The Journal of FINANCE
The Journal of THE AMERICAN FINANCE ASSOCIATION

THE JOURNAL OF FINANCE • VOL. LXXIX, NO. 6 • DECEMBER 2024

Equity Term Structures without Dividend Strips Data

STEFANO GIGLIO, BRYAN KELLY, and SERHIY KOZAK*

ABSTRACT

We use a large cross section of equity returns to estimate a rich affine model of equity prices, dividends, returns, and their dynamics. Our model prices dividend strips of the market and equity portfolios without using strips data in the estimation. Yet model-implied equity yields closely match yields on traded strips. Our model extends equity term-structure data over time (to the 1970s) and across maturities, and generates term structures for various equity portfolios. The novel cross section of term structures from our model covers 45 years and includes several recessions, providing a novel set of empirical moments to discipline asset pricing models.

THE TERM STRUCTURE OF DISCOUNT rates for risky assets plays an important role in many fundamental economic contexts. For example, pricing an asset with a specific horizon of cash flows and evaluating an investment opportunity with a specific maturity requires knowing the maturity-specific discount rate. Investment in climate change mitigation, where the maturity of the project is especially long and therefore the long end of the term structure is especially important, is one well-known example.

*Stefano Giglio is with the Yale School of Management, NBER, and CEPR. Bryan Kelly is with the Yale School of Management, NBER, and AQR. Serhiy Kozak is with the Robert H. Smith School of Business at the University of Maryland. We appreciate helpful comments from the editorial team and from Jules van Binsbergen; Niels Gormsen; Ralph Koijen; Jessica Wachter; participants at the Chicago Booth Asset Pricing Conference, SFS Finance Cavalcade, Virtual Derivatives Workshop, WFA; and seminar participants at Stockholm School of Business, Temple University, Tilburg University, Tsinghua University, University of Gothenburg, University of Hong Kong, University of Massachusetts at Amherst, and the Five Star conference. We have read *The Journal of Finance* disclosure policy. Stefano Giglio and Serhiy Kozak have no conflicts of interest to disclose. Bryan Kelly has received income from AQR Capital Management exceeding $10,000 over the past five years. AQR Capital Management is a global investment management firm, which may or may not apply similar investment techniques or methods of analysis as described herein. The views expressed here are those of the authors and not necessarily those of AQR.

Correspondence: Serhiy Kozak, Finance Department, Robert H. Smith School of Business, University of Maryland, College Park, MD 20742; e-mail: sekozak@umd.edu.

DOI: 10.1111/jofi.13394

In this paper, we specify and estimate a rich affine model of equity portfolios. The model describes the prices, dividends, and excess returns of a large cross section of portfolios, as well as their dynamics. While the model is driven by many parameters, we impose discipline on the model by imposing pricing restrictions, by appropriately choosing the state vector that drives the dynamics of the economy, and by imposing parameter restrictions that reflect recent findings in the literature on return predictability with large cross sections. We then use our model to generate a term structure of discount rates not only for aggregate cash flows (the S&P 500), but also for many other equity portfolios, obtaining a large panel of 102 term structures of discount rates with arbitrary maturity going up to infinity and with a time series going back to the 1970s. We validate the predictions of our model for the discount rates of risky cash flows by comparing the implied dividend strips from our model to the prices of actually traded dividend strips (from Bansal et al. (2021) (BMSY) and van Binsbergen and Koijen (2015)). We show that our model—estimated using no dividend strip data at all—well matches the prices of traded dividend strips on the S&P 500 of maturities of one, two, five, and seven years, observed since 2004, along a variety of dimensions (average slope, time series, etc.). After validating the model using observed strip data, we then use it to explore the properties of implied dividend strips extending the time series back to the 1970s, and to study the cross section of term structures of different portfolios.

To estimate the term structure of discount rates for risky assets, early work in the asset pricing literature has proposed extracting information about the term structure from the cross section of equity portfolios. The broad idea behind this approach is that if some stocks are mostly exposed to long-term cash-flow shocks, while others are mostly exposed to short-term shocks, the difference in risk premia between the two types of stocks can be attributed to a difference in how investors price shocks to cash flows of different maturities, that is, to the term structure of discount rates applied by investors (see Bansal, Dittmar, and Lundblad (2005), Lettau and Wachter (2007), Hansen, Heaton, and Li (2008), and Da (2009)).

While this literature has advanced our understanding of what the cross section of equity portfolios implies for the term structure of discount rates, it faces an important challenge: the term structure of discount rates depends on the entire dynamics of cash flows as well as investors' risk preferences and variations in these preferences over time. That is, the term structure of discount rates is an equilibrium result that depends on the interactions between a large number of factors. To identify and estimate the model, the papers in this literature impose strong assumptions on investors' risk preferences (e.g., Epstein-Zin preferences, constant risk aversion, restrictions on which shocks are priced by investors, etc.), on the dynamics of the economy and cash flows, or both.

Renewed interest in the study of term structures has come from the introduction of data on traded dividend claims (van Binsbergen, Brandt, and Koijen (2012), van Binsbergen et al. (2013), and van Binsbergen and Koijen

(2015)). The ability to directly observe the returns of finite-maturity dividend claims gives us a direct window into the risk premia that investors require to hold risks of different maturities, obviating the need to estimate investors' risk preferences and the dynamics of the economy. Studying term structures with traded dividend strip data, however, faces several challenges of its own. First, the time series is quite limited, since data are available starting around 2004, and it includes only one recession with associated recovery, the Great Recession. Second, there is little cross-sectional data, since only the aggregate market dividend strips (for the United States as well as other countries) start around 2004; much more limited data are available on individual firms. Third, typically only a part of the term structure is observed (a few maturities up to seven years). Fourth, there are concerns about the liquidity of these contracts, which could lead to measurement error.

In this paper, we return to the first approach to estimate term structures based on equity portfolios alone. We use our model to generate new (implied) term-structure data that expand existing (observed) data along each of the dimensions above. In particular, the term structures that we generate cover a large number of cross-sectional portfolios, in addition to the S&P 500, for a total of 102 portfolios. Moreover, they start in 1975 and therefore have a long time series that covers several recessions and booms. They also span all possible maturities, including very short and very long ends of the term structure.

While closely related to models used in prior literature to study term structures using equity portfolios, our model has a few distinct features that are crucial to generate realistic implied term structures that match those we observed from traded dividend claims. The model features rich dynamics that are motivated by recent empirical evidence in the literature and we believe are both economically reasonable and statistically parsimonious. Specifically, consistent with Kozak, Nagel, and Santosh (2020) and Giglio and Xiu (2021), who show that a few dominant principal components (PCs) of a large cross section of anomaly portfolio returns explain the cross section of expected returns well, our state vector includes four-factor returns (PCs) estimated from a large cross section of 51 anomalies.

Further motivation comes from Haddad, Kozak, and Santosh (2020), who demonstrate that valuation ratios strongly and robustly predict expected returns on these PCs, and their risk prices in the stochastic discount factor (SDF), in the time series. They argue that the resulting time variation in risk prices is critical to adequately capture dynamic properties of the pricing kernel. Chernov, Lochstoer, and Lundeby (2018) echo the importance of time variation in prices of risk by proposing that asset pricing models use multihorizon returns. Motivated by these findings, our specification also includes four-factor yields (dividend-to-price (D/P) ratios) associated with the factor returns—for a total of eight variables in the state vector—which allows us to capture the dynamics of conditional means and SDF risk prices.

Overall, our specification allows both dividend growth and risk premia to vary over time in minimally restricted ways, and has a general affine specification for the SDF in which shocks to factor returns are priced and their risk prices are captured by valuation ratios. The fact that our state variables are factor yields and factor returns means that, on the one hand, our state variables are forward-looking and can be expected to contain information about the evolution of the economy, and, on the other hand, the factors need to satisfy certain pricing restrictions, that better pin down their dynamics. It is this balance of a rich model with appropriately chosen restrictions that represents the core of our paper: it allows us to produce term structures of discount rates that well match the observed discount rates, which increases our confidence in extending term structures over time, maturities, and portfolios.

We start with a large cross section of test asset returns r_t on 102 portfolios. We specify and estimate a homoskedastic affine model in which the factors F_t that drive the dynamics of the test asset returns are chosen to be the linear combinations of assets' excess returns and dividend yields, that is, $F_t = [f_{r,t}\ f_{y,t}]'$, where factor returns, $f_{r,t} = Q'\left(r_t - r_{f,t}\right)$, and factor yields, $f_{y,t} = Q'y_t$, are the same linear combination of log excess returns and yields on the test assets, given by some matrix Q. We restrict the first factor to coincide with the excess log market return, and the rest to be based on the three largest PCs of 51 long-short anomaly portfolios constructed from the 102 long-only portfolios, as in Haddad, Kozak, and Santosh (2020) and Kozak (2024). This choice of model factors has two main advantages. First, it well captures the covariation of returns of the anomaly portfolios (together, the four-factor returns explain 93.3% of the total variation in excess returns in the panel of 102 long portfolios). Second, the factor returns are well predicted by their own dividend yields, which, conveniently, are also part of our state vector F_t. So our state vector F_t contains variables that are useful to predict excess returns (and dividends, which are related by an identity to prices and returns).

We restrict this affine model in only two ways. First, we impose that the innovations of the SDF in the economy depend only on the innovations in factor returns (and not on the innovations in factor yields), that is, we assume that the SDF innovations are fully spanned by the factor returns. Second, we impose that the four-factor yields $f_{y,t}$ contain all available information about the future (so that lagged returns do not help predict future yields and returns after controlling for lagged yields). In practice, it is well known that dividend yields have much stronger predictive power for future dividends and returns than lagged returns, and we simply impose this regularity by assumption in our statistical model.

We do not impose any other restrictions on the model, except of course the pricing restrictions that link the pricing kernel to prices, dividends, and returns. These restrictions imply, as is usual in these cases, that the dividend growth process for each portfolio is fully pinned down by the (nonlinear) identity that links prices, returns, and dividends.

We end up with an estimated model that, with its eight state variables, is rich enough to capture a variety of possible dynamics for prices, excess returns, and dividends. The model immediately produces implied term structures of dividend strips and forwards, that is, spot or forward claims on a specific dividend at some point in the future; it produces a different term structure for each of the four portfolios that are part of the state vector F_t. But the model does more: because the four factors well span the cross section of returns and dividend yields of the original 102 characteristic-sorted portfolios, it can easily generate implied term structures for any of these portfolios.

Our estimated model delivers a variety of novel empirical results. First, we extend the study of the term structure of aggregate dividend claims (on the S&P 500, as in van Binsbergen and Koijen (2015)) over time (back to the 1970s) and across maturities. In the sample starting in 2004 that was used in van Binsbergen and Koijen (2015), we match the time series of forward dividend yields very closely, and thus we, mechanically, also match the term structure of discount rates.[1] The term structure of forward risk premia in this sample appears to be mildly upward sloping, though not significantly so. We also confirm that our implied term structure inverts during the financial crisis, just like the observed term structure does.[2]

Extending the sample to the 1970s allows us to include several additional recessions in our sample; at the same time, the Great Recession carries less overall weight in the sample. It is interesting to see that all of the results of the post-2004 sample carry over to the longer sample. The term structure inverts in almost all of the additional recessions (e.g., in the early 1980s and 1990s), and the term structure of forward discount rates is still close mildly upward sloping on average, but not significantly so.

In addition, a decomposition of the movements of prices of short-term dividend claims into expected dividend growth and expected risk premia shows that the former varies substantially over time: investors expected low dividend growth during the 1980 and 1990 recessions, as well as during the Great Recession, and this moved the prices of the short-term dividend claims substantially.

The most important and novel results of our paper are the estimated term structures of discount rates for different portfolios, like value and growth firms, or small and large firms. Our model generates interesting differences in both the average term structure across portfolios and the time series. For example, term structures of long-short portfolios such as Small Minus Big (SMB) and High Minus Low value (HML) are fundamentally different: the SMB portfolio appears to exhibit a downward-sloping term structure, while the HML portfolio has an upward-sloping term structure, despite the fact that both portfolios have positive unconditional risk premium in our sample. We show that other long-short portfolios, such as profitability, growth, momentum, and

[1] We also verify that the ability of the model to match traded strip prices is robust to a variety of changes in our model specification.

[2] We note that because our implied dividend strip prices are based on equity portfolios, they are less susceptible to the potential criticism raised by Bansal et al. (2021) that traded dividend strips may be illiquid and that bid-ask spreads may affect conclusions about the slope of the term structure. The fact that we match those prices very closely using only (very liquid) equity portfolios suggests that liquidity does not drive the findings of van Binsbergen and Koijen (2015) in the first place.

idiosyncratic volatility, exhibit distinct term structures of expected returns. Theoretical models that aim to explain these different spreads (like those proposed to explain the size and value premia) can make use of these estimates to help refine and calibrate the economic mechanisms. Our estimates of the term structure of different portfolios provide us with additional empirical moments that can be compared to the corresponding models' moments.

Finally, there are interesting patterns in the time series of slopes of the yield term structure of different portfolios. For example, the slopes of small and large stocks tend to move together, with both term structures upward sloping during the 1990s and both downward sloping during the Great Recession. Yet only the term structure for small stocks inverted during the late-1990s cycle, marking an important divergence between the two portfolios that lasted several years. In contrast, no such divergence in the shape of the term structure can be seen for value and growth stocks in that period—instead, the largest difference in these stocks occurred during the recovery from the financial crisis: after 2008, the term structure of value-stock expected returns increased significantly, whereas expected returns of growth stocks showed no such pattern.

In sum, our model effectively processes a rich information set—the time-series and cross-sectional behavior of 102 portfolios spanning a wide range of equity risks—to produce "stylized facts"—the time-series and cross-sectional behavior of implied dividend term structures—that summarize a dimension of the data that is particularly informative about our economic models. Similar to how the introduction of vector autoregressions (VARs) by Sims (1980) provided new moments with which to evaluate structural macro models (the impulse-response functions that were generated by the VARs), the objective of this paper is to produce a realistic term structure of discount rates for different portfolios that closely resemble the actual dividend claims observed in the data, and that can be used by asset pricing models as a moment for evaluation and guidance.

It is important to note that our estimates of equity yields, discount rates, and returns on specific dividend claims are subject to two types of estimation uncertainty. First, uncertainty comes from the fact that the prices of the dividend strips that our model produces are not actually observed, but rather are obtained from model parameters, which are subject to estimation error. We refer to this source of uncertainty as parameter uncertainty. Second, uncertainty comes from the fact that some moments of interest (e.g., the average slope of the dividend term structure) are estimated from averages taken over finite samples. This last type of uncertainty, which we refer to as sampling uncertainty, is present even when prices are observable. What our model brings to the table is, in a sense, a trade-off between these two sources of uncertainty. Compared to using observed dividend strips, working with estimated dividend strips introduces parameter uncertainty. At the same time, our method allows us to dramatically expand the time series available, reducing sampling uncertainty. Which of the two forces dominates total uncertainty depends on the context. In our empirical setting, we find that sampling uncertainty dominates, so our method allows us to significantly reduce standard errors on many moments of interest. To help future researchers take this estimation uncertainty into account, we provide the standard errors of our estimated equity yields along with their point estimates on our website.[3]

Our methodology lends itself to a variety of applications. As we mentioned above, one direct application is to test and calibrate asset pricing models, by using our estimated model to produce additional empirical moments. In the paper, we illustrate this idea using our estimated term structures to test workhorse asset pricing models (like Campbell and Cochrane (1999), Bansal and Yaron (2004)). More interestingly, however, our rich set of term structures of different portfolios can be used to test further predictions of the models, such as the cross-sectional heterogeneity in the shapes of the term structures across portfolios. In addition, our term structures can be used for the valuation of projects with specific horizons. For example, Gupta and Van Nieuwerburgh (2021) use them to evaluate private equity investments. An alternative application is climate change mitigation investments, where the long end of the term structure is especially important (see, e.g., Giglio et al. (2015)).

Beyond the seminal literature using equity portfolios or dividend strips to calibrate and estimate empirical term structures, our paper also relates to a more recent literature that builds on these approaches to improve our understanding of term structures. This literature focuses largely on the term structure of aggregate dividend claims. Some papers explore the joint behavior of the aggregate stock market and Treasury bonds (Lettau and Wachter (2011), Ang and Ulrich (2012), Koijen, Lustig, and Van Nieuwerburgh (2017)), whereas others use traded dividend strip data in the estimation (Kragt, de Jong, and Driessen (2014), Yan (2015), Gomes and Ribeiro (2019)). Recently, Gupta and Van Nieuwerburgh (2021) use term structures of discount rates from a similar affine model to value private equity investment; given their different objective, they use specific portfolios in their model (small and large firms, real estate investment trusts (REITs), and infrastructure firms). In contrast, our objective is to select the state variables that best describe the full dynamics of the economy, so our choice of factors is determined by the ability to best represent a vast cross section of portfolios, and our main objective is to produce realistic dividend strips that best match the traded ones.

The paper also relates to a large number of studies that explore the term structure of risky assets, in addition to that of equity market dividend claims. Among these, prior studies focus on the term structure of currency risk (Backus, Boyarchenko, and Chernov (2018)), variance risk (Dew-Becker et al. (2017)), and housing risk (Giglio, Maggiori, and Stroebel (2014)). Chernov, Lochstoer, and Lundeby (2018) stress the usefulness of multiperiod returns to test asset pricing models. Several papers propose models that aim to explain observed patterns in the term structure of discount rates (Croce, Lettau, and

[3] See https://www.serhiykozak.com/data.

Ludvigson (2014), Gormsen (2021)). Methodologically, our paper is also related to Adrian, Crump, and Moench (2015), who propose a similar affine structure of the SDF but do not use it to explore the term structure of risky assets.

Finally, our paper relates to a third approach used in the literature to explore the term structure of discount rates using equity portfolios only. This approach, followed by Weber (2018), Gormsen and Lazarus (2023), and Gonçalves (2021), estimates the duration of portfolios directly (rather than by modeling the dynamics of dividends and estimating duration as exposures to dividend shocks of different horizons) and uses it to back out implied discount rates at different horizons.

The rest of the paper is organized as follows. In Section I, we develop our framework. In Section II, we describe the data. Section III presents the results. Section IV concludes.

I. The Model

A. Motivation

In this section, we introduce our reduced-form asset pricing model. The model specifies: (i) an SDF that depends on factors and their shocks, (ii) a full specification of the factor dynamics, and (iii) equations linking the prices and returns of any portfolios to those factors. This model is quite general; to make estimation and identification feasible, we impose different restrictions when we bring the model to the data.

The specification for the SDF and the choice of factors builds on the empirical evidence in Kozak, Nagel, and Santosh (2018) and Haddad, Kozak, and Santosh (2020), who construct asset pricing factors for the cross section using PCs of a large cross section of test portfolios, and discuss empirically successful restrictions on the determinants of risk prices (specifically, the valuation ratios of each PC). This specification strikes a good balance between fit and parsimony that was explored in the papers mentioned above. To this specification of the SDF, our model adds the factor dynamics, which allows us to solve for the equilibrium prices of securities with arbitrary maturity.

Our model therefore integrates an existing and empirically successful specification for the SDF with a VAR structure to capture predictability across multiple horizons and therefore obtain term-structure implications. When applied to pricing one-period returns, the model directly maps into the standard cross-sectional pricing literature, for example, Kozak, Nagel, and Santosh (2018, 2020). Prices of risk and expected returns on factor portfolios, however, are time-varying, which maps directly into the framework of Haddad, Kozak, and Santosh (2020). Our paper thus generalizes and combines existing approaches that aim to explain cross-sectional and time-series patterns, with the ultimate goal of extracting a cross section of term structures for different portfolios. The structure of the affine model allows us to make progress by decomposing dividend yields into discount rates and expected dividend growth at each horizon, and uses the entire estimated model (both the dynamics and the SDF, which are estimated simultaneously) to construct the various term structures.

B. The Setup

State-space dynamics: We begin by specifying a general factor model with k factors F_t, whose identity we will discuss later, that follow linear dynamics:

$$\underbrace{F_{t+1}}_{k\times 1} = \underbrace{c}_{k\times 1} + \underbrace{\rho}_{k\times k} F_t + u_{t+1}, \tag{1}$$

with $\text{var}_t\,(u_{t+1}) = \Sigma$ constant (i.e., we assume homoskedasticity). Furthermore, we assume that shocks u_{t+1} are normally distributed.

SDF: Denote the risk-free rate by $r_{f,t}$. We assume a log-linear SDF, where the priced shocks are u_{t+1}, with time-varying risk prices λ_t,

$$m_{t+1} = -r_{f,t} - \frac{1}{2}\lambda_t' \Sigma \lambda_t - \lambda_t' u_{t+1}. \tag{2}$$

Risk prices λ_t are assumed to be affine in F_t,

$$\underbrace{\lambda_t}_{k\times 1} = \underbrace{\lambda}_{k\times 1} + \underbrace{\Lambda}_{k\times k}\underbrace{F_t}_{k\times 1}. \tag{3}$$

The price P_t, dividend D_t, and gross return $R_{t+1} = \frac{P_{t+1}+D_{t+1}}{P_t}$ of any test asset satisfy the Euler equation

$$1 = \mathbb{E}_t\left[e^{m_{t+1}}\frac{P_{t+1}}{P_t}\left(1 + \frac{D_{t+1}}{P_{t+1}}\right)\right] = \mathbb{E}_t\left[e^{m_{t+1}+r_{t+1}}\right]. \tag{4}$$

Here,

$$r_{t+1} = \log\left[R_{t+1}\right] = \log\left[\frac{P_{t+1}}{P_t}\left(1 + \frac{D_{t+1}}{P_{t+1}}\right)\right] = \Delta p_{t+1} + y_{t+1}, \tag{5}$$

where we denote by r_{t+1} the log returns on the asset, which we decompose into the sum of log price changes, $\Delta p_{t+1} = \log\left(\frac{P_{t+1}}{P_t}\right)$, and the dividend yield, which we define as

$$y_t \equiv \log\left(1 + \frac{D_t}{P_t}\right). \tag{6}$$

Given this definition of the dividend yield, we do not use any approximations in the decomposition of log returns in equation (5). In other words, the identity linking price changes, dividend yields, and returns is always satisfied exactly in our paper given our definition of the dividend yield.

Under log-normality, by taking the log of both sides of equation (4), we obtain

$$0 = \mathbb{E}_t\left[m_{t+1}\right] + \mathbb{E}_t\left[\Delta p_{t+1} + y_{t+1}\right] + \frac{1}{2}\text{var}_t\left[m_{t+1} + \Delta p_{t+1} + y_{t+1}\right]. \tag{7}$$

We consider the cross section of n financial assets, which can be priced using equation (7).

Price dynamics: We directly specify the dynamics of log price changes in excess of the risk-free rate on financial assets

$$\Delta p_{t+1} - r_{f,t} = \gamma_0 + \gamma_1 F_t + \gamma_2 u_{t+1} + \epsilon_{p,t+1}, \tag{8}$$

where we assume that expected price changes are driven by state variables in F_t, and shocks include both shocks to the state vector u_{t+1} and asset-specific shocks $\epsilon_{p,t+1}$. Note that in macrofinance models, often the dynamics of *dividends* are specified first, together with the SDF, with prices then obtained by combining the two through the Euler equation. Here, we follow the alternative approach of specifying the dynamics of *prices* first, together with the SDF, with the dynamics of dividends backed out of the other equations (the Euler equation and the identity linking prices, returns, and dividends). This is similar to the setup of Campbell (1991), who eliminates consumption from the various equations and expresses the entire model in terms of the dynamics of wealth and returns. The relationship between the two approaches is discussed in detail in Section I.C.

Dividend yields: The Euler equation (4) implies that the price-dividend ratio of any asset can be expressed as an affine function of the state vector and stock-specific residuals $\epsilon_{y,t}$

$$y_t = b_0 + b_1 F_t + \epsilon_{y,t}, \tag{9}$$

where parameters b_0 and b_1 are of sizes $n \times 1$ and $n \times k$, respectively, as we show in Appendix A.[4]

As we discuss below, the dividend yields and returns of some linear combination of the original assets will be part of the state vector F_t itself. For those portfolios, equation (9) is automatically satisfied without residual, with an appropriate choice of coefficients b_0 (zeroes) and b_1 (linking each portfolio to its position in the state vector). For other assets, parameters b_0 and b_1 can be solved using equation (7), given the estimates of the risk price parameters.

Excess returns: Equations (5), (8), and (9) imply that excess returns are also affine in factors and shocks:

$$r_{t+1} - r_{f,t} = \beta_0 + \beta_1 F_t + \beta_2 u_{t+1} + \epsilon_{r,t+1}. \tag{10}$$

In this specification, $\mathbb{E}_t\left[r_{t+1} - r_{f,t}\right] = \beta_0 + \beta_1 F_t$ is the risk premium of the n assets, which satisfies the no-arbitrage condition in equation (7), parameters β_0 and β_1 are of sizes $n \times 1$ and $n \times k$, respectively, u_{t+1} and $\epsilon_{r,t+1}$ are the

[4] In principle, y_t could also depend on u_t, in addition to F_t. In such a case, one could express u_t in terms of F_t and F_{t-1} using equation (1), substitute it in, and satisfy the assumption by expanding the state vector to include F_{t-1}. To reduce estimation errors associated with large state spaces, we instead rely on a specific empirical choice of the state vector, which we discuss below, to guarantee that this assumption holds up well in the data (our state vector spans the yields on included financial assets with an average R^2 of 99%).

systematic and idiosyncratic shocks, respectively, and β_2 is the $n \times k$ matrix of exposures of n assets to k systematic risk factors such that

$$\beta_0 = \gamma_0 + b_0 + b_1 c, \tag{11}$$

$$\beta_1 = \gamma_1 + b_1 \rho, \tag{12}$$

$$\beta_2 = \gamma_2 + b_1, \tag{13}$$

$$\epsilon_{r,t+1} = \epsilon_{p,t+1} + \epsilon_{y,t+1}. \tag{14}$$

Note that the factor dynamics, the SDF specification, and the Euler equation impose additional restrictions on the coefficients of these equations. For example, the Euler equation links risk premia β_0 and β_1 to risk exposures β_2. Below, we specify the additional restrictions that lead us to an identified model.

Deflated dividends: We infer log *deflated* dividend growth (defined as log dividend growth net of the log risk-free rate, $\Delta d_{t+1} - r_{f,t}$) from the returns identity

$$r_{t+1} - r_{f,t} = y_{t+1} + pd_{t+1} - pd_t + (\Delta d_{t+1} - r_{f,t}), \tag{15}$$

where the log price-dividend ratio, pd_t, is a nonlinear function of y_t given by $pd_t = -\,log(\exp(y_t) - 1)$.

In general, dividend dynamics are not linear because of the nonlinearity of the relation between returns, dividends, and prices. Rather than impose approximate log-linearity via a Campbell-Shiller log-linearization, we work directly with the deflated dividends and their nonlinear dynamics, expressing them in closed form after the model has been fully solved using observable returns r and dividend yields y.

Specializing the state vector: We now take a stand on the state vector F_t. Our choice is motivated by the empirical findings of Kozak, Nagel, and Santosh (2020) and Haddad, Kozak, and Santosh (2020). In particular, Kozak, Nagel, and Santosh (2020) show that a few dominant PCs of a large cross section of anomaly portfolio returns explain the cross section of expected returns well. Haddad, Kozak, and Santosh (2020) further demonstrate that valuation ratios strongly and robustly predict expected returns on these PCs—and their risk prices in the SDF—in the time series. They argue that the resulting time variation in risk prices is critical to adequately capture the dynamic properties of the pricing kernel.

We therefore employ a specification that is motivated by both results. We assume that the dynamics of the economy are fully captured by $p = k/2$ linear combinations of *excess log returns* $f_{r,t} = Q'\left(r_t - r_{f,t}\right)$ and p linear combinations of *yields* $f_{y,t} = Q'y_t$ of the n assets, for some $n \times p$ matrix Q that constructs the same p linear combinations of these variables based on the assets:

$$F_t \equiv \begin{bmatrix} f_{r,t} \\ f_{y,t} \end{bmatrix} = \begin{bmatrix} Q'(r_t - r_{f,t}) \\ Q'y_t \end{bmatrix}. \tag{16}$$

We refer to the p linear combinations of excess log returns, $f_{r,t}$, as *factor (excess) returns* and to the p linear combinations of yields, $f_{y,t} = Q'y_t$, as *factor yields* (or factor D/P ratios).

Being a linear combination of log returns, the “factor returns” $f_{r,t}$ are not themselves log returns on a portfolio of assets with weights given by Q, that is, $f_{r,t}$ is not itself the log excess return of a tradable portfolio (except for the market, which we treat separately and include as the first factor). However, because each of the n assets is tradable, the Euler equation has to hold for each, so that the linear combination Q of the Euler equations has to hold as well—a restriction we impose in solving the model.

Our parsimonious specification allows us to be consistent with both Kozak, Nagel, and Santosh (2020) and Haddad, Kozak, and Santosh (2020): the p-factor returns explain the cross section of expected returns well, while the p-factor yields allow us to explain the cross section of dividend yields to satisfy equation (9), as well as capture the dynamics of conditional means and SDF risk prices.

This specification can be seen as an extension of the setup of Campbell (1991), in which the dynamics of the economy (also represented by a VAR) include returns and the D/P ratio of *one* portfolio (the market), plus additional predictors. Here, the state vector includes the pair $(f_{r,t}, f_{y,t})$ for p linear combinations of test assets and no additional predictors.

Given that our SDF is based on that of Kozak, Nagel, and Santosh (2018) and Haddad, Kozak, and Santosh (2020), our model naturally performs equally well in pricing the cross section of characteristic-sorted portfolios that we use as test assets. It is important to note, however, that this does not imply that our specification will mechanically be able to price *other* assets, including dividend strips on different underlying portfolios. The ability to price these other assets depends on both the SDF and the structure of the dynamics in the model. This is why, after estimating our model using available stock portfolios, we evaluate the ability of the model to generate realistic term structures by comparing the implied dividend term structures with those observed in the data. As we discuss below, we find that the model does a good job in matching traded S&P 500 dividend strip data.

Restrictions: We now introduce three restrictions on the model. First, we assume that there are only p priced risks, and that they are fully spanned by our p return factors, so m_t loads only on the p innovations in return factors (the first p elements of u_{t+1}). This means that only the first p elements of λ_t are nonzero. In turn, this implies that only the first p rows of λ and Λ can be nonzero. The remaining shocks in u_t drive the dynamics of the economy but are not priced by investors. This assumption is motivated by Kozak, Nagel, and Santosh (2018, 2020), who suggests that an SDF constructed from a small number (p) of diversified portfolio returns well prices the cross section of returns. We then simply allow the dynamics to also include additional, nonpriced shocks.

Second, we impose that only dividend yields (and not lagged returns) drive time variation in risk premia. This means that the matrix Λ will have the structure

$$\Lambda = \begin{bmatrix} 0_{p\times p} & \tilde{\Lambda} \\ 0_{p\times p} & 0_{p\times p} \end{bmatrix},$$

where the zeros in the second row are due to the fact that only return shocks are priced, the zeros in the top-left corner imply that variation in risk premia is entirely driven by factor yields $f_{y,t}$, and $\tilde{\Lambda}$ is a $p \times p$ matrix of risk price loadings on dividend yields.

As we show in greater detail below, these first two assumptions imply restrictions on the transition matrix ρ of the factors. In particular, they imply that expectations of factor returns are a function of lagged yields but not lagged realized returns. To this restriction, we add a third one, on the conditional mean of the yields: we impose that it is also a function of lagged yields but not lagged returns. We therefore assume that

$$\rho = \begin{bmatrix} 0_{p\times p} & \rho_{r,y} \\ 0_{p\times p} & \rho_{y,y} \end{bmatrix},$$

where $\rho_{r,y}$ could potentially be further restricted to be a $p \times p$ diagonal matrix, based on the evidence in Haddad, Kozak, and Santosh (2020) that own valuation ratios are the strongest predictors of factor returns. We do not currently impose the latter restriction to remain as flexible as possible. The former restriction—that lagged returns forecast neither returns nor yields—is relatively mild in our opinion, and consistent with the voluminous literature documenting low autocorrelation in equity returns.

Note that similar restrictions have been imposed in the term-structure literature, for instance, in Cochrane and Piazzesi (2008). They specify an SDF in which only shocks to bond yields are priced, and their SDF risk prices are fully driven by the Cochrane and Piazzesi (2005) factor alone.

Prices of risk: Using the definition of factor returns in (10), we use (16) to express them as

$$f_{r,t+1} = Q'(r_{t+1} - r_{f,t}) = \beta_{f,0} + \beta_{f,1}F_t + \beta_{f,2}u_{t+1}, \tag{17}$$

where

$$\beta_{f,0} = Q'\beta_0 = c_r, \tag{18}$$

$$\beta_{f,1} = Q'\beta_1 = [\mathbf{0}_{p\times p}, \rho_{r,y}], \tag{19}$$

$$\beta_{f,2} = Q'\beta_2 = [I_{p\times p}, \mathbf{0}_{p\times p}]. \tag{20}$$

The first equality in each of the equations above reflects the fact that "factor returns" $f_{r,t+1}$ are linear combinations of test assets' log excess returns, while the second equality holds because each factor is part of the state-space vector

F_t, which imposes specific restrictions on all of the parameters (and leads to the absence of idiosyncratic shocks in (17)).

We now use these restrictions to show how the factors and the Euler equation can be used to link the parameters β to the prices of risk in the SDF.

We plug the expressions in (2), (8), and (9) into (7) to obtain

$$\mathbf{0}_{n\times 1} = \beta_0 + \beta_1 F_t - \beta_2 \Sigma(\lambda + \Lambda F_t) + \frac{1}{2}\,\mathrm{diag}\left[\beta_2 \Sigma \beta_2' + \Sigma_\epsilon\right]. \tag{21}$$

Premultiplying by Q', expressing in terms of $\beta_{f,\cdot}$, and matching coefficients on F_t, we get two equations that can be used to solve for λ and Λ, given the parameters c_r and $\rho_{r,y}$ and estimates of the test assets' variance terms:

$$0 = \beta_{f,1} - \beta_{f,2}\Sigma\Lambda, \tag{22}$$

$$0 = \beta_{f,0} - \beta_{f,2}\Sigma\lambda + \frac{1}{2}Q'\,\mathrm{diag}\left[\beta_2 \Sigma \beta_2' + \Sigma_\epsilon\right]. \tag{23}$$

C. Discussion of the Setup

An alternative modeling setup adopted in the literature, which we refer to as the "alternative" modeling approach, is to specify a log-linear SDF and linear dynamics for log dividend growth, and to then obtain prices by imposing the Euler equation (e.g., see Brennan, Wang, and Xia (2004), Koijen, Lustig, and Van Nieuwerburgh (2017), and Gupta and Van Nieuwerburgh (2021)). The main difference between such a specification and our specification is that we model log prices (or, equivalently, excess log equity returns) as linear in factors F_t and shocks u_{t+1}, whereas the alternative specification models log dividend growth Δd_{t+1} as linear in F_t and u_{t+1}. Both assumptions are plausible and both are widely used in asset pricing. The alternative assumption of dividend growth linearity is primarily used in the macrofinance literature (e.g., in the context of Lucas economies). Our assumption of (log) equity price and return linearity is widespread in empirical asset pricing in the context of factor models, and in return-based models like the intertemporal capital asset pricing model (ICAPM). In this section, we discuss in detail the relationship between the two approaches.

Returns, dividends, and price-dividend ratios are related by the usual nonlinear identity in (15). Given that we aim to price both dividend strips and stocks, two Euler equations need to be satisfied, for strips and stocks, respectively,

$$\frac{P_t^{(n)}}{D_t} = E_t\left[M_{t+1}\frac{D_{t+1}}{D_t}\frac{P_{t+1}^{(n-1)}}{D_{t+1}}\right], \tag{24}$$

$$1 = E_t\left[M_{t+1}\frac{P_{t+1}}{P_t}\left(1 + \frac{D_{t+1}}{P_{t+1}}\right)\right] \equiv E_t\left[M_{t+1}R_{t+1}\right], \tag{25}$$

where $P_t^{(n)}$ denotes the price on an n-year dividend claim.

As mentioned above, the alternative approach assumes that $\Delta d_{t+1} = \log\left(\frac{D_{t+1}}{D_t}\right)$ is linear in F_t and u_{t+1}. If M_{t+1} is also log-linear in F_t and u_{t+1}, then it is immediate from (24) that the log price-dividend ratios of dividend strips of all portfolios, $pd_t^{(n)}$, will be linear in F_t. The stock's price-dividend ratio, however, cannot be linear in F_t at the same time. To facilitate estimation of the model for directly observable stock D/P ratios, the alternative approach typically relies on the Campbell-Shiller approximation (approximate log-linearization), under which pd_t is approximately linear in F_t.[5] Due to these approximations, the price of a stock under this approach is not equal to the sum of the prices of the strips.

Under our approach, from (25), if excess returns $r_{t+1} - r_{f,t}$ are linear in F_t and u_{t+1}, and M_{t+1} is log-linear, it follows that the equity—rather than strip—D/P ratios, $y_t = \log\left(1 + \frac{D_t}{P_t}\right)$, are linear in F_t. Equity returns are then linear too, whereas pd_t, $pd_t^{(n)}$, and $\Delta d_{t+1} - r_{f,t}$ will not be linear in F_t and u_{t+1}. Our assumptions have several advantages. First, based on the observable state space, our affine model can be estimated without imposing the Campbell-Shiller approximation (instead we impose the exact identity linking prices, dividends, and returns). Intuitively, this is because we observe D/Ps of stocks rather than strips in the data, so making stocks' D/Ps (rather than dividend strips' D/Ps) linear in states and including them in the state vector makes the estimation process much easier. Second, after we fully solve the affine model, we compute dividend strip yields and implied deflated dividend growth, allowing them to be exact nonlinear functions of F_t. As a result, we automatically satisfy the no-arbitrage condition that the price of a stock is exactly equal to the sum of the prices of all of its dividends.

Note that if one wants to impose the Campbell-Shiller approximation, then our approach is exactly equivalent to the alternative approach. To see this, note further that the Campbell-Shiller approximation of the returns equation (15) makes $\log\left(1 + e^{pd_t}\right)$ an approximately linear function of pd_t, that is,

$$\log\left(1 + e^{pd_t}\right) \simeq a + b \times pd_t. \tag{26}$$

Under the Campbell-Shiller approximation, both y_t and pd_t are (approximately) linear in F_t, and hence so are dividend growth and returns. To see this, note that the approximation in (26) and the identity $y_t = \log\left(1 + \frac{D_t}{P_t}\right) = \log\left(1 + e^{pd_t}\right) - pd_t$ imply that y_t is affine in pd_t. Therefore, under either of the two setups discussed above, both y_t and pd_t are linear in F_t—the two setups are equivalent.

To summarize, given the nonlinearity of the fundamental identity relating pd_t, r_{t+1}, and Δd_{t+1} in (15), if any of the three components is linear in some

[5] See, for example, Koijen, Lustig, and Van Nieuwerburgh (2017) and Gupta and Van Nieuwerburgh (2021). Alternatively, rather than specify an observable Gaussian state vector as we do in the paper, one could introduce a latent Gaussian state vector that is nonlinearly related to D/P ratios and solve the alternative approach exactly.

variables F_t, the others cannot all also be linear in F_t. The alternative approach assumes that Δd_{t+1} and $pd_t^{(n)}$ are linear (and so both pd_t and y_t are nonlinear), but the Euler equation cannot be solved analytically under this assumption alone. To close the model analytically, one needs to also assume the Campbell-Shiller approximation or approximately solve a system of nonlinear equations defining D/P ratios for each portfolio. Our approach assumes instead that $r_{t+1} - r_{f,t}$ and y_t are linear in F_t. These assumptions are sufficient to solve the model fully even without imposing the Campbell-Shiller approximate linearization. Only *after* our model is fully solved analytically, the remaining quantities (e.g., $\Delta d_{t+1} - r_{f,t}$, pd_t, $pd_t^{(n)}$) can be expressed as exact nonlinear functions of the state vector F_t and shock u_{t+1}. Finally, if one is willing to impose the Campbell-Shiller approximation, the two setups are equivalent.

We conclude by noting that the idea of modeling $y_t = \log\left(1 + \frac{D_t}{P_t}\right)$, as opposed to $pd_t = \log\left(\frac{P_t}{D_t}\right)$, has appeared in the macrofinance literature. For example, Martin (2013) uses it to study the importance of cumulants in consumption-based models, and Gao and Martin (2021) derive a log-linear approximation of y_t instead of pd_t similar to that of Campbell-Shiller.

D. Identification and Estimation

While the model contains a large number of parameters, many of them are related through no-arbitrage restrictions or through some of the additional restrictions imposed above. It will be useful to distinguish *reduced-form* parameters that directly enter moment conditions that can be estimated from the data, and *structural* parameters that determine the reduced-form parameters in turn.

Reduced-form parameters and moment conditions: The entire state vector F_t is fully observed. Equation (1) therefore implies a first set of moment conditions, which depend on the parameters c and ρ. We use ordinary least squares (OLS) moments, with the additional restriction that elements of ρ corresponding to loadings on lagged returns are all zero, that is, $\mathbb{E}\left(u_{t+1} \otimes \left[1, F_{y,t}\right]\right) = 0$, where $\otimes$ denotes a Kronecker product. Note that for factor returns, which are part of F_t, the parameters $\beta_{f,0}, \beta_{f,1}$, and $\beta_{f,2}$ from (17) are a direct function of c and ρ (because these return equations are just the first p rows of (1)).

For test assets, we have two sets of moment conditions: one contemporaneously relating their yields to the factors, (9), and another relating their returns to lagged factors, (10). These equations form a set of moment conditions that depend on parameters β_0, β_1, β_2, b_0, and b_1 for all test assets. We again use OLS moments, with parameter restrictions $\mathbb{E}\left(\varepsilon_{t+1} \otimes \left[1, F_{y,t}, u_{r,t+1}\right]\right) = 0$ and $\mathbb{E}\left(\epsilon_t \otimes \left[1, F_{y,t}\right]\right) = 0$.

To summarize, moment conditions from equations (1), (9), and (10) depend on the parameters c and ρ, as well as β_0, β_1, β_2, b_0, and b_1 for all assets.

Structural parameters: We now have a set of structural parameters that are linked to the reduced-form parameters by identities and arbitrage restrictions.

First, the parameters γ_0, γ_1, and γ_2 are related one-to-one to the reduced-form parameters ((11), (12), (13)); these parameters do not add any additional restriction to the system, and simply correspond to an alternative representation of (10), stated in terms of Δp_{t+1} instead of r_{t+1}.

More important are the other structural parameters λ and Λ. These parameters introduce restrictions on the reduced-form parameters of *all* portfolios, through valuation equation (7).

Estimation and inference: We estimate the model using generalized method of moments (GMM). We use the moment conditions described above to estimate the reduced-form and structural parameters (i.e., imposing the valuation restriction (7) using all test assets). We use a prespecified, diagonal weighting matrix for GMM where the factor moments are weighted by one, and the individual asset moments are all normalized by the square root of the number of test assets, $\sqrt{n}$, to keep their contribution to the GMM objective invariant to n.[6] This weighting matrix ensures a good balance between the two sets of moment conditions and yields reasonable estimates of the risk premia for all 102 portfolios, which is important in our context given the weaker factor structure of equities compared to other settings (e.g., bonds).

We estimate the dynamics at an annual horizon using monthly data (and therefore, using overlapping yearly observations). We derive standard asymptotic GMM standard errors for all reduced-form and structural parameters. To account for overlapping data, we use a spectral density covariance matrix of moments with 12 lags, following the approach in Hansen and Hodrick (1980).[7] Finally, we compute standard errors on derived quantities—such as model-implied yields and returns on dividend strips, which are nonlinear functions of structural parameters—using the delta method.

Note that there are two sources of uncertainty for the model-implied moments of interest. First, there is uncertainty stemming from the fact that the model parameters have to be estimated, and so some time series are not observed but rather estimated (e.g., the time series of the prices of dividend strips, which are not an input in our estimation and hence must be estimated). We refer to this type of uncertainty as *parameter uncertainty*. Second, there is the uncertainty stemming from the fact that some objects of interest are unconditional moments, and their estimation requires computing time-series averages. We refer to this type of uncertainty as *sampling uncertainty*. For example, suppose that we are interested in the unconditional slope of the dividend strip term structure. The GMM estimator directly gives us the estimate

[6] We also explored the use of efficient GMM, and found that it is numerically unstable in our setting (which features a large number of moment conditions and parameters). This result is similar to that in Campbell, Giglio, and Polk (2013), who use GMM to jointly estimate, in a lower-dimensional setting than ours, the dynamics of the model and the test asset moment conditions and also find efficient GMM to be numerically unstable. Rather than imposing restrictions on the parameter space (e.g., bounds on risk prices) as in Campbell, Giglio, and Polk (2013) to ensure convergence of the efficient GMM estimator, here we use a prespecified weighting matrix.

[7] We perform a nonparametric bootstrap exercise to validate our standard errors in Section III.G.6.

for this moment, together with standard errors that incorporate *all* types of estimation uncertainty. However, we can separate the two by employing the following estimation procedure. First, an estimate of the slope at each point in time is obtained by estimating our model with GMM, estimating the model-implied prices of long- and short-term dividend strips at each point in time, and computing the difference between the two in each period. Next, the unconditional average slope (the quantity of interest) is estimated by taking the time-series average of the slope estimated at each point in time. The two sources of uncertainty reflect the two steps in this procedure. Note that if dividend strip prices were directly observable, the parameter uncertainty would disappear and the sampling uncertainty would account for the entirety of the estimation uncertainty.

The GMM estimation approach that we describe above automatically takes *both* sources of uncertainty into account when estimating any moment. But comparing the two sources of uncertainty is useful to understand how much of our estimation uncertainty on objects of interest (like the unconditional slope of the dividend term structure) is due to the fact that we do not observe the dividend strip prices, and how much of that uncertainty would exist even if we did observe dividend strip prices, since it arises from the fact that we only have a finite sample to work with to estimate unconditional moments.

This estimation uncertainty decomposition allows us to understand the trade-off that our approach to estimating term structures faces, compared to working only with traded securities: our procedure adds *parameter uncertainty* to the estimation (since we need to estimate the model to generate prices of securities like dividend strips), but has the advantage of significantly lengthening the time series available for the study of unconditional moments, therefore reducing *sampling uncertainty* (since traded dividend strips are available only for a short time period). In practice, we find that parameter uncertainty is small relative to data sampling uncertainty for most moments we consider. This observation suggests that our procedure leads to the overall reduction of standard errors compared to methods that use traded strips but rely on a shorter sample.

We provide more details on the estimation in Appendix B. We return to the comparison of the sources of uncertainty when we present the empirical results.

E. Dividend Strips

We next derive the prices and returns of theoretical dividend strips in the model. These can be computed *after* having estimated the entire model because they are simple functions of the parameters we estimate via GMM. Consider first a fully diversified portfolio with dividend D_t and price P_t. Note first that since

$$P_{t+n} = P_t \exp\left[\sum_{i=1}^{n} \Delta p_{t+i}\right]$$

and

$$\frac{D_{t+n}}{P_{t+n}} = [\exp(y_{t+n}) - 1],$$

we have

$$\frac{D_{t+n}}{P_t} = \frac{D_{t+n}}{P_{t+n}} \exp\left[\sum_{i=1}^{n} \Delta p_{t+i}\right] = [\exp(y_{t+n}) - 1] \exp\left(\sum_{i=1}^{n} \Delta p_{t+i}\right),$$

and hence the price of a claim to D_{t+n}, $P_t^{(n)}$, as a fraction of the price of the portfolio, P_t, is

$$\frac{P_t^{(n)}}{P_t} \equiv w_t^{(n)} = \mathbb{E}^Q\left\{\frac{D_{t+n}}{P_t} \exp\left[-\sum_{i=1}^{n} r_{f,t+i-1}\right]\right\}$$

$$= \mathbb{E}^Q\left\{[\exp(y_{t+n}) - 1] \exp\left[\sum_{i=1}^{n} \left(\Delta p_{t+i} - r_{f,t+i-1}\right)\right]\right\} \tag{27}$$

$$= \mathbb{E}^Q\left\{\exp\left[y_{t+n} + \sum_{i=1}^{n} \left(\Delta p_{t+i} - r_{f,t+i-1}\right)\right]\right\} - \mathbb{E}^Q\left\{\exp\left[\sum_{i=1}^{n} \left(\Delta p_{t+i} - r_{f,t+i-1}\right)\right]\right\} \tag{28}$$

$$= \exp\left(a_{n,1} + d_{n,1}F_t\right) - \exp\left(a_{n,2} + d_{n,2}F_t\right). \tag{29}$$

The parameters in this equation can be shown to satisfy the recursions

$$a_{n,\cdot} = a_{n-1,\cdot} + \gamma_0^* + d_{n-1,\cdot}c^* + \frac{1}{2}\left(d_{n-1,\cdot} + \gamma_2\right)\Sigma\left(d_{n-1,\cdot} + \gamma_2\right)' + \frac{1}{2}\sigma_v^2, \tag{30}$$

$$d_{n,\cdot} = \gamma_1^* + d_{n-1,\cdot}\rho^*, \tag{31}$$

with initial values $a_{0,1} = b_0 + \frac{1}{2}\left(\sigma_r^2 - \sigma_v^2\right)$, $d_{0,1} = b_1$, $a_{0,2} = 0$, $d_{0,2} = 0$, $\sigma_r^2 = \text{var}(\varepsilon_t)$, and $\sigma_v^2 = \text{var}(v_t)$. In these formulas, stars indicate risk-neutral parameters.[8] In our estimation, we find that the risk-neutral dynamics of the model are stationary, which is sufficient to guarantee that the price of the stock market, which is the infinite sum of the prices of the dividend strips, is finite. Note that the prices of dividend strips are obtained as a function of observable excess returns and dividend yields and their dynamics (captured by F_t and their dynamics). This is because specifying these dynamics implicitly fully specifies the dynamics of deflated dividends, which are sufficient to pin down dividend strip prices. Note also that $w_t^{(n)}$ can be interpreted as cap-weights of each dividend within a portfolio of all dividends of a stock (i.e., the stock itself).

[8] Risk-neutral parameters are defined as $\gamma_0^* = \gamma_0 - \gamma_2\Sigma\lambda$, $\gamma_1^* = \gamma_1 - \gamma_2\Sigma\Lambda$, $c^* = c - \Sigma\lambda$, and $\rho^* = \rho - \Sigma\Lambda$.

Equity yields: We can also compute equity yields (for assets/portfolios with strictly positive dividends) at time t with maturity n, $e_{t,n}$, as defined in van Binsbergen et al. (2013):

$$e_{t,n} = \frac{1}{n} \log\left(\frac{D_t}{P_t^{(n)}}\right) = \frac{1}{n}\left[\log\left(\exp\left(y_t\right) - 1\right) - \log\left(\frac{P_t^{(n)}}{P_t}\right)\right]. \tag{32}$$

Forward equity can be easily computed using the (externally) observable bond yields

$$e_{t,n}^{f} = \frac{1}{n} \log\left(\frac{D_t}{F_{t,n}}\right) = e_{t,n} - y_{t,n}^{b}, \tag{33}$$

where $y_{t,n}^{b}$ is the nominal government bond yield with no default risk, and $F_{t,n}$ denotes the futures (or forward) price, which, under no arbitrage, is linked to the spot price by

$$F_{t,n} = P_{t,n} \exp\left(n y_{t,n}^{b}\right). \tag{34}$$

In Appendix A, we provide additional definitions and derivations for returns and expected returns on dividend strips in our model. In addition, we show that equity yields can be decomposed into log (expected) hold-to-maturity returns and log (expected) cumulative deflated dividend growth rates (see (A17) and (A19)).

II. Data

A. Stock Data

We focus on a broad set of 51 stock-specific characteristics and long and short legs of portfolio sorts based on these characteristics.[9] We construct these portfolios as follows. We use the universe of CRSP and Compustat stocks and sort them into three value-weighted portfolios for each of the characteristics studied in Kozak, Nagel, and Santosh (2020) and Kozak and Santosh (2020) and listed in Table IA.III in the Internet Appendix, for a total of 51 characteristics.[10] Portfolios include all NYSE, AMEX, and NASDAQ firms but the breakpoints use only NYSE firms as in Fama and French (2016). We obtain 102 portfolios, two for each anomaly (P1 and P3). Our sample consists of monthly returns from February 1973 to December 2020. For each portfolio, we

[9] Our data are available at https://www.serhiykozak.com/data.

[10] The Internet Appendix may be found in the online version of this article. We apply the following two automatic filters to characteristics as in Kozak, Nagel, and Santosh (2020) and Kozak and Santosh (2020) to arrive at 51 characteristics: (i) mark portfolio returns as missing if a portfolio contains fewer than 100 firms at any point in time, and (ii) remove characteristics for which more than 120 months of returns are missing, in either a short leg or a long leg.

construct a corresponding measure of its valuation based on D/P ratios of the underlying stocks. We then construct the yield, y_t, as defined in equation (6).[11]

We also construct long-short portfolios as differences between each anomaly's log return on portfolio 3 minus the log return on portfolio 1. Their valuation ratios are defined as the difference in yields y_t between the two legs. Most of these portfolio sorts exhibit a significant spread in average returns and CAPM alphas. This finding is documented in the vast literature on the cross section of returns and can be verified in Table IA.III. In our sample, most anomalies show a strong pattern in average returns across tercile portfolios, consistent with prior research.

B. Choice of Variables in the State Vector

Kozak, Nagel, and Santosh (2018, 2020) show that a few dominant PCs of a large cross section of anomaly portfolio returns explain the cross section of expected returns well. We use this insight to guide our choice of portfolios in the state-space dynamics in equation (1). In particular, we use excess log returns on the market and three PCs of 51 long-short portfolio returns based on the underlying long and short ends of each characteristic used in sorting as our choice for $f_{r,t}$ in equation (16). Formally, consider the eigenvalue decomposition of anomaly excess returns, $\text{cov}\left(r_{LS,t+1}\right) = \tilde{Q}\Lambda\tilde{Q}'$, where $\tilde{Q} = [\tilde{q}_1, \ldots, \tilde{q}_{51}]$ is the matrix of eigenvectors and Λ is the diagonal matrix of eigenvalues. The i^{th} PC portfolio is formed as $PC_{i,t+1} = \tilde{q}_i' r_{LS,t+1} = q_i'\left(r_{t+1} - r_{f,t}\right)$, where $\tilde{q}_i$ is the i^{th} column of $\tilde{Q}$, $q_i = [\tilde{q}_i, -\tilde{q}_i]'$, and $r_{t+1} - r_{f,t}$ are excess log returns on 102 portfolios underlying the long-short anomalies.

Table I shows that anomaly portfolio returns exhibit a moderately strong factor structure. Panel A focuses on long-short portfolios with the market, which captures the vast majority of all comovement across portfolios, effectively removed. It therefore focuses on explaining the remaining variation in portfolio returns once the market has been removed. The first PC accounts for one-fourth of the total variation. The first three PCs explain more than 50% of the variation not captured by the market. Panel B extracts PCs of all 102 long and short legs of each sort, not subtracting the market. It shows that the first four PCs (approximately the market and three cross-sectional PCs) capture more than 95% of the variation in returns across 102 portfolios.

Haddad, Kozak, and Santosh (2020) further demonstrate that valuation ratios strongly and robustly predict expected returns on these PCs—and their risk prices in the SDF—in the time series. They argue that the resulting time variation in risk prices is critical for adequately capturing dynamic properties of the pricing kernel. Motivated by this evidence, we pick four yields corresponding to the market and three PC factors as our choice for y_t in equation (16). Yields on PCs are constructed simply as a linear combination of log

[11] The D/P ratio of a portfolio is defined as the sum of all dividends paid within the last year relative to the total market capitalization of all firms in that portfolio. Equivalently, it is the market-capitalization-weighted average of individual stocks' D/P ratios.

Table I
Percentage of Variance Explained by Anomaly PCs

This table reports the percentage of variance explained by each PC of the 51 long-short anomaly portfolio returns (Panel A) and each PC of 102 returns on long and short legs of each characteristic sort (Panel B). PC1 in Panel B refers to the market portfolio.

	PC1 (1)	PC2 (2)	PC3 (3)	PC4 (4)	PC5 (5)	PC6 (6)	PC7 (7)	PC8 (8)	PC9 (9)	PC10 (10)
	Panel A: Long-Short (51 portfolios)									
% Var. explained	24.9	19.4	10.1	5.7	4.3	3.9	3.1	2.9	2.5	2.2
Cumulative	24.9	44.3	54.4	60.1	64.4	68.3	71.4	74.3	76.8	79.0
	Panel B: Long and Short Legs (102 portfolios)									
% Var. explained	88.8	4.1	1.6	0.9	0.6	0.5	0.5	0.3	0.3	0.2
Cumulative	88.8	92.9	94.5	95.5	96.1	96.5	97.0	97.3	97.6	97.8

yields on underlying portfolios, with weights determined by returns' eigenvectors, $y_{pc_i} = q_i' y_t$.

In summary, our state vector includes returns and yields on the market, three PCs of long-short log returns, and their valuation ratios for all test assets. We validate this choice by conducting an out-of-sample analysis in Section III.G.4.

C. Bond Data

Our model is designed to be orthogonal to the term structure of interest rates and does not require any bond data to estimate the model, other than the one-year risk-free rate $r_{f,t}$. To compute forward equity yields (for which bond yields are needed), we use data from Gurkaynak, Sack, and Wright (2006), who provide a long history of interpolated U.S. bond yields.

After merging the equity and bond data, we obtain a full sample of annual (overlapping) observations at the monthly frequency from February 1973 to December 2020.

III. Results

In this section, we report empirical results of our estimation. We begin by reporting the estimates and fit of the model. Specifically, we show how the model fits the returns of the 51 anomaly portfolios, their price-dividend ratios, and their dividends.

We next evaluate the performance of the model against data *not* used in the estimation: the term structure of S&P 500 dividend strips and futures. We show that the dividend strips implied by our estimated model closely match the prices of the traded strips. We therefore replicate the main facts of van Binsbergen, Brandt, and Koijen (2012), van Binsbergen et al. (2013), and van Binsbergen and Koijen (2015) on the term structure of S&P 500 dividend strips and forwards. We view this exercise mainly as an out-of-sample validation of our empirical setup.

Having validated our model, we explore the novel empirical facts that emerge from our model, along both the cross-sectional and the time-series dimensions. In the cross section, we show that our model produces a rich variety of shapes for the term structures for different portfolios. In the time series, we show that the slope and the shape of the term structure of different portfolios vary in interesting ways over time. Finally, we discuss various implications and applications of our results.

Taken together, our results show that the term structures of discount rates are heterogeneous across types of risks (captured by different portfolios) and vary significantly over time. The empirical patterns that we extract from the data in Section III.D (e.g., the difference in the term-structure behavior of value and growth stocks or of high- and low-profitability stocks) provide new moments that can help guide the construction and evaluation of asset pricing models.

A. Fit of the Model

As discussed in Section I.D, we estimate our model in one step via GMM. The core of our model are the time-series dynamics of the dividend yields and returns of the four factors (the market plus three PCs from 51 anomalies). Table II reports estimates of the dynamics of the factors F_t: c and ρ. As discussed in Section I, we impose the restriction that the conditional mean of returns and yields is a function only of lagged yields, not lagged returns, that is, $\rho = \left[0_{K\times p}, \rho_{\cdot,y}\right]$. Thus, in Table II we omit zeros and reports only estimates and standard errors of $\rho_{\cdot,y}$, as well as c. In parentheses, we report GMM standard errors using a spectral density matrix with 12 lags.

For comparison, Table III reports the risk-neutral parameters c^Q and ρ^Q implied by the model. Note that all conditional loadings of returns on lagged yields are zero, $\rho_{r,y} = 0$, since under the risk-neutral dynamics the expected excess return on any asset is equal to the risk-free rate. We therefore report only loadings of yields onto lagged yields, $\rho_{y,y}$, as well as the intercepts, c_r^Q and c_y^Q. The intercepts of the return regressions, c_r^Q, are nonzero and reflect a variance adjustment.

The last column of Table II also shows the R^2 for each of the eight equations of the dynamics of F_t. The first four regressions are effectively predictive regressions of yearly excess returns using the lagged dividend yields of the factor portfolios as predictors; the last four regressions are predictive regressions of dividend yields using lagged dividend yields as predictors, again for the four-factor portfolios. The results are consistent with those in Haddad, Kozak, and Santosh (2020), who show that PCs of anomaly returns are strongly and robustly predictable by their valuation ratios, with this predictability essential to capture dynamic properties of an SDF. The authors also show that this predictability survives out-of-sample, suggesting that our analysis should be robust to out-of-sample tests.

Table II

Estimates of the Dynamics of the Factors F_t

This table reports estimates of the dynamics of the factors F_t in equation (1): c and $\rho_{\cdot,y}$ in $\rho = [0_{K\times p}, \rho_{\cdot,y}]$, and the R^2 for each of the eight equations of the dynamics of F_t. Hansen-Hodrick GMM standard errors using a spectral density matrix with 12 lags are reported in parentheses. The dynamics are estimated at an annual horizon using monthly overlapping observations in the February 1973 to December 2020sample.

	c	$\rho_{\cdot,y}$				R^2
	(1)	(2)	(3)	(4)	(5)	(6)
r_{mkt}	0.07	−0.16	1.35	3.21	0.50	10.44
	(0.068)	(4.8)	(1.7)	(1.5)	(0.9)	–
r_{pc1}	−0.05	−16.13	11.39	4.23	−2.57	22.09
	(0.2)	(15)	(4.8)	(3.2)	(2.1)	–
r_{pc2}	0.26	−7.54	0.34	1.12	0.06	7.19
	(0.12)	(3.8)	(2.2)	(2.9)	(1.6)	–
r_{pc3}	0.08	−0.64	1.46	3.88	1.40	11.62
	(0.081)	(5)	(1.7)	(1.5)	(1)	–
y_{mkt}	0.00	0.93	−0.00	0.02	−0.01	97.58
	(0.0015)	(0.086)	(0.035)	(0.038)	(0.034)	–
y_{pc1}	−0.00	0.57	0.72	0.02	−0.23	96.98
	(0.0046)	(0.25)	(0.091)	(0.083)	(0.051)	–
y_{pc2}	−0.01	−0.25	0.03	0.44	−0.10	71.19
	(0.0062)	(0.27)	(0.17)	(0.25)	(0.072)	–
y_{pc3}	−0.01	−0.58	0.31	0.08	0.01	20.32
	(0.0053)	(0.27)	(0.088)	(0.11)	(0.11)	–

Table III

Risk-Neutral Estimates of the Dynamics of the Factors F_t

This table reports risk-neutral estimates of the dynamics of the factors F_t in equation (1): c_r^Q, c_y^Q, and $\rho_{y,y}^Q$ in $c^Q = \left[c_r^Q, c_y^Q\right]'$ and $\rho^Q = \left[0_{K\times p}, \rho_{\cdot,y}^Q\right]$, where $\rho_{\cdot,y}^Q = \left[0, \rho_{y,y}^Q\right]'$. The sample comprises annual overlapping observations at the monthly frequency from February 1973 to December 2020.

	c_r^Q	c_y^Q	$\rho_{y,y}^Q$			
	(1)	(2)	(3)	(4)	(5)	(6)
MKT	−0.01	0.00	0.92	0.03	0.09	−0.00
PC1	0.03	−0.00	0.33	0.94	0.19	−0.26
PC2	0.03	−0.00	−0.24	−0.16	0.35	−0.06
PC3	−0.00	−0.00	−0.35	0.18	0.14	0.10

One of the main objectives of the model is to fit the cross section of returns for the 51 anomalies (not just the four linear combinations that we use as factors). The average R^2 for the regression of portfolio returns in equation (10) is 93.3%. The average R^2 for the regression of dividend yields onto the factor dividend yields in equation (9) is 98.7%. Our model also well fits the cross section of risk premia for the 51 anomalies, reaching a cross-sectional R^2 of 42.9% in the full sample from February 1973 to December 2020. This compares favorably with

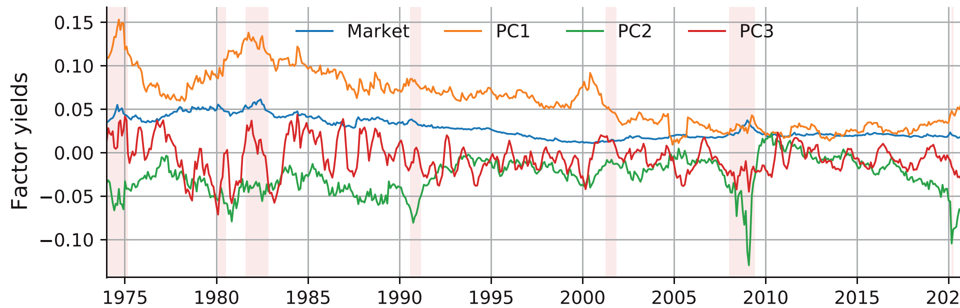


Figure 1. Time series of factor yields. This figure shows factor yields, $f_{y,t}$, for the aggregate market and three PCs of anomaly portfolios. Annual dividends; monthly overlapping observations.

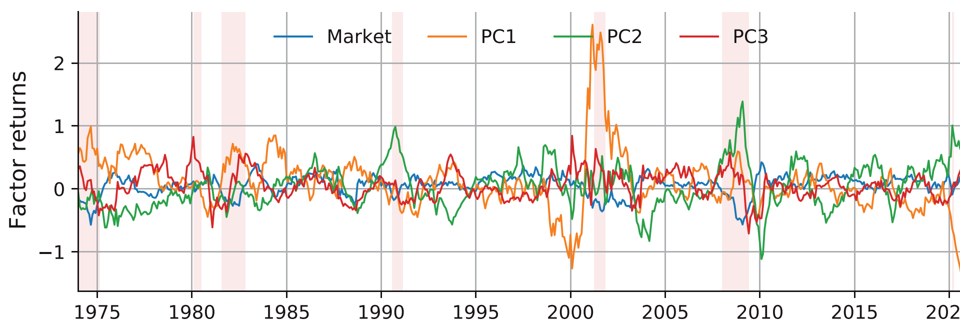


Figure 2. Time series of factor returns. This figure shows factor returns, $f_{r,t}$, for the aggregate market and three PCs of anomaly portfolios. Annual returns; monthly overlapping observations.

the R^2 of 73% one obtains on the much narrower cross section of 25 size and book-to-market-sorted portfolios when using the three Fama-French factors (Fama-French factors explain the cross section of 51 anomalies with an R^2 close to 0%).[12] This evidence demonstrates that our model performs well in fitting not only the dynamics of the PCs themselves, but also the time-series and cross-sectional properties of the 51 anomalies.

To get a sense of the factors' behavior in the data, Figures 1 and 2 plot the time series of factor returns and factor yields. Several interesting patterns are worth noting. First, while the first factor (the market) has an extremely persistent dividend yield, as is well known in the literature, the other factors' persistence is significantly lower. Second, the factors clearly capture different economic forces. For example, the fourth factor captures relatively high-frequency dynamics, whereas the third factor is most strongly associated with the financial crisis.

The fact that these factors display different dynamics is crucial for our identification. Haddad, Kozak, and Santosh (2020) argue that time-series

[12] Kozak, Nagel, and Santosh (2018) provide more detailed evidence on the relative performance of a PC-based model and the Fama-French model in explaining a wide cross section of anomaly returns, which is consistent with our findings.

predictability of these PC-based factors is critical to adequately capture the dynamic properties of the pricing kernel — this is the key motivation for our choice of the state-space vector. We can identify term structures of discount rates only because we can identify shocks to dividends and discount rates at different horizons—in other words, to estimate how investors price these shocks differently, we need to be able to estimate shocks that give rise to different dynamic responses of the economy. Intuitively, by studying how prices of portfolios respond to short-lived shocks to their dividends, we can learn about the discount rate at short horizons (controlling for discount rate movements that are correlated with the dividend shocks) and by studying how prices respond to shocks that affect dividends far in the future, we can learn about the long-term discount rate (again controlling for simultaneous discount rate changes). To identify different points along the term structure, we therefore need to observe different shocks at different horizons for dividends and discount rates. To empirically validate the idea that our model is able to capture both short- and long-term shocks to portfolios, we regress returns on the long-short duration sorted-portfolio on the variables in the state vector, F_t, and find that the R^2 of this regression is 92% (see Table IA.V). This result suggests that the factor innovations do indeed capture news that drives the wedge between short-duration and long-duration portfolios.

B. Fit to Traded S&P 500 Dividend Forwards

An important step forward in understanding the term structure of risk premia in the data and in the models comes from the study of traded claims to dividends of finite maturity (dividend strips and dividend futures), starting with the seminal work of van Binsbergen, Brandt, and Koijen (2012). Recent papers expand the sample to include the prices of dividend forwards traded over the counter and, more recently, on exchanges (e.g., van Binsbergen and Koijen (2015) and Bansal et al. (2021)).

One of the main goals of our paper is to provide a framework to recover implied dividend strip and forward prices from data that only include equity portfolios.[13] A simple and direct criterion to evaluate whether we can do so successfully is to verify whether our implied dividend forwards match those from the traded contracts (when the latter are available). In this section, we compare the time-series and cross-sectional moments of our implied dividend forwards against those reported in Bansal et al. (2021) and van Binsbergen and Koijen (2015) for the traded S&P 500 forwards. Given that the sample of traded dividend forwards used in these papers starts in 2004, we estimate our model in the full sample but focus here on the corresponding moments from the 2004 to 2020 sample, so that the sample moments are directly comparable.

Figure 3 plots the time series of the forward equity yields implied by our model (defined in equation (33)) against the most recent data from Bansal

[13] Our recovered synthetic dividend strip yields data are available at https://www.serhiykozak.com/data, together with the standard errors on the estimated prices.

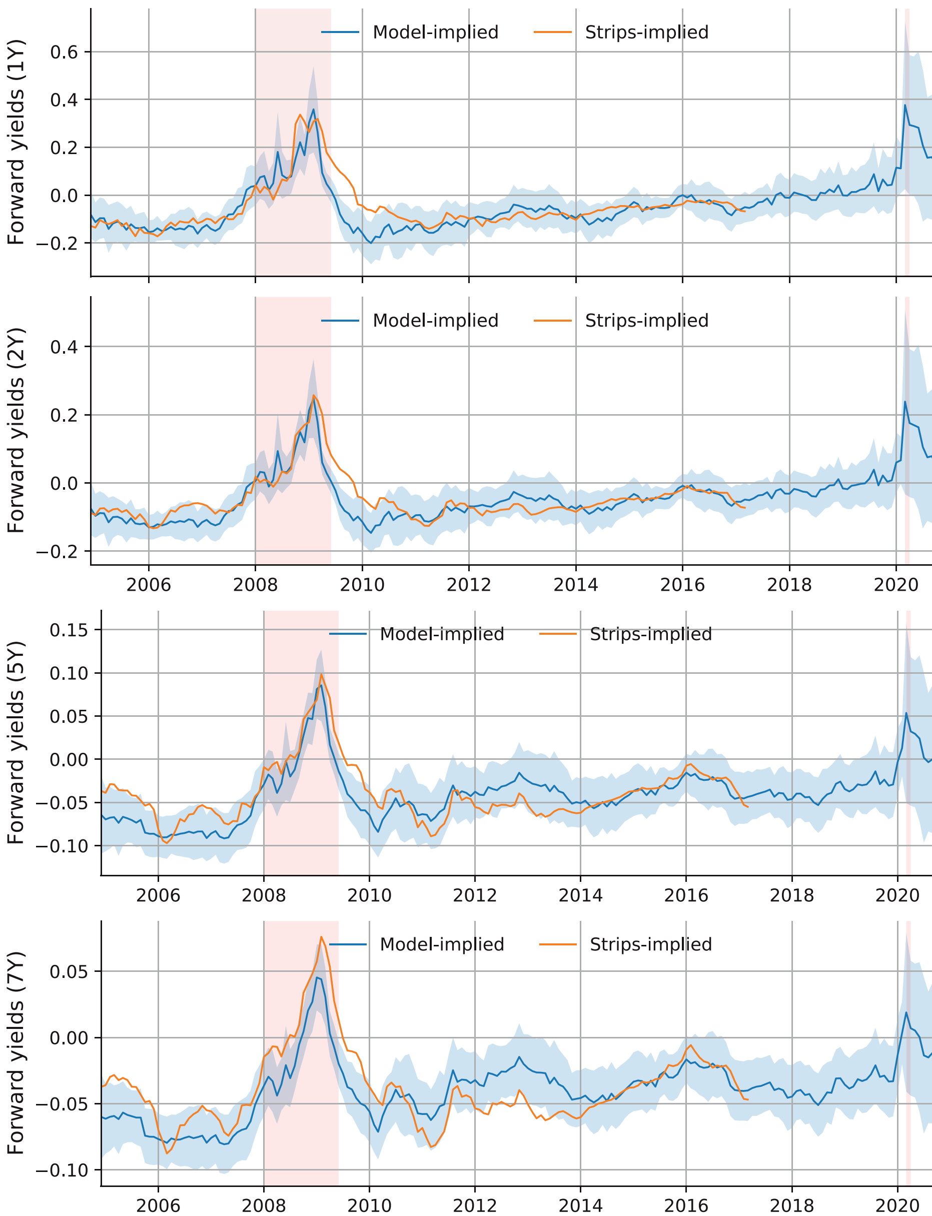


Figure 3. Model-implied forward equity yields versus forward equity yield data. This figure compares our model-implied forward yields for maturities one, two, five, and seven years to their empirical counterparts in Bansal et al. (2021). Shaded areas depict two-standard-deviation bands around point estimates. Model parameters are estimated in the full sample, from February 1973 to December 2020.

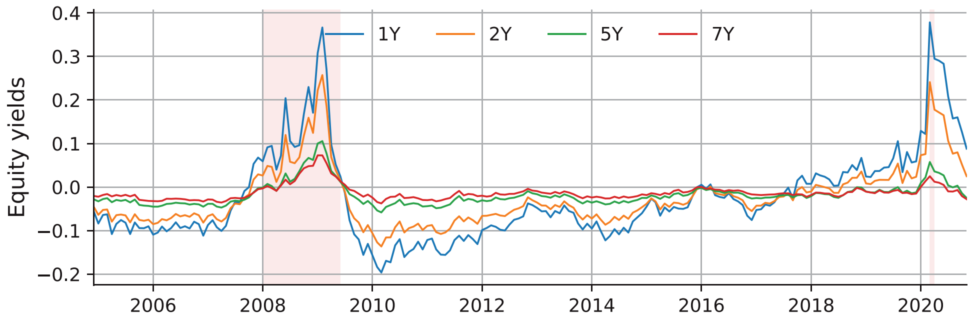


Figure 4. Dynamics of model-implied yields in the Bansal et al. (2021) **sample**. This figure shows equity yields constructed using the trailing 12-month dividend. Model parameters are estimated in the full sample, from February 1973 to December 2020.

et al. (2021). The first four panels report the equity yields and standard errors for maturities one, two, five, and seven years, respectively. The figure shows that our model does a good job overall in matching the traded forward equity yields. Individual model-implied yields are close to the dividend strip yields that we observe in the data, and most of the time they are within standard error bounds. That said, they are not an exact fit, just as we see in the bond literature.

Figure 4 summarizes the dynamics of the model-implied yields in this sample, and closely resembles figure 1 in Bansal et al. (2021). As can be seen from this figure, the shape of the term structure varies over time. It is sometimes relatively flat (e.g., between 2012 and 2019), sometimes upward sloping (as at the beginning of 2010), and sometimes steeply downward sloping (as during the financial crisis)—just like the traded forwards.

To be clear, this evidence should not be interpreted as implying that our synthetic dividend yields are precise estimates of true yields point-by-point, as the confidence bands reflecting parameter estimation uncertainty are fairly wide. Rather, by extending the effective length of the time series of dividend strips, our model delivers more statistical power that can be used to more precisely estimate interesting moments (e.g., the average slope of the term structure), compared to using the shorter time series of observable dividend strips.

Table IV reports statistical patterns of empirical yields based on BMSY data ("Actual"), the model estimated in the same sample ("Fitted"), and the fitting error ("Actual− Fitted") for the level (Panel A) and slope (Panel B) of the term structure. The table shows that the model fits yields well on average (first row). That said, the fit is not always precise in every time period, which can be seen in this table as the variability of pricing errors (second row). These findings carry over to slopes of dividend yields (Panel B).

Given these results, it should not be surprising that we also well match the unconditional moments. For example, we find that, consistent with van Binsbergen and Koijen (2015), the average term structure of forward premia is close to flat for the United States after 2004, with a slight downward

Table IV

Accuracy in Fitted versus Actual Dividend Yields

We present means and standard deviations (in %), as well as annual autocorrelation for observed BMSY forward yields ("Actual"), model-implied yields ("Fitted"), and the fitting errors ("Actual−Fitted") in levels (top panel) and slopes (bottom panel).

	1Y (1)	2Y (2)	5Y (3)	7Y (4)	1Y (5)	2Y (6)	5Y (7)	7Y (8)	1Y (9)	2Y (10)	5Y (11)	7Y (12)
Panel A: Level												
	Actual				Fitted				Actual−Fitted			
Mean	−5.09	−4.55	−3.88	−3.77	−6.80	−5.69	−4.45	−4.14	1.71	1.14	0.58	0.37
S.D.	9.85	6.71	3.36	2.94	9.19	6.44	3.16	2.39	5.58	3.54	1.96	1.89
AC	0.30	0.26	0.24	0.25	0.13	0.10	0.15	0.20	0.13	0.06	0.42	0.60
Panel B: Slope												
	Actual				Fitted				Actual−Fitted			
Mean	–	0.54	1.22	1.32	–	1.11	2.35	2.66	–	−0.57	−1.13	−1.33
S.D.	–	3.56	6.86	7.26	–	2.96	6.54	7.38	–	2.96	4.91	5.13
AC	–	0.31	0.30	0.29	–	0.25	0.21	0.20	–	0.31	0.18	0.14

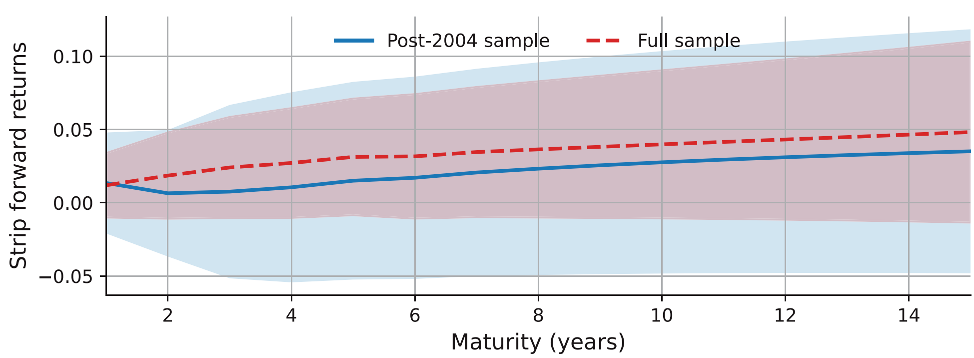


Figure 5. Estimated term structure of forward risk premia. This figure shows mean realized returns on dividend strip forward contracts in the 2005 to 2020 sample (solid blue) and the full sample (dashed red). Shaded areas for each line depict two-standard-deviation error bands around point estimates. Model-implied GMM standard errors using a spectral density matrix with 12 lags for full-sample errors (narrow red bands; model and sampling uncertainty). In the 2005 to 2020 sample, we show empirical HAC robust standard errors associated with the estimate of the mean of realized strip returns (sampling uncertainty; wide blue bands).

slope at the short end and a slight upward slope at the long end. The solid blue line in Figure 5 shows the estimated term structure of forward risk premia for this (post-2004) sample. The figure also shows the average term structure estimated using the full sample, which begins in 1973, along with two-standard-deviation bounds for both lines. We use model-implied GMM standard errors, which incorporate all model uncertainty, for the full-sample

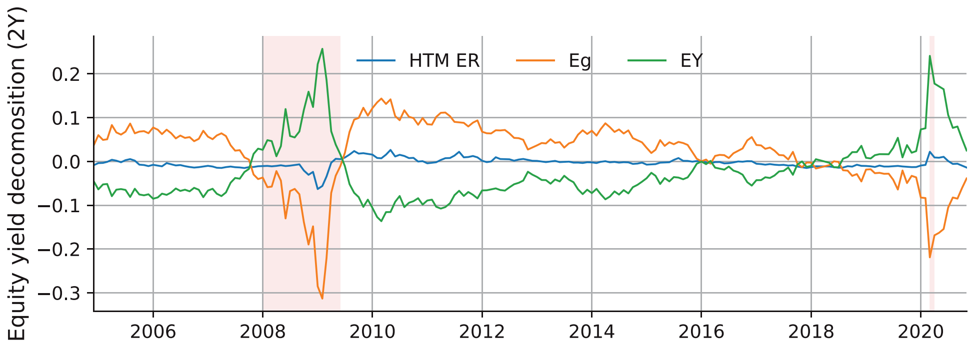


Figure 6. Decomposition of the two-year equity yield in the Bansal et al. (2021) **sample**. The figure decomposes the two-year equity yield on the market into log HTM risk premium and log dividend growth net of the risk-free rate.

estimates, and empirical heteroskedasticity and autocorrelation consistent (HAC) robust standard errors associated with the estimate of the mean of realized strip returns, which only reflect sampling uncertainty, for the BMSY sample estimates. The figure shows that our standard errors are substantially narrower even though they incorporate all model estimation uncertainty. This is due to our ability to estimate means in the longer sample enabled by our use of stock returns data rather than dividend futures data.

We explore the two sources of uncertainty further in Figure IA.5. Panels A and B focus on the full sample and the BMSY sample, respectively. In each figure, the blue shaded area shows GMM two-standard-error bounds, which incorporate model parameter and sampling uncertainty, while the red area reflects sampling uncertainty only. The figure shows that the entirety of the variation here comes from the sampling uncertainty, which would exist *even if prices were perfectly observable*. That is, while model uncertainty due to the fact that prices have to be estimated might be nontrivial, it is dominated and washed out by the sampling uncertainty associated with estimating mean returns or average yields. Going back to Figure 5, this explains why the standard errors over the 2004+ sample (in blue) are close to those implied by the estimates in table 4 in Bansal et al. (2021), whereas the standard errors that we obtain over the full sample (in red) are significantly tighter.

Another dimension along which our data well match the results in van Binsbergen et al. (2013) is the decomposition of equity yields into cash-flow and discount rate components. van Binsbergen et al. (2013) show that, surprisingly, expected cash-flow variation is a major driver of the movement in the equity yields; figure 7 in their paper shows that the sharp increase in the equity yield for the S&P 500 during the financial crisis can be attributed almost entirely to a sharp decline in expected dividend growth. We perform the same decomposition in our model. Figure 6 shows that the results are extremely similar to those in van Binsbergen et al. (2013).

The Internet Appendix shows the similarity between our implied yields and those from Bansal et al. (2021) and van Binsbergen and Koijen (2015) along

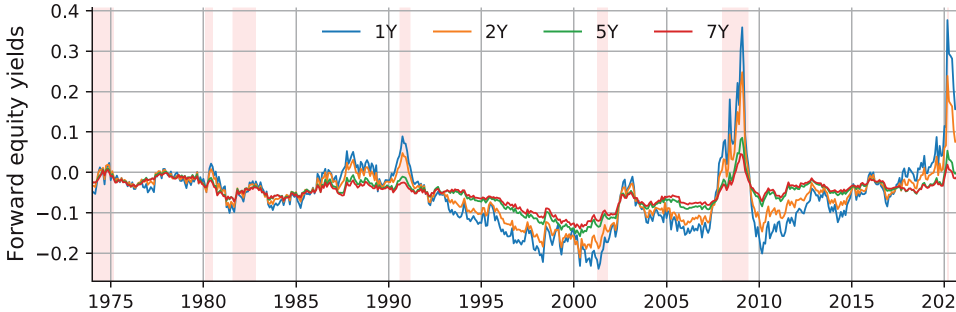


Figure 7. Dynamics of model-implied forward equity yields for the aggregate market for different maturities. The figure plots dynamics of model-implied forward equity yields of maturities one, two, five, and seven years. Equity yields are constructed using the trailing 12-month dividend. The sample is from February 1973 to December 2020.

additional dimensions. Specifically, in Figure IA.1, the average market beta of the strips of different maturity is similar to that reported in van Binsbergen and Koijen (2015, figure 4). Figure IA.2 further shows that the one-year returns of the one- and two-year dividend forwards in the data and in our model match well.

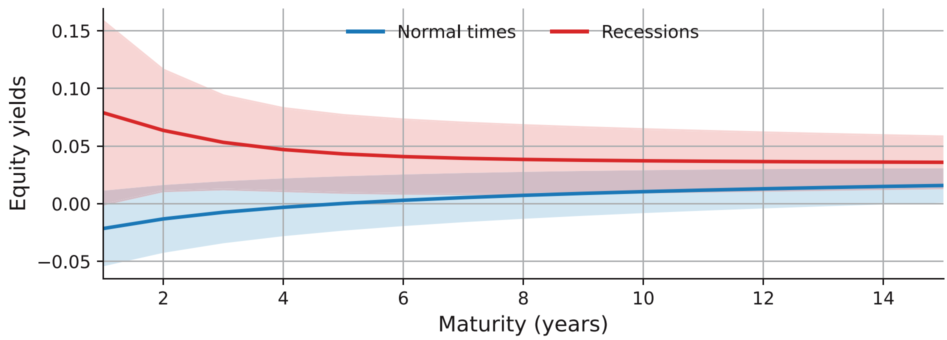
To conclude, when we compute theoretical dividend forwards and strips from our model (estimated using only equity portfolios) and compare them with the prices of actually traded dividend strips, we find that they match well along several dimensions. These results provide external validation of the ability of our model to capture the dynamics of risk and cash flows and investor risk preferences along the term structure.

We next use the model to explore the behavior of the equity term structure over the longer sample (beginning in 1975), after which we turn to the cross section of term structures for different risks.

C. The Time Series of the Equity Term Structure since the 1970s

Figure 7 shows the forward equity yields for the aggregate market for different maturities, as estimated from our model. The results confirm many of the patterns reported above for the post-2004 data: the term structure is generally close to flat, with periods of positive slope during booms and periods of inversion during busts. Interestingly, the term structure appeared significantly more stable up to the early 1990s, remaining essentially flat for almost two decades. The brief recession of the 1990s led to an inversion of the curve, which was followed by a period in which the slope changed sign several times.

The changes in the slope of the term structure are strongly correlated with the macroeconomic cycle. To see this more clearly, Figure 8 shows the term structures conditional on being or not being in an NBER recession. Outside recessions (blue line) the term structure is mildly upward sloping on average. In recessions, in contrast, the term structure is downward sloping.



Maturity:	2	3	4	5	7	10	15
Level diff.	0.08 (2.73)	0.06 (2.79)	0.05 (2.66)	0.04 (2.47)	0.03 (2.17)	0.03 (1.92)	0.02 (1.71)
Slope diff.	0.00 -	−0.02 (−1.71)	−0.03 (−1.80)	−0.03 (−1.87)	−0.04 (−1.97)	−0.05 (−2.08)	−0.06 (−2.22)

Figure 8. Term structures of equity yields conditional on NBER recessions. The figure shows the term structures conditional on being (red) or not being (blue) in an NBER recession by maturity. Shaded areas depict two-standard-deviation error bands around point estimates. The table reports differences of equity yields ("Level diff.") and differences of equity yield slopes (relative to a two-year equity yield; "Slope diff.") in recession minus expansion, and their associated t-statistics based on GMM standard errors with a spectral density matrix with 12 lags.

The table under the plot shows the differences between these two lines and their associated t-statistics ("Level diff."). It also shows the difference in the slopes (difference between each maturity and maturity 2, "Slope diff"). Equity yields are statistically significantly higher in recessions than expansions for dividend claims of up to eight years. Equity yield slopes are significantly different at the 5% level for maturities above seven years (i.e., 7-2 slopes) and at the 10% level for all maturities. Standard errors in this plot reflect both parameter and sampling uncertainty.

One of the advantages of our method is the ability to study the behavior of the term structure over a much longer sample that includes several recessions (contrary to the post-2004 sample, in which the only recession is the financial crisis). We can therefore ensure that the patterns we find are not due entirely to the specialness of the Great Recession. For example, we see inversion in the yield curve at several other times in history, as confirmed in Figure 8 which averages across *all* recessions.

It is worth recalling that the slope of the term structure of equity yields has a direct interpretation in terms of the term structure of discount rates for different horizons. As pointed out by Dew-Becker et al. (2017) and Backus, Boyarchenko, and Chernov (2018)

$$\frac{1}{n}\mathbb{E}\log\left(R_{t,t+n}\right) - \mathbb{E}\log\left(R_{t,t+1}\right) = \mathbb{E}\,e_{t,n} - \mathbb{E}\,e_{t,1}.$$

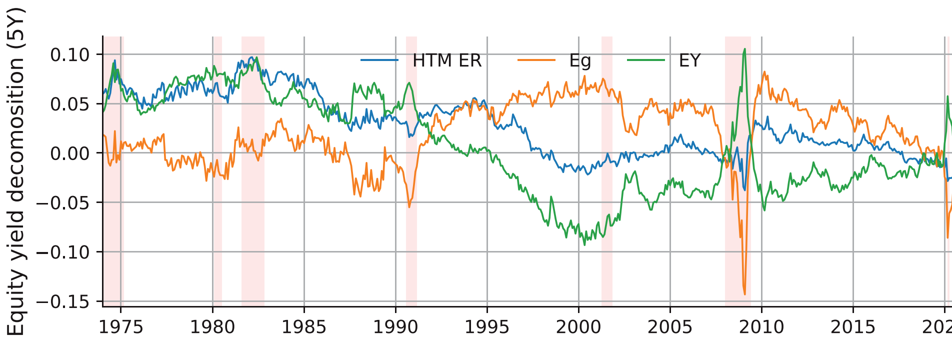


Figure 9. Decomposition of the five-year equity yield. The figure decomposes the five-year equity yield on the market into log HTM risk premium and expected log dividend growth net of the risk-free rate in full sample, February 1973 to December 2020.

This result highlights how the main facts about the conditional and unconditional term structure of equity yields from van Binsbergen, Brandt, and Koijen (2012) extend to our full sample starting in the 1970s.

Finally, in Figure 9, we show the decomposition of the five-year equity yield into expected annual dividend growth over five years and expected hold-to-maturity (HTM) excess returns. This figure confirms that a large fraction of the variation in equity yields is driven by expected dividend growth as opposed to discount rate variation.

Extending the equity term-structure data to the 1970s yields several new insights. First, as mentioned above, the term structure appeared much less volatile before the 1990s. Second, it appeared to have been effectively flat for decades. Third, there had been times before the financial crisis in which markets strongly anticipated negative dividend growth; for example, during the recession of the early 1980s and early 1990s, these movements appear to have been reflected in the prices of equities and (implied) equity strips.

D. Cross Section of Term Structures of Different Risks

The most interesting advantage of our model is that it can generate term structures for different types of risks, as captured by different portfolios. For example, it can produce a term structure of discount rates for value firms and a term structure for growth firms, one for small firms and one for large firms, and so on. These term structures can be used in turn to test the implications of structural models that have cross-sectional implications.

As an example, different structural models have been proposed to explain the value premium. But these models mainly confront the *average risk premium* of value versus growth portfolios. These models also have implications for the *term structure* of discount rates on value and growth stocks, which will be an especially important moment for models in which the dynamics of shocks play a role in determining risk premia.

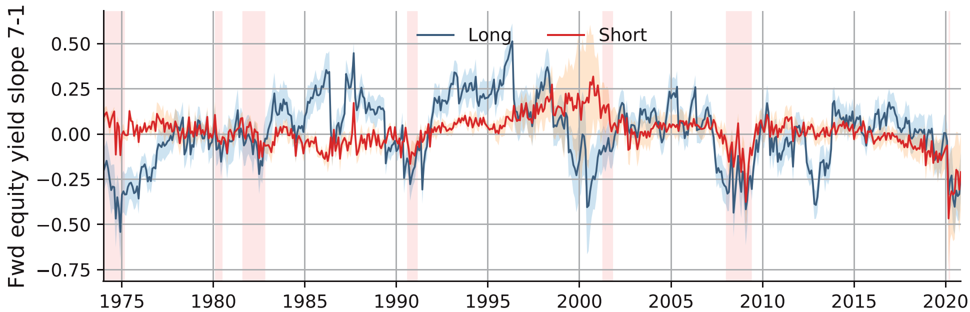


Figure 10. Slope 7-1 of forward equity yields for small and large stocks. The figure shows time series of the estimated slope of forward equity yields (7y − 1y) from our model for diversified portfolios of small (long) and large (short) stocks in the full sample, February 1973 to December 2020.

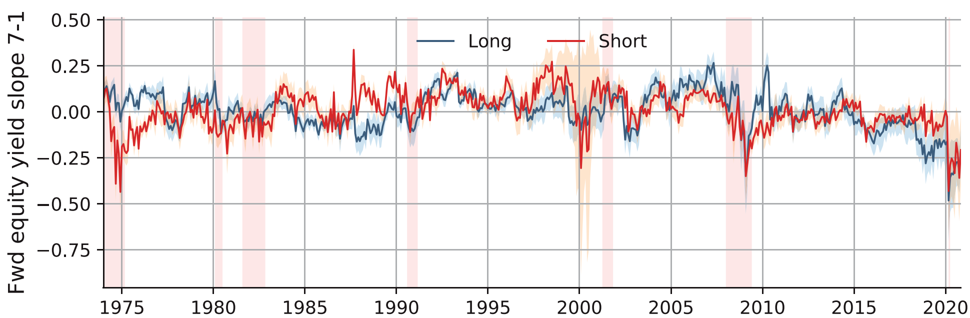


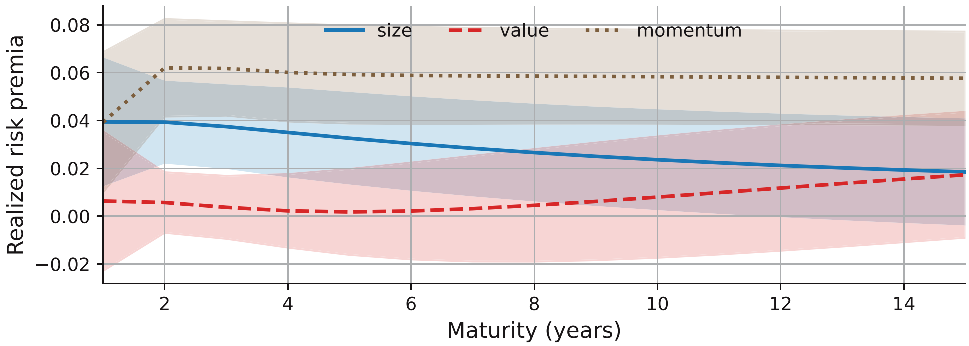
Figure 11. Slope 7-1 of forward equity yields for value and growth stocks. The figure shows time-series of estimated slope of forward equity yields (7y − 1y) from our model for diversified portfolios of value (long) and growth (short) stocks in the full sample, February 1973 to December 2020.

Figures 10 and 11 show the time series of the estimated slope of forward equity yields (7y − 1y) from our model, for small stocks (long) and large stocks (short) and for value stocks (long) and growth stocks (short), respectively. Many interesting patterns can be seen.

Large firms' and small firms' term structures moved in similar ways in several periods (e.g., during the financial crisis), but their equity yields went in opposite directions during the tech boom and bust. Moreover, whereas the equity term structure inverted for small firms during the tech boom, it did not do so for large firms. During the financial crisis and the Covid episode, however, both curves inverted. In contrast, no such divergence in the behavior of the term structure can be seen for value and growth stocks in that period—the largest difference in that case occurred in the recovery from the financial crisis: after 2008, the slope of the term structure increased significantly for value but not for growth stocks.

Figure 12 depicts the term structure of risk premia for a number of long-short portfolios. Panel A shows the difference in risk premia between small

Panel A: SMB, HML, and MOM risk premia with standard error bounds



Panel B: Select long-short portfolios

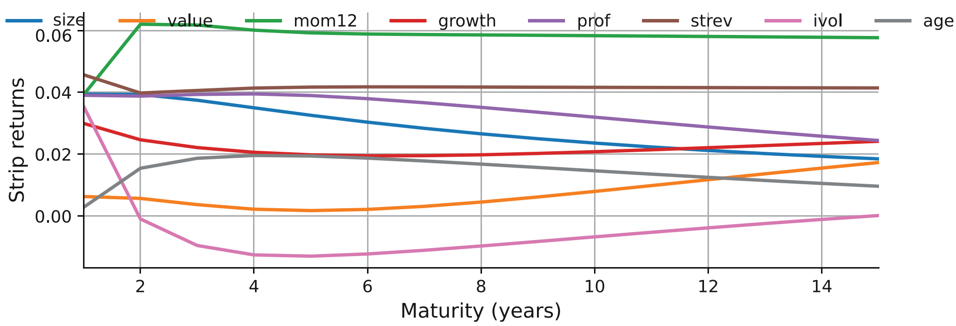


Figure 12. Term structure of risk premia for long-short portfolios. Panel A shows the difference of realized risk premia between small and large stocks, value and growth stocks, and high versus low momentum stocks. Shaded areas depict one-standard-deviation error bands around point estimates. Panel B shows the difference of realized risk premia across select anomaly portfolios: SMB, HML, MOM, INV, PROF, STREV, IVOL, and Age.

and large stocks, value and growth stocks, and high versus low momentum stocks, for different horizons. Shaded areas depict one-standard-deviation error bands around point estimates (for readability). Note that these long-short portfolios have positive risk premia in this sample. Yet the term structures of these risk premia are quite heterogeneous. For example, the SMB portfolio appears to exhibit a downward-sloping term structure, while the HML portfolio has an upward-sloping term structure. Theoretical models that aim to explain the value and size spreads can make use of these estimates to help refine and calibrate the economic mechanisms (taking into account, of course, the substantial estimation uncertainty associated with these unconditional moments).

It is important to note that our term structures correspond to term structures of actively managed portfolios that are rebalanced at an annual or monthly frequency. This choice of basis assets is motivated by two observations. First, it directly parallels the construction of traded claims on dividend strips of aggregate indices, such as the S&P 500 in the United States. Indeed, dividend strips on the S&P 500 are claims to the dividends that future constituents of the S&P 500 index will pay in the future, rather than dividends on today's constituents of the index. Second, Keloharju, Linnainmaa, and Nyberg (2019) show that differences in expected returns across firms decay within five years. Their findings suggest that term structures of any portfolio sorted on today's characteristics should approximately converge to the term structure of the aggregate market at that horizon. We avoid such convergence in our results by focusing on term structures of actively managed portfolios. To the extent that firm characteristics capture economically relevant risk exposures (e.g., if book-to-market captures distress risk), it is meaningful to try to understand the pricing of potential risks that affect stocks with those characteristics at various horizons, for example, the pricing of shocks that will affect distressed stocks in the future.[14]

Many other portfolios exhibit interesting term structures. Panel B of the figure shows the difference of realized risk premia across a few additional anomaly portfolios: SMB, HML, momentum, investment, profitability, short-term reversals, idiosyncratic volatility, and age. Some have strikingly different shapes. For example, the term structure of risk premia for the portfolio sorted on idiosyncratic volatility is sharply downward sloping up to around five years, then it becomes mildly upward sloping. The term structure of risk premia for the age portfolio is hump-shaped.

We conclude that our model generates rich predictions about the differential behavior of term structures across portfolios, both in the time series and on average. These results present new moments that structural asset pricing models can try to match, in addition to the term structure of the aggregate S&P 500 dividend that has previously been studied using traded dividend forwards.

E. Implied Dividend Growth

In our model, deflated dividend growth (i.e., net of the risk-free rate) is pinned down exactly by equation (15). Since we model both returns and D/P ratios as inputs, dividends are just a deterministic function of these observable variables and thus perfectly match the data, month by month. Note that we do not use log-linearizations in this expression; dividends are therefore given by an exact (nonlinear) function of returns and equity D/P ratios. The dynamics of dividends are thus pinned down entirely by the dynamics of these variables.

We now come back to the discussion of dividend growth predictability in our model. We observe relatively strong model-implied dividend predictability across the board, for many stock portfolios as well as for the aggregate market (Panel A of Figure 13). Dividend growth predictability is strongest, however, for medium- to high-frequency strategies, such as momentum, in which case dividend growth can be predicted with R^2 exceeding 80%. Panel B of Figure 13

[14] That said, if one wanted to construct term structures of dividends of a fixed portfolio of stocks, they could be easily approximated by overlaying the term structures of the managed portfolios we report (at short horizons) and that of the aggregate market (at longer horizons).

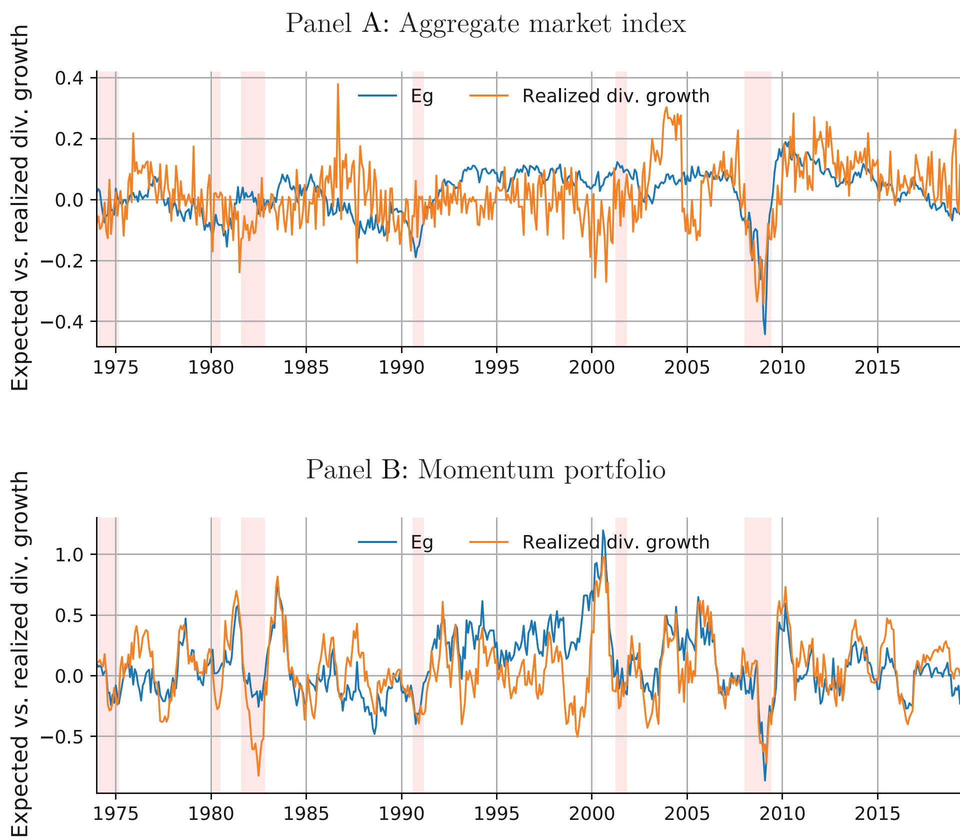


Figure 13. Expected and realized dividend growth for the market index and momentum stocks. Panel A plots time series of the model-implied expected (blue) and realized (orange) dividend growth for the aggregate market index, net of the risk-free rate. Panel B plots results for the momentum stock portfolio.

shows time series of the model-implied expected (blue) and realized (orange) deflated dividend growth for the momentum stock portfolio. Similar to Figure 6, we find that most of the variation in equity yields on high-momentum stock is driven by variation in expected dividend growth. This, in turn, implies that dividend growth on high-momentum stocks is highly predictable by their equity yields, consistent with the findings of van Binsbergen et al. (2013).

Dividends can thus be used to provide additional validation of the equity yields constructed using our model. To motivate this idea further, first note that log gross dividend growth at an n-year horizon, $\Delta d_{t,t+n}$ is given by

$$\Delta d_{t,t+n} = r_{t,t+n} - n e_{t,n}, \tag{35}$$

where $r_{t,t+n} = \log\left(\frac{D_{t+n}}{P_t^{(n)}}\right)$ is the log of the HTM return on an n-year dividend strip from t to $t+n$, and $e_{t,n}$ is the equity yield on an n-year dividend strip at

Table V

Dividend Growth Predictability

Panel A reports average R^2 in the cross section of 102 portfolios from regressions of n-year dividend growth on: (i) the equity's own dividend-to-price ratio (D/P), (ii) the model-implied dividend yield on an n-year dividend strip, (iii) the BMSY dividend yield on an n-year dividend strip (Panel B only), and (iv) model-implied expected dividend growth net of the risk-free rate at each horizon, $\mathbb{E}[\Delta d - r_f]$. The last row in each panel shows p-values of a one-sided t-test of the equality of the two means of R^2 distributions for D/P and EY (model) regressions. Panel A reports results for the full sample; Panel B focuses on the BMSY sample.

Horizon	1-Year	2-Year	3-Year	4-Year	5-Year	6-Year	7-Year
Panel A: Full Sample							
D/P	3.7	3.5	3.7	4.4	5.1	6.0	7.0
EY (model)	17.8	12.9	9.0	6.9	6.1	6.5	7.3
$\mathbb{E}[\Delta d - r_f]$	19.9	16.6	14.6	13.1	13.4	13.4	16.0
$p(R^2_{\text{D/P}} = R^2_{\text{EY}})$	0.00	0.00	0.00	0.00	0.10	0.28	0.35
Panel B: BMSY Sample							
D/P	36.0	25.6	16.6	9.7	10.3	11.5	12.9
EY (model)	41.5	36.4	25.1	13.3	12.9	13.5	14.8
EY (BMSY)	20.7	18.5	–	–	7.9	–	9.1
$\mathbb{E}[\Delta d - r_f]$	42.6	44.9	36.6	21.5	21.5	23.5	25.4
$p(R^2_{\text{D/P}} = R^2_{\text{EY}})$	0.01	0.00	0.00	0.01	0.06	0.12	0.15

time t as defined in equation (32) (see Appendix A for the derivation). This is an exact formula that does not require any approximations.

Second, notice that this equation holds in expectations, that is,

$$\mathbb{E}_t\left[\Delta d_{t,t+n}\right] = \mathbb{E}_t\left[r_{t,t+n}\right] - ne_{t,n}. \tag{36}$$

Third, recall from Figure 6 that most of the variation in short-term equity yields comes from expected dividend growth rather than discount rates.

Equity yields, therefore, are natural predictors of dividend growth rates at all horizons, especially so at short horizons. We use this insight to investigate dividend growth time-series predictability using model-implied equity yields in the full sample. As a benchmark, we use dividend growth predictability based on a stock's own D/P ratio as is standard in the literature.[15]

In Table V, we report means of R^2 computed in the cross section of 102 portfolios from regressions of n-year dividend growth on several variables (all univariate regressions). Panel A presents the results in the full sample, Panel B in the post-2004 sample. In the first row, we use as the predictor the equity's own D/P ratio. We find limited predictability of dividends using this variable, consistent with the literature.

[15] We also considered a specification with an additional predictor—lagged dividend growth—with similar results.

In the second row, motivated by the analysis above, we use the model-implied dividend yield on an n-year dividend strip as a predictor. We observe substantially higher predictability—a statistically significant difference in predictability up to four years in maturity, as can be seen from the p-values of a one-sided t-test of the equality of the two means of R^2 distributions for D/P and EY (model) regressions (the last row in the panel). In the third row, we use the model-implied expected dividend growth net of the risk-free rate at each horizon, $\mathbb{E}[\Delta d - r_f]$, as a predictor. This is the same object depicted in blue in Figure 13. We see that this variable predicts realized dividends even better than the dividend yield.[16]

Lastly, we compare the predictability between our model-implied equity yields and their empirical counterparts based on dividend strip data from BMSY. To do so, we limit our sample to theirs. We report all the results in Panel B. We now include an additional row "EY (BMSY)," which includes R^2 averages of regressions that use BMSY equity yields as a predictor. Perhaps somewhat surprisingly, in spite of the visual similarity of BMSY and our equity yields depicted in Figure 3, we find that our model-implied EY performs significantly better than the observed BMSY yields in predicting cumulated realized dividend growth in the data. Similarly to the full-sample case, our model-implied expected dividend growth (fourth row of Panel B) predicts realized dividend growth even more robustly than model-implied equity yields, even though we do not directly target dividend predictability in our estimation.

We conclude that dividends up to four years are significantly more predictable using equity yields extracted from the model than when using a stock's D/P. This fact serves as additional indirect validation of the model in that it is able to extract useful information embedded in equity yields that helps predict dividends on a wide cross section of test asset portfolios. Figure IA.8 provides distributions of R^2 for both predictors across portfolios.

F. Applications

We briefly discuss in this section three potential applications of our findings.

F.1. New Moments for Model Evaluation

As mentioned in the introduction, we first view our methodology as a way to extract novel moments of the term structure of discount rates (its conditional and unconditional slope) for a variety of portfolios. Rather than work directly with the rich dynamics of the economy and the large cross section of available equity portfolios, researchers can use the generated term structures, to calibrate and evaluate models and compute valuations of investments. We also

[16] We use expected deflated dividend growth as opposed to expected dividend growth because the former is an output of our model; since we do not model the process of the risk-free rate, our model does not provide estimates for expected nondeflated dividend growth at horizons beyond one.

produce standard errors on our estimates, which can be used together with the point estimates to properly account for the estimation uncertainty.

One natural application of our term structures is the evaluation and testing of asset pricing models. Many asset pricing models, for example, the long-run risk, habit formation, and rare disaster models, have strong predictions about the moments we estimate in the data. For example, van Binsbergen, Brandt, and Koijen (2012) and van Binsbergen and Koijen (2015) show that the slope of the aggregate dividend term structure observed in the post-2004 sample is on average too low to be explained by the long-run risk and habit formation models, but their conclusions rely on a small sample that includes the financial crisis.

Using our approach, we can evaluate these models against the unconditional slope of the forward yields curve estimated since 1975. Having 45 years of data, and being able to look at maturities beyond seven years, gives our approach more power for these tests and brings data from new economic cycles. In Section I of the Internet Appendix, we illustrate this exercise by simulating different models and evaluating them against our estimated term-structure moments. Specifically, we look at the slope of the term structure of equity yields and forward equity yields, the risk premia associated with strips and forwards of different maturity, and the cyclicality of the slope of the term structure of equity yields. We study five models—the long-run risk models of Bansal and Yaron (2004) (BY) and Bansal, Kiku, and Yaron (2012) (BKY), the habit-formation model of Campbell and Cochrane (1999) (habit), the model of Lettau and Wachter (2007) (LW), and the rare-disaster model of Gabaix (2012) (disaster)—and show that each model can match some, but not all, of the rich set of new moments that we bring to the table. This exercise represents an illustration of the type of additional model evaluations that can be conducted using our estimates.

In addition to extending the results on aggregate dividend strips, our paper can also be used to directly test the implications of models about the term structure of discount rates of specific portfolios. For example, Hansen, Heaton, and Li (2008) discuss the implied term structure of value and growth stocks in their model. Production-based models (e.g., Belo (2010) and Kogan and Papanikolaou (2013, 2014)) specify the dynamics of the SDF and the risks of different types of firms and therefore have direct implications about the term structure of discount rates for value firms, growth firms, high- and low-investment firms, and so on, for which we provide estimates. We leave evaluation of the cross-sectional term-structure implications of these models to future research.

F.2. Valuation

Our model can also be used to value investments whose cash flows pay off at various horizons. Once the riskiness of any investment with a specified maturity is determined (i.e., the relation between investment returns and our factors F_t and their innovations), either through economic theory or empirically, our model provides the appropriate term structure of discount rates for that investment. For example, it can be used to value private equity investments as in Gupta and Van Nieuwerburgh (2021), or to study long-term discounting for climate change mitigation investments as in Giglio et al. (2015), who used the term structure of long-run discount rates on housing only.

F.3. Hedging Portfolio Construction

Our work identifies four priced factors and associated dynamics and pricing. The model can therefore be used by investors that would like to get exposure to specific types of risks. The main advantage of our model relative to standard models that are specified only in terms of shocks is types of risks can be more easily identified. For example, the model can be used to build a portfolio that isolates cash-flow or discount rate risk at any specified horizon.

In addition, as we explored in Section II of the Internet Appendix, the model can be used to explore the links between characteristics sorts (e.g., duration) and risk exposures. As an example, in that section we show that sorting portfolios by the level of the term structure of discount rates yields a portfolio that is principally exposed to the first PC (the second factor after the market), which commands a Sharpe ratio of 0.71 in our model.

G. Robustness

G.1. Counterfactual Analysis: Alternative Models of the SDF

In this section, we investigate whether simpler models (in which the factors are not based on the PCs of a large set of anomalies) could also produce estimates for the equity strip prices that match the data from the traded contracts. We consider two benchmarks.

Our first benchmark is the conditional CAPM model. That is, we assume that there is a single priced source of risk—the aggregate market—and its risk price time variation is driven by the aggregate D/P ratio. We replicate our main analysis above for this choice of state vector, which now contains only two variables.

Our second benchmark is the five-factor model of Fama and French (2016) supplemented by the momentum factor from Carhart (1997). Like our main specification, we consider the setup with four returns (the market and three PCs of the five cross-sectional factors), that is, our state vector contains eight variables.[17] We replicate our analysis for this choice of state vector.

We display the results in Figure IA.6. Panel A shows results for the CAPM benchmark. Panel B shows results for the Fama and French five-factor model, supplemented with the momentum factor. The figure plots the time series of yields in the BMSY sample. It is evident from the plots that the dynamics of

[17] This specification performs better than a specification based on the six unrotated factor returns. We therefore report the former specification.

implied dividend strip yields for both benchmarks are rudimentary: all yields move almost synchronously with no yield curve inversions occurring in this sample for the CAPM model. The yields implied by the Fama and French five-factor model look very different from dividend strip yields observed in the data.

We conclude that our rich specification is key to capture the dynamics of equity strips in the data. This is because our specification successfully explains both the cross section and the time series of U.S. equity returns. In contrast, SDFs implied by stylized models, such as the CAPM or the Fama and French five-factor model, do not feature rich enough dynamics to match the empirically observed patterns in equity yields.

G.2. Additional Out-of-Sample Analysis

As we discuss in the previous sections, the main way the model is evaluated is in its ability to match the time series of the traded dividend strip prices. In this sense, our main analysis is already out of sample because traded dividend strip prices are not used in the estimation. In this section, we present two additional out-of-sample analyses.

First, we split our sample into two parts: 1973 to 2005, and 2005 to the present. We chose the split point in 2004 because this is when the BMSY data become available, which provides a natural out-of-sample test for our purposes. We estimate the model using data in the first part of the sample only. We then fix the parameters and apply them to estimate equity yields in the 2005 to present sample. Importantly, the PCs and their associated eigenvectors are also estimated only in the pre-2005 sample.

We report the root mean squared errors (RMSEs) of model-implied relative to empirical equity yields in Panel A of Table VI (second line; split sample). We depict the dynamics of the estimated equity yields in the 2005 to present sample in Figure IA.9. Overall, compared to the full-sample RMSEs in the first line of Panel A and to Figure 4, the out-of-sample yields are very similar to their in-sample counterparts. Parameter estimates of the state dynamics are also similar (not reported). This evidence gives us confidence that our results are not driven by merely overfitting the data in sample.

As a second test, we also implement a full rolling estimation. The third line of Panel A shows RMSE of the model which starts with the pre-2005 training sample and updates it annually on a rolling basis. Figure IA.10 depicts the model-implied yields based on this rolling estimation procedure. Qualitatively, the patterns in yields are similar to those estimated both in full sample and in pre-2005 sample.

G.3. Alternative portfolio sorts

While our main specification focuses on a cross section of 51 portfolios which were used in prior work (Kozak, Nagel, and Santosh (2020), Kozak (2024)), we consider here two alternative sets of portfolios to assess the robustness of our results.

Table VI
Robustness

The table reports RMSEs for forward equity yields of maturities of one, two, five, and seven years and the average RMSE (last column) across several alternative specifications. Panel A reports RMSEs of the main specification (51 anomalies) in and out of sample. Panel B explores alternative portfolio sets and reports RMSE in full sample. Panel C reports out-of-sample RMSEs for models with varying number of PC factors. Panel D adds bond factors. Panel E adds volatility factors.

	1-Year (1)	2-Year (2)	5-Year (3)	7-Year (4)	Average (5)
	Panel A: Main Specification: 51 Anomalies				
Full sample	0.058	0.037	0.020	0.019	0.034
Split sample	0.057	0.037	0.022	0.020	0.034
Rolling estimation	0.058	0.038	0.021	0.020	0.034
	Panel B: Alternative Portfolio Sets				
GHZ	0.075	0.046	0.027	0.024	0.043
WRDS financial ratios	0.070	0.042	0.019	0.017	0.037
	Panel C: Varying Number of PCs Out of Sample				
MKT only	0.094	0.065	0.033	0.029	0.055
MKT + 1 PC	0.084	0.056	0.027	0.024	0.048
MKT + 2 PCs	0.064	0.045	0.027	0.024	0.040
MKT + 4 PCs	0.088	0.058	0.025	0.022	0.048
MKT + 5 PCs	0.088	0.061	0.033	0.028	0.052
MKT + 6 PCs	0.087	0.068	0.045	0.039	0.06
	Panel D: Bonds				
MKT+2PCs + 1 bond PC	0.073	0.053	0.032	0.028	0.047
MKT+3PCs + 1 bond PC	0.061	0.043	0.024	0.022	0.037
MKT+2PCs + 2 bond PCs	0.070	0.051	0.043	0.044	0.052
MKT+3PCs + 2 bond PCs	0.083	0.058	0.033	0.028	0.051
	Panel E: Volatility				
MKT+2PCs + vol (mkt)	0.073	0.053	0.033	0.029	0.047
MKT+3PCs + vol (mkt)	0.065	0.044	0.025	0.022	0.039
MKT+2PCs + vol (PC)	0.067	0.049	0.031	0.028	0.044
MKT+3PCs + vol (PC)	0.092	0.073	0.046	0.037	0.062

The first set uses return predictive signals from Green, Hand, and Zhang (2014) (GHZ) to form sorted portfolios, which are widely used in the literature.[18] This set contains 102 anomaly characteristics (therefore 204 between low- and high-tercile portfolios).

[18] Green, Hand, and Zhang (2014) provide SAS code to generate their characteristics, which we have used to construct our own portfolio sorts.

The second set uses WRDS financial ratios constructed by Wharton Research Data Services. This set was also used in Kozak, Nagel, and Santosh (2020). The set contains over 70 different financial ratios, categorized based on economic intuition into the following seven groups: (i) Capitalization: measures the debt component of a firm's total capital structure (e.g., Capitalization Ratio, Total Debt-to-Invested Capital Ratio), (ii) Efficiency: captures the effectiveness of firm's usage of assets and liability (e.g., Asset Turnover, Inventory Turnover), (iii) Financial Soundness/Solvency: captures the firm's ability to meet long-term obligations (e.g., Total Debt to Equity Ratio, Interest Coverage Ratio), (iv) Liquidity: measures a firm's ability to meet its short-term obligations (e.g., Current Ratio, Quick Ratio), (v) Profitability: measures the ability of a firm to generate profit (e.g., Return on Asset (ROA), Gross Profit Margin), (vi) Valuation: estimates the attractiveness of a firm's stock (overpriced or underpriced; e.g., P/E ratio, Shiller's CAPE ratio), and (vii) Others: Miscellaneous ratios (e.g., R&D-to-Sales, Labor Expenses-to-Sales).

In Panel B of Table VI, we report full-sample RMSEs for traded dividend strip yields of the model estimated using these alternative data sets. The table shows that using these alternative data sets produces qualitatively similar results as the main specification (first line of Panel A), in terms of both general fit metrics and implications for dividend yields. Figures IA.3 and IA.4 compare yields at each maturity to BMSY yields. Overall, the results indicate that our setup is not particularly sensitive to the choice of test assets and that alternative choices deliver qualitatively comparable results.

G.4. Number of PC Factors

We next explore the robustness of our results to using a different number of PCs in the state space. There is an inherent trade-off between choosing a number that is too low versus too high. Figure IA.7 shows that in cases in which the number of factors is one (market) or two (market + PC1), implied strip dividend yields exhibit extremely simplistic dynamics that are visibly incompatible with the prices of dividend futures that we observe in the data. A model that has two cross-sectional PCs (and the market) performs only slightly worse than our primary specification. Increasing the number of PCs to four or five still performs reasonably well, but the out-of-sample (split sample) performance starts to deteriorate as can be seen in Panel C of Table VI. The reason is that the number of parameters we need to estimate for the physical dynamics (VAR) grows as a square of the number of PCs. Predicting returns in the time series is notoriously difficult, so the increased number of free parameters proves challenging for the estimation.

A model that includes the market and three PCs—our primary specification (second line in Panel A)—strikes a good balance between flexibility (bias) and overfitting (variance).

G.5. Additional Factors and Predictors of Risk Prices

The goal of this paper is to produce a parsimonious specification that can generate empirically plausible patterns of dividend strip yields. As shown above, we find that our specification that includes only valuation (D/P) ratios as predictors is sufficiently rich to achieve that objective. D/P ratio is a somewhat special predictor because it is a part of the return, it must be included in the state space to make the model solvable. Conveniently, Haddad, Kozak, and Santosh (2020) show that valuation ratios alone are powerful predictors of SDF risk prices. However, it is interesting to explore the potential of additional predictors to improve the fit of the model. Importantly, adding predictors does not necessarily improve model performance out of sample (e.g., to fit the dividend strip prices) because it adds parameters to be estimated, which can lead to overfit. In this section, we explore this question empirically.

Specifically, we explore modifications to our specification in which we add additional variables: (i) rolling one-year volatility (of the market or the first PC of factor volatilities), and (ii) the risk-free bond (the first two PCs of bond excess returns across maturities). For each of the new variables, we supplement the state space by a corresponding factor return and a predictor. For volatility, we use a volatility-mimicking portfolio return (as a return) and volatility itself (as a predictor).[19] For bond PCs, we use a bond return PC (as a return) and its corresponding forward spread (as a predictor). We report the results in Panels D and E of Table VI, respectively. Overall, adding these variables leads to a deterioration of fit, which suggests that the benefits of more restrictive parsimonious specifications (without bond and volatility) generally outweigh the costs of the bias due to the exclusion of these variables.

G.6. Nonparametric Bootstrap

We conduct a nonparametric bootstrap exercise to compute standard errors that we then compare to the GMM standard errors used in our main specification. We draw 48 years with replacement, each including 12 monthly observations, and include the 12 preceding monthly observations to maintain the time-series structure in estimating the dynamics in our state vector. We then use this bootstrapped sample to estimate the main parameters of the model ($c, \rho, b_0, b_1, \beta_2$) in our GMM and compute the rest of quantities based on these estimates. Next, we use the actual time series to compute and construct equity yields, returns, and other variables of interest. We repeat this entire process 500 times. Finally, we compute standard deviations of equity yields across all 500 bootstrapped estimates of forward equity yields and produce an analog of Figure 3 in which standard errors are now based on this bootstrap procedure.

[19] Volatility is partially spanned by the variables included in the state vector ($R^2 = 30\%$) and is negatively correlated with the slope of equity yields (-0.3).

Figure IA.11 shows the results. Standard errors are remarkably close to our GMM-based standard errors with Hansen-Hodrick correction. To assess the relative magnitudes of standard errors across the two methods quantitatively, we also report time-series averages of standard errors (in %) for forward equity yields of maturities of one to seven years in Table IA.IV. In Panel A we focus on the post-2004 (BMSY) sample used in Figure IA.11. As in the figure, we can see that bootstrap standard errors are very close to the model-implied GMM standard errors. Our standard errors tend to be somewhat smaller than the bootstrapped ones (up to 30% smaller for the seven-year yield) only at high maturities. Panel B shows the same results for the entire sample. The conclusion remains unchanged.

IV. Conclusion

Our model effectively processes a rich information set (the time-series and cross-sectional behavior of 102 portfolios spanning a wide range of equity risks) to produce "stylized facts"—the time-series and cross-sectional behavior of implied dividend term structures—that summarize a dimension of the data that is particularly informative about our economic models. Similar in spirit to the way in which the introduction of VARs by Sims (1980) provided new moments against which to evaluate structural macro models (the impulse-response functions that were generated by the VARs), the objective of this paper is to produce realistic term structures of discount rates for different portfolios that closely resemble the actual dividend claims that we observe in the data and that can be used by asset pricing models as a moment for evaluation and guidance.

Our approach uses only a cross section of equity portfolios to produce new (implied) term-structure data that expand the existing (observed) data along each of those dimensions. The term structures that we generate cover a large number of cross-sectional portfolios, in addition to the S&P 500: value, size, profitability, momentum, etc., for the 102 portfolios. Moreover, they have a long time series, starting in 1975, and therefore cover several recessions and booms. They also have all possible maturities, including the very short and the very long ends of the term structure.

The main building block of our specification is the carefully crafted state-space vector, which includes four excess returns (on the market and three PCs of anomaly portfolios) and four valuation ratios corresponding to these portfolios. This choice is motivated by recent empirical evidence in Kozak, Nagel, and Santosh (2020) and Haddad, Kozak, and Santosh (2020) that (i) an SDF constructed from dominant PCs of a large cross section of characteristic-sorted portfolios explains the cross section of expected returns well, and (ii) SDF risk prices corresponding to these factors are highly predictable in the time series by their valuation ratios, and this variation in risk prices is essential to adequately capture dynamic properties of the SDF. It is this choice of the state-space vector that represents the core of our paper: it allows us to produce term structures of discount rates that well match the observed term structures, and increases confidence in extending them over time, maturities, and portfolios.

We derive several novel empirical results. First, we extend the study of the term structure of aggregate dividend claims (on the S&P 500, as in van Binsbergen and Koijen (2015)) over time (back to the 1970s) and across maturities. In the sample starting in 2004 used by van Binsbergen and Koijen (2015), we match the time series of dividend forward prices very closely, and mechanically we also match the term structure of discount rates. The term structure of discount rates appears to be slightly downward sloping in this sample, in contrast to the predictions of many models like the long-run risk model that instead predict that it should be steeply upward sloping. Extending the sample to the 1970s allows us to include several additional recessions in our sample, and the Great Recession carries less overall weight in the sample. It is interesting to see that many of the results of the post-2004 sample carry over to the longer sample. For example, the term structure inverts in almost all of the additional recessions (e.g., in the early 1980s and 1990s), and the term structure of forward discount rates is still close to flat (it is mildly upward sloping on average, but not significantly so).

Most importantly, the model generates interesting differences in the average term structures *across portfolios*. For example, we show that some portfolios (e.g., size-sorted portfolios or idiosyncratic volatility-sorted portfolios) have downward-sloping average term structures of risk premia, whereas others (e.g., book-to-market sorted portfolios) have upward-sloping term structures. These results give us new moments that can be used to evaluate structural asset pricing models that have direct implications about the risk premia of these portfolios (as well as any of the other 102 portfolios that we include in our analysis).

There are also interesting patterns in the *time series* of slopes of the term structure of different portfolios. For example, the slopes of small and large stocks often move closely together, with both term structures were upward sloping during the 1990s and downward sloping during the Great Recession, but only the term structure for small stocks inverted during the late-1990s stock cycle, marking an important divergence between the two portfolios that lasted several years. In contrast, no such divergence in the shape of the term structure can be seen for value and growth stocks in that period. Instead, the largest difference in that case occurred in the recovery from the financial crisis: after 2008, the term structure of value-stock expected returns increased significantly, whereas this pattern did not occur for growth stocks.

Overall, stylized facts that we document provide new conditional moments that asset pricing and structural macro models should seek to match. While we have explored some of the term-structure implications of a small number of models, more work remains to be done to fully explore the implications of our cross section of equity yield term structures.

Finally, we emphasize that our estimates of equity yields, discount rates, and returns on specific dividend claims are subject to estimation uncertainty. Compared to the traditional approach of using observed prices of traded dividend strips to learn about the term structure of risky claims, our methodology has important advantages and disadvantages. On the one hand, it requires estimating the full model to obtain estimated prices for the dividend strips. This introduces parameter uncertainty. On the other hand, it allows us to dramatically expand the available time series, which reduces the sampling uncertainty that is common to both approaches. This trade-off, and, more generally, overall uncertainty, need to be taken into account when bringing empirical moments from these estimation methods to the theoretical models. For this reason, we provide full data on standard errors of our synthetic equity yields along with their point estimates on our website.[20]

Initial submission: March 31, 2021; Accepted: April 7, 2023
Editors: Stefan Nagel, Philip Bond, Amit Seru, and Wei Xiong

Appendix A: The Model

Dividend yields: We start with the Euler equation in (7) and guess the solution of the form in (9). Plugging the expressions (1) to (3), (8), and (9) into (7), we obtain

$$0 = -r_{f,t} - \frac{1}{2}\lambda_t' \Sigma_t \lambda_t + \left(r_{f,t} + \gamma_0 + \gamma_1 F_t + b_0 + b_1 F_{t+1}\right) + \frac{1}{2}\operatorname{var}_t\left[-\lambda_t' u_{t+1} + (b_1 + \gamma_2)u_{t+1} + \epsilon_{r,t+1}\right], \tag{A1}$$

$$0 = (\gamma_0 + \gamma_1 F_t) + \left[b_0 + b_1(c + \rho F_t)\right] - (b_1 + \gamma_2)\Sigma\lambda_t + \frac{1}{2}\operatorname{diag}[\Omega], \tag{A2}$$

where $\operatorname{diag}[\Omega] = \operatorname{diag}\left[(b_1 + \gamma_2)\Sigma(b_1 + \gamma_2)' + \Sigma_\epsilon\right]$. Matching coefficients on the constant and F_t, we have

$$0 = (\gamma_0 - \gamma_2 \Sigma\lambda) + b_0 + b_1(c - \Sigma\lambda) + \frac{1}{2}\operatorname{diag}[\Omega], \tag{A3}$$

$$0 = (\gamma_1 - \gamma_2 \Sigma\Lambda) + b_1(\rho - \Sigma\Lambda). \tag{A4}$$

Given the estimated risk price parameters λ and Λ, price dynamics parameters γ_0, γ_1, and γ_2, state-space dynamics parameters c and ρ, and variances Σ and Σ_ϵ, this gives us two equations to solve for two parameters determining dividend yield parameters b_0 and b_1.

Expected returns on dividend strips: Excess (level) returns on dividend strips can be computed using

$$R_t^{(n)} = \frac{P_{t+1}^{(n-1)}/P_{t+1}}{P_t^{(n)}/P_t}\frac{P_{t+1}}{P_t} \times \exp\left(-r_{f,t}\right) = \frac{P_{t+1}^{(n-1)}/P_{t+1}}{P_t^{(n)}/P_t}\exp\left[\left(r_{t+1} - r_{f,t}\right) - y_{t+1}\right]. \tag{A5}$$

Note that

[20] See https://www.serhiykozak.com/data.

$$\frac{P_t^{(n)}}{P_t} = \mathbb{E}_t^{\mathbb{Q}}\left[\frac{P_{t+1}^{(n-1)}}{P_{t+1}}\frac{P_{t+1}}{P_t}e^{-r_{f,t}}\right] \tag{A6}$$

$$= \mathbb{E}_t^{\mathbb{Q}}\left[\left(\exp\left(a_{n-1,1}+d_{n-1,1}F_{t+1}\right) - \exp\left(a_{n-1,2}+d_{n-1,2}F_{t+1}\right)\right)e^{\Delta p_{t+1}-r_{f,t}}\right] \tag{A7}$$

$$= \exp\left(a_{n,1}+d_{n,1}F_t\right) - \exp\left(a_{n,2}+d_{n,2}F_t\right), \tag{A8}$$

where $a_{n,\cdot}$ and $d_{n,\cdot}$ are given by iteration equations (30) and (31).

We can now compute

$$\mathbb{E}_t\left[\frac{P_{t+1}^{(n-1)}}{P_t}R_{f,t}^{-1}\right] = \mathbb{E}_t\left[\frac{P_{t+1}^{(n-1)}}{P_{t+1}}\frac{P_{t+1}}{P_t}e^{-r_{f,t}}\right] \tag{A9}$$

$$= \mathbb{E}_t\left[\left(\exp\left(a_{n-1,1}+d_{n-1,1}F_{t+1}\right) - \exp\left(a_{n-1,2}+d_{n-1,2}F_{t+1}\right)\right)e^{\Delta p_{t+1}-r_{f,t}}\right] \tag{A10}$$

$$= \exp\left(\tilde{a}_{n,1}+\tilde{d}_{n,1}F_t\right) - \exp\left(\tilde{a}_{n,2}+\tilde{d}_{n,2}F_t\right), \tag{A11}$$

which is exactly the same equation as above, except that expectation is taken under physical dynamics. The solution for $\tilde{a}_{n,\cdot}$ and $\tilde{d}_{n,\cdot}$, therefore, is given by the recursion equations (30) and (31), where we simply replace γ_0^*, γ_1^*, c^*, and ρ^* by their physical counterparts.

Log expected excess returns on the strip are

$$\log\left(\mathbb{E}_t\left[R_{t+1}^{(n)}\right]\right) - r_{f,t} = \log \mathbb{E}_t\left[\frac{P_{t+1}^{(n-1)}}{P_t}R_{f,t}^{-1}\right] - \log\left[\frac{P_t^{(n)}}{P_t}\right] \tag{A12}$$

$$= \log\left[\exp\left(\tilde{a}_{n,1}+\tilde{d}_{n,1}F_t\right) - \exp\left(\tilde{a}_{n,2}+\tilde{d}_{n,2}F_t\right)\right] \tag{A13}$$

$$- \log\left[\exp\left(a_{n,1}+d_{n,1}F_t\right) - \exp\left(a_{n,2}+d_{n,2}F_t\right)\right]. \tag{A14}$$

Equity yield decomposition into hold-to-maturity (HTM) returns and expected growth rates: Let $R_{t,t+n} = \frac{D_{t+n}}{P_t^{(n)}}$ be the HTM return from t to $t+n$. Excess returns are then given by

$$\frac{R_{t,t+n}}{R_{f,t,t+n}} = \frac{D_{t+n}/P_t}{P_t^{(n)}/P_t}R_{f,t,t+n}^{-1}. \tag{A15}$$

Taking expectation, we have

$$\mathbb{E}_t\left[\frac{R_{t,t+n}}{R_{f,t,t+n}}\right] = \frac{\mathbb{E}\left(R_{f,t,t+n}^{-1}D_{t+n}/P_t\right)}{P_t^{(n)}/P_t}, \tag{A16}$$

where the numerator is just equation (27) computed using physical parameters, and the denominator is the same expression computed under risk-neutral dynamics.

Taking logs of the above and dividing through by n gives an annualized log expected HTM return.

Furthermore, note that we can write

$$R_{t,t+n} = \frac{D_{t+n}}{P_t^{(n)}} = \frac{D_t}{P_t^{(n)}} G_{t,t+n}, \tag{A17}$$

where $G_{t,t+n}$ is the cumulative growth rate. We then get

$$\mathbb{E}_t\left[\frac{R_{t,t+n}}{R_{f,t,t+n}}\right] = \frac{D_t}{P_t^{(n)}} \times \mathbb{E}_t\left[\frac{G_{t,t+n}}{R_{f,t,t+n}}\right], \tag{A18}$$

or

$$\frac{1}{n}\log \mathbb{E}_t\left[\frac{R_{t,t+n}}{R_{f,t,t+n}}\right] = \frac{1}{n}\log\left(\frac{D_t}{P_t^{(n)}}\right) + \frac{1}{n}\log \mathbb{E}_t\left[\frac{G_{t,t+n}}{R_{f,t,t+n}}\right], \tag{A19}$$

that is, the dividend yield can be decomposed into the annualized log expected HTM return and the log cumulative growth rate (both excess of the risk-free rate).

Note that log returns on dividend strips can be expressed as

$$\frac{1}{n}\log\left(R_{t,t+n}\right) = \frac{1}{n}\log\left(\frac{D_{t+n}}{P_t^{(n)}}\right) = \frac{1}{n}\log\left(\frac{D_t}{P_t^{(n)}} G_{t,t+n}\right) \tag{A20}$$

$$= e_{t,n} + \frac{1}{n}\log\left(G_{t,t+n}\right). \tag{A21}$$

As in Backus, Boyarchenko, and Chernov (2018), taking unconditional expectations of this equation for strips of maturities n and one, and subtracting them, gives

$$\frac{1}{n}\mathbb{E}\log\left(R_{t,t+n}\right) - \mathbb{E}\log\left(R_{t,t+1}\right) = \mathbb{E}\, e_{t,n} - \mathbb{E}\, e_{t,1}, \tag{A22}$$

that is, the unconditional slope of the dividend yield curve is equal to the slope of the expected log strip return curve.

Appendix B: Estimation

The system of equations consisting of state-space dynamics (1), test asset returns (10), and yields (9) is estimated jointly to obtain the estimates of the vector of parameters

$$\theta = [c, \rho_{r,y}, \rho_{y,y}, \beta_{2,u_r}, b_0, b_{1,y}]', \tag{B1}$$

where β_2 is restricted to load only on shock to returns, that is, $\beta_2 = [\beta_{2,u_r}, \mathbf{0}_{n\times p}]$, and $b_1 = [\mathbf{0}_{n\times p}, b_{1,y}]$ is restricted to load only on valuation ratios. Parameters $\rho_{r,y}$ and $\rho_{y,y}$ denote subblocks of the matrix ρ that correspond to return and D/P loadings on y, respectively.

The system is solved using an iterative Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm. At each iteration, we perform the following steps, which we repeat until convergence:

(i) Given current values of c, $\rho_{r,y}$ and $\rho_{y,y}$, the transition matrix ρ is reconstructed, shocks u_{t+1} in (1) are computed, and Σ is estimated.
(ii) Given current values of β_{2,u_r} and $b_{1,y}$, the variables β_2 and b_1 are constructed for test assets by imposing the restrictions that β_2 loads only on shocks to returns and b_1 loads only on valuation ratios, respectively.
(iii) For state variables (market and principal components [PCs]), the βs are given by equations (18) to (20).
(iv) For given parameter values at any step, we compute prices of risk parameters λ and Λ in equation (3) via a simple solution to the system of $p = \frac{1}{2}k$ equations for variables in the state space given by equations (A4) and (A3). Note that parameters λ and Λ inherit restrictions on physical dynamics. Specifically, only the first p elements of λ are nonzero, since only shocks to PC returns are priced. Similarly, Λ takes the form

$$\Lambda = \begin{bmatrix} 0_{p\times p} & \tilde{\Lambda} \\ 0_{p\times p} & 0_{p\times p} \end{bmatrix},$$

where $\tilde{\Lambda}$ is a $p \times p$ matrix of risk price loadings on D/Ps, which are all fully pinned down by the physical dynamics as well.
(v) Given the estimates of parameters λ and Λ, we can now compute β_0 and β_1 for all test assets using an analog of equations (A4) and (A3) applied to test assets.
(vi) Lastly, we stack state and test asset returns and D/P equations, as well as all parameters. We then use them to estimate the shocks, instruments, and moment conditions. The moment conditions are as follows:
  (a) Using the state vector dynamics in equation (1), we construct shocks u_{t+1} and interact them with instruments that contain a vector of ones and valuation ratios F_t. This corresponds to standard OLS moments with the additional restrictions that the elements of ρ corresponding to loadings on lagged returns are all zero,

$$\mathbb{E}\left(u_{t+1} \otimes \left[1, F_{y,t}\right]\right) = 0,$$

  where $\otimes$ denotes a Kronecker product.
  (b) Using return equation (5), we construct shocks to returns, $\epsilon_{r,t+1}$, and interact them with instruments that contain a vector of ones, valuation ratios in F_t, as well as contemporaneous return shocks in u_{t+1},

$$\mathbb{E}\left(\epsilon_{r,t+1} \otimes \left[1, F_{y,t}, u_{r,t+1}\right]\right) = 0.$$

  (c) Using the price-dividend equation (9), we construct residuals $\epsilon_{y,t}$ and interact them with instruments contain a vector of ones and contemporaneous D/P ratios in the state vector,

$$\mathbb{E}\left(\epsilon_{y,t} \otimes \left[1, F_{y,t}\right]\right) = 0.$$

We use a prespecified generalized method of moments (GMM) weighting matrix in the GMM objective, where the time-series moments have weight one and individual assets' moments are all weighted by the inverse of the square root of the number of test assets, $\sqrt{n}$, to keep their contribution to GMM objective invariant of n.

$$\hat{\theta} = \arg\min_{\theta} g_T(\theta)' W g_T(\theta). \tag{B2}$$

The following parameters are estimated by GMM:

(i) State-space vector variables:
   (a) Intercept c: k parameters.
   (b) Loadings ρ: k loadings onto D/P ratios, $k \times p$ parameters.
(ii) Asset-specific parameters:
   (a) Intercepts of D/P equations b_0: n parameters.
   (b) Loadings of assets' D/P ratios onto state-space D/P ratios: $n \times p$ (loadings on returns are restricted to zeros).
   (c) Loadings of assets' returns onto state-space shocks to returns: $n \times p$ (loadings on D/Ps are restricted to zeros). All other loadings of test assets' returns are pinned down by no arbitrage and SDF risk prices.

The spectral density covariance matrix of moments uses 12 lags and follows the approach in Hansen (1982).

Lastly, to compute standard errors on means of yields or returns (risk premia) that account for both model parameter uncertainty and data sampling uncertainty, we expand the set of moments to include moments corresponding to a regression of these variables on a constant. We then use the structure of our GMM estimator to obtain correct standard errors on the intercept, which is the estimate of interest.

REFERENCES

Adrian, Tobias, Richard K. Crump, and Emanuel Moench, 2015, Regression-based estimation of dynamic asset pricing models, *Journal of Financial Economics* 118, 211–244.

Ang, Andrew, and Maxim Ulrich, 2012, Nominal bonds, real bonds, and equity, Working paper, Columbia Business School.

Backus, David, Nina Boyarchenko, and Mikhail Chernov, 2018, Term structures of asset prices and returns, *Journal of Financial Economics* 129, 1–23.

Bansal, Ravi, Robert F. Dittmar, and Christian T. Lundblad, 2005, Consumption, dividends, and the cross-section of equity returns, *Journal of Finance* 60, 1639–1672.

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, Shane Miller, Dongho Song, and Amir Yaron, 2021, The term structure of equity risk premia, *Journal of Financial Economics* 142, 1209–1228.

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

Belo, Frederico, 2010, Production-based measures of risk for asset pricing, *Journal of Monetary Economics* 57, 146–163.

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, 1743–1776.

Campbell, John Y., 1991, A variance decomposition for stock returns, *Economic Journal* 101, 157–179.

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, and Christopher Polk, 2013, Hard times, *Review of Asset Pricing Studies* 3, 95–132.

Carhart, Mark M., 1997, On persistence in mutual fund performance, *Journal of Finance* 52, 57–82.

Chernov, Mikhail, Lars A. Lochstoer, and Stig R. H. Lundeby, 2018, Conditional dynamics and the multi-horizon risk-return trade-off, Working paper, National Bureau of Economic Research.

Cochrane, John H., and Monika Piazzesi, 2005, Bond risk premia, *American Economic Review* 95, 138–160.

Cochrane, John H., and Monika Piazzesi, 2008, Decomposing the yield curve, Working paper, Graduate School of Business, University of Chicago.

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.

Dew-Becker, Ian, Stefano Giglio, Anh Le, and Marius Rodriguez, 2017, The price of variance risk, *Journal of Financial Economics* 123, 225–250.

Fama, Eugene F., and Kenneth R. French, 2016, Dissecting anomalies with a five-factor model, *Review of Financial Studies* 29, 69–103.

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

Gao, Can, and Ian W. R. Martin, 2021, Volatility, valuation ratios, and bubbles: An empirical measure of market sentiment, *Journal of Finance* 76, 3211–3254.

Giglio, Stefano, Matteo Maggiori, and Johannes Stroebel, 2014, 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, National Bureau of Economic Research.

Giglio, Stefano, and Dacheng Xiu, 2021, Asset pricing with omitted factors, *Journal of Political Economy* 129, 1947–1990.

Gomes, Leandro, and Ruy Ribeiro, 2019, Term structure(s) of the equity risk premium. https://doi.org/10.2139/ssrn.3144028

Gonçalves, Andrei, 2021, The short duration premium, *Journal of Financial Economics* 141, 919–945.

Gormsen, Niels Joachim, 2021, Time variation of the equity term structure, *Journal of Finance* 76, 1959–1999.

Gormsen, Niels Joachim, and Eben Lazarus, 2023, Duration-driven returns, *Journal of Finance* 78, 1393–1447.

Green, Jeremiah, John Hand, and Frank Zhang, 2014, The remarkable multidimensionality in the cross-section of expected US stock returns, Working paper, University of North Carolina at Chapel Hill.

Gupta, Arpit, and Stijn Van Nieuwerburgh, 2021, Valuing private equity investments strip by strip, *Journal of Finance* 76, 3255–3307.

Gurkaynak, Refet S., Brian Sack, and Jonathan H. Wright, 2006. The U.S. treasury yield curve: 1961 to the present, Federal Reserve Board Finance and Economics Discussion Series Paper 2006-28.

Haddad, Valentin, Serhiy Kozak, and Shrihari Santosh, 2020, Factor timing, *Review of Financial Studies* 33, 1980–2018.
Hansen, Lars Peter, 1982, Large sample properties of generalized method of moments estimators, *Econometrica* 50, 1029–1054.
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 Robert J. Hodrick, 1980, Forward exchange rates as optimal predictors of future spot rates: An economic analysis, *Journal of Political Economy* 88, 829–854.
Keloharju, Matti, Juhani T. Linnainmaa, and Peter Nyberg, 2019, Long-term discount rates do not vary across firms, Working paper, National Bureau of Economic Research.
Kogan, Leonid, and Dimitris Papanikolaou, 2013, Firm characteristics and stock returns: The role of investment-specific shocks, *Review of Financial Studies* 26, 2718–2759.
Kogan, Leonid, and Dimitris Papanikolaou, 2014, Growth opportunities, technology shocks, and asset prices, *Journal of Finance* 69, 675–718.
Koijen, Ralph S. J., Hanno Lustig, and Stijn Van Nieuwerburgh, 2017, The cross-section and time series of stock and bond returns, *Journal of Monetary Economics* 88, 50–69.
Kozak, Serhiy, 2024, Kernel trick for the cross section, Working paper, University of Maryland.
Kozak, Serhiy, Stefan Nagel, and Shrihari Santosh, 2018, Interpreting factor models, *Journal of Finance* 73, 1183–1223.
Kozak, Serhiy, Stefan Nagel, and Shrihari Santosh, 2020, Shrinking the cross-section, *Journal of Financial Economics* 135, 271–292.
Kozak, Serhiy, and Shrihari Santosh, 2020, Why do discount rates vary?, *Journal of Financial Economics* 137, 740–751.
Kragt, Jac, Frank de Jong, and Joost Driessen, 2014, The dividend term structure, *Journal of Financial and Quantitative Analysis* 55, 1–72.
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.
Martin, Ian W. R., 2013, Consumption-based asset pricing with higher cumulants, *Review of Economic Studies* 80, 745–773.
Sims, Christopher A., 1980, Macroeconomics and reality, *Econometrica* 48, 1–48.
van Binsbergen, Jules H., Michael W. Brandt, and Ralph Koijen, 2012, On the timing and pricing of dividends, *American Economic Review* 102, 1596–1618.
van Binsbergen, Jules, Wouter Hueskes, Ralph Koijen, and Evert Vrugt, 2013, Equity yields, *Journal of Financial Economics* 110, 503–519.
van Binsbergen, Jules H., and Ralph S. J. Koijen, 2015, The term structure of returns: Facts and theory, Working paper, National Bureau of Economic Research.
Weber, Michael, 2018, Cash flow duration and the term structure of equity returns, *Journal of Financial Economics* 128, 486–503.
Yan, Wen, 2015, Essays on the term structures of bonds and equities, Ph. D. thesis, The London School of Economics and Political Science.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

Appendix S1: Internet Appendix.
Replication Code.