Equity Term Structures without Dividend Strips Data

Stefano Giglio  
*Yale SOM and NBER*

Bryan Kelly  
*Yale SOM, NBER, and AQR*

Serhiy Kozak  
*University of Maryland*

October 16, 2019

PRELIMINARY AND INCOMPLETE

— PLEASE DO NOT CIRCULATE —

Abstract

We use a large cross-section of equity returns to estimate a rich affine model of equity prices, dividends, returns and their dynamics. Using the model, we can price dividend strips of the aggregate market index, as well as any other well-diversified equity portfolio. We do not use any dividend strips data in the estimation of the model; however, model-implied equity yields generated by the model match closely the equity yields from the traded dividend forwards reported in the literature. Our model can therefore be used to extend the data on the term structure of discount rates in three dimensions: (i) over time, back to the 1970s; (ii) across maturities, since we are not limited by the maturities of actually traded dividend claims; and most importantly, (iii) across portfolios, since we generate a term structure for any portfolio of stocks (e.g., small or value stocks). The new term structure data generated by our model (e.g., separate term structures for value, growth, investment and other portfolios, observed over a span of 45 years that covers several recessions) represent new empirical moments that can be used to guide and evaluate asset pricing models.

*Giglio: stefano.giglio@yale.edu; Kelly: bryan.kelly@yale.edu; Kozak: skozak@rhsmith.umd.edu.*
1 Introduction

The term structure of discount rates for risky assets plays an important role in many fundamental economic contexts. For example, pricing an asset or 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 case.

In this paper we specify and estimate a rich affine model of equity portfolios, that describes the prices, dividends, and 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 in several ways: by imposing pricing restrictions, by choosing appropriately the state vector that drives the dynamics of the economy, and by imposing parameter restrictions that reflect recent findings in the literature on returns predictability with large cross-sections. We then use our model to generate term structure of discount rates not only for aggregate cash flows (the S&P 500), but also for all of the other equity portfolios, thus obtaining a large panel of 100 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. (2017) and van Binsbergen and Kojien (2015)). We show that our model – estimated using no dividend strip data at all – manages to match well the prices of observed dividend strips on the S&P 500 of maturities 1, 2, 5 and 7 years, observed since 2004, along a variety of dimensions (average slope, time-series, and so on). After validating the model using observed strip data, we then use the model to explore the properties of implied dividend strips extending the time series back to the 1970s, and studying the cross-section of term structures of different portfolios.

The asset pricing literature has tackled the task of estimating the term structure of discount rates for risky assets in a variety of ways. An earlier literature has proposed extracting information about the term structure from the cross-section of equity portfolios. The broad idea behind this method is that if some stocks are mostly exposed to long-term cash-flow shocks, and other firms are mostly exposed to short-term shocks, the difference in risk premia between the two stocks can be reconducted 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 et al. (2005), Lettau and Wachter (2007), Hansen et al. (2008), Da (2009)).

While this literature made substantial progress in understanding what the cross-section of equity portfolios implied for the term structure of discount rates, it has faced an important hurdle: the term structure of discount rates depends on the entire dynamics of cash flows as well as the risk preferences of investors (and their variation over time). That is, the term structure of discount rates is an equilibrium result that depends on the interaction of a large number of forces. To identify and estimate the model, the papers in this literature have imposed strong assumptions on the preferences (e.g., Epstein–Zin preferences, constant risk aversion, restrictions on which shocks are by investors, etc.), or on the dynamics of the economy and cash flows, or both.
New impetus in the study of term structures has come from the introduction of data on traded dividend claims (van Binsbergen et al. (2012b), Van Binsbergen et al. (2013) and van Binsbergen and Koijen (2015)). The ability to directly observe the returns of finite-maturity dividend claims gives a direct window into the risk premia investors require to hold risks of different maturities, and obviates the need to estimate the dynamics of the economy and the preferences of investors. Studying term structures with traded dividend strip data, however, has a few shortcomings of its own. First, the time series is quite limited, since data starts around 2004, and includes only one recession, specifically the Great Recession. Second, there is little cross-sectional data, since only the aggregate market dividend strips – for the US and other countries – start around 2004; much more limited data is instead available on individual firms. Third, only a portion of the term structure is typically observed (some maturities up to 7 years). Fourth, there are concerns about liquidity of these contracts, which could potentially lead to measurement error.

In this paper, we return to the first approach to estimate term structures, based on equity portfolios alone; we use the model to effectively produce new (implied) term structure data that expands the existing (observed) data along each of those dimension. The term structures we generate cover a large number of cross-sectional portfolios, in addition to the S&P 500: value, size, profitability, momentum, etc., for the total of 100 portfolios. They have a long time series, starting in 1975 and therefore covering several recessions and booms. They have all possible maturities, including the very short and the very long ends of the term structure.

While closely related to the models that had been used in this 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 the ones we observe from traded dividend claims. The model features rich dynamics which are motivated by recent empirical findings in the literature, and we believe are both economically reasonable and statistically parsimonious. In particular, consistent with Kozak et al. (2019), 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” (PC) portfolios, estimated from a large cross-section of 50 anomalies.

Our further motivation comes from Haddad et al. (2019), who 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. Chernov et al. (2018) echoes the importance of time-variation in prices of risk by proposing to test asset-pricing models using multi-horizon returns. Motivated by these findings, our specification also includes four dividend yields (D/P ratios) associated with the “factor” portfolios—for the 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 stochastic discount factor in which shocks to “factor” returns are priced and their risk prices are driven by valuation ratios. The fact that our state variables are yields and 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, it implies that pricing restrictions apply directly to the factors and help better pin
down the dynamics. It is this balance of a rich model with appropriately-chosen restrictions that rep-
resents the core of our paper: it allows us to produce term structures of discount rates that match well
the observed ones – and gives us confidence in extending them over time, maturities, and portfolios.

More specifically, we estimate a standard homoskedastic affine model, in which the factors $F_t$
that drive the dynamics of the economy are chosen to be the returns and dividend yields of four
appropriately-chosen portfolios: that is, $F_t = [r_t, y_t]'$, where $r_t$ are the returns of the four portfolios,
and $y_t$ are their dividend yields. These portfolios are the market and three principal components
of 50 anomaly portfolios, as in Haddad et al. (2019); Kozak (2019). This choice of factors for the
model has two main advantages. First, it captures very well the covariation of returns of the anomaly
portfolios (together, the four factors explain 94.0% of the total variation in returns in the panel of 100
long portfolios). Second, the returns of the factors (each of which is a long-short portfolio) 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 returns (and, through the Campbell-Shiller
identity, dividends).

We restrict this affine model in only two ways. First, we impose that the innovations in the
stochastic discount factor in the economy depends only on the innovations in $r_t$ (and not on the
innovations in $y_t$): that is, we are effectively assuming that the SDF is fully spanned by the returns
of the four factors. This assumption is effectively assuming an APT model for returns, in which the
four factors are the only factors.

Second, we impose that the four yields $y_t$ 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 by assumption in our statistical model.

We do not impose any other restrictions to the model, except of course the pricing restrictions that
link the pricing kernel to prices, dividends and returns. These restrictions imply, as usual in these
cases, that the process of dividend growth for each portfolio is fully pinned down.¹

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, returns and dividends. The model immediately produces
implied term structures of dividend strips and forwards, that is, spot or forwards claims to 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 span
extremely well the cross-section of returns and dividend yields of all the original 100 characteristic-
sorted portfolios, it can easily build implied term structures of any of these portfolios.

Our model derives 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))

¹We do not actually need to use the Campbell-Shiller loglinearization to obtain dividends. As we show in the paper,
since we model returns $r_t$ and dividend yields $y_t = \log(1 + D_t/P_t)$, dividends will be a non-linear function of $r_t$ and $y_t$,
that we do not need to linearize since we always just work with $r_t$ and $y_t$. An advantage of working with $y_t$ instead of the
more standard dividend yield $\log(D_t/P_t)$ is that in our model the log yields of dividend strips will sum up exactly (without
approximations) to the market dividend yield, thus eliminating the need of approximations when relating dividend strips
to the aggregate claim.
over time (back to the 70s) and maturities. In the sample starting in 2004 that was used in 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 in this sample slightly downward sloping, in contrast with the predictions of many models like the long-run risk model, that instead predict that it should be steeply upward-sloping.\textsuperscript{2} We also confirm that our implied term structure inverts during the financial crisis, just like the observed one does.\textsuperscript{3}

Extending the sample to the 1970s allows us to include several additional recessions to our sample; at the same time, the Great Recession carries less overall weight in the sample. It is interesting to see that all the results of the post-2004 sample carry over to the longer sample. The term structure inverts in almost all of the additional recessions (for example, in the early 80s and 90s). And the term structure of forward discount rates is still close to flat (it is mildly upward sloping on average, but it is not significantly so, and is sufficiently flat to still reject the Bansal and Yaron (2004) and Bansal et al. (2012) models in simulations).

In addition, a decomposition of the movements of prices of short-term dividend claims into expected dividend growth and returns 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.

We next study the term structures of discount rates for different portfolios, like value and growth firms, and small and large firms. Our model generates interesting differences in the average term structure across portfolios, and in the time series. For example, we show that value stocks have a strongly increasing term structure of discount rates, whereas growth stocks have a decreasing one; small stocks have an effectively flat discount rate, whereas large stocks have a mildly increasing one. These results give us new moments that can be used to evaluate structural asset pricing models that have direct implication about the risk premia of these portfolios (as well as any of the other 100 we include in our analysis).

Finally, there are interesting patterns in the time series of slopes of the term structure of different portfolios. For example, the slopes of small and large stocks tend to move closely together; both term structures were upward sloping during the 1990s, and both were downward sloping during the Great Recession. Yet, only the term structure for small stocks inverted during the late 90s stock cycle, marking an important divergence between the two portfolios that lasted several years. On the contrary, 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 didn’t happen for growth stocks.

\textsuperscript{2}We confirm by simulating the Bansal and Yaron (2004) and Bansal et al. (2012) models that the null hypothesis that these models generate the observed term structures can be strongly rejected statistically in this sample.

\textsuperscript{3}Incidentally, we note that because our implied dividend strip prices are obtained from equity portfolios, they are less susceptible to the potential criticism raised by Bansal et al. (2017) that traded dividend strips might be illiquid and that bid-ask spreads might affect the conclusions about the slope of the term structure. The fact that we actually match those prices very closely using only (very liquid) equity portfolios suggests that liquidity is not a driver of the findings of van Binsbergen and Koijen (2015) in the first place.
To sum up: our model effectively processes a rich information set (the time-series and cross-sectional behavior of 100 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. Similarly in spirit to the way in which the introduction of vector autoregressions (VAR) 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 structure of discount rates for different portfolios that closely resemble the actual dividend claims we observe in the data, and that can be used by asset pricing models as a moment for evaluation and guidance.

In addition to using our cross-sectional term structures to test asset pricing models, our term structures can, of course, also be used for the valuation of projects with specific horizons. For example, Gupta and Van Nieuwerburgh (2019) use them to evaluate private equity investments. An alternative application is to climate change mitigation investments, where the long end of the term structure is especially important (see, for example, Giglio et al. (2015)).

In addition to 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 has also built on those approaches to improve our understanding of term structures. This literature for the most part focuses 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 et al. (2017)), whereas others use the traded dividend strip data in the estimation (Kragt et al. (2014), Gomes and Ribeiro (2019), Yan (2015)). Recently, Gupta and Van Nieuwerburgh (2019) use term structures of discount rates from a similar affine model to value private equity investments; given their different objective, they use specific portfolios in their model (small and large firms, REITs, and infrastructure firms). Our objective is instead to select the state variables that best describe the whole 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 have explored the term structure of risky assets, in addition to that of equity market dividend claims. Among these, different studies have focused on the term structure of currency risk (Backus et al. (2018)), variance risk (Dew-Becker et al. (2017)), housing risk (Giglio et al. (2014)). Chernov et al. (2018) have stressed the usefulness of multi-period returns to test asset pricing models. Several papers have proposed models that aim to explain observed patterns in the term structure of discount rates, among which Croce et al. (2014), Gormsen (2018).

Finally, our paper relates to a third approach used in the literature to explore the term structure of discount rates using only equity portfolios. This approach, followed by Weber (2018) and Gormsen and Lazarus (2019), is based on estimating the duration of portfolios directly (instead of by modeling the dynamics of dividends and effectively estimating duration as exposures to dividend shocks of different horizons), and using it to back out implied discount rates at different horizons.
2 The Model

State space dynamics. We begin by specifying a general factor model with \( k \) factors \( F_t \), with linear dynamics:

\[
F_{t+1} = c + \rho F_t + u_{t+1},
\]

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

SDF. Denote the risk-free rate as \( r_{f,t} \). We assume a log-linear SDF, where the priced shocks are \( u_{t+1} \), with time-varying risk prices \( \lambda_t \):

\[
m_{t+1} = -r_{f,t} - \frac{1}{2} \lambda_t^\prime \Sigma \lambda_t - \lambda_t^\prime u_{t+1}.
\]

Prices of risk follow:

\[
\lambda_t = \lambda + \Lambda F_t.
\]

The price \( P_t \) and dividend \( D_t \) of any asset satisfies the Euler equation:

\[
1 = E_t \left[ e^{m_{t+1}} \left( \frac{P_{t+1} + D_{t+1}}{P_t} \right) \right] = E_t \left[ e^{m_{t+1}} \frac{P_{t+1}}{P_t} \left( 1 + \frac{D_{t+1}}{P_{t+1}} \right) \right] = E_t \left[ e^{m_{t+1} + \Delta p_{t+1} + y_{t+1}} \right],
\]

where we define the dividend yield as

\[
y_t \equiv \log \left( 1 + \frac{D_t}{P_t} \right).
\]

Note that, given this definition of the dividend yield, we do not use any approximations in the formula above.

Under log-normality, by taking the log of both sides, we obtain:

\[
0 = E_t m_{t+1} + E_t \Delta p_{t+1} + E_t y_{t+1} + \frac{1}{2} \mathbb{V}_t \left[ m_{t+1} + \Delta p_{t+1} + y_{t+1} \right].
\]

Price dynamics. The usual approach to model asset prices is to specify the dividend process \( \Delta d_{t+1} \) and its dynamics, and then derive the equilibrium prices. In this paper we follow an equivalent approach, in which we directly specify the dynamics of log price changes, \( \Delta p_{t+1} \), of assets in the cross-section. We then impose the Euler equation, which in turn implies equations for the dividend process and returns. One advantage of doing so is that by following this route, we do not have to make any approximations or use the Campbell-Shiller loglinearization.

We consider first \( p \) assets that are fully diversified and exposed to the \( u_{t+1} \) shocks. For these, assets, we assume that log price changes in excess of the risk-free rate follow:
\[
\frac{\Delta p_{t+1}}{p_{t+1}} - r_{f,t} = \gamma_0 + \gamma_1 F_t + \gamma_2 u_{t+1},
\]  
(7)

The set of latent factors \( F_t \) determines the conditional mean of price changes, and it is assumed to contain all relevant information about future expected dividend growth and expected returns for these assets.

**Dividend-price ratio.** Specifying the price process as in Eq. (7) is sufficient (together with the SDF) to pin down the implied price of the asset and the dividend process. In particular, the assumptions made so far will guarantee that the yield \( y_t \) will be linear in the factors \( F_t \) (recall that \( y_t \equiv \log \left( 1 + \frac{D_t}{P_t} \right) \)) for assets that are only exposed to the shocks \( u_{t+1} \):4

\[
y_t = b_0 + b_1 F_t.
\]  
(8)

The benefit of this representation is that \( y_t \) approximately equals the level of the dividend-to-price ratio and aggregates properly. That is, D/P on a value-weighted portfolio is approximately just the weighted average of D/Ps; it is also approximately equal to the sum of dividends divided by sum of market caps. Moreover, D/P on a long-short portfolio will often have negative D/P, which is now explicitly permitted by our representation.

**Specializing the state vector.** So far, we have written the setup in a general way. We now take a specific stand on the state vector \( F_t \). We assume that the dynamics of the economy are fully captured by the **returns** and **prices** of \( p = k/2 \) portfolios.

\[
F_t = \begin{bmatrix} r_t \\ y_t \end{bmatrix}
\]  
(9)

So the vector \( F_t \) will contain a total of \( k = p + p \) variables, half of them the returns \( r \) of \( p \) portfolios, and the remaining half yields \( y_t \) of the remaining portfolios. This can be seen as an extension of the setup of Campbell (1991), in which the dynamics of the economy (also represented by a VAR) included returns and the dividend-price ratio of one portfolio (the market), plus additional predictors. Here, the state vector includes a pair \((r_t, y_t)\) for each of \( p \) diversified portfolios, and no additional predictors.

This specification is also motivated by the empirical findings of Kozak et al. (2019) and Haddad et al. (2019). In particular, Kozak et al. (2019) 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. Haddad et al. (2019) 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. Our parsimonious specification allows us to be consistent with both results: \( p \) returns will be chosen in a way that explains the cross-section of expected returns well, while the \( p \) yields

4We show this result formally below.
allow us to capture the dynamics of conditional means and SDF risk prices.

**Restrictions.** We now introduce three restrictions on the model. First, we assume that there are only \( p \) priced risks, and they are fully spanned by our \( p \) portfolios: so \( m_t \) loads only on the \( p \) innovations in returns (the first \( p \) elements of \( u_{t+1} \)). This means that only the first \( p \) elements of \( \lambda_t \) are non-zero. In turn, this implies that the only the first \( p \) rows of \( \lambda \) and \( \Lambda \) 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 et al. (2018); Kozak et al. (2019) who suggests that an SDF constructed from a small number of \( (p) \) diversified portfolio returns prices well the cross-section of returns. We then simply allow the dynamics to also include additional, non-priced shocks.

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

\[
\Lambda = \begin{bmatrix}
0_{p \times p} & \bar{\Lambda} \\
0_{p \times p} & 0_{p \times p}
\end{bmatrix}
\]

where the zeros in the second row are due to the fact that only returns shocks are priced, and the zeros in the top-left corner imply that risk premia variation is entirely driven by \( y_t; \) \( \bar{\Lambda} \) is a \( p \times p \) matrix of risk price loadings on dividend yields.

As we show in greater detail below, these assumptions imply restrictions on the transition matrix \( \rho \) of the factors: it implies that conditional expected excess returns are a function of lagged yields but not lagged realized returns. To this restriction, we add a third restriction on the conditional mean of the yields: we impose that it is also only a function of the lagged yields but not of the 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 et al. (2019) that own valuation ratios are the strongest predictors of PC 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 voluminous literature documenting low autocorrelation in equity returns.

Similar restrictions have been explored 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 only by the Cochrane and Piazzesi (2005) factor.

### 2.1 Prices, returns, and dividends

We now derive the prices, returns, and dividends of the \( p \) portfolios included in \( F_t \).
2.1.1 Prices of the $p$ portfolios

The parameters of the model in the setup above are linked through the valuation equation. Note that for the $p$ portfolios that enter $F_t$, we have:

$$V_t[m_{t+1} + \Delta p_{t+1} + y_{t+1}] = V_t[-\lambda'_t u_{t+1} + (b_1 + \gamma_2) u_{t+1}]$$
$$= \lambda'_t \Sigma \lambda_t + (b_1 + \gamma_2) \Sigma (b_1 + \gamma_2)' - 2(b_1 + \gamma_2) \Sigma \lambda_t,$$

We can use this to write the Euler equation as:

$$0 = (\gamma_0 + \gamma_1 F_t) + [b_0 + b_1 (c + \rho F_t)] - (b_1 + \gamma_2) \Sigma \lambda_t + \frac{1}{2} \text{diag} [\Omega], \quad (10)$$

where $\text{diag} [\Omega] = \text{diag} [(b_1 + \gamma_2) \Sigma (b_1 + \gamma_2)']$. Matching coefficients, we have:

$$0 = (\gamma_0 - \gamma_2 \Sigma \lambda) + b_0 + b_1 (c - \Sigma \lambda) + \frac{1}{2} \text{diag} [\Omega]. \quad (11)$$

Matching coefficients on $F_t$:

$$0 = (\gamma_1 - \gamma_2 \Sigma \Lambda) + b_1 (\rho - \Sigma \Lambda). \quad (12)$$

2.1.2 Returns of the $p$ portfolios

Log returns can be expressed as

$$r_{t+1} = \Delta p_{t+1} + y_{t+1}. \quad (13)$$

Combining equations for $\Delta p_{t+1}$ and $y_{t+1}$ gives an expression for excess returns:

$$\begin{bmatrix} r_{t+1} \\ F_{t+1} \\ u_{t+1} \end{bmatrix} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix} \begin{bmatrix} m_{t+1} \\ \Delta p_{t+1} \\ y_{t+1} \end{bmatrix}, \quad (14)$$

where

$$\begin{align*}
\beta_0 &= \gamma_0 + b_0 + b_1 c \\
\beta_1 &= \gamma_1 + b_1 \rho \\
\beta_2 &= \gamma_2 + b_1.
\end{align*} \quad (15, 16, 17)$$

Note that because the returns of the $p$ portfolios are part of $F_t$, the matrix $\beta_2$ will have nonzero elements only in the first $p$ columns (that is, the first $p$ of the $u_{t+1}$ shocks are the $p$ portfolios’ returns).
2.1.3 Dividends of the p portfolios

Our model specifies the joint dynamics of log dividend yields $y_t$ and returns $r_t$. In turn, this implies dynamics for the dividends. These dynamics are, in general, not linear, because of the nonlinearity of the relation between returns, dividends, and prices. Rather than imposing log-linearity via a Campbell-Shiller loglinearization, we work directly with the dividends and their nonlinear dynamics, solving for them numerically.

We infer log dividend growth from the returns equation:

$$\Delta p_{t+1} = y_{t+1} + pd_{t+1} - pd_t + \Delta d_{t+1}, \quad (18)$$

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

2.1.4 Prices, returns, and dividends of other portfolios

The equations above hold for the $p$ portfolios that are included in $F_t$. In practice, there will be many diversified portfolios that also span the same risks, and that are, in this sense, redundant relative to those portfolios. Similarly to the bond pricing literature, we assume these portfolio’s prices, price changes, and returns are measured with error, respectively $e_t$, $v_t$, and $\epsilon_t = \epsilon_t + v_t$:

$$y_t = b_0 + b_1 F_t + \epsilon_t \quad (19)$$

$$\Delta p_{t+1} - r_{f,t} = \gamma_0 + \gamma_1 F_t + \gamma_2 u_{t+1} + v_{t+1} \quad (20)$$

$$r_{t+1} - r_{f,t} = \beta_0 + \beta_1 F_t + \beta_2 u_{t+1} + \epsilon_{t+1}. \quad (21)$$

We allow the measurement error on the return and price equation to be arbitrarily correlated, and we solve for the returns, prices, and dividends of these portfolios (the equations above apply with minor modifications since $v$ and $\epsilon$ are assumed to be idiosyncratic errors).

2.2 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. In particular, it is useful to distinguish reduced-form parameters that directly enter moment conditions that can be estimated from the data, and structural parameters that in turn determine the reduced-form parameters.

Reduced-form parameters and moment conditions. Consider first that once $p$ portfolios are selected to be in $F_t$, we observe their yields $y_t$ and returns $r_t$; hence, the entire state vector $F_t$ is observed. Eq. (1) therefore implies a first set of moment conditions, which depend on the parameters $c$ and $\rho$. Note that for the $p$ portfolios in $F_t$, the reduced-form parameters $b_0$ and $b_1$ from Eq. (8) are pinned down by the fact that $y_t$ is part of $F_t$, and the parameters $\beta_0$, $\beta_1$ and $\beta_2$ from Eq. (14) are directly a function of $c$ and $\rho$ (because these return equations are just the first $p$ rows of Eq. (1)). For any other fully-diversified portfolio beyond the $p$ that we assume to observe without error, we will have two sets of moment conditions: one relating contemporaneously their yields to the factors,
Eq. (19), and another relating returns to lagged factors, Eq. (21); given that both Eq. (19) and Eq. (20) are assumed to have measurement error, the returns equation will also have measurement error. These equations form a set of moment conditions that depend on parameters $\beta_0, \beta_1, \beta_2, b_0$ and $b_1$ for all diversified portfolios.

To sum up, moment conditions from equations Eq. (1), (8), (14) depend on the parameters $c$ and $\rho$, as well as $\beta_0, \beta_1, \beta_2, b_0$ and $b_1$ for all diversified portfolios.

Structural parameters. We then have a set of structural parameters that are linked to the reduced-form parameters by identities and arbitrage restrictions. First, the parameters $\gamma_0, \gamma_1$ and $\gamma_2$ are related one-to-one to the reduced-form parameters (Eq. (15), (16), (17)); these parameters do not add any additional restriction to the system, and simply correspond to an alternative representation of Eq. (14), in terms of $\Delta p_{t+1}$ instead of $r_{t+1}$.

More important are the other structural parameters $\lambda$ and $\Lambda$. These parameters are introduce restrictions on the reduced-form parameters of all portfolios (not just the $p$ portfolios that enter $F_t$), through the valuation equations Eq. (11) and Eq. (12).

Estimation. We estimate the model using GMM. We use the moment conditions described above to estimate the reduced-form and structural parameters (i.e., imposing the valuation restrictions Eq. (11) and Eq. (12)) using all available portfolios, and assuming that only the $p$ portfolios included in $F_t$ are observed without error. Individual assets’ moments are all normalized by the square root of the number of test assets, $\sqrt{n}$, to keep their contribution to GMM objective invariant to $n$. We use an identity GMM weighting matrix in the estimation.

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

2.3 Dividend strips

We next derive the prices and returns of theoretical dividend strips in the model. These can be computed after having estimated the model, because they are simple functions of the parameters we estimate via GMM. Consider first a fully-diversified portfolio with its 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

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

we then have that the price of a claim to \( D_{t+n} \), \( P_t^{(n)} \), as a fraction of the price of the portfolio, \( P_t \), is:

\[
\frac{P_t^{(n)}}{P_t} = w_t^{(n)} = \mathbb{E}^Q \left\{ D_{t+n} \exp \left[ - \sum_{i=1}^{n} r_{f,t+i-1} \right] \right\} = \mathbb{E}^Q \left\{ \exp \left( y_{t+n} \right) \exp \left[ \sum_{i=1}^{n} \left( \Delta p_{t+i} - r_{f,t+i-1} \right) \right] \right\} = \mathbb{E}^Q \left\{ \exp \left[ y_{t+n} + \sum_{i=1}^{n} \left( \Delta p_{t+i} - r_{f,t+i-1} \right) \right] \right\} - \mathbb{E}^Q \left\{ \exp \left[ \sum_{i=1}^{n} \left( \Delta p_{t+i} - r_{f,t+i-1} \right) \right] \right\} = \exp \left( a_{n,1} + d_{n,1} F_t \right) - \exp \left( a_{n,2} + d_{n,2} F_t \right),
\]

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

\[
a_{n,-} = a_{n-1,-} + \gamma^*_0 + d_{n-1,-} c^* + \frac{1}{2} (d_{n-1,-} + \gamma_2) \Sigma \left( d_{n-1,-} + \gamma_2 \right)' + \frac{1}{2} \sigma^2_v,
\]

\[
d_{n,-} = \gamma^*_1 + d_{n-1,-} \rho^*,
\]

and initial values by \( a_{0,1} = b_0 + \frac{1}{2} (\sigma_r^2 - \sigma_v^2) \), \( d_{0,1} = b_1 \), \( a_{0,2} = 0 \), \( d_{0,2} = 0 \), \( \sigma_r^2 = \var (\varepsilon_t) \), and \( \sigma_v^2 = \var (\nu_t) \). In these formulas, stars indicate risk-neutral parameters.\(^5\) Note that the prices of dividend strips are obtained as a function of observable returns and dividend yields and their dynamics (captured by \( F_t \) and their dynamics). This is because specifying these dynamics implicitly fully specifies the dividend dynamics as well. 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). Finally, a nice feature of our setup is that the log dividend yields on the strips are exact linear functions of the factors – without the need of an approximation.

**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):

\[
\begin{align*}
    e_{t,n} &= \frac{1}{n} \log \left( \frac{D_t}{P_t^{(n)}} \right) = \frac{1}{n} \left[ \log \left( \exp \left( y_t \right) \right) - \log \left( \frac{P_t^{(n)}}{P_t} \right) \right].
\end{align*}
\]

No approximation are required for these calculations. Forward equity can be easily computed using the observable bond yields:

\[
\begin{align*}
    e^f_{t,n} &= \frac{1}{n} \log \left( \frac{D_t}{F_t^{(n)}} \right) = e_{t,n} - y_{t,n},
\end{align*}
\]

where \( F_{t,n} \) denotes the futures (or forward) price, which, under no arbitrage, is linked to the spot price

\[^5\text{Risk-neutral parameters are defined as: } \gamma_0^* = \gamma_0 - \gamma_2 \Sigma \lambda, \gamma_1^* = \gamma_1 - \gamma_2 \Sigma \lambda, c^* = c - \Sigma \lambda, \rho^* = \rho - \Sigma \lambda.\]
by

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

(29)

where $y_{t,n}$ is the nominal government bond yield with no default risk.

**Returns on dividend strips.** Excess (level) returns on dividend strips can be computed using

$$R_{t}^{(n)} = \frac{P_{t+1}^{(n-1)}/P_{t+1} - P_{t}^{(n)}/P_{t}}{\exp (-r_{f,t})} = \exp \left( r_{t+1} - r_{f,t} \right) - y_{t+1} .$$

(30)

**Expected returns on dividend strips.** In Section A.1 we show that log expected excess returns on an $n$-year dividend strip are given by

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

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

$$- \log \left[ \exp \left( a_{n,1} + d_{n,1} F_{t} \right) - \exp \left( a_{n,2} + d_{n,2} F_{t} \right) \right],$$

(31)

(32)

(33)

where $\tilde{a}_{n,\cdot}$ and $\tilde{d}_{n,\cdot}$ are given by the recursion in Eq. (25)-(26), where we simply replace $\gamma_{*}, \gamma_{1}^{*}, c^{*},$ and $\rho^{*}$ by their physical dynamics counterparts.

**Equity yield decomposition into HTM returns and expected growth rates.** Section A.2 shows that an $n$-year dividend yield can be decomposed into annual hold-to-maturity (HTM) expected excess returns over $n$ years and annual expected dividend growth rates over the $n$-year period, net of the risk-free rate:

$$\frac{1}{n} \log \mathbb{E}_{t} \left[ \frac{R_{t,t+n}}{R_{f,t,t+n}} \right] = \frac{1}{n} \log \left( \frac{D_{t}}{P_{t}^{(n)}} \right) + \frac{1}{n} \log \mathbb{E}_{t} \left[ \frac{G_{t,t+n}}{R_{f,t,t+n}} \right].$$

(34)

**Relationship between the slope of the strip yield curve and strip expected returns.** Note that log returns on dividend strips can be expressed as

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

$$= e_{t,n} + \frac{1}{n} \log (G_{t,t+n})$$

(35)

(36)

As in Backus et al. (2018), taking unconditional expectations of this equation for strips of maturities $n$ and 1, and subtracting them, gives:

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

14
that is, the unconditional slope of the dividend yield curve is equal to the slope of the expected log strip return curve.

3 Data

3.1 Stock data

We focus on a broad set of fifty stock-specific characteristics and long and short legs of portfolio sorts based on these characteristics. This set of portfolios effectively summarizes heterogeneity in expected returns, as shown by Kozak et al. (2019). 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 50 characteristics studied in Kozak et al. (2019); Kozak and Santosh (2019) and listed in Table A.1. Portfolios include all NYSE, AMEX, and NASDAQ firms; however, the breakpoints use only NYSE firms as in Fama and French (2016). We obtain 100 portfolios, two for each anomaly (P1 & P3). Our sample consists of monthly returns from August 1975 to November 2016. For each portfolio we construct a corresponding measure of its valuation based on dividend-to-price ratios of the underlying stocks. We then construct the yield, $y_t$, as defined in Eq. (5).

We also construct long-short portfolios as differences between each anomaly’s log return on portfolio 3 minus the 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 has been documented in the vast literature on the cross-section of returns and can be verified in Appendix Table A.1. In our sample, most anomalies show a strong pattern in average returns across tercile portfolios, consistent with prior research.

3.2 Choice of variables in the state vector

Kozak et al. (2018); Kozak et al. (2019) 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. We use this insight to guide our choice of portfolios in the state space dynamics in Eq. (1). In particular, we use returns on the market and three PCs of 50 long-short portfolios based on the underlying long and short ends of each characteristic used in sorting, as our choice for $r_t$ in Eq. (9). Formally, consider the eigenvalue decomposition of anomaly excess returns, \( \text{cov} (F_{t+1}) = Q \Lambda Q' \), where \( Q \) is the matrix of eigenvectors and \( \Lambda \) is the diagonal matrix of eigenvalues. The \( i \)-th PC portfolio is formed as \( PC_{i,t+1} = q_i' F_{t+1} \) where \( q_i \) is the \( i \)-th column of \( Q \).

Table 1 shows that anomaly portfolio returns exhibit a moderately strong factor structure. The top panel focuses on long-short portfolios with the market—which captures the vast majority of all co-movement across portfolios—effectively removed. Table 1, 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

---

6The dividend-to-price 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 1: Percentage of variance explained by anomaly PCs

Percentage of variance explained by each PC of the 50 long-short anomaly portfolio returns (top panel) and each PC of 100 returns on long and short legs of each characteristic sort (bottom panel).

<table>
<thead>
<tr>
<th>PC1</th>
<th>PC2</th>
<th>PC3</th>
<th>PC4</th>
<th>PC5</th>
<th>PC6</th>
<th>PC7</th>
<th>PC8</th>
<th>PC9</th>
<th>PC10</th>
</tr>
</thead>
<tbody>
<tr>
<td>% var. explained</td>
<td>24.9</td>
<td>18.6</td>
<td>10.9</td>
<td>5.7</td>
<td>4.4</td>
<td>4.2</td>
<td>3.0</td>
<td>2.8</td>
<td>2.4</td>
</tr>
<tr>
<td>Cumulative</td>
<td>24.9</td>
<td>43.5</td>
<td>54.4</td>
<td>60.1</td>
<td>64.5</td>
<td>68.7</td>
<td>71.7</td>
<td>74.5</td>
<td>76.9</td>
</tr>
</tbody>
</table>

| % var. explained | 88.4 | 4.1  | 1.6  | 1.1  | 0.6  | 0.5  | 0.4  | 0.3  | 0.3  |
| Cumulative      | 88.4 | 92.5 | 94.1 | 95.2 | 95.8 | 96.4 | 96.9 | 97.2 | 97.5 | 97.8 |

captured by the market. The bottom panel extracts PCs of all 100 long and short legs of each sort, leaving the market intact. It shows that the first four PCs (approximately the market and three cross-sectional PCs) capture more than 95% of variation in returns across 100 portfolios.

Haddad et al. (2019) 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 $p$ yields corresponding to the market and three PC factors as our choice for $y_t$ in Eq. (9). Yields on PCs are constructed simply as a linear combination of log yields on underlying portfolios, with weights determined by returns’ eigenvectors, $y_{pc_i} = q_i'y_t$.

In summary, our state vector is given by $F_t = [r_{mkt,t}, r_{pc1}, r_{pc2}, r_{pc3}, y_{mkt,t}, y_{pc1}, y_{pc2}, y_{pc3}]'$.  

3.3 Bond data

Our model is designed to be orthogonal to the term structure of interest rates and does not require any bond data in model estimation, other than the one-year risk-free rate $r_{f,t}$. To directly compare our results to Bansal et al. (2017); Van Binsbergen et al. (2013); van Binsbergen et al. (2012a), we do, however, need to use yields on government bond of maturities 1Y-15Y to convert our model-implied strip equity yields into forward yields, reported by these papers. Our data comes from Gurkaynak et al. (2006), who provide a long history of interpolated US bond yields.

After merging the equity and bond data, we obtain a full sample of annual (overlapping) observations at monthly frequency from August 1975 to November 2016.

4 Results

In this section we report the empirical results of our estimation. We begin by reporting the estimates and the fit of the model; specifically, we show how the model fits the returns of the 50 anomaly portfolios, their price-dividend ratios, and their dividends.

Next, we evaluate the performance of the model against data that was not used in the estimation:
Table 2: Estimates of the dynamics of the factors $F_t$

We report estimates of the dynamics of the factors $F_t$: $c$ and $\rho_{y}$ in $\rho = [0_{K \times p}, \rho_{y}]$, and the $R^2$ for each of the eight equations of the dynamics of $F_t$. GMM standard errors using a spectral density matrix with 12 lags in parentheses. The dynamics are estimated at an annual horizon using monthly overlapping observations in the August 1975 to November 2016 sample.

<table>
<thead>
<tr>
<th></th>
<th>$c$</th>
<th>$\rho_{y}$</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r_{mkt}$</td>
<td>0.02</td>
<td>1.87</td>
<td>2.17</td>
</tr>
<tr>
<td></td>
<td>(0.06)</td>
<td>(3.04)</td>
<td>(1.36)</td>
</tr>
<tr>
<td>$r_{pc1}$</td>
<td>0.10</td>
<td>-17.49</td>
<td>11.10</td>
</tr>
<tr>
<td></td>
<td>(0.17)</td>
<td>(11.66)</td>
<td>(3.50)</td>
</tr>
<tr>
<td>$r_{pc2}$</td>
<td>-0.11</td>
<td>1.85</td>
<td>-1.03</td>
</tr>
<tr>
<td></td>
<td>(0.10)</td>
<td>(3.45)</td>
<td>(2.27)</td>
</tr>
<tr>
<td>$r_{pc3}$</td>
<td>0.05</td>
<td>-1.50</td>
<td>4.28</td>
</tr>
<tr>
<td></td>
<td>(0.06)</td>
<td>(4.07)</td>
<td>(1.45)</td>
</tr>
<tr>
<td>$y_{mkt}$</td>
<td>0.00</td>
<td>0.94</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
<td>(0.07)</td>
<td>(0.04)</td>
</tr>
<tr>
<td>$y_{pc1}$</td>
<td>-0.01</td>
<td>0.45</td>
<td>0.76</td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
<td>(0.16)</td>
<td>(0.05)</td>
</tr>
<tr>
<td>$y_{pc2}$</td>
<td>-0.00</td>
<td>0.64</td>
<td>0.20</td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td>(0.20)</td>
<td>(0.16)</td>
</tr>
<tr>
<td>$y_{pc3}$</td>
<td>-0.01</td>
<td>-0.71</td>
<td>0.31</td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
<td>(0.22)</td>
<td>(0.09)</td>
</tr>
</tbody>
</table>

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 et al. (2012b), Van Binsbergen et al. (2013) and van Binsbergen and Koijen (2015) on the term structure of S&P 500 dividend strips and forwards.

We then explore the novel empirical facts that emerge from our model, along 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 shape of the term structure of different portfolios vary in interesting ways over time. We then 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 (for example, the difference in the term structure behavior of value and growth stock, or high- and low-profitability stocks) provide new moments that can help guide the construction and evaluation of asset pricing models.

4.1 Fit of the model

As discussed in section Section 2.2, 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
We report risk-neutral estimates of the dynamics of the factors $F_t$: $c^Q_r, c^Q_y$, and $\rho^Q_{y,y}$ in $c^Q = [c^Q_r, c^Q_y]'$ and $\rho^Q = [0_{K \times p}, \rho^Q_{y,y}]$, where $\rho^Q_{y,y} = [0, \rho^Q_{y,y}]'$. Annual overlapping observations at monthly frequency from August 1975 to November 2016.

<table>
<thead>
<tr>
<th></th>
<th>$c^Q_r$</th>
<th>$c^Q_y$</th>
<th>$\rho^Q_{y,y}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>MKT</td>
<td>-0.01</td>
<td>0.00</td>
<td>0.96</td>
</tr>
<tr>
<td>PC1</td>
<td>-0.08</td>
<td>-0.00</td>
<td>0.24</td>
</tr>
<tr>
<td>PC2</td>
<td>-0.06</td>
<td>-0.00</td>
<td>0.56</td>
</tr>
<tr>
<td>PC3</td>
<td>-0.02</td>
<td>-0.01</td>
<td>-0.47</td>
</tr>
</tbody>
</table>

plus three principal components from 50 anomalies). Table 2 reports the estimates of the dynamics of the factors $F_t$: $c$ and $\rho$. As discussed in Section 2, we impose that conditional mean of returns and yields is only a function of the lagged yields but not of the lagged returns, that is, $\rho = [0_{K \times p}, \rho_{y,y}]$. In Table 2 we, therefore, omit zeros and report only estimates and standard errors of $\rho_{y,y}$, as well as $c$. In parentheses, we report GMM standard errors using a spectral density matrix with 12 lags.

For comparison, Table 3 reports the risk-neutral parameters $c^Q$ and $\rho^Q$ implied by the model. Note that all conditional loading of returns on lagged yields are zero, $\rho_{r,y} = 0$, since under the risk-neutral dynamics 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, $c^Q_r$ and $c^Q_y$. The intercepts of the return regressions, $c^Q_r$, are non-zero and reflect a variance adjustment.

The last column of Table 2 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 et al. (2019), who show that principal components of anomaly returns are strongly and robustly predictable by their valuation ratios, and that this predictability is essential for capturing dynamic properties of an SDF. That paper also shows that this predictability survives in the out-of-sample exercise, suggesting that our analysis should also be robust to out-of-sample tests.

One of the main objectives of the model is to fit the cross-section of returns for the 50 anomalies (and not just the four linear combinations we use as factors). The average $R^2$ for the regression of portfolio returns in Eq. (21) is 94.0%. The average $R^2$ for the regression of dividend yields onto the factor dividend yields in Eq. (19) is 98.7%. Our model also fits well the cross-section of risk premia for the 50 anomalies, reaching a cross-sectional $R^2$ of 49.1% in the full sample from August 1975 to November 2016. This compares favorably with 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 50 anomalies with an $R^2$ close to 0%).

Kozak et al. (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.
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 50 anomalies.

To get a sense of the factors’ behavior in the data, Fig. 1 and Fig. 2 report the time-series of the dividend yield for the four factors and their returns. There are several interesting patterns to note. First, while the first factor (the market) has an extremely persistent dividend yield, as 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 et al. (2019) argue that time-series predictability of these factors is critical to adequately capturing 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, we need to be able to estimate shocks that give rise to different dynamic responses of the economy, to be able to estimate how investors
price these shocks differently.

4.2 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 has come from the study of traded claims to dividends of finite maturity (dividend strips and dividend futures), starting with the seminal work of van Binsbergen et al. (2012b). Recent papers have expanded the sample to include the prices of dividend forwards traded over the counter and, more recently, on exchanges (e.g. Bansal et al. (2017) and van Binsbergen and Koijen (2015)).

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 includes equity portfolios. 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. (2017) 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 in this section on the corresponding moments from the sample 2004–2019, so that the sample moments are directly comparable.

Fig. 3 plots the time series of forward equity yields implied by our model (defined in Eq. (28)) against the most recent data from Bansal et al. (2017). The first four panels report the equity yields and standard errors for maturities 1, 2, 5, and 7 years, respectively. The bottom panel reports the time-series of the slope between the 7 and 1 year maturity yields. The figure shows that our model does an excellent job in matching the traded forward equity yields.

Fig. 4 summarizes the dynamics of the model-implied yields in this sample, and resembles closely Figure 1 in Bansal et al. (2017). As clear from this figure, the shape of the term structure varies over time. It is sometimes relatively flat (for example, after 2012), sometimes upward sloping (as at the beginning of 2010), and sometimes steeply downward sloping (as during the financial crisis) – just like the traded forwards.

Given that our implied equity forward yields match closely those of the traded claims, it should not be surprising that we also 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 US. Fig. 5 reports the estimated term structure of forward risk premia for this sample.

Another dimension in which our data matches 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 expected cash-flow variation is, surprisingly, 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 almost entirely attributed to a sharp decline in expected dividend growth. We can perform the same decomposition in our model; Fig. 6 shows that the results looks extremely similar to those in Van Binsbergen et al. (2013).

The Appendix shows the similarity between our implied yields and the ones from Bansal et al. (2017) and van Binsbergen and Koijen (2015) along additional dimensions: the average market beta
Figure 3: Model-implied forward equity yields vs. forward equity yield data. We compare our model-implied forward yields to their empirical counterparts in Bansal et al. (2017). Shaded areas depict two-standard-deviation bands around point estimates. Model parameters are estimated in full sample.
of the strips of different maturity in Fig. A.16 is similar to the one reported in van Binsbergen and Koijen (2015) (Figure 4); it also plots the one-year returns of the 1- and 2-year dividend forwards in the data and in our model in Fig. A.17, showing that they 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 see that they match well along several dimensions. This provides an external validation of the ability of our model to capture the dynamics of risk and cash flows and investor preferences for risks along the term structure.

Next, we use the model to explore the behavior of the equity term structure over the longer sample (beginning in 1975), and we then turn to the cross-section of term structures for different risks.

### 4.3 The time series of the equity term structure since the 1970s

Fig. 7 shows the forward equity yields for the aggregate market for different maturities, as estimated...
Figure 6: Decomposition of the 2-year equity yield in the Bansal et al. (2017) sample. The figure decomposes the 2-year equity yield on the market into log HTM risk premium and log dividend growth net of the risk-free rate.

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 1, 2, 5, and 7 years. Equity yields are constructed using the trailing 12-month dividend. The sample is from August 1975 to November 2016.

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 induced an inversion of the curve, followed by a period in which the slope changed sign several times.

These changes in the slope of the term structure are strongly correlated with the macroeconomic cycle. To see this more clearly, Fig. 8 shows the term structures conditional on being or not in an NBER recession. Outside recessions (blue line) the term structure is essentially flat on average. In recessions, instead, the term structure is steeply downward sloping.

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 where the only recession is the financial crisis). We can therefore ensure that the patterns we find are not entirely
Figure 8: Term structures of forward equity yields conditional on NBER recessions. The figure shows the term structures conditional on being (red) or not (blue) in an NBER recession. Shaded areas depict two-standard-deviation error bands around point estimates. GMM standard errors using a spectral density matrix with 12 lags.

Figure 9: Decomposition of the 5-year equity yield. The figure decomposes the 5-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, August 1975 to November 2016.

due to the specialness of the Great Recession. So for example, we see inversion in the yield curve at several other times in history, as confirmed by figure Fig. 8 that 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 term structure of discount rates for different horizons: as pointed out in Dew-Becker et al. (2017) and Backus et al. (2018),

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

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

Finally, we report in Fig. 9 the decomposition of the 5-year equity yield into expected annual dividend growth over 5 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
Figure 10: Slope 7-1 of forward equity yields for small and large stocks. The figure shows time-series of estimated slope of forward equity yields (7y - 1y) from our model for diversified portfolios of small (long) and large (short) stocks in full sample, August 1975 to November 2016.

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.

4.4 Term structures of different risks

Perhaps the biggest 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 one for growth firms, one for small firms and one for large firms, and so on. In turn, these term structures can be used to test the implications of structural models.

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

Fig. 10 and Fig. 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 emerge.

Large firms' and small firms’ term structures moved in similar ways for long stretches of time (including during the 1992 recession). But their equity yields went in opposite directions during the tech boom and bust. Whereas the equity term structure inverted for small firms during the tech boom, it didn’t do so for large firms. During the recent financial crisis, instead, both curves inverted. On the contrary, 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 didn’t happen for growth stocks.

Fig. 12 and Fig. 13 show the full-sample term structure of forward risk premia for the same four portfolios. The figures do not merely differ in their levels: the term structure is on average strongly positively sloped for value firms, mildly positively sloped for large firms, flat for small firms, and negatively sloped for growth firms.

We conclude that our model generates very 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&P500 dividend that has already been studied using traded dividend forwards).

4.5 Applications

We briefly discuss here potential applications of our findings. As mentioned in the introduction, we 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 working directly with the rich dynamics of the economy and the large cross-section of available equity portfolios, researchers can use the generated term structures directly, to calibrate and evaluate models and compute valuations of investments.

One immediate application of our term structures is in the evaluation and testing of asset pricing models. Many asset pricing models, for example, those built upon the long-run risk models of Bansal and Yaron (2004) (BY) and Bansal et al. (2010) (BKY), have strong predictions about the term structure of discount rates for risky assets. van Binsbergen et al. (2012b) and van Binsbergen and Koijen (2015) show that the slope of the aggregate dividend term structure observed in the data is on average too steep to be explained by the BY and BKY models – but their conclusions relied on a small sample which includes the financial crisis.
Figure 12: Term structure of forward risk premia for small and large portfolios. Mean realized returns on forward contracts in full sample, August 1975 to November 2016, by maturity. Shaded areas depict two-standard-deviation error bands around point estimates.

Using our method, 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 7 years, gives our test substantially more power. As reported above, in the longer sample the curve is estimated to be mildly upward sloping. However, we find that it still is substantially less steep than what the model predicts. For example, using the unconditional slopes 7-1 and 15-1 as moments, we find that the BY and BKY models are both statistically rejected in the 2004+ sample as well as in the full sample starting in 1975. These results confirm the difficulty of the long-run risk models in matching the term structures of discount rates observed in aggregate dividend forwards by van Binsbergen et al. (2012b) and van Binsbergen and Koijen (2015), in variance swaps by Dew-Becker et al. (2013), in bonds by Beeler and Campbell (2012), and in housing by Giglio et al. (2014).

In addition to extending the results on aggregate dividend strips, our paper can also be used to test directly the implications of models about the term structure of discount rates of specific portfolios. For example, Hansen et al. (2008) directly discuss the implied term structure of value and growth stocks in their model. Many production-based models (for example, Belo (2010); Kogan and Papanikolaou
Finally, our model can be used to evaluate investments of different horizons. Once the riskiness of any investment with specified maturity is determined (that is, the relation between the 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 investment as in Gupta and Van Nieuwerburgh (2019), 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).

4.6 Counterfactual analysis

In this section we investigate whether the dynamics of equity strip yields produced by our model and observed in the data can be obtained by specifying a simpler benchmark model, instead of relying on

Figure 13: Term structure of forward risk premia for value and growth portfolios. Mean realized returns on forward contracts in full sample, August 1975 to November 2016, by maturity. Shaded areas depict two-standard-deviation error bands around point estimates.
the wide cross-section of 50 characteristics we use. We consider two benchmarks.

Our first benchmark is the conditional CAPM model. That is, we assume there is a single priced source of risk—the aggregate market—and its risk price time-variation is driven by the aggregate dividend-price ratio. We replicate the analysis in the paper for this choice of the 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. We replicate the analysis in the paper for this choice of a state vector.

We report the results in Fig. 14 and Fig. 15. Panel a in the figures shows results for the CAPM benchmark. Panel b shows result for the Fama and French 5-factor model, supplemented with the momentum factor. Fig. 14 plots the time-series of yields in the Bansal et al. (2017) sample. It is evident from the plots that the dynamics of implied dividend strip yields for both benchmarks are rudimentary: all yields move almost synchronously with no yield curve inversions occurring in this sample.

Fig. 15 directly compares forward strip equity yields implied by the two benchmarks to the most recent data from Bansal et al. (2017). The figure plots the time-series of the slope between the 7 and 1 year maturity yields. Neither of the two benchmarks is able to capture the dynamics of the slope observed in the data.

We conclude that our specification choice is key to capturing the dynamics of equity strips in the data. This is because our specification is successful in explaining both the cross-section and the time-series of US equity returns. On the contrary, SDFs implied by stylized models, such as the CAPM or the FF 5-factor model, do not feature rich enough dynamics to reconcile empirically observed patterns in equity yields.

5 Concluding Remarks

Our model effectively processes a rich information set (the time-series and cross-sectional behavior of 100 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. Similarly in spirit to the way in which the introduction of vector autoregressions (VAR) 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 structure of discount rates for different portfolios that closely resemble the actual dividend claims 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 expands the existing (observed) data along each of those dimension. The term structures

---

8This specification performs better than a specification based on the six unrotated factor returns. We, therefore, report the former specification.
Figure 14: Dynamics of benchmark-implied equity strip yields for the aggregate market for different maturities. The figure plots dynamics of yields implied by the CAPM (Panel a) and Fama-French 5-factor model supplemented with the momentum factor (Panel b) for maturities 1, 2, 5, and 7 years.
Figure 15: Benchmark-implied forward equity yields vs. forward equity yields data. We compare the slope between the 7 and 1 year maturity forward yields implied by the CAPM (Panel a) and Fama-French 5-factor model supplemented with the momentum factor (Panel b) to their empirical counterparts in Bansal et al. (2017).
we generate cover a large number of cross-sectional portfolios, in addition to the S&P 500: value, size, profitability, momentum, etc., for the total of 100 portfolios. They have a long time series, starting in 1975 and therefore covering several recessions and booms. They 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 space state 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 et al. (2019) and Haddad et al. (2019), who show that (i) an SDF constructed from dominant PCs of a large cross-section of characteristics-sorted portfolios explains the cross-section of expected returns well, and (ii) that SDF risk prices corresponding to these factors are highly predictable in the time-series by their valuation ratios and that this variation in risk prices is essential to adequately capturing 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 match well the observed ones – and gives us confidence in extending them over time, maturities, and portfolios.

We derive 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 70s) and maturities. On the sample starting in 2004 that was used in 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 in this sample slightly downward sloping, in contrast with 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 to our sample; at the same time, the Great Recession carries less overall weight in the sample. It is interesting to see that all the results of the post-2004 sample carry over to the longer sample. The term structure inverts in almost all of the additional recessions (for example, in the early 80s and 90s). The term structure of forward discount rates is still close to flat (it is mildly upward sloping on average, but it is not significantly so, and is sufficiently flat to still reject the Bansal and Yaron (2004) and Bansal et al. (2012) models in simulations).

The model generates interesting differences in the average term structure across portfolios, and in the time series. For example, we show that value stocks have a strongly increasing term structure of discount rates, whereas growth stocks have a decreasing one; small stocks have an effectively flat discount rate, whereas large stocks have a mildly increasing one. These results give us new moments that can be used to evaluate structural asset pricing models that have direct implication about the risk premia of these portfolios (as well as any of the other 100 we include in our analysis).

Finally, there are interesting patterns in the time series of slopes of the term structure of different portfolios. For example, the slopes of small and large stocks tend to move closely together; both term structures were upward sloping during the 1990s, and both were downward sloping during the Great Recession. Yet, only the term structure for small stocks inverted during the late 90s stock cycle, marking an important divergence between the two portfolios that lasted several years. On the contrary, 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 didn’t happen for growth stocks.

Overall, this set of stylized facts provide new conditional moments that asset pricing and structural macro models should seek to match. We have briefly examined some of the existing models, such as Bansal and Yaron (2004) and Bansal et al. (2012), but more work remains to be done on this front.
References


Gomes, L. and R. Ribeiro (2019). Term structure (s) of equity risk premia.


Gormsen, N. J. and E. Lazarus (2019). Duration-driven returns. Available at SSRN.


Appendix

A Derivations

A.1 Expected returns on dividend strips

Note that

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

(A1)

(A2)

(A3)

where \(a_{n,1}, d_{n,1}\) are given by iterations Eq. (25)-(26).

We can now compute

\[
\mathbb{E}_t \left[ \frac{F_{t+1}^{(n-1)} - R_{f,t}^{-1}}{P_t} \right] = \mathbb{E}_t \left[ \frac{P_{t+1}^{(n-1)} P_{t+1} e^{-r_{f,t}}}{P_{t+1} P_t} \right] = \mathbb{E}_t \left[ (\exp(a_{n-1,1} + d_{n-1,1} F_{t+1}) - \exp(a_{n-1,2} + d_{n-1,2} F_{t+1})) e^{\Delta p_{t+1} - r_{f,t}} \right] = \exp(\tilde{a}_{n,1} + \tilde{d}_{n,1} F_{t}) - \exp(\tilde{a}_{n,2} + \tilde{d}_{n,2} F_{t}),
\]

(A4)

(A5)

(A6)

which is exactly the same equation as above, except that expectation is taken under physical dynamics. The solution for \(\tilde{a}_{n,1}, \tilde{d}_{n,1}\), therefore, is given by the recursion Eq. (25)-(26), where we simply replace \(\gamma^*_0, \gamma^*_1, c^*, \rho^* \) by their physical counterparts.

Log expected excess returns on the strip are

\[
\log \left( \mathbb{E}_t \left[ P_{t+1}^{(n)} \right] \right) - r_{f,t} = \log \mathbb{E}_t \left[ \frac{P_{t+1}^{(n-1)} P_{t+1} e^{-r_{f,t}}}{P_{t+1} P_t} \right] = \log \left[ \frac{P_{t+1}^{(n-1)} P_{t+1} e^{-r_{f,t}}}{P_{t+1} P_t} \right] = \log \left[ \exp(\tilde{a}_{n,1} + \tilde{d}_{n,1} F_{t}) \right] - \log \left( \exp(\tilde{a}_{n,2} + \tilde{d}_{n,2} F_{t}) \right) - \log \left( \exp(a_{n,1} + d_{n,1} F_{t}) \right) + \log \left( \exp(a_{n,2} + d_{n,2} F_{t}) \right).
\]

(A7)

(A8)

(A9)

A.2 Equity yield decomposition into HTM returns and expected growth rates

Let \(R_{t,t+n} = \frac{D_{t+n}}{P_t} \) be the hold-to-maturity (HTM) return from \(t\) to \(t+n\). Excess return is then:

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

(A10)

Take expectations:

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

(A11)

where the numerator is just Eq. (22) computed using physical parameters, and the denominator is the same expression computed under risk-neutral dynamics.

Then, taking logs of the above and dividing through by \(n\), gives an annualized log expected HTM return.

Further, note that we can write

\[
R_{t,t+n} = \frac{D_{t+n}}{P_t^{(n)}} = \frac{D_{t}}{P_t^{(n)}} G_{t,t+n},
\]

(A12)

where \(G_{t,t+n}\) is the cumulative growth rate. We then get that

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

(A13)
Figure A.16: Average market beta of dividend strips by maturity. We plot average betas between returns on a dividend strip of any given maturity and the aggregate market index. Shaded areas depict two-standard-deviation bands around point estimates.

or

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

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

B Additional results

B.1 Properties of anomaly portfolio returns

Table A.1 shows annualized mean excess returns on the fifty anomaly long-short portfolios as well as the underlying characteristic-sorted tercile portfolios.
Table A.1: Anomaly portfolios mean excess returns, %, annualized

Columns Long and Short show mean annualized returns (in %) on each anomaly portfolio long (P3) and short ends (P1) of a sort, respectively, net of risk-free rate. The column L-S lists mean returns on the strategy which is long portfolio 3 and short portfolio 1. Portfolios include all NYSE, AMEX, and NASDAQ firms; however, the breakpoints use only NYSE firms. Monthly data from August 1975 to November 2016.

<table>
<thead>
<tr>
<th></th>
<th>Short</th>
<th>Long</th>
<th>L-S</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size</td>
<td>6.5</td>
<td>8.2</td>
<td>1.8</td>
</tr>
<tr>
<td>Value (A)</td>
<td>6.1</td>
<td>8.8</td>
<td>2.8</td>
</tr>
<tr>
<td>Gross Profitability</td>
<td>5.4</td>
<td>7.6</td>
<td>2.2</td>
</tr>
<tr>
<td>Value-Profitablity</td>
<td>4.5</td>
<td>11.3</td>
<td>6.8</td>
</tr>
<tr>
<td>F-score</td>
<td>6.1</td>
<td>7.1</td>
<td>1.0</td>
</tr>
<tr>
<td>Debt Issuance</td>
<td>6.3</td>
<td>7.9</td>
<td>1.6</td>
</tr>
<tr>
<td>Share Repurchases</td>
<td>6.1</td>
<td>7.7</td>
<td>1.6</td>
</tr>
<tr>
<td>Net Issuance (A)</td>
<td>5.6</td>
<td>8.7</td>
<td>3.1</td>
</tr>
<tr>
<td>Accruals</td>
<td>5.1</td>
<td>7.5</td>
<td>2.5</td>
</tr>
<tr>
<td>Asset Growth</td>
<td>5.9</td>
<td>8.9</td>
<td>3.0</td>
</tr>
<tr>
<td>Asset Turnover</td>
<td>5.7</td>
<td>7.8</td>
<td>2.0</td>
</tr>
<tr>
<td>Gross Margins</td>
<td>6.8</td>
<td>6.3</td>
<td>-0.4</td>
</tr>
<tr>
<td>Earnings/Price</td>
<td>4.8</td>
<td>9.6</td>
<td>4.7</td>
</tr>
<tr>
<td>Cash Flows/Price</td>
<td>5.7</td>
<td>9.7</td>
<td>4.0</td>
</tr>
<tr>
<td>Net Operating Assets</td>
<td>4.8</td>
<td>7.9</td>
<td>3.1</td>
</tr>
<tr>
<td>Investment/Assets</td>
<td>5.6</td>
<td>8.0</td>
<td>2.4</td>
</tr>
<tr>
<td>Investment/Capital</td>
<td>5.8</td>
<td>7.9</td>
<td>2.1</td>
</tr>
<tr>
<td>Investment Growth</td>
<td>5.9</td>
<td>8.5</td>
<td>2.6</td>
</tr>
<tr>
<td>Sales Growth</td>
<td>6.4</td>
<td>7.5</td>
<td>1.1</td>
</tr>
<tr>
<td>Leverage</td>
<td>5.9</td>
<td>8.0</td>
<td>2.0</td>
</tr>
<tr>
<td>Return on Assets (A)</td>
<td>6.0</td>
<td>6.8</td>
<td>0.8</td>
</tr>
<tr>
<td>Return on Book Equity (A)</td>
<td>5.5</td>
<td>6.9</td>
<td>1.4</td>
</tr>
<tr>
<td>Sales/Price</td>
<td>5.6</td>
<td>9.9</td>
<td>4.4</td>
</tr>
<tr>
<td>Growth in LTNOA</td>
<td>6.4</td>
<td>7.4</td>
<td>1.0</td>
</tr>
<tr>
<td>Momentum (6m)</td>
<td>7.9</td>
<td>7.4</td>
<td>-0.5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Short</th>
<th>Long</th>
<th>L-S</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value-Momentum</td>
<td>6.4</td>
<td>8.3</td>
<td>2.0</td>
</tr>
<tr>
<td>Value-Momentum-Prof.</td>
<td>6.3</td>
<td>9.8</td>
<td>3.5</td>
</tr>
<tr>
<td>Short Interest</td>
<td>6.6</td>
<td>5.4</td>
<td>-1.3</td>
</tr>
<tr>
<td>Momentum (12m)</td>
<td>3.6</td>
<td>9.4</td>
<td>5.8</td>
</tr>
<tr>
<td>Industry Momentum</td>
<td>5.0</td>
<td>7.9</td>
<td>2.9</td>
</tr>
<tr>
<td>Momentum-Reversals</td>
<td>6.2</td>
<td>8.5</td>
<td>2.4</td>
</tr>
<tr>
<td>Long Run Reversals</td>
<td>6.7</td>
<td>8.8</td>
<td>2.1</td>
</tr>
<tr>
<td>Value (M)</td>
<td>5.8</td>
<td>8.6</td>
<td>2.8</td>
</tr>
<tr>
<td>Net Issuance (M)</td>
<td>5.8</td>
<td>8.4</td>
<td>2.7</td>
</tr>
<tr>
<td>Earnings Surprises</td>
<td>4.6</td>
<td>8.1</td>
<td>3.5</td>
</tr>
<tr>
<td>Return on Book Equity (Q)</td>
<td>3.7</td>
<td>7.6</td>
<td>3.9</td>
</tr>
<tr>
<td>Return on Market Equity</td>
<td>3.0</td>
<td>10.8</td>
<td>7.9</td>
</tr>
<tr>
<td>Return on Assets (Q)</td>
<td>4.2</td>
<td>7.4</td>
<td>3.2</td>
</tr>
<tr>
<td>Idiosyncratic Volatility</td>
<td>4.7</td>
<td>7.2</td>
<td>2.5</td>
</tr>
<tr>
<td>Beta Arbitrage</td>
<td>4.7</td>
<td>8.1</td>
<td>3.4</td>
</tr>
<tr>
<td>Seasonality</td>
<td>5.0</td>
<td>8.6</td>
<td>3.5</td>
</tr>
<tr>
<td>Short-Term Reversals</td>
<td>5.1</td>
<td>7.7</td>
<td>2.6</td>
</tr>
<tr>
<td>Industry Rel. Reversals</td>
<td>3.7</td>
<td>10.4</td>
<td>6.8</td>
</tr>
<tr>
<td>Industry Rel. Rev. (L.V.)</td>
<td>4.0</td>
<td>11.5</td>
<td>7.5</td>
</tr>
<tr>
<td>Ind. Mom-Reversals</td>
<td>5.0</td>
<td>10.3</td>
<td>5.3</td>
</tr>
<tr>
<td>Composite Issuance</td>
<td>5.1</td>
<td>8.7</td>
<td>3.7</td>
</tr>
<tr>
<td>Price</td>
<td>6.5</td>
<td>6.5</td>
<td>0.0</td>
</tr>
<tr>
<td>Share Volume</td>
<td>5.9</td>
<td>6.8</td>
<td>0.9</td>
</tr>
<tr>
<td>Duration</td>
<td>6.0</td>
<td>8.9</td>
<td>3.0</td>
</tr>
<tr>
<td>Firm age</td>
<td>6.6</td>
<td>6.5</td>
<td>-0.1</td>
</tr>
</tbody>
</table>
Figure A.17: One-year returns of the 1- and 2-year dividend forwards in the data and in our model. We compare our model-implied returns on dividend forwards to their empirical counterparts in Bansal et al. (2017). Shaded areas depict two-standard-deviation bands around point estimates. Model parameters are estimated in full sample.