Jäckel \(1999\)](#) and is only necessary in some specifications when risk prices are estimated. Asymptotically, the second step is not performed since the spectral density matrix estimator based on pairwise complete observations is consistent, producing a positive semidefinite matrix.

C.4 VAR Calibration for the ICAPM with a Strategic Investor

The ICAPM with a strategic investor relies on the VAR structure in equation [20](#). I calibrate the VAR to capture the annual dynamics of z_t and the long-term predictability in r_f and xr_e . In this subsection, I explain this calibration procedure. To simplify the explanation, I treat all variables as demeaned variables (so that all terms related to the VAR intercept drop from the expressions provided). However, my analysis applies the calibration procedure directly to the raw data and adjusts all expressions accordingly.

I start by running OLS regressions (with [Amihud and Hurvich \(2004\)](#) correction for Stambaugh bias) for all variables in s_t (except r_f and xr_e) onto z_t (the predictive variables). After this calibration, there are 14 terms in the VAR matrix, Φ_1 , that are not specified (the predictive coefficients of r_f and xr_e onto z_t). I calibrate these 14 coefficients to match long-term dynamics in expected returns.

From the VAR specification, we have $\sum_{h=1}^{h=10} \rho^h \mathbb{E}_t[r_{f,t+h}] = \left(\sum_{h=1}^{h=10} \rho^h \mathbf{1}'_{r_f} \Phi_1^h \right)' s_t$ (similarly for

xr_e). In contrast, the analysis in Section 3 relies on long-term predictive regressions and implies $\sum_{h=1}^{h=10} \rho^h \mathbb{E}_t [r_{f,t+h}] = \left(\sum_{h=1}^{h=10} \rho^h b_{ir}^{(h)} \right)' s_t = B'_{ir} s_t$ (similarly for xr_e) where b s and B s are augmented to have zero coefficients on s_t variables not present in z_t . I select the 14 coefficients by solving the non-linear system:

$$\left(\sum_{h=1}^{h=10} \rho^h \mathbf{1}'_{rf} \Phi_1^h \right) = B_{ir} \tag{A.10}$$

$$\left(\sum_{h=1}^{h=10} \rho^h \mathbf{1}'_{xr} \Phi_1^h \right) = B_{ep} \tag{A.11}$$

In words, I calibrate the VAR-implied predictability of r_f and xr_e to match the long-term predictability in the data. Subsection D.2 explores the same model when the VAR is calibrated to match short-term predictability in r_f and xr_e .

After calibrating the dynamics of r_f and r_e , I get the implied Δd dynamics that assure Campbell and Shiller (1989) approximation continue to hold in expectation (it holds ex-post by construction since I measure Δd directly from the approximation). In specifications in which Δd is part of the z_t vector, this affects the left hand side of the above non-linear system. In this case, I repeat the procedure iteratively until convergence.

Appendix D Other Results

D.1 Robustness Analysis

a) Robustness of Key Results in Section 3

I ask whether the key dividend and bond term structure results presented in Section 3 are robust to variation in the empirical choices I make. The main finding is that there is some variation in the quantitative results based on important empirical choices, but the qualitative results are similar for a large range of alternative empirical designs (I explore a total of 345 specifications).

The following list outlines some of the key empirical decisions and alternative options for them:

- **Sample Period:** (i) Main (1952-2016), (ii) Long (1928-2016), (iii) Post Depression (1935-2016), (iv) Postwar (1946-2016), (v) Compustat Period (1963-2016), (vi) No Great Recession/Depression (1935-2006);
- **Model:** (i) ICAPM in Section 2, (ii) ICAPM in Section 2, but no λ restrictions, (iii) ICAPM in Section 2 + misspecification adjustment, (iv) ICAPM in Section 2 + volatility news;²⁸
- **Dividends:** (i) Based on regular dividends, (ii) Incorporate cash-based M&A activity;²⁹
- **Dividend Growth:** (i) From Campbell and Shiller (1989) approximation, (ii) $\Delta d_t = \ln(D_t/D_{t-1})$;
- **Risk Free Rate:** (i) Based on the one year Treasury yield, (ii) Treasury bill rate compounded over 12 months;

²⁸In specification (iii), I address the criticism in Chen and Zhao (2009) that the ICAPM treats the residual information as cash flow news. Formally, we can write $\tilde{r}_e = N_{cf} - N_{ep} - N_{ir}$. By pricing N_{ep} and N_{ir} , the ICAPM implicitly assumes $N_{cf} = \tilde{r}_e + N_{ep} + N_{ir}$, which is not priced (beyond its effect on \tilde{r}_e) according to the ICAPM SDF. I relax this assumption by defining $N_{resid} = \tilde{r}_e + N_{ep} + N_{ir} - N_{cf}$ (with N_{cf} measured based on my predictive regressions using an expression analogous to equation A7b) and using it to adjust the equity premium and interest rate risk factors. I assume the misspecification error is split equally across the three N s components so that $N_{ep}^* = N_{ep} - \frac{1}{3}N_{resid}$ and $N_{ir}^* = N_{ir} - \frac{1}{3}N_{resid}$ are the misspecification adjusted equity premium and interest rate risk factors. I then apply the ICAPM normally.

In specification (iv), the structure of volatility shocks is similar to Campbell et al. (2017) except that I use long-run predictive regressions to get shocks (approach similar to the one for N_{ep}) and do not impose any restriction on the volatility risk price other than being negative.

²⁹In the main analysis, I ignore the effect of M&A activity (mergers and acquisitions) on the dividend measurement. However, Sabbatucci (2015) shows that M&A's paid in cash have an economically important effect both on dividend growth and on dividend yield. To address this issue, this robustness specification accounts for M&A activity when measuring dividend growth. The measurement details can be found in subsection B.1.

- **Value Spread:** (i) Based on HML portfolios, (ii) Small value spread in [Campbell and Vuolteenaho \(2004\)](#);
- **State Variables:** (i) Δd , dy , $epoy$, ty , TS , CS , and VS , (ii) Excluding one state variable at a time;³⁰
- **Testing Assets:** (i) r_e and r_{Tb} , (ii) r_e , r_{cb} and r_{Tb} , (iii) r_e , r_{BM} , r_{cb} and r_{Tb} , (iv) r_e , r_{BM} , r_{cb} and r_{Tb} , but replacing r_{Tb} by seven CRSP zero-coupon nominal Treasury bond portfolios with maturities of 1, 2, 5, 7, 10, 20 and 30 years, (v) r_e , r_{BM} , r_{cb} and r_{Tb} , but replacing r_{BM} by the book-to-market quintiles from Kenneth French's data library;
- **GMM Weighting Matrix:** (i) Identity, (ii) Diagonal with T_j (iii) Diagonal with T_j/σ_j^2 ;³¹
- **Maximum h :** (i) 10 years, (ii) 15 years;
- **Predictability after maximum h :** (i) None, (ii) Based on a Vector Autoregressive System;
- **Correcting for Stambaugh Bias?** (i) Yes, (ii) No.

The first alternative of each empirical decision defines the baseline specification, which is also studied in the main text. I vary empirical decisions at most two at a time while maintaining the others fixed at the baseline specification in order to keep the analysis manageable. This procedure gives a total of 345 specifications, which are then used to plot the cross-specification distribution for several statistics related to the dividend and bond term premiums: the three ICAPM β s for $r_{pv}^{(1)} - r_{pv}^{(10)}$ and $r_b^{(10)} - r_b^{(1)}$, the ICAPM-based risk premiums for $r_{pv}^{(1)} - r_{pv}^{(10)}$ and $r_b^{(10)} - r_b^{(1)}$, and CAPM α s for $r_{pv}^{(1)} - r_{pv}^{(10)}$.

Figure [A.4](#) provides the results with all statistics relying on the risk price estimates from the ICAPM. In the models in which the link between rra and λ s is imposed, I restrict rra to $0 < rra < 15$. This approach assures the results consider only ICAPM specifications that reasonably capture the given testing assets and, at the same time, do not require unreasonable risk aversion.

³⁰I do not drop dy from the model to be consistent with Section [4](#), which requires dy in the state variables to impose [Campbell and Shiller \(1989\)](#) approximation. However, a previous analysis before the inclusion of [4](#) produced similar results even though dy was not part of the state vector.

³¹The first alternative is a diagonal matrix with the number of years with available data for the respective testing asset (T_j) on each diagonal entry. With this weighting matrix, more attention is paid to testing assets with longer time-series. The second alternative is a diagonal matrix with $1/(\sigma_j^2/T_j)$ on each diagonal entry. This matrix effectively weights each testing asset based on the estimation precision associated with its average excess return when we ignore autocorrelation and heteroskedasticity, but treat all months observed within a year as not providing independent information (i.e., it is conservative). This is in the spirit of efficient GMM as it considers the statistical reliability of each testing asset.

All results from Figure A.4 confirm (qualitatively) the findings presented in previous sections. First, the one-year dividend claim (in relation to the 10-year dividend claim) has lower market β , higher equity premium β , and lower interest rate β in all specifications. Similarly, the longest-term bond portfolio (relative to the shortest-term bond portfolio) has higher market and equity premium β s and lower interest rate β s in almost all specifications. Second, the ICAPM-based dividend term premium is negative (1-year higher than 10-year) and the bond term premium is positive for the vast majority of specifications. Finally, the term spread in CAPM α s of dividend claims is very negative for the vast majority of specifications.

One interesting result that cannot be seen from Figure A.4 is that, for some specifications, the short end of the dividend term structure (one to five years) is upward sloping while longer-term dividend claims (six to ten years) remain with much lower risk-premia. This finding echoes the results in Section 4.

b) Robustness of Key Result in Section 4

The key result in Section 4 is the hump-shaped term structure of dividend strips, with very long-term dividend strips having much lower risk-premia than mid-term contracts.

I focus on 12 robustness checks that address key empirical decisions that could affect the results. I follow this approach both to keep the analysis manageable and to be able to provide a full picture of the dividend term structure for each specification considered.

Figure A.7 displays robustness checks that: (i) exclude one state variable at a time from the analysis (except dy as it is needed to impose Campbell and Shiller (1989) approximation); (ii) consider lower rra or higher ies relative to the baseline specification; (iii) ignore Stambaugh bias correction; (iv) keep only r_f and r_e in the opportunity set of the marginal investor; (v) measure dividend growth directly (as opposed to measuring it from Campbell and Shiller (1989)'s log-linear approximation); and (vi) change the measurement of the risk-free asset to the 3-month treasury bill.

The result is easy to summarize: the shape of the dividend term structure is highly robust to alternative specifications. In particular, for all specifications considered, the term structure is upward sloping over the first few years (typically one to five) and downward sloping for later years, with very long-term dividend strips having relatively low risk premia.

D.2 A VAR Calibrated to Match the Short-Term Predictability in r_f and r_e

This subsection calibrates the VAR in Section 4 to match short-term dynamics in r_f and r_e as opposed to the calibration to long-term dynamics performed in my main analysis. I refer to these as Short-term and Long-term VAR in this subsection.

Figure A.5 displays the dividend term structures using a short-term VAR (analogous to Figure 7). Panel (a) shows that longer-term dividend strips have higher market risk, but are also better hedges for reinvestment risk. This is one of the key results in this paper. However, in contrast to the results in previous sections, the market risk effect dominates the reinvestment risk effect, producing an *upward sloping* dividend term structure of risk premia. The term structures of dividend strip α s and Sharpe ratios remain hump-shaped (but with less pronounced differences between short- and long-term dividend strips).

Why would a calibration to short-term $\mathbb{E}[r]$ dynamics produce an economically different result? The answer is simple: the short-term VAR produces too little volatility for long-term $\mathbb{E}[r]$.

Consider the long-run expected return $\sum_{h=1}^{10} \rho^h \mathbb{E}[r]$. The short-term VAR produces 45.1% of the variability in this quantity relative to the long-run predictive regressions used in Section 3. Since the long-run predictive regressions I rely on are unbiased, this indicates that a short-term VAR substantially understates the variability in $\mathbb{E}[r]$. This translates into an understatement of the volatility of the ICAPM reinvestment risk factor. For instance, shocks to $\sum_{h=1}^{10} \rho^h \mathbb{E}[r]$ have a volatility of 20.4% when the long-run regressions in Section 3 are used. In contrast, the same volatility based on the short-term VAR is 9.8%. Even if we consider all shocks with horizon going to infinity the total volatility is still only 13.1%.

The direct consequence is that the term structure of reinvestment risk is less pronounced when using a short-term VAR because $Cov(r_{ds,t}^{(H)}, N_{\mathbb{E}r}) - Cov(r_{ds,t}^{(h)}, N_{\mathbb{E}r}) = \sigma_{\mathbb{E}r} \cdot Cor(r_{j,t}, N_{\mathbb{E}r}) \cdot [\sigma_{ds,H} - \sigma_{ds,h}]$. Since the risk price for reinvestment risk is fixed (depends on rra), the impact of reinvestment risk on risk premia is understated when relying on a short-term VAR. This effect is not present for market risk because the market volatility is the same whether we use long-run predictive regressions or a short-term VAR.

The long-term VAR approach is a (admittedly simple) solution to this problem as it restricts the variability in $\sum_{h=1}^{10} \rho^h \mathbb{E}[r]$ to be identical to the one produced by long-run predictive regressions. The

economic result is that reinvestment risk becomes more important and its term structure dominates the market risk term structure, producing a downward sloping dividend term structure.

Despite being particularly important in my analysis, the fact that VAR models calibrated to short-term dynamics fail to capture long-term dynamics of relevant variables (such as expected returns) is not a new result. [Jordà \(2005\)](#) proposes a local projection approach to study impulse response functions in macroeconomic models precisely for this reason. Moreover, [Bianchi and Tamoni \(2016\)](#) argues that “...low-order autoregression models for short-term expected returns imply long-term dynamics that have a (too) fast vanishing persistence when compared with the evidence from long-horizon predictive regressions.” Of course, low persistence translate into low volatility of long-term discount rates, which is the fundamental problem for my analysis.

The literature has also demonstrated that short-term VAR models fail to capture long-term dynamics of cash flow prices. In particular, [Giglio and Kelly \(2017\)](#) find that VAR term structure models calibrated to match the price dynamics of short maturity claims produce too little volatility for long maturity claims. In my case, calibration to the dynamics of short-term discount rates produces too little volatility for long-term discount rates.

Overall, the conclusion is that a short-term VAR produces a risk term structure for dividend strips that is qualitatively similar to my main analysis, but a risk premia term structure that is qualitatively different. Moreover, this is a consequence of understating the volatility of the ICAPM reinvestment risk factor, which depends on long-term $\mathbb{E}[r]$ dynamics. Finally, the inability of a short-term VAR to capture long-term discount rate dynamics is largely consistent with the literature.

Appendix E Supplementary Tables and Figures

E.1 Supplementary Tables

Table A.1
Shocks to Risk Factors and Dividend PV Returns

Panel A reports the loadings of reinvestment risk factors (N_{ep} , N_{ir}) and dividend PV unexpected returns in excess of the equity market ($\tilde{r}_{pv}^{(h)} - \tilde{r}_e$) on z_t . Multiplying the loadings by z_t shocks would recover the respective risk factors and unexpected excess returns (even though my methodology does not require measuring z_t shocks). The values reported are based on the estimation of the system in equations 14a to 14d over my main sample (1952-2016). Each coefficient is normalized to be in standard deviation units. For instance, the coefficient of N_{ep} on Δd is multiplied by $\sigma_{\Delta d}/\sigma_{N_{ep}}$. The state variables are the log dividend growth (Δd), dividend yield (dy), equity payout yield ($epoy$), one year treasury yield (ty), term spread (TS), credit spread (CS), and value spread (VS). Details about the construction of risk factors and dividend PV returns can be found in Section 3.1. Equations 14a to 14c in the estimated system are predictive regressions. Panel B reports the cumulative R^2 's from such predictive regressions (e.g., the 2-year R^2 for Δd is the R^2 from predicting the dividend growth from t to $t + 2$).

	PANEL A: Normalized Loadings							PANEL B: Cumulative Predictive R^2			
	Δd	dy	$epoy$	ty	TS	CS	VS	(Horizon)	Δd	$r_m - r_f$	r_f
$r_{pv}^{(1)} - r_e$	0.24	0.40	0.36	0.31	0.63	-0.02	-0.07	(1 year)	31%	17%	36%
$r_{pv}^{(2)} - r_e$	0.28	0.37	0.35	0.35	0.66	-0.04	-0.11	(2 years)	34%	28%	44%
$r_{pv}^{(3)} - r_e$	0.25	0.32	0.32	0.30	0.61	0.07	-0.11	(3 years)	36%	37%	48%
$r_{pv}^{(4)} - r_e$	0.29	0.27	0.34	0.24	0.53	0.16	-0.02	(4 years)	37%	39%	51%
$r_{pv}^{(5)} - r_e$	0.38	0.23	0.37	0.24	0.52	0.18	0.03	(5 years)	35%	39%	53%
$r_{pv}^{(6)} - r_e$	0.27	0.11	0.49	0.22	0.43	0.25	0.00	(6 years)	35%	35%	56%
$r_{pv}^{(7)} - r_e$	0.20	0.04	0.36	0.11	0.24	0.50	0.00	(7 years)	30%	33%	59%
$r_{pv}^{(8)} - r_e$	0.20	0.08	-0.02	-0.02	-0.02	0.76	0.06	(8 years)	26%	33%	62%
$r_{pv}^{(9)} - r_e$	0.15	0.03	-0.38	-0.10	-0.36	0.83	0.15	(9 years)	30%	37%	64%
$r_{pv}^{(10)} - r_e$	-0.07	-0.22	-0.49	-0.17	-0.58	0.29	0.19	(10 years)	38%	37%	66%
N_{ep}	0.07	0.08	0.35	-0.26	0.34	0.37	-0.36				
N_{ir}	0.36	0.59	-0.45	0.95	0.65	-0.11	-0.04				

E.2 Supplementary Figures

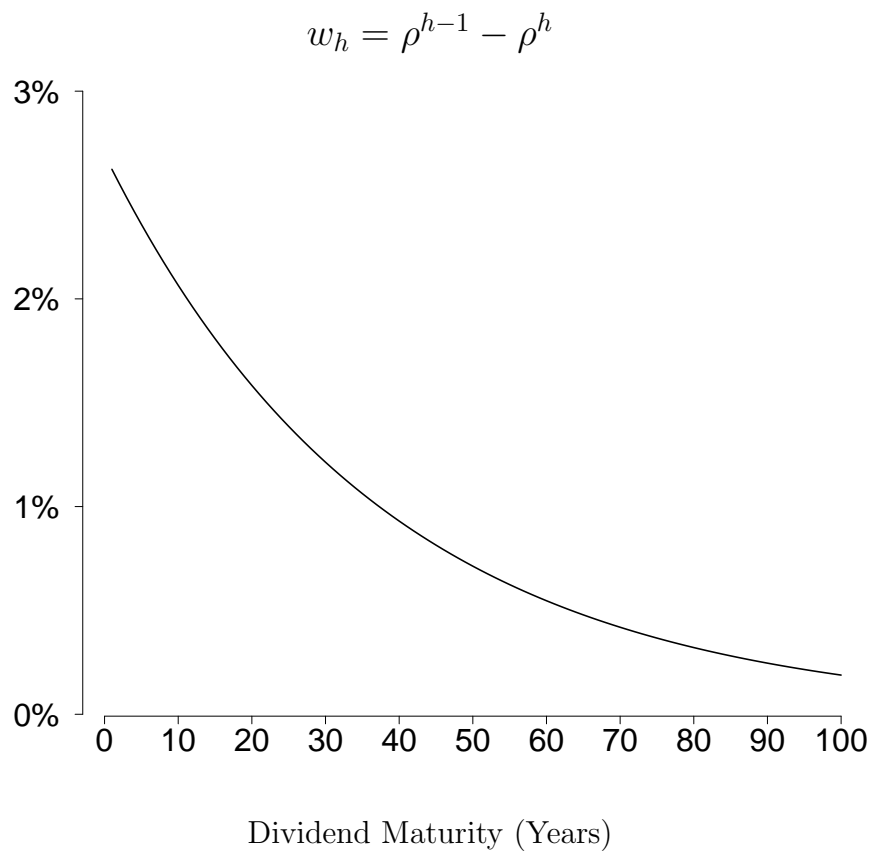


Figure A.1
Dividend PV Return Weights in Equation 5 ($\rho = 0.97$)

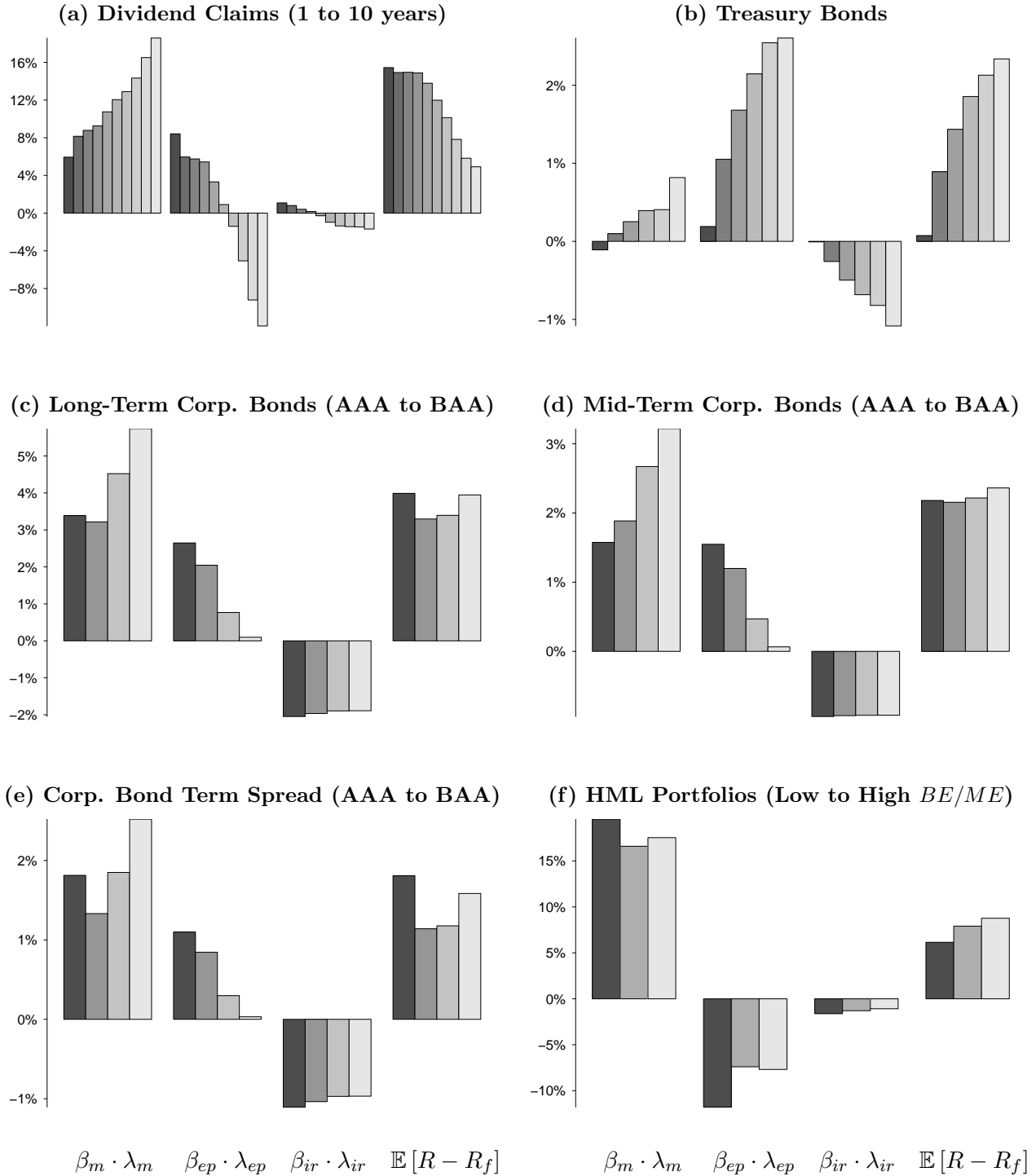


Figure A.2

Decomposing Risk Premia of Testing Assets Based on ICAPM Risk Factors

The graphs decompose the risk premia of dividend claims (from 1-year to 10-year), Treasury bond portfolios (from shortest to longest duration), mid-term and long-term corporate bond portfolios as well as their term spread (all from highest to lowest rating), and book-to-market sorted portfolios (from lowest to highest book-to-market) into each asset compensation for its exposure to market risk ($\beta_{j,m} \cdot \lambda_m$), equity premium risk ($\beta_{j,ep} \cdot \lambda_{ep}$), and interest rate risk ($\beta_{j,ir} \cdot \lambda_{ir}$). Results are based on equation 13 with risk aversion estimated in column (2b) of Table 2 and β_s estimated from the system in equations A7a to 14d (studied in subsection 3.3). Empirical details can be found in subsections 3.1 and 3.5.

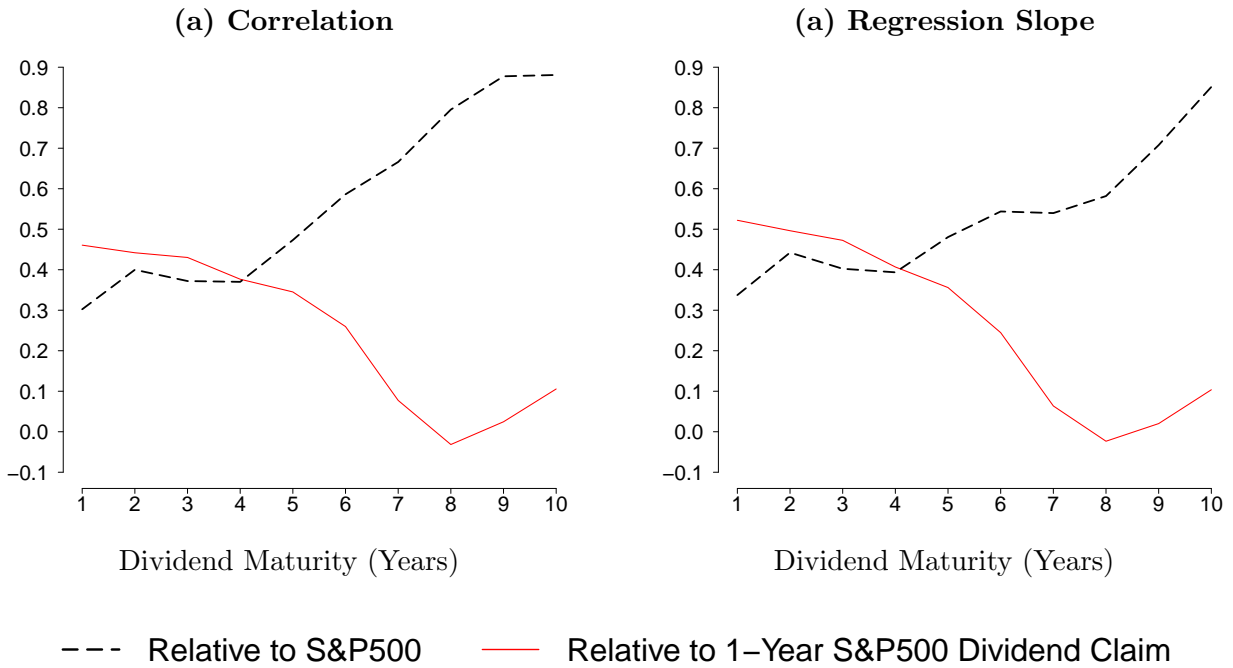


Figure A.3
Relation Between Dividend PV Excess Returns and
Excess Returns on the S&P 500 Index and 1-Year Dividend Strip

The graphs report 1997-2016 correlations (and regression slopes) of dividend PV excess log returns relative to excess log returns on the S&P 500 and a portfolio of short-term dividend futures ($R_{df}^{(st)}$ in subsection 3.6). The red solid (black dashed) line refers to the 1-year dividend future (S&P 500). All statistics are obtained using dividend PV returns estimated over my main sample period (1952-2016). Empirical details about the estimation of dividend PV returns are in subsection 3.1 while the information about $R_{df}^{(st)}$ is in subsections 3.6 and B.3.

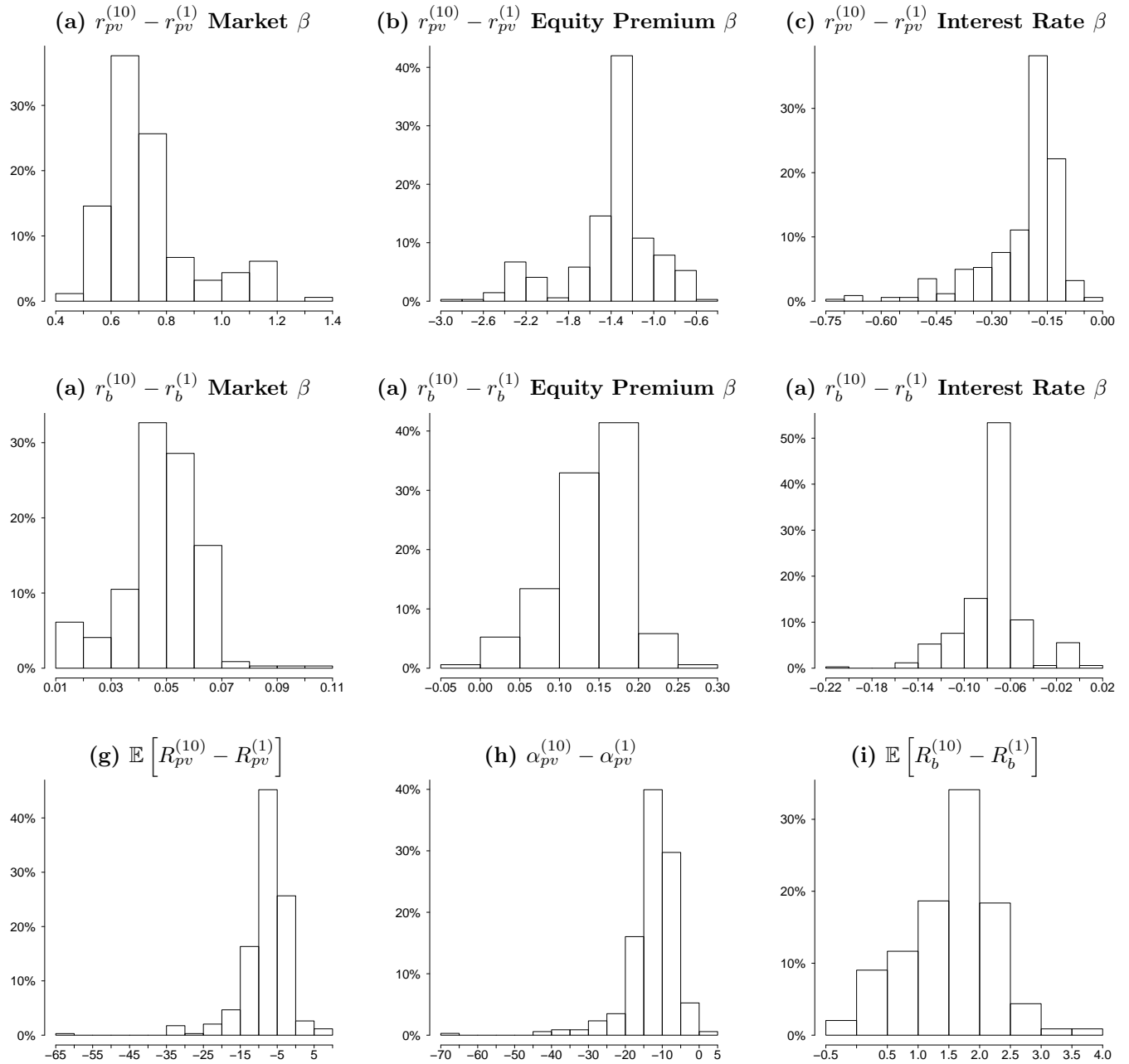


Figure A.4

Robustness: Term Premium Statistics in the ICAPM with a Passive Investor

The graphs report the cross-specification distribution for several statistics related to the dividend and bond term premia: the three ICAPM β s, the ICAPM-based risk premia, and CAPM α s. Results are based on varying empirical decisions at most two at a time while keeping the others fixed at the baseline (composed by the first alternative of each empirical decision in the list provided in Appendix D.1 (a)). This procedure produces a total of 345 specifications, which are used towards the distributions plotted. All β s are in market β units (i.e., covariance normalized by market variance). All risk premia and CAPM α s rely on the risk price estimates for the respective specification. Empirical details about the statistics can be found in subsection 3.1 and details about the robustness analysis (the specifications) are provided in Appendix D.1.

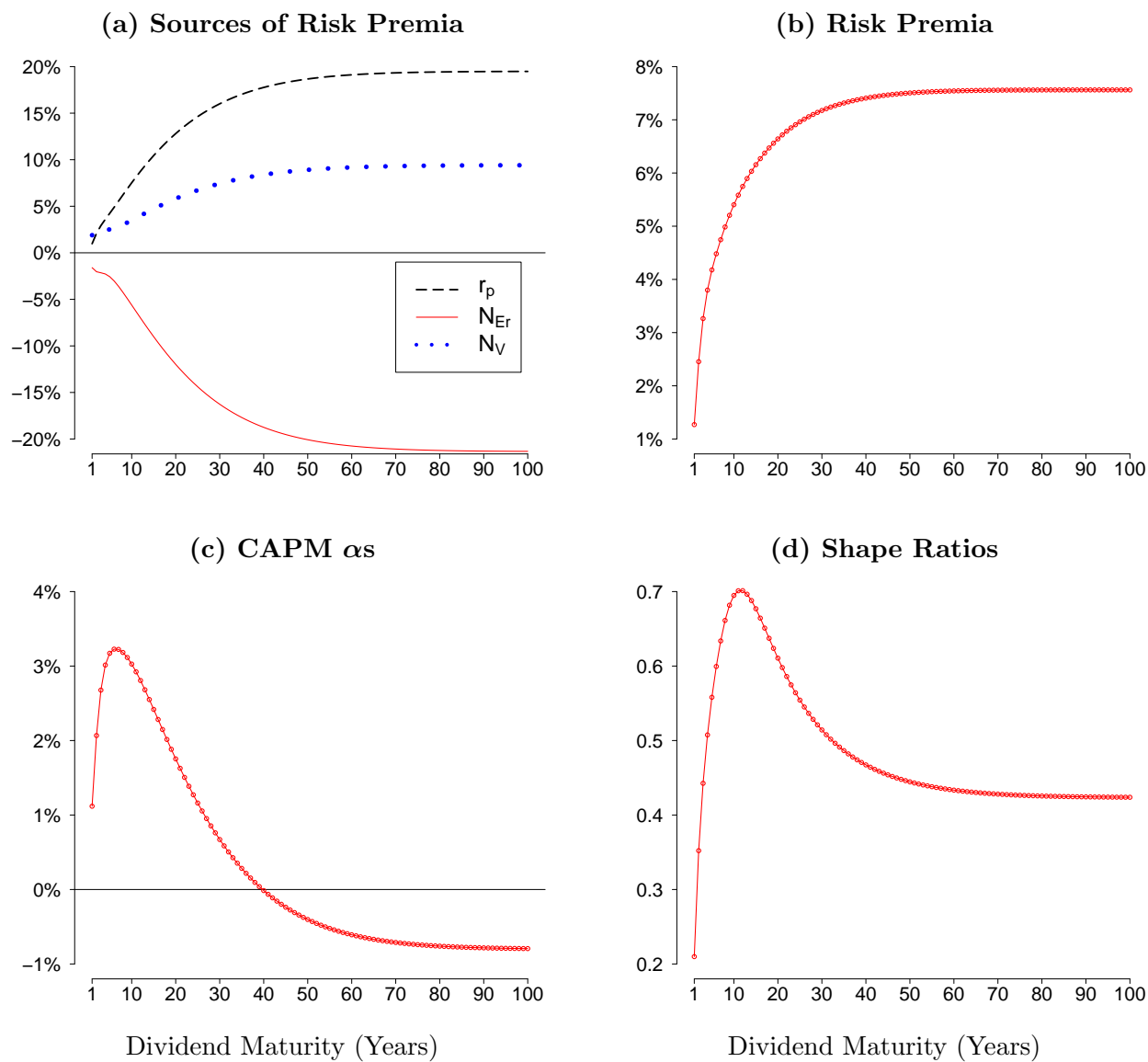


Figure A.5
Dividend Term Structures in the ICAPM with a Strategic Investor
(VAR Calibrated to Match s_t Short-Term Dynamics)

The graphs report the dividend term structures of (i) risk premia, (ii) CAPM α s, and (iii) Sharpe ratios within the ICAPM with an investor who responds to changes in reinvestment rates (i.e., a strategic investor). The Risk premia term structure can be decomposed into the effect of market risk (r_p), reinvestment risk (N_{Er}), and volatility risk (N_V) and panel (a) provides such decomposition (based on equation 24b). All results are obtained based on my main sample period (1952-2016) and details for the ICAPM and calibration can be found in subsection 4.1. The only difference relative to the original calibration is that the procedure for the VAR calibration matches 1-year r_f and r_e dynamics (as opposed to long-term dynamics), and thus is based on a simple OLS.

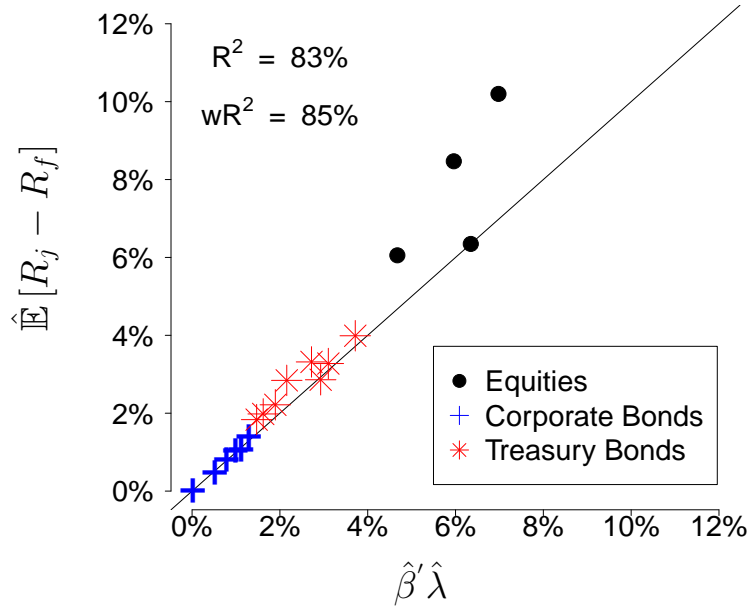


Figure A.6
Predicted vs Realized Risk Premia for Testing Assets
(ICAPM with a Strategic Investor)

The graphs display predicted (based on the right hand side of equation 24a) and realized (based on the left hand side of equation 24a) risk premia for the testing assets. The testing assets are excess returns on equity and bond portfolios: the market portfolio, the three (size controlled) book-to-market sorted portfolios in Fama and French (1993) (i.e., HML portfolios), the Barclays' mid-term and long-term AAA, AA, A, and BAA corporate bond portfolios and the six CRSP Treasury bond portfolios with maturities up to 1, 2, 3, 4, 5 and 10 years. $R^2 = 1 - \sum \alpha_j^2 / \sum (\hat{\mathbb{E}}[R_j - R_f] - \bar{\mathbb{E}}[R_j - R_f])^2$ captures the fraction of average return variability explained by the model and $wR^2 = 1 - \sum (\alpha_j - \bar{\alpha}_{class})^2 / \sum (\hat{\mathbb{E}}[R_j - R_f] - \bar{\mathbb{E}}_{class}[R_j - R_f])^2$ captures the same quantity, but for average return variability within each asset class (equities, corporate bonds, and Treasury bonds). All quantities are estimated from their respective available observations within the main sample period (1952-2016). The model used is the ICAPM with a strategic investor, with all necessary details described in subsection 4.1.

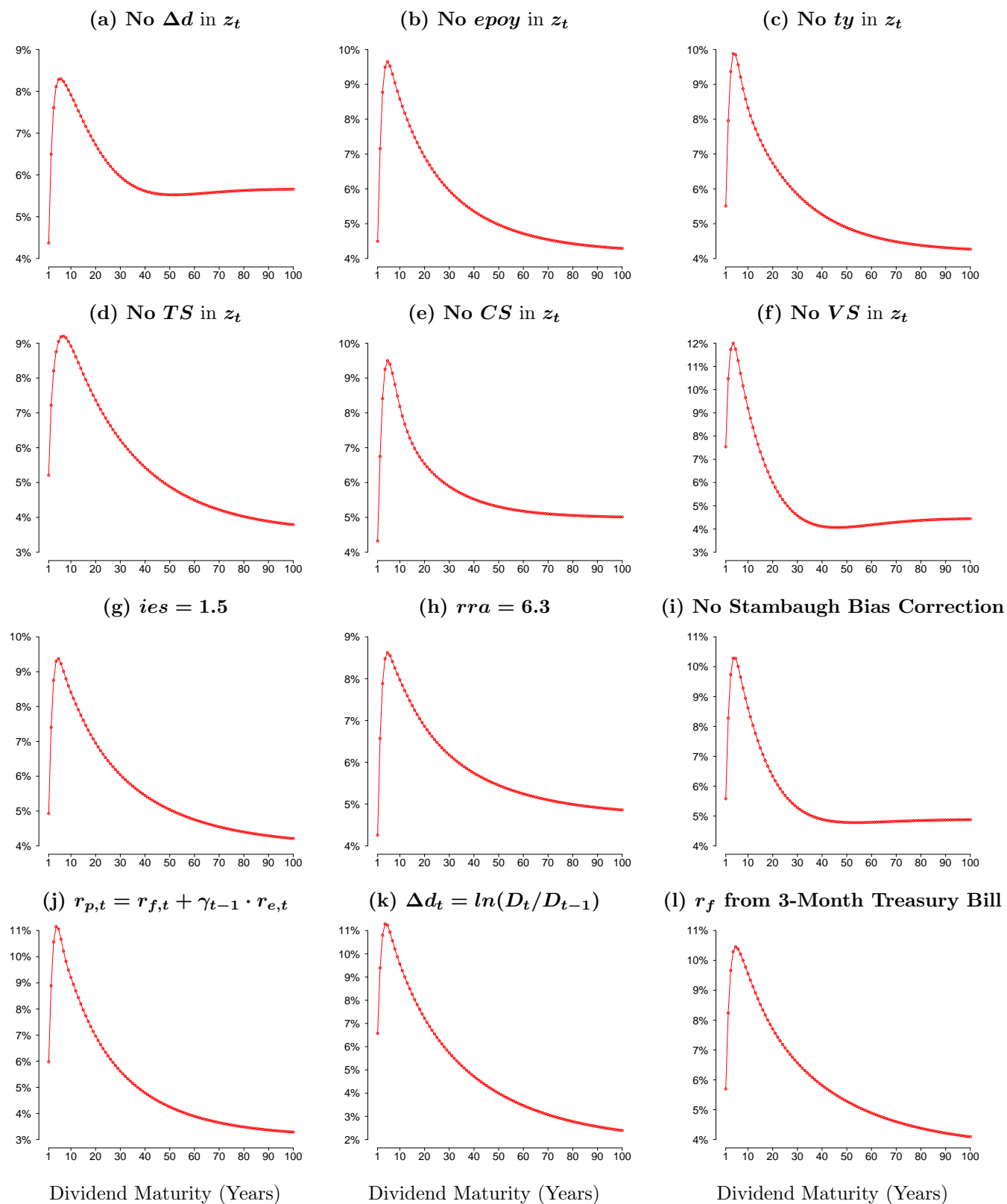


Figure A.7

Robustness: Dividend Risk Premia in the ICAPM with a Strategic Investor

The graphs report the dividend risk premia term structure within the ICAPM with an investor who responds to changes in reinvestment rates (i.e., a strategic investor). Sixteen alternative specifications are considered and they are described in Appendix D.1. Details for the ICAPM and calibration (under baseline specification) can be found in subsection 4.1.