

Slicing an Asset to Learn about Its Future: A New Perspective on Return and Cash-Flow Forecasting

Ye Li*

Chen Wang[†]

June 26, 2024

Abstract

Slicing an asset by payout horizons unseals information about its future returns and cash flows. As an example, we slice an equity market index into granular pieces (dividend strips) and show that valuation ratios of its strips span the underlying state variables of the index. Strip valuation ratios form a term structure. The level and slope strongly predict the index dividends. The slope alone is sufficient for forecasting the index return. The steepening and flattening of valuation term structure reflect discount-rate variations rather than information on the cash-flow trajectory, because market participants have very limited information about long-term cash flows.

*We are grateful to Nick Barberis, Geert Bekaert, Jules van Binsbergen, Zhi Da, Kent Daniel, Carter Davis, Joost Driessen, Stefano Giglio, Will Goetzmann, Benjamin Golez, Hui Guo, Dashan Huang, Bryan Kelly, Ralph Koijen, Kai Li, Lars Lochstoer, Shane Miller, Alan Moreira, Riccardo Sabbatucci, Tano Santos, José Scheinkman, Jessica Wachter, Michael Weber, Colin Ward, and conference and seminar participants at the Columbia Business School, Federal Reserve Board, Florida State University Truist Beach Conference, GMO, HKUST Finance Symposium, Johns Hopkins University, MFA, Office of Financial Research, Paris December Finance Conference, RAPS Conference at Baha Mar, UT Dallas Finance Conference, and Yale SOM for helpful comments.

*University of Washington Foster School of Business. E-mail: liye@uw.edu

[†]University of Notre Dame Mendoza College of Business. E-mail: chen.wang@nd.edu

1 Introduction

Asset pricing theories suggest a great variety of state variables that drive asset prices through expected cash flows and returns (discount rates). These state variables reflect different parts of the economy, such as consumption, production, financial intermediation, and government policies, and cover many economic forces from behavioral biases to institutional frictions. With numerous signals available, data-driven approaches have become popular for forecasting. The state variables form a signal base and are aggregated via statistical models of nonlinear and complex mappings to the forecasting targets, such as cash flows and returns (e.g., [Kelly, Malamud, and Zhou, 2024](#)).

We propose a simple alternative that does not require the buildup of a large set of signals (“big data”) or state-of-art statistical models. Importantly, the source of predictive power is easy to interpret. We take a market-driven approach. For any asset, information relevant to its future cash flows and returns is likely to be in its market price and the prices of its payout at different horizons. Therefore, slicing the asset by payout horizons unlocks the information we need for forecasting.

Consider the S&P 500 index.¹ It consists of dividend strips at different maturities ([Binsbergen, Brandt, and Koijen, 2012](#)). In an exponential-affine model that generalizes the setup in [Lettau and Wachter \(2007\)](#), we show that the logarithm of strip prices scaled by realized dividend are linear functions of state variables. As strips with different maturities differ in their state-variable loadings, their valuation ratios are linearly independent combinations of state variables and map out the state space as long as there are as many strips as the state variables. Thus, by slicing the asset along payout horizons into sufficiently granular pieces, we obtain strip valuation ratios that proxy for the state variables and reveal information needed for forecasting return and dividends of the asset.

Our paper provides four findings. First, by empirically analyzing the valuation ratios of strips across maturities, we find the state space of the equity index has a low dimensionality. These strip valuation ratios form a term structure. The level and slope span all the valuation ratios, and for forecasting dividend growth and return of the index, we only need these two state-variable proxies.²

¹We focus on the S&P 500 as its derivatives (futures), which are used to construct strip prices, have liquid markets. In the appendix, we show that our results hold for the Fama-French market portfolio (CRSP value-weighted index).

²The first two principal components of valuation ratios account for 96.3% of total variance, which indicates two

The level is the log price-dividend ratio of the asset itself (pd). The slope is the difference between pd and valuation ratio of one-year strip. While there are many economic forces that affect the stock market index, the level and slope of its valuation term structure efficiently summarize these forces.

Second, the slope alone is sufficient for forecasting return. It delivers an in-sample R^2 of 24.8% and an out-of-sample R^2 of 14.6% and subsumes the predictive power of the level (pd). Augmenting the slope with other predictors from the literature does not improve the forecasting performances.³ What motivates our analysis is the linear mapping between strip valuation ratios and state variables in an exponential affine model.⁴ To address the concern of nonlinearity, we show that the slope outperforms machine learning algorithms that aggregate many predictors nonlinearly (Kelly, Malamud, and Zhou, 2024).⁵ Many mechanisms have been proposed to explain the dynamics of expected return. The slope summarizes these forces. The steepening (flattening) of valuation term structure predicts negative (positive) returns. The expected return is highly volatile: a decrease of the slope by one standard deviation adds 7.7% to the expected annual return.

Third, the return predictive power of the slope has a simple interpretation: market participants have very limited information on cash-flow growth beyond the very next year; therefore, when the valuation term structure steepens, it is not driven by an improving expectation of long-run growth but due to a lower discount rate that benefits the valuation of long-duration cash flows more than that of near-term cash flows. Specifically, we find that forecasting cash flows over the next year can be done fairly accurately using our state variable proxies (the level and slope) or survey expectations (analyst forecasts); however, these variables cannot forecast cash flows beyond one year. Both theoretically and empirically, we demonstrate a tight connection between the return predictor power of the slope of valuation term structure and market participants' lack of information on long-term growth.

Lastly, we find the role of traditional price-dividend ratio (pd)—the level of valuation term state variables are sufficient, such as the level and slope. When forecasting returns and dividend growth, the level and slope perform as well as any pair of valuation ratios or combinations of three or more valuation ratios.

³The slope outperforms other predictors, including those summarized in Goyal and Welch (2007) and in recently published papers, across evaluation metrics, such as Hodrick (1992) adjustment for standard errors, Stambaugh (1999) adjustment for small-sample bias, and out-of-sample tests such as encompassing (ENC) and Clark-West (CW) tests.

⁴The linear mapping can also be obtained via log-linearization (e.g., Binsbergen and Koijen, 2010).

⁵Kelly, Malamud, and Zhou (2024) consider the predictors in Goyal and Welch (2007). To enhance the predictive power of their models, we include more return predictors from the more recent literature.

structure—in spanning the state space is not to predict returns (as the slope subsumes its return predictive power) but to augment the slope in forecasting near-term (annual) dividends. Both the level and slope predict dividend growth poorly on a standalone basis, but together, they deliver an in-sample R^2 of 38.6% and an out-of-sample R^2 of 31.9%. Expected return and cash-flow news can be correlated (Lettau and Ludvigson, 2005; Kothari, Lewellen, and Warner, 2006). Therefore, when forecasting dividend growth, it is important to control for variation in the expected return (the slope). The literature has coalesced around using pd to predict returns. For this reason, it has always been included as a state variable in asset pricing and macro-finance studies. However, our findings show that the slope of valuation term structure is actually the state variable that corresponds to the expected return, while the level forecasts dividend growth when we control for the slope.

Overall, our paper illustrates a new method of identifying and combining information for forecasting returns and cash flows of financial assets. First, we slice an asset along payout horizons and compute valuation ratios of its cash-flow strips. Next, we examine the strip valuation ratios that proxy for state variables to determine the state space dimensionality. For a two-dimensional state space, the level and slope contain sufficient information for forecasting returns and cash flows. In particular, the slope of an asset’s valuation term structure is a powerful return predictor. For the S&P 500 index, the slope is calculated as the difference between its own price-dividend ratio and valuation ratio of the one-year dividend strip. The one-year horizon is chosen because we find market participants are not informed about cash flows beyond the next year. For a different asset, a similar analysis can be performed to identify the cash-flow “information cliff.”

Our paper shares with the bond literature the use of valuation ratios (i.e., yields for bonds) as state-variable proxies (e.g., Litterman and Scheinkman, 1991; Duffie and Kan, 1996; Dai and Singleton, 2000; Duffee, 2002) but offers a different approach to return and cash-flow prediction. As pointed out by Kelly, Malamud, and Pedersen (2023), traditionally, an asset’s own valuation ratio is used to predict its return and cash flows. Bond yields are used to predict returns of the same bonds (e.g., Campbell and Shiller, 1991). The price-dividend ratio and other valuation ratios of a stock market index are used to predict the index return and cash flows (e.g., Campbell and Shiller,

1988; Fama and French, 1988a; Lewellen, 2004; Cochrane, 2011), and the same has been done for individual stocks (e.g., Vuolteenaho, 2002). More recently, valuation ratios of index dividend strips have been used to predict returns of these strips (e.g., Binsbergen et al., 2013). Kelly, Malamud, and Pedersen (2023) propose using valuation ratios of multiple assets to jointly predict their returns.

Our paper challenges this traditional approach by showing that an asset's own valuation ratio is not the optimal predictor for returns or cash flows; instead, one should slice the asset into strips by payout horizons and explore information in its valuation term structure. For stock market indices, we use the slope to predict index returns rather than returns of dividend strips, and we use both the level and slope to predict cash flows.⁶ Our method cannot be applied to single-payoff assets, such as zero-coupon bonds and dividend strips (the “base assets”). In our method, their valuation ratios predict returns and cash flows of the assets they aggregate up to, not those of themselves.

We emphasize that an asset has many valuation ratios, its own valuation ratio and those of its strips. Kelly and Pruitt (2013) view an equity index as the sum of stocks rather than its own strips. Using partial least squares, they extract signals on index returns and cash flows from firm-level valuation ratios. Our approach differs in both the signal base and signal aggregation. In contrast to firm-level valuation ratios, valuation ratios of the strips only contain information about the index they aggregate up to and do not contain firm-level noise. Thus, they directly map out state variables for the index. Moreover, while the method in Kelly and Pruitt (2013) applies to indices that consist of assets with individual market valuations, our method applies to any asset whose strip prices can be obtained from the market or indirectly from its derivatives. Second, we do not rely on statistical models to aggregate signals. After determining the dimensionality of strip valuation ratios, we form state-variable proxies (the level and slope of valuation term structure) that can be measured in real time from market prices, less prone to estimation error. These variables forecast returns and cash flows. For return prediction, we further consolidate signals: after identifying the cash-flow information cliff, we compute a univariate predictor, the slope around the (one-year) cutoff.

Identifying the cash-flow information cliff is a critical step in our method. Next, we provide

⁶For bonds, the analogy is to compute the slope of term structure of bond yields and use the slope factor to predict returns of the aggregate bond market portfolio rather than returns of the individual bonds.

more details on the findings that market participants are not well informed about cash-flow growth beyond the very next year. To derive a test for this hypothesis, we set up a two-dimensional state space model, motivated by our findings on the state space dimensionality. The two state variables can be rotated to represent the conditional expectations of annual return and of annual dividend growth rate that both follow AR(1) processes (Lettau and Wachter, 2007; Binsbergen and Koijen, 2010). In our model, the level and slope of valuation term structure are linear combinations of these state variables, and vice versa, as we have found empirically. At any time t , information about expected future cash flows is embedded in the current expectation of dividend growth over the next year. Given that the expected dividend growth rate follows an AR(1) process, having the information cliff at the one-year horizon is equivalent to having a zero autoregressive coefficient.

To estimate this autoregressive coefficient (persistence) of expected dividend growth, we use analyst forecasts to proxy for cash-flow growth expectations.⁷ The results across empirical specifications consistently show that growth expectation lacks persistence in line with De La O and Myers (2021).⁸ We also prove analytically that forecasting errors from using the slope as a return predictor should comove with value of the autoregressive coefficient. To test this prediction of our model, we conduct a rolling-window estimation. In each window, we estimate the autoregressive coefficient of expected cash-flow growth and compute the return forecasting error. Both the in- and out-of-sample forecasting errors comove with the value of the autoregressive coefficient.

In our model, having a zero autoregressive coefficient of the expected cash-flow growth rate is equivalent to the fact that market participants are not informed about growth beyond the next year. This equivalence condition allows for a sharp statistical test, and our model setup is grounded by our empirical analysis of the state space. Next, we step outside our model and provide further evidence on the cash-flow information cliff at a one-year horizon by directly examining market

⁷For robustness, we fit a state-space model to dividend data to estimate the persistence of expected dividend growth and find results consistent with the estimation based analyst forecasts: the autoregressive coefficient is close to zero.

⁸When the autoregressive coefficient is zero, agents' belief on cash-flow dynamics in our model is the same as that in De La O and Myers (2021). Our paper differs in that agents' cash-flow belief is not the focus but guides the construction of our return predictor and that we jointly model cash-flow belief and agents' price of risk following Lettau and Van Nieuwerburgh (2007). Moreover, we show theoretically and empirically that the estimated persistence of expected cash-flow growth is closely associated with return forecasting errors of the slope of valuation term structure.

participants' cash-flow expectations. We show that earnings and dividend growth within one year are highly predictable, by both our state variables (the level and slope) and by analysts' forecasts. The R^2 is 73% (45%) from forecasting earnings (dividend) growth over the next year. In contrast, cash-flow predictability beyond one year is weak. For earnings growth from the next one to two years, the R^2 drops to 8%. It further declines to 7% for growth between the second and third years.

Beyond discount-rate variation, return predictability often has an alternative interpretation based on mispricing. Consider a market timing strategy for the S&P 500 index that we call "betting against the slope". The slope has a negative coefficient for predicting return, so the strategy reduces exposure when the valuation term structure steepens and increases exposure when the term structure flattens. Its Sharpe ratio of 0.58 is 55% higher than that of buy-and-hold strategy (Campbell and Thompson, 2008). Traditionally, market timing strategies bet against the level: it scales up exposure when the price-dividend ratio is low and scales down when pd is high, betting against overpricing relative to fundamentals (Lewellen, 2004). Our findings indicate that market participants are well informed of near-term cash flows, so mispricing is likely in the valuation of long-term cash flows. Betting against the slope is betting on market participants' lack of information on long-term cash flows. When the valuation of long-term cash flows rises relative to that of near-term cash flows, it signals exuberance about long-term growth; it falls when market participants are pessimistic.

Literature. Duffie and Kan (1996) point out that state variables of the bond market can be linearly mapped to zero-coupon bond yields. This observation is critical for estimating term structure models (Duffee, 2013). The equity counterparts of zero-coupon bonds are dividend strips. In theory, strip valuation ratios proxy for state variables of the equity market (e.g., Lettau and Wachter, 2007).⁹

Several papers applied this insight empirically. Binsbergen et al. (2013) use valuation ratios of dividend strips to forecast strip returns, dividends, and macroeconomic variables.¹⁰ Different

⁹The bond literature widely applied affine models of stochastic discount factors (e.g., Duffie and Kan, 1996; Dai and Singleton, 2000; Duffie, Pan, and Singleton, 2000; Duffee, 2002; Ang and Piazzesi, 2003; Bikbov and Chernov, 2010). Dividend processes are added to build no-arbitrage equity models that are more flexible than fully specified equilibrium models (e.g., Bekaert and Grenadier, 1999; Pan, 2002; Brennan, Wang, and Xia, 2004; Eraker, 2008; Koijen, Lustig, and Van Nieuwerburgh, 2015; Backus, Boyarchenko, and Chernov, 2018; Kragt, de Jong, and Driessen, 2020).

¹⁰Binsbergen et al. (2013) also model the market betas of dividend strips as functions of their valuation ratios.

from [Binsbergen et al. \(2013\)](#), our focus is on the asset that strips aggregate up to (equity index) rather than strips themselves. Our goal is to illustrate a method of return and cash-flow forecasting based on slicing assets into strips, and we provide several findings. State space for the equity index is two-dimensional.¹¹ The level and slope of valuation term structure predict index returns and cash flows. For return prediction, the striking difference in near- and long-term cash-flow predictability leads to a particular combination of strip valuation ratios (the slope) as the composite signal.

Using bond yields and equity market price-dividend ratio as rotated state variables, [Cieslak and Pang \(2021\)](#) extract state-variable shocks, such as monetary, growth, and risk premium news. Our paper focuses on forecasting and stays agnostic about the interpretation of state variables.¹² Return and cash-flow forecasting often relies on theoretically motivated state variables. Our approach follows the other tradition of latent state modeling ([Cochrane, 2008a](#); [Binsbergen and Koijen, 2010](#); [Kelly and Pruitt, 2013](#); [Rytchkov, 2012](#)). Our contribution is to use the observable strip valuation ratios to map out the latent state variables rather than to rely on statistical filtering.

Strip valuation ratios based on market prices are powerful “summary statistics” of various economic forces (state variables) from macroeconomic dynamics to belief distortions. Our exercises rely on these market prices, and for the S&P 500 strips, we impute their prices from prices of S&P 500 index futures, which are among the most actively traded futures. [Giglio, Kelly, and Kozak \(forthcoming\)](#) analyze the dual problem, that is to calculate strip prices from empirically specified and observed dynamics of state variable when market data on strip prices are unavailable.¹³

Our paper does not study the term structure of equity risk premium (the difference in average returns between short- and long-horizon dividend strips), which has attracted arguably most attention among studies on dividend strips.¹⁴ We only use dividend strip prices for information on state

¹¹Our findings confirm the two-dimensional setup of state space models of the aggregate stock market (e.g., [Lettau and Wachter, 2007](#); [Cochrane, 2008a](#); [Binsbergen and Koijen, 2010](#); [Rytchkov, 2012](#); [Kragt et al., 2020](#)). Low dimensionality is also found in bond markets (e.g., [Litterman and Scheinkman, 1991](#); [Dai and Singleton, 2000](#)) and in the cross-section of equity factors (e.g., [Feng et al., 2020](#); [Kozak et al., 2020](#)).

¹²While we do not take a stand on the interpretation of state variables, we report the results from using the level and slope to predict macroeconomic variables in Table A.4 the online appendix.

¹³A large literature discusses how to measure strip prices from market data (e.g., [Binsbergen et al., 2012](#); [Binsbergen and Koijen, 2017](#); [Cejnek and Randl, 2016, 2020](#); [Cejnek et al., 2021](#); [Gormsen and Lazarus, 2023](#); [Golez and Jackwerth, 2024](#)) and the associated challenges ([Schulz, 2016](#); [Song, 2016](#); [Boguth et al., 2022](#)).

¹⁴There is an extensive literature on the term structure of equity risk premium (e.g., [Lettau and Wachter, 2007](#);

variables of the asset they add up to. The term structure that is relevant for our analysis is the one of strip valuation ratios rather than the term structure of strip average returns.

Traditionally, an asset's own valuation ratio is used to predict its returns. We demonstrate that an asset has many valuation ratios, its own valuation ratio and those of its strips that form a valuation term structure. Importantly, it is the slope rather than an asset's own valuation ratio, i.e., the level of valuation term structure, that predicts its return. Our findings of the strong return predictability of equity market index contributes to the voluminous literature on return predictability, and we compare thoroughly the performance of our return predictor against those in the literature.¹⁵ The construction of our predictor is simple, and it outperforms the other predictors across various metrics. Moreover, from both theoretical and empirical perspectives, we provide an explanation of the connection between the slope of valuation term structure and the conditional expected return. Our explanation is based on the fact that market participants lack information on long-term growth.

In the bond literature, two issues have been raised regarding the use of yields (valuation ratios for bonds) to predict returns: nonlinear relationship between valuation ratios and state variables and unspanned state variables.¹⁶ To address the first concern, we show that our predictor, the slope, outperforms the machine learning algorithms in [Kelly, Malamud, and Zhou \(2024\)](#) that nonlinearly aggregate many predictors. To address the second concern, we run spanning tests: when forecasting returns, augmenting the slope with other predictors does not reliably improve the performance.

Using strip valuation ratios to map out state variables has broad applications beyond what

[Hansen et al., 2008](#); [Lettau and Wachter, 2011](#); [Binsbergen et al., 2013](#); [Belo et al., 2015](#); [Hasler and Marfè, 2016](#); [Ai et al., 2018](#); [Backus et al., 2018](#); [Miller, 2018](#); [Bansal et al., 2021](#); [Gonçalves, 2021](#); [Gormsen, 2021](#); [Boguth et al., 2022](#); [Hasler and Khapko, 2023](#)). The difference in average returns of short- and long-term dividend strips led to decomposing returns of the market and investment strategies to the short-duration or long-duration component ([Gonçalves, 2019](#); [Gormsen and Koijen, 2020](#); [Binsbergen, 2021](#); [Knox and Vissing-Jørgensen, 2022](#)).

¹⁵Our approach to return prediction is based on slicing assets into strips along payout horizons and adds the growing literature on return predictability (e.g., [Fama and French, 1988b](#); [Campbell and Shiller, 1988](#); [Baker and Wurgler, 2000](#); [Lettau and Ludvigson, 2001](#); [Lewellen, 2004](#); [Cochrane, 2008b](#); [Ang and Bekaert, 2007](#); [Goyal and Welch, 2007](#); [Lettau and Van Nieuwerburgh, 2007](#); [Campbell and Thompson, 2008](#); [Rapach et al., 2010](#); [Kelly and Pruitt, 2013](#); [Rapach et al., 2013](#); [Golez, 2014](#); [Rapach et al., 2016](#); [Martin, 2017](#); [Golez and Koudijs, 2018](#); [Johnson, 2019](#); [Kozak and Santosh, 2020](#); [Chen et al., 2022](#); [Hillenbrand and McCarthy, 2022](#); [Kelly et al., 2024](#); [Bordalo et al., 2024](#)).

¹⁶The literature on unspanned factors emphasize stochastic volatility and macro factors (e.g., [Collin-Dufresne and Goldstein, 2002](#); [Li and Zhao, 2006](#); [Collin-Dufresne et al., 2008](#); [Cooper and Priestley, 2008](#); [Bikbov and Chernov, 2009](#); [Collin-Dufresne et al., 2009](#); [Ludvigson and Ng, 2009](#); [Andersen and Benzoni, 2010](#); [Duffee, 2011](#); [Joslin et al., 2014](#); [Cieslak and Povala, 2015, 2016](#); [Bauer and Rudebusch, 2016](#); [Feldhütter et al., 2016](#)).

we have done in this paper. Forecasting return and cash flows are critical for understanding what drives asset prices (Cochrane, 2008b, 2011; Kojien and Van Nieuwerburgh, 2011; Pruitt, 2023). Identifying state variables is an important task in macro finance as it lays the foundation for various research topics, for example, vector autoregression models (e.g., Sims, 1980; Campbell and Ammer, 1993; Patelis, 1997; Bernanke and Kuttner, 2005; Larrain and Yogo, 2008; Cieslak and Pang, 2021).

Finally, our paper makes contributions to the literature on equity cash-flow predictability.¹⁷ Characterizing the dynamics of cash-flow expectations is at the center of asset pricing literature (e.g., Bansal and Yaron, 2004; Beeler and Campbell, 2012; Belo, Collin-Dufresne, and Goldstein, 2015; Collin-Dufresne, Johannes, and Lochstoer, 2016). Our state-variable proxies and analyst forecasts strongly predict near-term cash flows (dividends and earnings) but do not predict cash flows beyond one year, suggesting that the steepening of valuation term structure is not driven by improving expectation of long-term growth but due to a lower discount rate or errors in long-term growth expectation. The mispricing interpretation is in line with Bordalo et al. (2024) and similar insights found in the cross-section of stocks (Da and Warachka, 2011).¹⁸ In our state space model, the lack of information on long-term cash-flow growth translates into a zero autoregressive coefficient of agents' expectation of cash-flow growth, which leads to a model of belief formation in line with De La O and Myers (2021). Our findings echo recent studies on the importance of agents' perceived persistence of state variables (Gabaix, 2019; Wang, 2020). Our use of analyst forecasts to proxy for agents' expectations follows a growing body of research on analyzing subjective expectations based on survey data (see reviews by Adam and Nagel, 2023; D'Acunto and Weber, 2024).¹⁹

¹⁷The lack of cash-flow predictability motivates using valuation ratios to predict returns Cochrane (2011). The recent literature has found more evidence of cash-flow predictability (Binsbergen and Kojien, 2010; Lacerda and Santa-Clara, 2010; Kojien and Van Nieuwerburgh, 2011; Binsbergen et al., 2013; Chen et al., 2013; Golez, 2014; Jagannathan and Liu, 2018; Pettenuzzo et al., 2020; Gao and Martin, 2021; Golez and Koudijs, 2023; Sabbatucci, 2022; Pruitt, 2023).

¹⁸Errors in inflation expectation also contribute to mispricing of long-duration claims (De la O and Myers, 2024).

¹⁹Studies analyze firm-level cash-flow expectations (La Porta, 1996; Dechow and Sloan, 1997; Copeland et al., 2004; Da and Warachka, 2011; Piotroski and So, 2012; Bordalo et al., 2019; Bouchaud et al., 2019; Binsbergen et al., 2022), expectations of index cash flows (Chen et al., 2013; Gao and Martin, 2021; McCarthy and Hillenbrand, 2021; Nagel and Xu, 2022; Charles et al., 2023; Schmidt-Engelbertz and Vasudevan, 2023; De la O and Myers, 2024), and expectations in bond markets and macroeconomy (Amromin and Sharpe, 2014; Coibion and Gorodnichenko, 2015; Piazzesi et al., 2015; Crump et al., 2016; Bordalo et al., 2020; Giglio et al., 2021; Pang, 2023; Farmer et al., forthcoming).

2 Valuation Ratios and the State Space

We set up a state space model to illustrate the idea that the level and slope of equity valuation term structure are linear combinations of the underlying state variables. Given a sufficiently granular set of strips with different maturities, their valuation ratios fully span the state space.

We consider a dynamic economy where the information filtration is driven by a Markov process. Specifically, the state of an economy at time t is summarized by X_t , a K -by-1 vector of state variables. We assume that the law of motion of X_t is given by a first-order vector autoregression

$$X_{t+1} = \Pi X_t + \sigma_X^\top \epsilon_{t+1}, \quad (1)$$

where ϵ_{t+1} is a N -by-1 vector of shocks that capture all the news at $t + 1$ and are independent over time with normal distribution $N(\mathbf{0}, \Sigma)$. Note that since any higher-order vector autoregression can be written as a first-order vector autoregression by expanding the number of state variables, the AR(1) specification is without loss of generality. The autoregressive coefficients are given by Π , a constant K -by- K matrix, and σ_X is a N -by- K matrix of shock loadings.

The growth rate of dividend from t to $t + 1$ has a N -by-1 shock-loading vector σ_D ,

$$\ln \left(\frac{D_{t+1}}{D_t} \right) = g_t + \sigma_D^\top \epsilon_{t+1}, \quad (2)$$

where the time-varying expected dividend growth rate is given by

$$g_t = \phi^\top X_t + \bar{g} - \frac{1}{2} \sigma_D^\top \Sigma \sigma_D. \quad (3)$$

We allow the state-variable loadings, ϕ , to be any K -by-1 vector.

No arbitrage condition implies the existence of a stochastic discount factor

$$M_{t+1} = \exp \left\{ -r_f - \frac{1}{2} \lambda_t^\top \Sigma \lambda_t - \lambda_t^\top \epsilon_{t+1} \right\}, \quad (4)$$

where r_f is the one-period risk-free rate and the N -by-1 vector of risk prices, λ_t , is given by

$$\lambda_t = \bar{\lambda} + \theta^\top X_t. \quad (5)$$

We do not impose any restrictions on θ , the state-variable loadings of the risk prices, λ_t . As pointed

out by [Kozak, Nagel, and Santosh \(2018\)](#), price of risk can also be interpreted as belief distortion.

Let P_t^n denote the time- t price of the dividend paid at $t+n$. The no-arbitrage pricing functional gives a recursive equation for the prices of dividend strips: for $n \geq 1$,

$$P_t^n = \mathbb{E}_t \left[M_{t+1} P_{t+1}^{n-1} \right], \quad (6)$$

with the boundary condition $P_t^0 = D_t$. The log price-dividend ratio of the dividend strip with maturity n is given by

$$s_t^n \equiv \ln \left(\frac{P_t^n}{D_t} \right) = A(n) + B(n)^\top X_t, \quad (7)$$

where $A(n)$ and $B(n)$ are deterministic functions of n given by a system of recursive equations (A.4)-(A.5) in the appendix with the initial conditions $A(0) = 0$, and $B(0) = 0$.

Given K log price-dividend ratios of strips, $\{s_t^{n_i}\}_{i=0}^K$, with a full-rank loading matrix, $\mathbf{B} \left(\{n_i\}_{i=1}^K \right) \equiv [B(n_1), B(n_2), \dots, B(n_K)]^\top$, the state space is recovered by

$$X_t = \mathbf{B} \left(\{n_i\}_{i=1}^K \right)^{-1} \left[s_t^{n_1} - A(n_1), \dots, s_t^{n_K} - A(n_K) \right]^\top \quad (8)$$

When the rank condition fails, these valuation ratios can still recover part of the state space. Let J ($< K$) denote the maximum number of log price-dividend ratios with linearly independent loadings $B(n)$ and $\{n_i\}_{i=1}^J$ denote the corresponding set of maturities. We can write (8) as follows

$$\mathbf{B} \left(\{n_i\}_{i=1}^J \right) X_t = \left[s_t^{n_1} - A(n_1), \dots, s_t^{n_J} - A(n_J) \right]^\top. \quad (9)$$

The market of strips acts as a linear mapping, i.e., $\mathbf{B} \left(\{n_i\}_{i=1}^J \right)$, that compresses the K -dimensional state space of X_t into a J -dimensional space generated by the log price-dividend ratios. In sum, a collection of log price-dividend ratios of dividend strips (partially) span the state space. The revealed (rotated) state variables may represent different forces in the economy, including various frictions and belief distortions. Empirically, these valuation ratios proxy for the state variables.

Next, we introduce the level and slope of equity valuation term structure. Let P_t denote the total stock market capitalization (i.e., the market price of dividends across all maturities). In

Appendix I, we solve the log price-dividend ratio of the market (the level of valuation term structure)

$$pd_t = \ln(P_t/D_t) = A + B^\top X_t, \quad (10)$$

where A and B are constants defined in Appendix I. The slope, denoted by dr_t , is defined as the difference between pd_t and the log ratio of one-year dividends to the realized dividend:

$$dr_t \equiv \ln(P_t/D_t) - \ln(P_t^1/D_t) = pd_t - s_t^1 = A - A(1) + (B - B(1))^\top X_t. \quad (11)$$

As will be made clear, we label the slope as “ dr_t ” as it predicts returns and thus reflects the “discount rate”. The level and slope of valuation term structure are linear combinations of state variables X_t . By having different coefficients of X_t , they reveal different information about the state space.

3 An Empirical Analysis of the State Space

3.1 Variable construction and summary statistics

Dividend strip prices. Let P_t^n denote the price of the dividend paid in year n . First, we calculate P_t^{n+} , the price of dividends that are paid after the first n years. Under the risk-neutral measure,

$$P_t^{n+} = e^{-nr_f} \mathbb{E}_t^{RN} \left[\sum_{\tau=1}^{+\infty} e^{-\tau r_f} D_{t+n+\tau} \right] = e^{-nr_f} \mathbb{E}_t^{RN} \left[\mathbb{E}_{t+n}^{RN} \left[\sum_{\tau=1}^{+\infty} e^{-\tau r_f} D_{t+n+\tau} \right] \right], \quad (12)$$

where the expectation operator, $\mathbb{E}_{t+n}^{RN}[\cdot]$, was inserted under the law of iterated expectations. Note that the (ex-dividend) stock price at $t+n$ is

$$P_{t+n} = \mathbb{E}_{t+n}^{RN} \left[\sum_{\tau=1}^{+\infty} e^{-\tau r_f} D_{t+n+\tau} \right], \quad (13)$$

so we have

$$P_t^{n+} = e^{-nr_f} \mathbb{E}_t^{RN} [P_{t+n}]. \quad (14)$$

The first component, e^{-nr_f} , is ZCB_t^n , the price of a zero-coupon bond with maturity n . The second component is the risk-neutral expectation of stock price, i.e., the futures price, F_t^n (Duffie, 2001).

We construct the price of dividend strips using zero-coupon bond prices and equity index

futures prices. First, we calculate P_t^1 , the price of the dividend paid in the next year,

$$P_t^1 = P_t - P_t^{1+}, \quad (15)$$

as the difference between the price of all dividends, P_t (i.e., stock price), and the price of dividends paid after one year. Following the same method, we calculate the price of dividends paid in the next six months, $P_t^{0.5}$ from $P_t - P_t^{0.5+}$. In our empirical analysis, we use the valuation ratios of dividend strips with maturity 1 and 0.5, i.e., $s_t^1 = \ln(P^1/D_t)$ and $s_t^{0.5} = \ln(P^{0.5}/D_t)$, and the valuation ratio of dividends paid beyond one year, $s_t^{1+} = \ln(P^{1+}/D_t)$.²⁰ Our analysis in the previous section shows that these valuation ratios are different linear combinations of state variables.

Data and sample. For futures prices, we use S&P 500 index futures, which are the most actively traded equity futures. The futures prices are from Datastream.²¹ The zero-coupon bond prices are from the Fama-Bliss database. The return and level of the S&P 500 index are obtained from CRSP. The dividend data is from S&P Global and obtained from the updated dataset of [Goyal and Welch \(2007\)](#). Our sample starts in January 1988 for high-quality dividend data and, importantly, a sufficiently liquid futures market without structural changes.²² After the market crash of October 1987, regulators overhauled several trade-clearing protocols.²³ Our sample ends in December 2019. Lastly, Fama-French factors at the monthly frequency are obtained from Ken French’s website.

The level of valuation term structure, pd_t , is the traditional price-dividend ratio, i.e., the logarithm of the ratio of S&P 500 market capitalization to realized dividend in the last year. The slope, dr_t , is calculated as pd_t minus s_t^1 , the logarithm of the ratio of one-year dividend strip

²⁰There is no collinearity: $s_t^{1+} + s_t^1$ is the sum of two ratios in logarithms, $\ln(P^1/D_t) + \ln(P^{1+}/D_t)$, which is not pd_t .

²¹We obtain the daily settlement prices for the S&P 500 futures. For return and cash-flow prediction at the monthly frequency, we use the settlement price of the last trading day of each month. The maturities of the traded futures contracts vary over time, so to obtain futures prices with constant maturity, we apply the shape-preserving piecewise cubic interpolation to complete the futures curve. The results using linear interpolation are similar.

²²[Wang, Michalski, Jordan, and Moriarty \(1994\)](#) identify structural changes of liquidity in the S&P 500 futures market in the pre-1987 period, during the market crash, and in the post-1987 period.

²³The stock market crash in October 1987 reveals anomalous trading in the futures market that was primarily driven by portfolio insurance ([Brady Report \(1988\)](#)). According to the New York Stock Exchange: “In response to the market breaks in October 1987 and October 1989, the New York Stock Exchange instituted circuit breakers to reduce volatility and promote investor confidence. By implementing a pause in trading, investors are given time to assimilate incoming information and the ability to make informed choices during periods of high market volatility.”

Table 1 Summary Statistics

This table reports the number of observations, mean, standard deviation, minimum, maximum, quartiles, and monthly autocorrelation (ρ) of the main variables in this paper, including our main return predictor, dr (“slope”); the price-dividend ratio pd of the S&P 500 index; the filtered series for demeaned expected returns and dividend growth following [Binsbergen and Koijen \(2010\)](#) μ^F and g^F ; the single predictive factors extracted from 100 book-to-market and size portfolios from [Kelly and Pruitt \(2013\)](#) for return and dividend growth, respectively, KP and KP^{CF} ; short-term dividend strip price to dividend ratio (0.5 year and 1 year) $s^{0.5} = \log(P^{0.5}/D)$ and $s^1 = \log(P^1/D)$; long-term dividend strip price to dividend ratio (beyond 1 year) $s^{1+} = \log(P^{1+}/D)$; one-month and one-year log returns of the S&P 500 index $r_{t+1/12}$ and r_{t+1} ; one-month and one-year log market returns from Fama-French market portfolio $r_{t+1/12}^{MKT}$ and r_{t+1}^{MKT} ; and the 1-year dividend growth rate of S&P 500 index and the Fama-French market portfolio $\Delta d_{t+1} = \log(D_{t+1}/D_t)$ and $\Delta d_{t+1}^{MKT} = \log(D_{t+1}^{MKT}/D_t^{MKT})$. Our sample is monthly observations 1988:01–2019:12.

	mean	std	min	25%	50%	75%	max	ρ
dr_t	4.027	0.494	2.952	3.727	4.044	4.208	6.632	0.919
pd_t	3.883	0.289	3.239	3.656	3.930	4.047	4.524	0.985
μ_t^F	-0.039	0.024	-0.091	-0.051	-0.041	-0.024	0.010	0.991
KP_t	-0.504	0.073	-0.725	-0.562	-0.482	-0.450	-0.378	0.955
$s_t^{0.5}$	-0.819	0.281	-2.629	-0.883	-0.768	-0.666	-0.280	0.604
s_t^1	-0.142	0.280	-2.241	-0.210	-0.098	0.016	0.393	0.766
s_t^{1+}	3.863	0.297	3.204	3.629	3.913	4.030	4.521	0.985
$r_{t+1/12}$	0.009	0.041	-0.184	-0.015	0.013	0.034	0.108	0.022
r_{t+1}	0.095	0.157	-0.568	0.046	0.126	0.187	0.429	0.929
$r_{t+1/12}^{MKT}$	0.009	0.042	-0.187	-0.016	0.014	0.036	0.108	0.051
r_{t+1}^{MKT}	0.096	0.159	-0.554	0.036	0.128	0.194	0.440	0.924
Δd_{t+1}	0.059	0.070	-0.237	0.025	0.068	0.112	0.168	0.994
Δd_{t+1}^{MKT}	0.058	0.081	-0.207	0.018	0.051	0.107	0.262	0.962
KP_t^{CF}	-0.385	0.068	-0.605	-0.422	-0.389	-0.338	-0.220	0.953
g_t^F	0.019	0.059	-0.233	-0.002	0.031	0.056	0.132	0.939

price to the realized dividend. Table 1 reports the summary statistics of pd_t , dr_t , and valuation ratios of the dividend strips ($s_t^{0.5}$, s_t^1 , and s_t^{1+}), the monthly return of S&P 500 ($r_{t+1/12}$ where in the subscript 1/12 denotes one month or 1/12 of a year), the annual return of S&P 500 (r_{t+1}), and for comparison, the monthly and annual returns of the Fama-French market portfolio (MKT) ($r_{t+1/12}^{MKT}$ and r_{t+1}^{MKT}). Our sample includes monthly observations until 2019, i.e., before the market turmoil during the Covid-19 pandemic. Our baseline analysis focuses on the returns and dividends of the S&P 500 index because we construct the strip prices using the S&P 500 futures data.²⁴ For robustness, we also report results using Fama-French market portfolio returns and dividends. In

²⁴Previous studies of return predictability (e.g., [Ang and Bekaert, 2007](#)) also use S&P 500 Index as a market proxy.

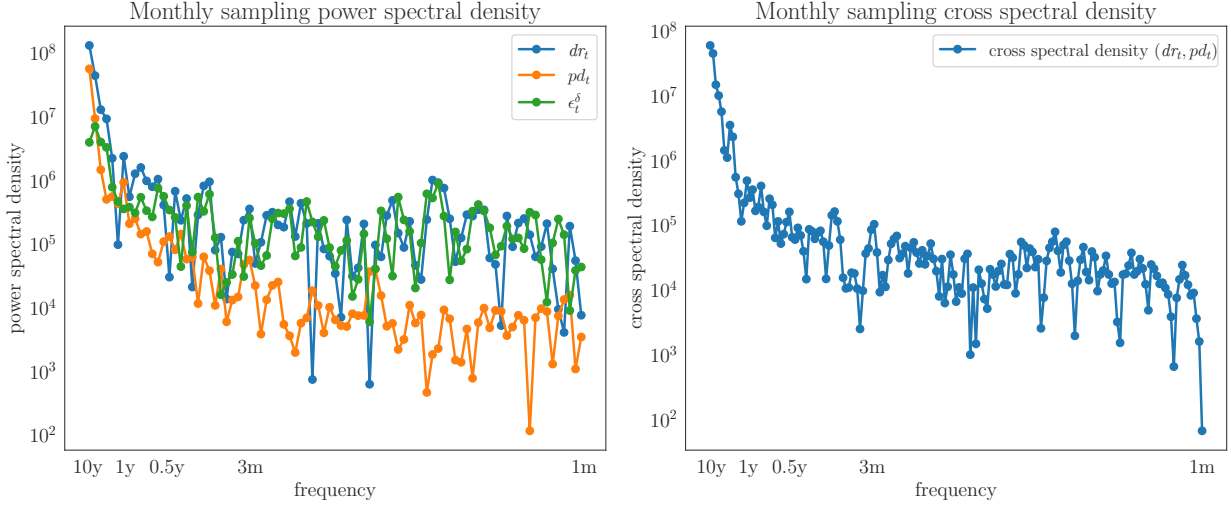


Figure 1 Spectrum and Cross-spectrum of the Slope and Level

The left panel shows the spectral densities of dr_t , pd_t , and the residuals of dr_t after projecting on pd_t (ϵ_t^{δ}). The integral of spectral density is equal to the variance. The horizontal line starts from zero and ends at π , but is labeled with the corresponding length of a cycle. The right panel shows the cross-spectral density between dr_t and pd_t . The integral of cross-spectral density is equal to the covariance.

Table 1, we include μ_t^F (g_t^F) and KP_t (KP_t^{CF}), the return (dividend growth) predictors in [Binsbergen and Koijen \(2010\)](#) and [Kelly and Pruitt \(2013\)](#), respectively. These variables are constructed to filter out (latent) state variables. To highlight our contribution, we benchmark against their results.

The log difference between the valuation ratio of the index and valuation ratio of its one-year dividend strip is the log ratio of total market value to one-year strip price:

$$dr_t = pd_t - s_t^1 = \ln(P_t/D_t) - \ln(P_t^1/D_t) = \ln(P_t/P_t^1). \quad (16)$$

The slope of valuation term structure can be interpreted easily. A mean of 4.027 translates into a ratio of index value to that of its one-year strip equal to $56 = \exp(4.027)$, meaning that the total value is 56 times the value of dividends in the next year. When the valuation term structure steepens (dr_t increases), a greater fraction of value comes from beyond the next year. dr_t varies widely, with a minimum of 2.952 in Nov. 1988 (before the 1990-1991 recession) and a maximum of 6.632 near the end of the dot-com boom (Nov 2000). dr_t has a lower monthly autocorrelation (ρ) than pd_t .

dr_t and pd_t are correlated but contain distinct information. As shown in the cross-spectrum in Figure 1, the correlation of 0.87 is mainly from low-frequency movements. Panel A of Figure 1

shows the spectrum of dr_t , pd_t , and ϵ_t^{dr} (the residual from linearly projecting dr_t on pd_t). The area under the spectrum curve is the variance, so the figure provides a variance decomposition in the frequency domain. On the horizontal axis, instead of showing the frequencies from zero to π , we mark the corresponding length of the cycle for easier interpretation. Once orthogonalized to pd_t , dr_t 's residual varies mainly at annual or higher frequencies. Panel B plots the cross-spectrum of dr_t and pd_t . The integral is the covariance between dr_t and pd_t . The correlation between dr_t and pd_t is mainly from low frequencies. This indicates that it is the high-frequency variation in dr_t that brings information distinct from that revealed by pd_t . This is consistent with the findings in [Kragt, de Jong, and Driessen \(2020\)](#) about different state variables fluctuating at different frequencies. Figure A.2 in the Appendix shows the spectrum analysis based on daily data with similar results.

As shown in Section 2, dr_t and pd_t are essentially different combinations of state variables. Our state-space approach is closely related to [Binsbergen and Koijen \(2010\)](#). [Binsbergen and Koijen \(2010\)](#) use the realized returns and dividends to estimate a latent-state model and filter out the conditional expected return, μ_t^F , and the conditional expected dividend growth rate, g_t^F . These filtered variables are also combinations of state variables (subject to estimation errors). We replicate the analysis of [Binsbergen and Koijen \(2010\)](#) and compare our state-space representation via observable valuation ratios with information from the filtered μ_t^F and g_t^F . [Kelly and Pruitt \(2013\)](#) also take a state-space approach and use the cross-section of market-to-book ratios of individual stocks to extract the expected return and dividend growth of the aggregate market. We have also replicated [Kelly and Pruitt \(2013\)](#) and include their state variables (predictors) for comparison.

3.2 Analyzing the state space

According to Section 2, information about the state space is embedded in valuation ratios, including those of the dividend strips, pd , and dr . In our empirical analysis, we use valuation ratios of six-month strip ($s^{0.5}$), one-year strip (s^1), and dividends paid beyond one year (s^{1+}) as the futures data at 0.5 and 1 year maturities are the most liquid. In Panel A of Figure 2, we report the results from principal component analysis (PCA). The first two components account for 96.3% of total

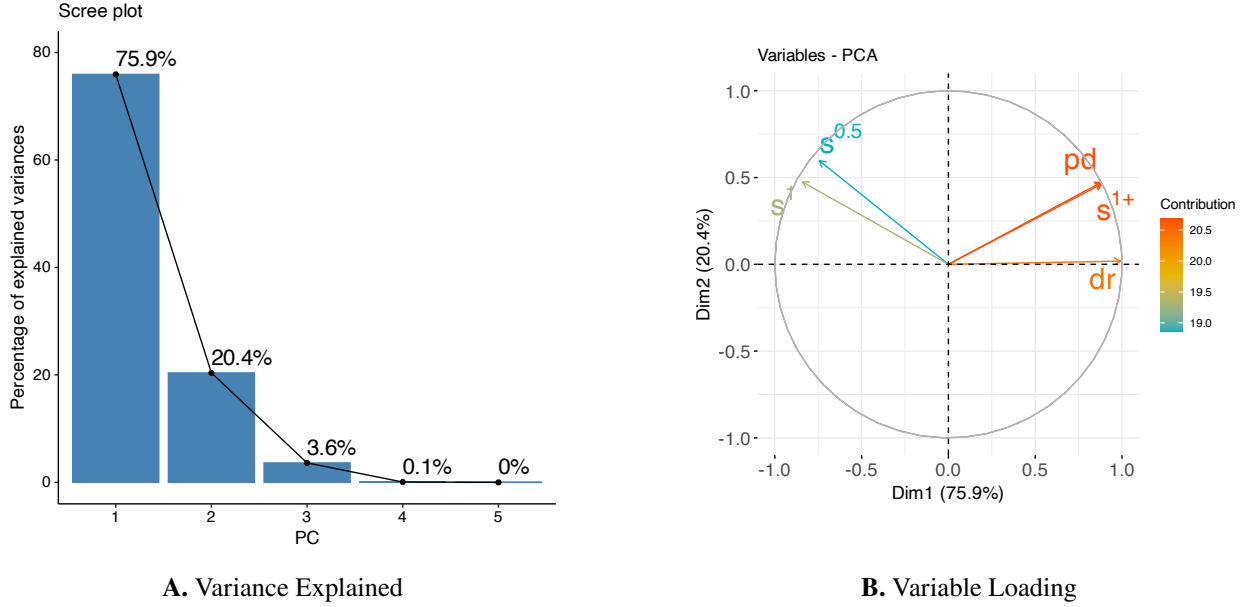


Figure 2 Principal Component Analysis of Valuation Ratios

This figure reports the PCA results for dr_t , pd_t , $s_t^{0.5}$, s_t^1 , and s_t^{1+} . Panel A plots the variance explained by each principal component. Panel B plots the loadings of each variable on the first two principal components.

variance. In line with our analysis in Section 2, we show in Panel B of Figure 2 that valuation ratios load differently on the two principal components, labeled as Dim1 (dimension 1) and Dim2 (dimension 2). These results indicate that the state space, mapped out by the valuation ratios, is two-dimensional. Thus, we may pick two valuation ratios or pd and dr to span the state space.

However, as pointed out by Kelly and Pruitt (2015), a shortcoming of PCA analysis is that information embedded in the principal components may not be the most relevant for objects of interest, which are the expected return and cash-flow growth. Next, we take a predictive regression approach. The expected return and expected dividend growth rate are driven by the state variables. By projecting future returns and dividend growth rates on the valuation ratios, we are able to evaluate which valuation ratios are the most informative on the cash-flow and return dynamics.

In Figure 3, we report R^2 of predicting S&P 500 dividend growth over the next year using different sets of valuation ratios. A round dot represents adjusted in-sample R^2 (reported with its 95% confidence interval) and a triangle represents out-of-sample R^2 . We report the detailed regression results in Table A.2. Our predictive regression is run on monthly observations. In the first specification, we include dr and pd , which achieve the highest in-sample and out-of-sample

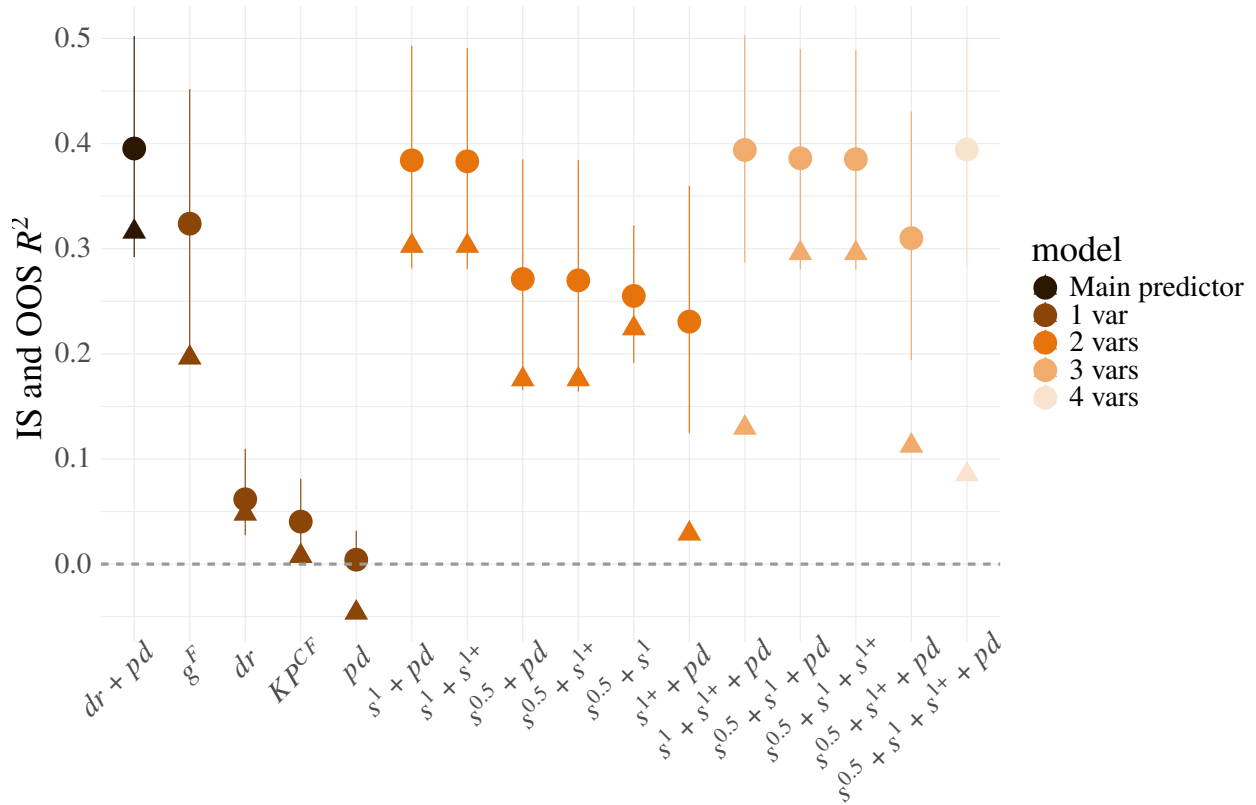


Figure 3 In-Sample and Out-of-Sample R^2 from Dividend Growth Predictive Regressions

This figure reports in-sample and out-of-sample R^2 for predicting annual S&P 500 Index dividend growth using various predictors. The predictors include our main predictor ‘slope’ dr , the price-dividend ratio pd , $dr + pd$, predictor from [Binsbergen and Koijen \(2010\)](#) (g^F), predictor from [Kelly and Pruitt \(2013\)](#) (KP^{CF}), and different combinations of pd , $s^{0.5}$ (price-dividend ratio of six-month strip), s^1 (price-dividend ratio of one-year strip) and s^{1+} (price-dividend ratio of dividends beyond one year). Each round dot represents in-sample R^2 with a 95% bootstrapped confidence interval. Each triangle represents out-of-sample R^2 by recursively forecasting returns beginning in 1998:01.

R^2 . Next, we show the R^2 of g^F , the predictor from [Binsbergen and Koijen \(2010\)](#), which is the conditional expectation filtered from a latent state model, and in the fourth specification is the R^2 of the predictor from [Kelly and Pruitt \(2013\)](#) based on the valuation ratios of individual stocks.

In the third specification in Figure 3, we only include the slope of valuation term structure, dr , and find very limited predictive power in comparison with the combination of the slope and level. In the fifth specification, we show that the level of valuation term structure, pd , also has rather weak predictive power. The subsequent specifications show that predictive power varies across different pairs of valuation ratios, indicating the importance of taking a predictive regression approach rather than simply relying on the PCA of valuation ratios. Any given pair of valuation

ratios fully spans the two principal components, as indicated in Panel B of Figure 2, but they contain different information about return and cash-flow dynamics. Finally, in the last five specifications in Figure 3, we show that three or four valuation ratios do not outperform two valuation ratios in forecasting dividend growth. Overall, our results indicate that two valuation ratios (in particular, the combination of dr and pd) are sufficient for forecasting dividend growth.

The finding that dr has very limited predictive power is somewhat surprising. When the valuation term structure steepens (i.e., dr increases), the stock market derives a greater fraction of value from long-duration cash flows, so one would expect the expected growth rate of dividends to rise accordingly (i.e., the term structure of cash-flow growth should steepen as well). However, this is not the case, suggesting that what drives variation in the slope of valuation term structure is not the trajectory of expected cash-flow growth but the discount rate (expected return). Next, we analyze the state space by forecasting returns with a particular focus on dr as a return predictor.

In Figure 4, we report the R^2 of predicting annual returns of the S&P 500 with different sets of valuation ratios. Our regression is run monthly. We report the detailed regression results in Table A.3. The conclusion is similar to that from cash-flow prediction: having three or more valuation ratios does not improve predictability relative to the best-performing pairs of valuation ratios. Combining pd and dr again achieves one of the highest R^2 s, suggesting that to capture information on state variables, we can rely on the level and slope of valuation term structure.

Interestingly, the predictive power of slope alone is comparable to that of pd and dr combined. Moreover, in the second, third, and fourth specifications in Figure 4, we find that μ^F , the predictor in Binsbergen and Koijen (2010), KP , the predictor in Kelly and Pruitt (2013), and the traditional price-dividend ratio all underperform dr . Our results suggest a tight link between the slope of valuation term structure and expected return. This explains why it is essential to combine pd and dr for forecasting cash-flow growth. The expected return and cash-flow growth can be correlated (Lettau and Ludvigson, 2005; Kothari, Lewellen, and Warner, 2006). Overall, this link between the slope and expected return is critical for understanding what drives the shape of valuation term structure and, vice versa (what drives the discount rate). We will explore further in Section 4.

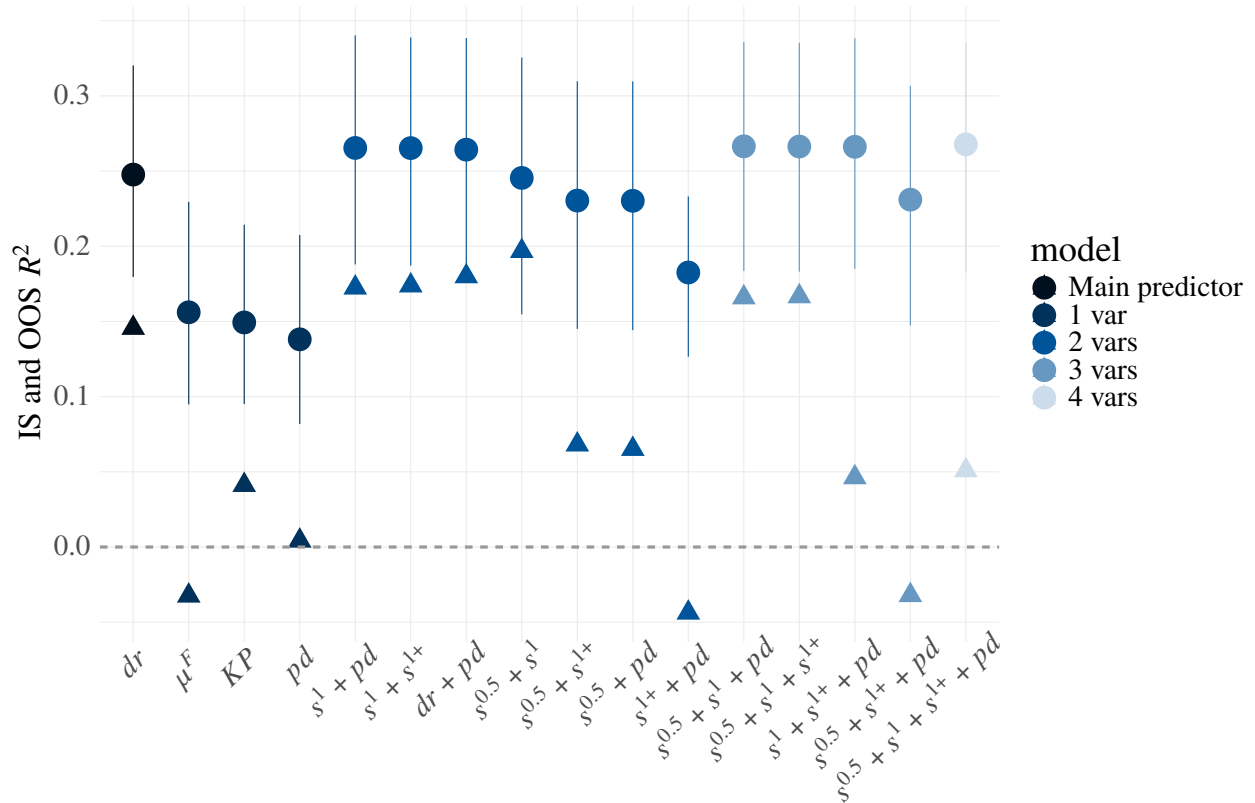


Figure 4 In-Sample and Out-of-Sample R^2 from Return Predictive Regressions

This figure reports in-sample and out-of-sample R^2 for predicting annual S&P 500 Index returns using various predictors. The predictors include our main predictor ‘slope’ dr , the price-dividend ratio pd , $dr + pd$, predictor from [Binsbergen and Koijen \(2010\)](#) (μ^F), predictor from [Kelly and Pruitt \(2013\)](#) (KP), and different combinations of pd , $s^{0.5}$ (price-dividend ratio of six-month strip), s^1 (price-dividend ratio of one-year strip) and s^{1+} (price-dividend ratio of dividends beyond one year). Each round dot represents in-sample R^2 with a 95% bootstrapped confidence interval. Each triangle represents out-of-sample R^2 by recursively forecasting returns beginning in 1998:01.

Before we move on to Section 4, we briefly discuss the rich information embedded in the level and slope of valuation term structure beyond that on expected return and cash-flow growth. In Table A.4 in the Appendix, we report the R^2 from forecasting a variety of macroeconomic and financial-market variables with dr and pd , only dr , only pd , and for illustration purposes, dr in combination with s^1 (the log price-dividend ratio of one-year dividend strip), and only s^1 . The pair dr and pd demonstrate the strongest predictive power when forecasting variables related to financial intermediaries’ balance-sheet capacity (with the R^2 ranging from 30% to 40%).²⁵ When forecasting macroeconomic variables related to business-cycle dynamics, dr and pd have an R^2

²⁵This implies a link between financial intermediation and equity state variables (e.g., [He, Kelly, and Manela, 2017](#)).

consistently above 20%. Moreover, the R^2 from forecasting sentiment proxies is consistently above 10%. Finally, combining dr and pd outperforms other specifications, suggesting that the level and slope of valuation term structure should both be included in empirical models of macroeconomic dynamics, such as the vector autoregression models (e.g., [Sims, 1980](#); [Campbell and Ammer, 1993](#); [Patelis, 1997](#); [Bernanke and Kuttner, 2005](#); [Larrain and Yogo, 2008](#); [Cieslak and Pang, 2021](#)).

4 Expected Return and the Slope of Valuation Term Structure

A striking finding from our forecasting exercises is that the slope of valuation term structure, dr , strongly predicts returns. Augmenting dr with pd or other valuation ratios does not meaningfully improve the performance. In [Section 2](#), we show that dr is a linear combination of the state variables. Our empirical findings suggest that dr and the conditional expected return may coincide in their state-variable loadings; in other words, dr corresponds to the conditional expected return (time-varying discount rate). Next, we provide further evidence on the return predictive power of dr and analyze the economic mechanism behind the tight link between dr and the discount rate.

4.1 Return prediction

Predictive regression. We provide a thorough analysis of the return predictive power of dr , the slope of valuation term structure. We start with standard predictive regression for annual index returns:

$$r_{t+1} = \alpha + \beta dr_t + \epsilon_{t+1}, \quad (17)$$

Because we use overlapping monthly data, we adopt [Newey and West \(1987\)](#) standard errors with 18 lags to account for the moving-average structure induced by overlap ([Cochrane and Piazzesi, 2005](#)). We also calculate [Hodrick \(1992\)](#) standard errors. [Hodrick \(1992\)](#) shows that GMM-based autocovariance correction (e.g., [Newey and West, 1987](#)) may have poor small-sample properties. Under the serial correlation in the error term, another concern is the bias induced by the persistence

of the predictor.²⁶ While dr has an autocorrelation below that of the traditional price-dividend ratio, pd , we still report the adjusted estimate of β following [Stambaugh \(1999\)](#). In the appendix (Table A.5), we also report the IVX-Wald test of predictive power ([Kostakis, Magdalinos, and Stamatogiannis, 2014](#)) that explicitly accounts for the persistence of the return predictor.

Adjusted R^2 measures in-sample forecasting performance. Following the literature on the discrepancy between in- and out-of-sample performances ([Bossaerts and Hillion, 1999](#); [Goyal and Welch, 2007](#)), we report the out-of-sample R^2 and two tests of out-of-sample performance. We form out-of-sample forecasts as a real-time investor, using data up to time t in the regression to estimate β , which is then multiplied by the time- t value of the predictor to form the forecast. Out-of-sample forecasting starts from Dec. 1997 when we have at least ten years of data. Out-of-sample R^2 is

$$R_{OOS}^2 = 1 - \frac{\sum_t (r_{t+1} - \hat{r}_{t+1})^2}{\sum_t (r_{t+1} - \bar{r}_t)^2},$$

where \hat{r}_{t+1} is the forecast value and \bar{r} is the average of twelve-month returns (the first is January-December 1998). The out-of-sample R^2 lies in the range $(-\infty, 1]$, where a negative number means that a predictor provides a less accurate forecast than the historical mean.

We report the p -value of two out-of-sample performance tests, “*ENC*” and “*CW*”. *ENC* is the encompassing forecast test derived by [Clark and McCracken \(2001\)](#), which is widely used in the literature. We test whether the predictor has the same out-of-sample forecasting performance as the historical mean and compare the value of the statistic with critical values calculated by [Clark and McCracken \(2001\)](#) to obtain a p -value range. [Clark and West \(2007\)](#) adjust the standard MSE t-test statistic to produce a modified statistic (*CW*) that has an asymptotic distribution well approximated by the standard normal distribution, so for *CW*, we report the precise p -value.

Table 2 presents the results. Column (1) shows that the slope of valuation term structure, dr , demonstrates a striking degree of return predictive power. The in-sample estimation generates a predictive R^2 reaching 24.8%.²⁷ Out-of-sample forecasts deliver an R^2 of 14.6%, significantly

²⁶The persistence of a return predictor can cause small-sample bias ([Nelson and Kim, 1993](#); [Stambaugh, 1999](#)) and spurious regression ([Ferson, Sarkissian, and Simin, 2003](#)).

²⁷[Foster, Smith, and Whaley \(1997\)](#) discuss the potential data mining issues that arise from researchers searching among potential regressors. They derive a distribution of the maximal R^2 when k out of m potential regressors are used as predictors and calculate the critical value for R^2 , below which the prediction is not statistically significant. For

Table 2 Return Prediction

This table reports the return prediction results. The dependent variable of the regression is the log annual return of the S&P 500 index, r_{t+1} . We consider the following predictors: ‘slope’ of valuation term structure dr_t , price-dividend ratio pd_t , filtered series for expected return following [Binsbergen and Koijen \(2010\)](#) μ^F , and the predictive factor extracted from 100 book-to-market and size portfolios from [Kelly and Pruitt \(2013\)](#) KP . The β estimate is followed by [Hodrick \(1992\)](#) t -statistic, [Newey and West \(1987\)](#) t -statistic (with 18 lags), and the β coefficient adjusted for [Stambaugh \(1999\)](#) bias. Starting from 1998:01, we form out-of-sample forecasts of return in the next twelve months by estimating the regression with data up to the current month and use the forecasts to compute out-of-sample R^2 , ENC test ([Clark and McCracken, 2001](#)), and the p -value of CW test ([Clark and West, 2007](#)). Our monthly sample is 1988:01–2019:12.

	r_{t+1}				
	(1)	(2)	(3)	(4)	(5)
dr_t	-0.156				-0.228
Hodrick t	[-3.354]				[-2.924]
Newey-West t	(-4.499)				(-3.517)
Stambaugh bias adjusted β	-0.146				
pd_t		-0.199			0.141
		[-2.367]			[1.721]
		(-2.747)			(1.209)
		-0.189			
μ_t^F			2.584		
			[2.313]		
			(2.804)		
			2.594		
KP_t				0.895	
				[2.960]	
				(2.857)	
				0.905	
N	372	372	372	372	372
R^2	0.248	0.138	0.156	0.149	0.264
OOS R^2	0.146	0.004	-0.032	0.041	0.180
ENC	2.968	0.833	0.651	2.978	5.985
$p(ENC)$	<0.05	>0.10	>0.10	<0.05	<0.01
$p(CW)$	0.022	0.200	0.303	0.031	0.021

outperforming the historical mean as shown by the p -values of ENC and CW .²⁸ The predictive coefficient is also large in magnitude, indicating high volatility of the conditional expected return. A decrease of dr by one standard deviation adds 7.7% to the expected return. Both Newey-West and Hodrick t -statistics are significant at least at the 1% level. The negative predictive coefficient of dr suggests that one can form a market timing strategy betting against the slope of valuation term structure: reduce market exposure when dr increases. An out-of-sample R^2 of 14.6% in column

instance, when $m = 50$, $k = 5$, and the number of observations is 250, the 95% critical value for R^2 is 0.164.

²⁸In our calculation of out-of-sample R^2 starts from Dec. 1997 (after the first ten years of data). Figure A.8 in the Appendix reports the out-of-sample R^2 for different start dates and compares the OOS R^2 of dr with that of pd .

(1) of Table 2 implies that the Sharpe ratio of this strategy is 0.58, which is much higher than the Sharpe ratio of 0.37 from the buy-and-hold strategy in [Campbell and Thompson \(2008\)](#).²⁹

Column (2) of Table 2 reports the results for pd , the most commonly adopted return predictor. Its predictive power is much weaker than that of dr across all metrics. Its in-sample R^2 is almost half of that of dr , and pd barely exhibits any out-of-sample predictive power with R^2 equal to 0.4%. In both ENC and CW tests, pd fails to beat the historical mean with any statistical significance. Its coefficient is smaller in magnitude than that of dr . A decrease in pd by one standard deviation leads to an increase of expected return by 5.8%, implying a less volatile expected return than the one from dr . The IVX-Wald test of [Kostakis, Magdalinos, and Stamatogiannis \(2014\)](#) in Table A.5 in the appendix also supports the significant predictive power of dr while rejecting that of pd .

Next, we compare dr with two return predictors that are conceptually related. [Binsbergen and Koijen \(2010\)](#) extract information about state variables that drive the conditional expected return and expected cash-flow growth by estimating a latent-state model. Our approach differs as we do not estimate or filter the state variables but instead rely on observable state-variable proxies, such as dr , pd , and valuation ratios of dividend strips. In Column (3) of Table 2, we follow the procedure in [Binsbergen and Koijen \(2010\)](#) to construct their return predictor, μ^F . While μ^F outperforms pd , its predictive power is weaker than that of dr across different metrics in our sample period.

[Kelly and Pruitt \(2013\)](#) deploy another filtering method that utilizes the cross-section of market-to-book ratios of individual stocks. These valuation ratios are correlated with state variables, but, as we show in Appendix I.2, they contain firm-level noise that is orthogonal to the expected return. [Kelly and Pruitt \(2013\)](#) use partial least squares to reduce noise. Our approach differs as dr , pd , and valuation ratios of index dividend strips do not contain firm-level noise. In our approach, the challenge is to find the combination of valuation ratios whose state-variable loadings coincide with those of the conditional expected return. In the next subsection, we characterize the necessary and sufficient condition for dr , the slope of valuation term structure, to be this optimal return predictor. Following the procedure in [Kelly and Pruitt \(2013\)](#), we construct their return predictor,

²⁹In the appendix, we show how to calculate the Sharpe ratio based on the out-of-sample R^2 .

Table 3 Correlation between Return Predictors

This table reports the correlation matrix of four main return predictors. dr is the slope of S&P 500 valuation term structure. pd is the price-dividend ratio of the S&P 500 index. μ^F is the filtered series for demeaned expected returns, following Binsbergen and Kojen (2010), KP is the return predictive factor extracted from 100 book-to-market and size portfolios from Kelly and Pruitt (2013). Our monthly sample period is 1988:01–2019:12.

	dr	pd	μ^F	KP
dr	1			
pd	0.873	1		
μ^F	-0.892	-0.967	1	
KP	-0.565	-0.496	0.468	1

denoted by KP . In column (4) of Table 2, we report the results. KP significantly outperforms pd but underperforms dr_t across metrics such as Newey-West t -statistic, Hodrick t -statistic, in-sample R^2 , out-of-sample R^2 , ENC , CW , and IVX-Wald test reported in Table A.5 in the online appendix.

These return predictors are correlated as shown in Table 3 but differ significantly in predictive power, with the slope of valuation term structure outperforming other predictors. In the appendix, we demonstrate the robustness of our results by repeating the analysis for alternative forecasting targets. In Table A.6, we replace the S&P 500 annual return with the excess annual return. In Table A.7 and A.8, we consider the Fama-French market portfolio return and excess return, respectively.³⁰ Note that our method forecasts returns of an asset with the slope of its valuation term structure. Thus, changing the forecasting target is for statistical robustness and not theoretically motivated.

Comparing the slope and other predictors. We have compared the return predictive power of dr with that of pd , μ^F , and KP . Figure 5 compares dr with more predictors from the literature, including the default yield spread (dfy), the inflation rate ($infl$), stock variance ($svar$), the cross-section premium (csp), the dividend payout ratio (de), the long-term yield (lty), the term spread (tms), the T-bill rate (tbl), the default return spread (dfr), the dividend yield (dy), the long-term rate of return (ltr), the earnings-to-price ratio (ep), the book to market ratio (bm), the investment-to-capital ratio (ik), the net equity expansion ratio ($ntis$), the percent equity issuing ratio ($eqis$), and

³⁰As will be made clear in the next subsection, our goal is to predict annual returns. However, we also show that dr demonstrates superior return predictive power at a monthly horizon. Our baseline results are reported in Table A.9, and see Table A.10 for results on predicting monthly S&P 500 excess return. Table A.11 and Table A.12 report the results on predicting the monthly Fama-French market portfolio return and excess return, respectively.

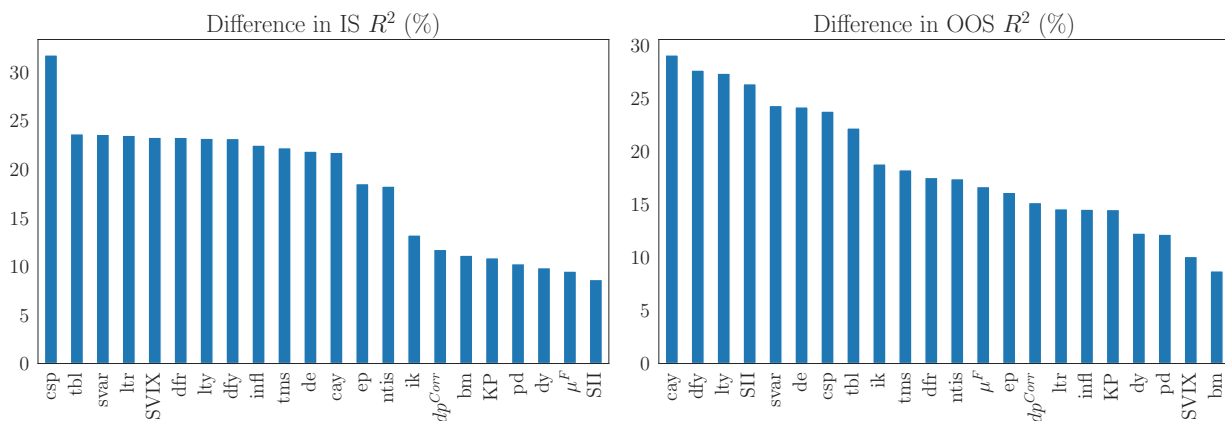


Figure 5 In-Sample and Out-of-Sample R^2 Wedge between dr and Other Return Predictors

This figure compares annual return predictive R^2 between dr_t and other commonly studied predictors. Panels A and B report, respectively, the differences in in-sample (IS) and out-of-sample (OOS) R^2 between dr and an alternative predictor. A positive value signifies that dr has a stronger predictive power than the alternative within the same sample period. Most predictors are from Goyal and Welch (2007) and include the price-dividend ratio (pd), the default yield spread (dfy), the inflation rate (infl), stock variance (svar), the cross-section premium (csp, available in 1988-2002), the dividend payout ratio (de), the long-term yield (lty), the term spread (tms), the T-bill rate (tbl), the default return spread (dfr), the dividend yield (dy), the long-term rate of return (ltr), the earnings-to-price ratio (ep), the book to market ratio (bm), the investment-to-capital ratio (ik), the net equity expansion ratio (ntis), the percent equity issuing ratio (eqis), and the consumption-wealth-income ratio (cay). KP is the predictive factor extracted from 100 book-to-market and size portfolios from Kelly and Pruitt (2013). dp^{Corr} is the dividend-price ratio corrected for option-implied dividend growth in Golez (2014) (available in 1994-2011). μ^F is the filtered series for expected returns following Binsbergen and Koijen (2010). SII is the short interests index from Rapach, Ringgenberg, and Zhou (2016) (available in 1988-2014). SVIX is an option-implied lower bound of annual equity premium in Martin (2017) (available in 1996-2012).

the consumption-wealth-income ratio (cay), which are summarized in Goyal and Welch (2007), and others that are proposed more recently, such as adjusted dividend yield, dp^{Corr} (Golez, 2014), short interest index, SII (Rapach, Ringgenberg, and Zhou, 2016), and SVIX (Martin, 2017).³¹ We also include pd , μ^F , and KP . In Figure 5, we report in- and out-of-sample R^2 of dr minus those of other predictors. All columns are in the positive region, indicating dr performs better.

In Table A.15, we report the correlation between dr and the other predictors. Besides pd , μ^F , and KP , all the other predictors with correlation above 0.5 or below -0.5 are all valuation ratios, such as the market-to-book ratio and price-earnings ratio. This is consistent with the insight in Kelly and Pruitt (2013) and with our emphasis on using valuation ratios to capture information about state variables. Finally, in the appendix, we repeat the exercise in Figure 5 for alternative forecasting

³¹Note that the dividend yield (dy) is not the inverse of price-dividend ratio (pd) because in the denominator of dy is the lagged market value (not the current value).

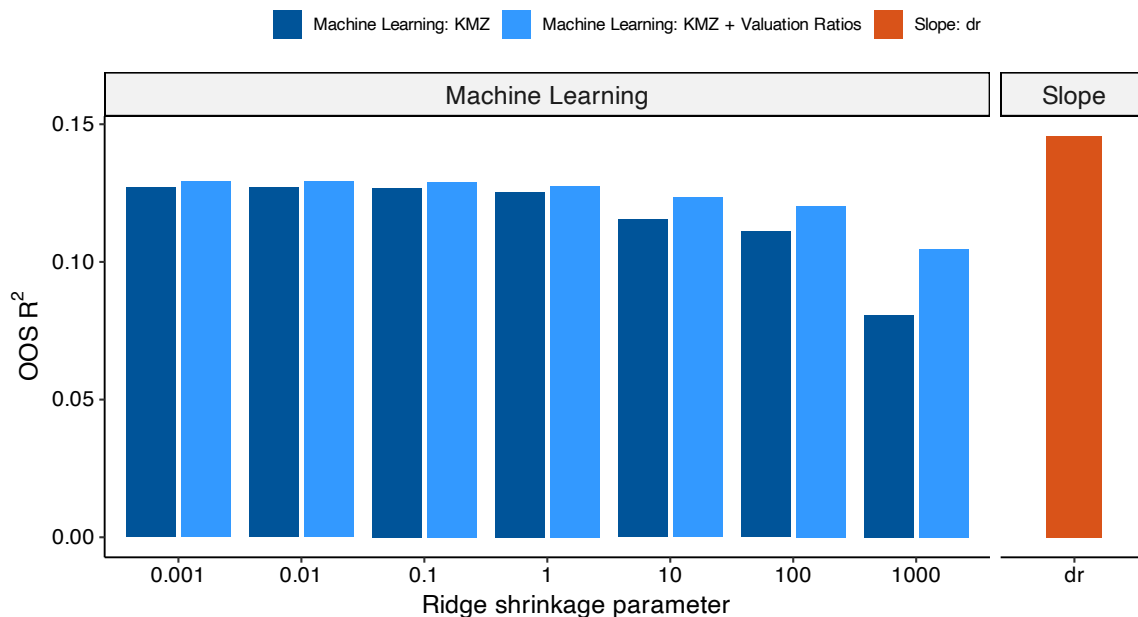


Figure 6 Return Prediction: The Role of Nonlinearity

This figure shows the out-of-sample (OOS) R^2 of the slope of S&P 500 valuation term structure dr and machine-learning (ML) models in Kelly, Malamud, and Zhou (2024). We forecast annual S&P 500 returns at the monthly frequency, with OOS prediction beginning in 1998:01 and OOS R^2 computed following Goyal and Welch (2007). The ML models use a 12-month training window, $\gamma = 2$, and a Random Fourier Features (RFF) count P ranging from 2 to 12,000. The darker blue bars represent the ML models using 15 predictor variables (as in Kelly, Malamud, and Zhou, 2024). For the lighter blue bars, we augment the signal set with valuation ratios (i.e., $s^{0.5}$, s^1 , s^{1+} , dr , and pd). The figure compares the best OOS R^2 for each shrinkage parameter against the OOS R^2 achieved using dr (the orange bar).

targets, such as S&P 500 excess annual return (Figure A.3), Fama-French market portfolio annual return (Figure A.4), and Fama-French market portfolio excess annual return (Figure A.5).

The role of nonlinearity. Following Lettau and Wachter (2007), our theoretical framework is an exponential-affine model where valuation ratios of dividend strips and the level and slope of valuation term structure are linear functions of state variables. This framework motivates our empirical analysis and facilitates the interpretation of our results. One concern over this type of model is nonlinearity: the valuation ratios, pd , and dr may no longer be linear functions of state variables (or vice versa), and accordingly, the forecasting exercises should account for nonlinearity.

Kelly, Malamud, and Zhou (2024) develop a method based on ridge regressions to account for nonlinearity. Given a set of predictors (signals), their forecasting models can be expanded progressively to incorporate nonlinear terms (“model complexity”). In the appendix, we replicated

their analysis: given a value of ridge shrinkage parameter that indexes a class of forecasting models, we plot the out-of-sample R^2 against the degree of model complexity (see Figure A.6). In Figure 6, we report the maximum R^2 under each value of ridge shrinkage parameter and compare it against the R^2 obtained from the univariate predictive regression with dr as the predictor. Under each value of ridge shrinkage parameter, we consider two cases, one with a signal base including all of our valuation ratios and other predictors and the other including only the other predictors.

The machine learning model is essentially a signal aggregator with the optimal degree of complexity and nonlinearity. The fact that the simple OLS with dr delivers an out-of-sample R^2 above that of the nonlinear model suggests that the linear structure generated from the exponential-affine model is an adequate approximation.³² Overall, our analysis has two implications. First, in terms of raw signals, pd , dr , and valuation ratios of the dividend strips constitute a sufficiently rich set. Therefore, we may not need to seek “big data” (i.e., alternative signals) for forecasting returns. Second, combining these signals in a nonlinear fashion does not improve forecasting performances. Therefore, the linear mapping between state variables and pd , dr , and strip valuation ratios offer a solid foundation for motivating our empirical analysis of state space and the forecasting exercises.

Spanning tests. In the bond literature, unspanned state variables have attracted enormous attention. In particular, variables that drive the expected bond returns may not be revealed by bond yields (as reviewed by Duffee, 2013). Next, we show that this is not the case in our analysis of expected equity return. Specifically, we perform the following spanning tests in the appendix. In Table A.13, we conduct bivariate predictive regressions with dr as one predictor and the other being one of the alternative predictors. Across all bivariate predictive regressions, dr , always has a coefficient that is statistically significant at 1% level, while almost all the other predictors are driven out, showing an insignificant coefficient. The short interest index has a significant coefficient but, as shown in Figure 5, its out-of-sample R^2 is deep in the negative territory (below the 15% out-of-sample R^2 of dr by more than 25%). Inflation also has a significant coefficient in Table A.13 but also an out-of-sample

³²Note that due to estimation errors, the machine learning model may underperform our simple OLS with dr as the predictor even when dr and other valuation ratios are included as signals.

R^2 close to zero.³³ Table A.14 reports an alternative spanning test. We run trivariate predictive regressions with dr , pd , and the third predictor being one of the alternative predictors. As in the bivariate predictive regressions, the coefficients of all the alternative predictors are insignificant at 1% (except SII). Moreover, the predictive coefficient of pd is insignificant across all specifications.

4.2 Understanding the slope as a return predictor

Next, we explain the tight between dr , the slope of valuation term structure of the equity index, and the expected index return through the lens of state-space model. By analyzing the model, we develop tests for the following hypothesis: steepening (flattening) of the valuation term structure is discount-rate driven because market participants lack information on long-term growth.

A two-dimensional state space model. Our analysis in Section 3 suggests that the state space is two-dimensional. In the following, we reduce the dimensionality of state space of the model in Section 2. As in Lettau and Wachter (2007) and Binsbergen and Kojen (2010), we can rotate the state variables so that one drives the expected dividend growth rate while the other drives the expected return through the price of risk. Specifically, the dividend growth, g_t , is given by

$$g_t = z_t + \bar{g} - \frac{1}{2}\sigma_D^\top \Sigma \sigma_D, \quad (18)$$

where z_t has the following law of motion

$$z_{t+1} = \rho_z z_t + \sigma_z^\top \epsilon_{t+1}. \quad (19)$$

The second state variable, y_t , with a law of motion

$$y_{t+1} = \rho_y y_t + \sigma_y^\top \epsilon_{t+1}, \quad (20)$$

drives the price of risk λ_t , so equation (5) becomes

$$\lambda_t = \bar{\lambda} + y_t, \quad (21)$$

³³De la O and Myers (2024) point out that errors in inflation expectations are important drivers of asset prices.

and the stochastic discount factor (SDF) is given by

$$M_{t+1} = \exp \left\{ -r_f - \frac{1}{2} \lambda_t^2 (\sigma_\lambda^\top \Sigma \sigma_\lambda)^2 - \lambda_t \sigma_\lambda^\top \epsilon_{t+1} \right\}. \quad (22)$$

The price of risk for the n -th shock is $\lambda_t \sigma_\lambda(n)$, where $\sigma_\lambda(n)$ is the n -th element of σ_λ . The N -by-1 shock vector ϵ_{t+1} contains news at $t + 1$. The variables' shock loadings may differ, for example, $\sigma_z \neq \sigma_y$. z_t and y_t can be correlated through their overlapping exposure to shocks.

In Appendix I, we solve the log price-dividend ratio of the aggregate market

$$pd_t = A_{pd} + B_{pd}y_t + C_{pd}z_t, \quad (23)$$

where A_{pd} , B_{pd} , and C_{pd} are constant, and the log price-dividend ratio of the one-year strip,

$$s_t^1 = A_1 + B_1y_t + C_1z_t. \quad (24)$$

Therefore, the slope of valuation term structure is given by

$$dr_t = pd_t - s_t^1 = A_{pd} - A_1 + (B_{pd} - B_1)y_t + (C_{pd} - C_1)z_t, \quad (25)$$

where $B_{pd} \neq B_{pd} - B_1$ and $C_{pd} \neq C_{pd} - C_1$. Since dr_t and pd_t have different loadings on z_t and y_t , we can solve z_t and y_t from dr_t and pd_t . In particular, dr_t and pd_t contain all the necessary information for forecasting return and dividend growth as we find empirically in Section 3.

The next proposition shows that when ρ_z , the autoregressive coefficient of expected dividend growth rate z_t , is zero, we have dr_t and $\mathbb{E}_t[r_{t+1}]$ being univariate functions of one another.

Proposition 1 (Discount rate and the slope of valuation term structure) *The expected return at time t (time-varying discount rate) is a linear function of y_t , i.e., $\mathbb{E}_t[r_{t+1}] = A_{er} + B_{er}y_t$, where A_{er} and B_{er} are constant. Under $\rho_z = 0$, $\mathbb{E}_t[r_{t+1}]$ is a univariate linear function of dr_t .*

Market participants' information about future cash flows is summarized by z_t , which determines the expected growth rate between t and $t + 1$. If z_t lacks persistence ($\rho_z = 0$), market participants do not have information about growth beyond $t + 1$.³⁴ Their expectation of growth from $t + 1$ onward is constant. Therefore, under this condition, when the valuation term structure

³⁴Under $\rho_z = 0$, our model of cash-flow expectations is in line with the belief model in [De La O and Myers \(2021\)](#).

Table 4 Summary Statistics: Cash-Flow Growth Forecasts

This table reports the number of observations, mean, standard deviation, minimum, maximum, quartiles, and monthly autocorrelation (ρ) of measures of cash-flow growth expectations. $\mathbb{E}_t^A \Delta e_{t,t+1}$, $\mathbb{E}_t^A \Delta e_{t+1,t+2}$, and $\mathbb{E}_t^A \Delta e_{t+2,t+3}$ are forecasts of 1-year earnings growth for fiscal year 1, 2, and 3 provided by IBES Global Aggregate (IGA). IGA Δe_t and Compustat Δe_t are the realized annual earnings growth from IGA and Compustat, respectively. $\mathbb{E}_t^A \Delta e_{t,t+1}$ and LTG_t are forecasts of 1-year and long-term earnings growth that we self-aggregate from the IBES Unadjusted US Summary Statistics File. Data sample: 1988:01–2019:12.

	obs	mean	std	min	25%	50%	75%	max	ρ
$\mathbb{E}_t^A \Delta e_{t,t+1}$	384	0.103	0.096	-0.167	0.056	0.103	0.154	0.425	0.897
$\mathbb{E}_t^A \Delta e_{t+1,t+2}$	384	0.134	0.043	-0.069	0.104	0.127	0.157	0.269	0.830
$\mathbb{E}_t^A \Delta e_{t+2,t+3}$	384	0.130	0.036	0.052	0.100	0.122	0.159	0.217	0.953
IGA $\Delta e_{t,t+1}$	384	0.072	0.135	-0.380	-0.008	0.092	0.148	0.425	0.929
Compustat $\Delta e_{t,t+1}$	384	0.068	0.481	-2.175	-0.042	0.122	0.187	2.190	0.976
LTG_t	384	0.125	0.018	0.093	0.115	0.120	0.129	0.187	0.986

steepens (i.e., dr_t increases), a greater fraction of market value is from cash flows from $t + 1$ onward not due to an improved expectation of long-run growth but due to a lower discount rate (a lower y_t); similarly, when the valuation term structure flattens (i.e., dr_t decreases), it is because of a higher discount rate rather than negative information on long-run growth. Next, we estimate ρ_z and demonstrate both empirically and theoretically that the return predictive power of dr_t is tied to ρ_z .

Estimating the persistence of growth expectations. When estimating ρ_z , we use analyst forecasts as a proxy for market participants' expectations. As the coverage of dividend forecasts started in 2003, we follow the literature and use analysts' earnings forecasts (available since 1976). The following accounting identity connects the earnings and dividends: $D_t = \text{Earnings}_t \times (1 - \text{plowback rate}_t)$. As documented by Pástor, Sinha, and Swaminathan (2008) and Chen, Da, and Zhao (2013), the plowback rate is quite stable. Therefore, the dividend growth rates are close to those of earnings.

IBES Global Aggregates (IGA) provides a forecast of earnings growth for the S&P 500 index based on firm-level earnings forecasts. The aggregation procedure weighs individual companies by their market capitalization.³⁵ To transform earnings forecasts to forecasts of growth rates, IGA takes the ratio of forecast for period $t + k$ to forecast for $t + k - 1$. We consider forecasting horizons

³⁵To deal with the fact that companies have different fiscal year-end, IGA calendarizes all company-level data to a December calendar year before aggregation. This approach follows the Compustat rule. Please refer to "Thomson Reuters Datastream IBES Global Aggregates Reference Guide" for more detail.

of one, two, and three years (i.e., $k = 1, 2, 3$).³⁶ The data is available at a weekly frequency. We consider both weekly and monthly frequencies. For estimation at monthly frequency, we take the last weekly observation of each month. In a different estimation method, we utilize analyst forecasts for long-term growth (LTG). We aggregate firm-level LTG from IBES to the index level.³⁷ This data is available at the firm level and at the monthly frequency. It is aggregated to the index level via the same aggregation procedure described above. Table 4 provides summary statistics.

Next, we map analyst expectations to the model counterparts and derive a system of equations for estimating ρ_z . Analyst forecasts may not perfectly capture market participants' expectations. Therefore, we add a noise term between the analyst expectations and expectations in our model:

$$\mathbb{E}_t^A (\Delta e_{t+k}) = \mathbb{E}_t (\Delta e_{t+k}) + \varepsilon_{t,k}^A, \quad (26)$$

where we consider $k = 1, 2, 3$, and $\mathbb{E}_t (\cdot)$ represents the market participants' expectation as in the model. From equation (18) in the model, we obtain

$$\begin{aligned} \mathbb{E}_t^A (\Delta e_{t+1}) &= g + z_t + \varepsilon_{t,1}^A \\ \mathbb{E}_t^A (\Delta e_{t+2}) &= g + \mathbb{E}_t (z_{t+1}) + \varepsilon_{t,2}^A = g + \rho_z z_t + \varepsilon_{t,2}^A \\ \mathbb{E}_t^A (\Delta e_{t+3}) &= g + \mathbb{E}_t (z_{t+2}) + \varepsilon_{t,3}^A = g + \rho_z^2 z_t + \varepsilon_{t,3}^A. \end{aligned}$$

Using the first equation to substitute out z_t in the second and third equations, we obtain a system:

$$\underbrace{\begin{bmatrix} \mathbb{E}_t^A (\Delta e_{t+2}) \\ \mathbb{E}_t^A (\Delta e_{t+3}) \end{bmatrix}}_{\equiv \mathbf{y}_t^A} = (1 - \rho_z) g + \rho_z \underbrace{\begin{bmatrix} \mathbb{E}_t^A (\Delta e_{t+1}) \\ \mathbb{E}_t^A (\Delta e_{t+2}) \end{bmatrix}}_{\equiv \mathbf{x}_t^A} + \underbrace{\begin{bmatrix} \varepsilon_{t,1}^A - \rho_z \varepsilon_{t,0}^A \\ \varepsilon_{t,2}^A - \rho_z \varepsilon_{t,1}^A \end{bmatrix}}_{\equiv \boldsymbol{\varepsilon}_t^A}. \quad (27)$$

Therefore, we can estimate ρ_z by regressing \mathbf{y}_t^A on \mathbf{x}_t^A . The identification assumption is that under the econometricians' belief, the expectation of the $\boldsymbol{\varepsilon}_t^A$ is zero conditional on \mathbf{x}_t^A . Note that the deviations of analysts' expectations from market participants' expectations are allowed to be correlated across the starting dates of annual dividend growth, i.e., t , $t + 1$, and $t + 2$.³⁸

³⁶Note that for $k = 1$, the growth rate is simply calculated as the forecast divided by realized earnings.

³⁷IBES firm-level forecasts of the annualized average growth rate of earnings over the next three to five years have been adopted in the recent literature (e.g., Nagel and Xu, 2022, Bordalo et al., 2024, and De la O and Myers, 2024).

³⁸The identification of ρ_z is robust to the correlation between $\boldsymbol{\varepsilon}_t^A$ and the structural shocks $\boldsymbol{\varepsilon}_t$ in the model, i.e., the

Table 5 Estimating the Persistence of Expected Cash-Flow Growth (Analyst Forecasts)

This table reports the estimates of ρ_z , the autoregressive coefficient of expected cash-flow growth, based on equation (27). The estimation uses aggregate earnings growth forecasts of the S&P 500 Index obtained from IGA. Columns (1) and (3) report the estimates of ρ_z using monthly data, while columns (2) and (4) report the estimates of ρ_z using weekly data. Columns (1) and (2) use earnings growth forecasts for 1, 2, and 3 years ahead (“Y1:Y3”) to estimate the two-equation system (27), while columns (3) and (4) only use earnings growth forecasts for 1 and 2 years ahead (“Y1:Y2”) to estimate the first equation in (27). t -statistics based on Driscoll-Kraay standard errors with autocorrelation of up to 18 lags are reported in parentheses. Data sample: 1988:01–2019:12.

	(1)	(2)	(3)	(4)
$(1 - \rho_z)g$	0.129 (13.995)	0.122 (16.906)	0.141 (15.536)	0.133 (16.745)
ρ_z	0.028 (0.690)	0.015 (0.381)	-0.071 (-1.379)	-0.073 (-1.295)
N	768	1887	384	943
R^2	0.003	0.001	0.025	0.028
Sample Periods	Monthly Y1:Y3	Weekly Y1:Y3	Monthly Y1:Y2	Weekly Y1:Y2

The results are reported in Panel A of Table 5. We estimate equation (27) with both monthly (columns 1 and 3) and weekly observations (columns 2 and 4) of analyst forecasts. In columns (1) and (2), our estimation includes both equations in (27), while in Column (3) and (4), we only include the first equation, i.e., only using forecasts at one- and two-year horizons for better data quality. Across the specifications, the estimate $\hat{\rho}_z$ is statistically indistinguishable from zero.

Next, we consider an alternative way to estimate ρ_z by exploring the relationship between forecasts of short- and long-term earnings growth (LTG). Given the autoregressive structure (19), the expected growth rate from period n to $n + 1$ depends on the expected growth rate over the very next period via a coefficient ρ_z^n . If ρ_z is zero, then ρ_z^n is zero, which implies that the average growth rate over three years and beyond does not depend on the expected growth rate over the next year. Therefore, we regress monthly observations of LTG forecast on the near-term expected growth rate, i.e., $\mathbb{E}_t^A [\Delta e_{t+1}]$, and denote the regression coefficient by ρ_z^{LT} . In Table 6, our estimate is statistically indistinguishable from zero, which implies $\rho_z = 0$, consistent with our findings in Table 5.

Finally, we conduct a rolling-window estimation of ρ_z following the method in Table 5. A rolling window contains three years of weekly observations of analyst forecasts.³⁹ We summarize

structural shocks to realized dividend, market participants’ beliefs on cash-flow dynamics, and their price of risk.

³⁹The results are similar if we use alternative window lengths from one to five years (available upon request). Our

Table 6 Estimating the Persistence of Expected Cash-Flow Growth (LTG Forecasts)

This table reports the estimates of ρ_z^{LT} , the regression coefficient in

$$\log(1 + LTG_t) = \text{const} + \rho_z^{LT} \mathbb{E}_t^A [\Delta e_{t+1}] + \varepsilon_t,$$

where LTG_t is the long-term growth forecasts (LTG) of the S&P 500 Index, self-aggregated from stock-level LTG forecasts from the IBES Unadjusted Summary File. The short-term forecast, $\mathbb{E}_t^A [\Delta e_{t+1}]$, is the IGA 1-year earnings growth forecast (IGA $\mathbb{E}_t^A [\Delta e_{t+1}]$). t -statistics based on Newey-West standard errors with autocorrelation of up to 18 lags are reported in parentheses. Our monthly observations are from 1988:01 to 2019:12.

	(1)
	$\log(1 + LTG_t)$
Intercept	0.116 (28.615)
$\mathbb{E}_t^A [\Delta e_{t+1}]$	0.017 (0.711)
N	384
R^2	0.011

the statistics of the rolling-window estimates in Table 7. Naturally, the model underlying agents' belief formation may vary over time, so the estimate of $\hat{\rho}_z$ fluctuates. However, the mean and median of the rolling-window estimates are close to zero, in line with the full-sample estimate in Table 5. Our findings suggest that the expectations of cash-flow growth lack persistence, i.e., ρ_z is close to zero, which implies a one-to-one mapping between dr and the expected return (see Proposition 1). In the next proposition, we show that the value of ρ_z is directly linked to the forecasting error from using dr to predict returns. The proof is in the appendix.

Proposition 2 (ρ_z and return forecast errors) *Let v_{t+1} denote the forecast error when predicting r_{t+1} with dr_t , and let ρ_z denote the autoregressive coefficient of expected cash flow growth z_t in equation (19). If $\rho_z > 0$, then v_{t+1} is positive. If $\rho_z < 0$, then v_{t+1} is negative.*

The proposition implies that in a subsample where the estimate of ρ_z is positive (negative), we would expect the return forecasting error to be positive (negative). In Figure 7, we plot the rolling-window estimate of ρ_z estimates against the return forecasting residuals (denoted by ε_t) from the corresponding rolling window with dr as the predictor. The two time series track each other closely. The first estimate of ρ_z uses three years of IGA data starting in 1985.

Table 7 Rolling Estimates of the Persistence of Expected Cash-Flow Growth

This table reports the summary statistics of the rolling-window estimates of ρ_z using weekly observations, where each rolling window spans three years. In each window, we estimate ρ_z following equation (27), using aggregate earnings growth forecasts of the S&P 500 Index obtained from IGA. Data sample: 1988:01–2019:12.

	count	mean	std	min	25%	50%	75%	max	ρ
$\hat{\rho}_{z,t}$	384	0.025	0.157	-0.260	-0.066	0.001	0.074	0.791	0.975

other closely, with a correlation of 0.42.⁴⁰ This provides further evidence of the connection between return predictability from dr and the lack of persistence in the expected cash-flow growth rates.

In our model, under $\rho_z = 0$, market participants do not have information about growth beyond the very next year. Under this condition, the slope of valuation term structure, dr , reflects the discount rate. When dr increases, the steepening of the valuation term structure suggests that a greater fraction of market value comes from cash flows at longer horizons. If market participants are not informed about growth beyond the very next year, the steepening must be driven by a decline in the discount rate that boosts valuation of long-duration cash flows more than that of near-term cash flows simply because the valuation of long-duration cash flows is more sensitive to discount-rate variation. When dr decreases, the flattening of valuation term structure is driven by a higher discount rate. This is the mechanism behind the return predictive power of dr . In Section 4.3, we will step outside of our model given by equations (18) and (19) and conduct tests in broader settings on whether market participants are informed about long-term growth.

Discussion: Alternative methods to estimate ρ_z . As previously discussed, our tests allow deviation of analyst forecasts from market participants’ expectations. The empirical literature finds that such deviation tends to be small.⁴¹ For robustness, we consider an alternative method to estimate ρ_z without using analyst forecasts. In Appendix II, we fit the latent state model given by

⁴⁰In Figure A.7 in the Appendix, we plot $\rho_{z,t}$ against the out-of-sample forecast errors and obtain a similarly positive correlation. We also regress the rolling-window return prediction errors, both in-sample and out-of-sample, on the rolling-window estimate of ρ_z and find a positive regression coefficient (see Table A.16 in the Appendix).

⁴¹Analyst forecasts reflect their beliefs as compensation are linked to forecast precision, and their forecasts are likely to reflect market participants’ beliefs broadly (e.g., Mikhail, Walther, and Willis, 1999; Cooper, Day, and Lewis, 2001; Bradshaw, 2004; McCarthy and Hillenbrand, 2021). Forecasts may be distorted due to behavioral, incentive, and institutional frictions (e.g., Gu and Wu, 2003; Malmendier and Shanthikumar, 2007, 2014; Binsbergen et al., 2022). Bias is contained as long as such frictions do not correlate systematically with analysts’ true beliefs.

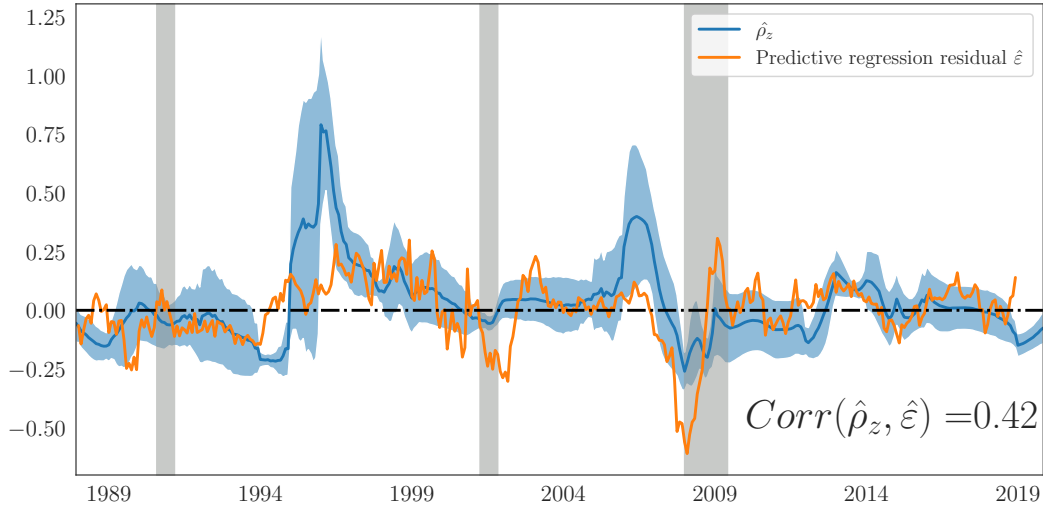


Figure 7 Rolling Estimates of Persistence of Expected Growth and Return Prediction Errors

This figure plots the rolling estimate of the autoregressive coefficient of expected cash flow growth, $\hat{\rho}_{z,t}$, and the return prediction errors using the slope of S&P 500 valuation term structure (dr_t) as the predictor. $\hat{\rho}_{z,t}$ is estimated using analyst forecasts of S&P 500 aggregate earnings in rolling regressions with a three-year window. This figure also plots the predictive residuals (denoted by ε_t) from the rolling-sample predictive regressions. The correlation between the two time series is also reported on the graph. Our monthly sample is 1988:01–2019:12.

(18) and (19) to dividend data to statistically filter out the expected cash-flow growth rate and its persistence. The results corroborate our findings on ρ_z being close to zero.

4.3 Cash-flow growth predictability: Short horizon vs. long horizon

Previously, our focus is on estimating ρ_z , the autoregressive coefficient of expected cash-flow growth rate. In our model, $\rho_z = 0$ implies that the market participants do not have information about cash-flow growth beyond the next year. While our analysis of state space in Section 3 supports the model specification, we step outside of our model and provide more evidence in this subsection on the fact that market participants are not well informed about long-term growth. We do so by characterizing cash-flow predictability at different horizons. If the market participants are informed about cash-flow growth at a certain horizon, we should be able to predict cash-flow growth with information revealed in analysts' forecasts and the state variable proxies such as dr and pd .

In Table 8, we show strong predictability of near-term growth. In column (1), we simply

Table 8 Forecasting Earnings and Dividend Growth over the Next Year

This table reports the results of dividend and earnings growth prediction. The dependent variables are the 1-year-ahead realized earnings growth from IGA (columns 1-3), realized dividend growth from Bloomberg (columns 4-6), and realized earnings growth from Compustat (columns 7-9). The independent variables are analyst forecasts of one-year earnings growth from IGA ($\mathbb{E}_t^A(\Delta e_{t+1})$), duration dr , and the price-dividend ratio pd . t -statistics are calculated based on Newey-West standard errors with 18 lags and are reported in parentheses. Data sample: 1988:01–2019:12.

	IGA Δe_{t+1}			Δd_{t+1}			Compustat Δe_{t+1}		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Intercept	-0.056 (-4.127)	-0.426 (-1.357)	-0.328 (-2.413)	0.028 (1.300)	-0.312 (-1.526)	-0.295 (-1.663)	0.122 (0.814)	0.478 (0.374)	0.394 (0.352)
$\mathbb{E}_t^A(\Delta e_{t+1})$	1.204 (20.101)		1.119 (15.193)	0.326 (2.537)		0.204 (2.824)	-0.591 (-0.629)		-0.963 (-1.067)
dr_t		-0.248 (-2.328)	-0.110 (-2.625)		-0.181 (-3.535)	-0.156 (-4.188)		-0.333 (-1.129)	-0.452 (-2.016)
pd_t		0.385 (2.354)	0.187 (2.599)		0.285 (2.782)	0.249 (3.095)		0.238 (0.405)	0.409 (0.958)
N	372	372	372	372	372	372	372	372	372
R^2	0.731	0.199	0.769	0.193	0.386	0.454	0.014	0.053	0.086

regress the realized one-year growth rate of aggregate earnings from firms covered by IGA on the ex ante analysts' forecast. The R^2 is 0.73, so analysts and market participants in general are able to forecast near-term cash-flow growth very well. In column (2), we use our pair of state variables, dr and pd , to forecast cash-flow growth and obtain a R^2 of 0.20. Combining the information in dr and pd with the analysts' forecast in column (3), the in-sample prediction R^2 rises to 0.77. In the other columns of Table 8, we change the forecasting target. Naturally, cash-flow predictability declines in the other cases because the predictor (IGA analyst forecast, in particular) targets earnings growth of the IGV-covered firms rather than dividend growth (columns 4-6) or earnings growth of Compustat firms (columns 7-9). Overall, our findings suggest that market participants are likely to be well-informed about near-term cash-flow growth. One explanation is that firms tend to provide forward guidance on their earnings outlook, increasingly so in recent years.⁴²

Next, we examine the predictability of growth at longer horizons. We add the LTG forecast as a predictor, i.e., forecast of growth over three to five years, and also include $\mathbb{E}_t^A(\Delta e_{t+1})$, pd , and dr from Table 8. The results are reported in Table 9.⁴³ For comparison, we first predict earnings

⁴²Firms typically offer guidance on near-term performances (Penman, 1980; Miller, 2002; Hutton et al., 2003).

⁴³In Figure A.9, we report the results of alternative predictive models that take advantage of more valuation ratios

Table 9 Forecasting Earnings Growth across Horizons

This table reports the results of regressions that predict earnings growth at various horizons. The dependent variables are realized earnings growth from IGA of next year (columns 1-3), between the first and second years (columns 4-6), and between the second and third years (columns 7-9). The independent variables are analysts' forecasts of one-year earnings growth from IGA ($\mathbb{E}_t^A(\Delta e_{t+1})$), the self-aggregated long-term earnings growth forecasts (LTG_t) of the S&P 500 Index, the slope of S&P 500 valuation term structure dr , and the level pd_t . t -statistics calculated based on Newey-West standard errors with 18 lags are reported in parentheses. Data sample: 1988:01–2019:12.

	IGA $\Delta e_{t,t+1}$			IGA $\Delta e_{t+1,t+2}$			IGA $\Delta e_{t+2,t+3}$		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Intercept	0.053 (1.360)	-0.426 (-1.357)	-0.309 (-2.786)	0.325 (2.446)	-0.063 (-0.211)	-0.037 (-0.141)	0.185 (1.596)	0.513 (1.577)	0.496 (1.685)
$\mathbb{E}_t^A(\Delta e_{t+1})$	1.222 (20.312)		1.174 (15.749)	0.030 (0.192)		-0.047 (-0.348)	-0.319 (-2.594)		-0.274 (-1.947)
LTG_t	-0.889 (-3.171)		-1.344 (-3.092)	-2.155 (-2.186)		-2.295 (-1.578)	-0.711 (-0.768)		-0.295 (-0.234)
dr_t		-0.248 (-2.328)	-0.077 (-1.568)		-0.144 (-2.878)	-0.103 (-2.319)		0.096 (1.449)	0.068 (1.215)
pd_t		0.385 (2.354)	0.189 (2.797)		0.181 (1.690)	0.207 (2.351)		-0.216 (-1.666)	-0.166 (-1.399)
N	372	372	372	360	360	360	348	348	348
R^2	0.74	0.20	0.78	0.08	0.08	0.14	0.07	0.07	0.11

growth over the next year from columns (1) to (3), and consistent with Table 8, predictability is strong. In columns (4) to (6), we predict earnings growth from $t + 1$ to $t + 2$ and find that predictability declines dramatically. Comparing columns (1) and (4), the R^2 declines from 0.74 to 0.08. Our forecasting exercise in columns (7) to (9) delivers the same message. The predictability of growth over an even longer horizon, i.e., from $t + 2$ to $t + 3$, is even weaker.

Our findings implies a cash-flow information cliff at the one-year horizon. Therefore, the slope of valuation term structure computed around this cutoff reveals the expected return: when the term structure steepens or flattens, it is driven by discount-rate variation that affects valuation of long-term cash flows more than that of short-term cash flows and not news on long-term growth. that may contain richer information on state variables than the pair pd and dr . Our conclusion remains robust.

5 Conclusion

An asset can be sliced into strips across payout horizons. Strip valuation ratios form a term structure and map out the underlying state variables that drive the expected return and cash-flow growth of this asset. In particular, the level and slope of valuation term structure form a pair of state-variable proxies. For the equity market index, our approach reveals strong return predictability and cash-flow predictability at the annual horizon. We also find that beyond the very next year, our state-variable proxies and survey expectations fail to predict cash-flow growth. Such information cliff suggests steepening (flattening) of the valuation term structure does not reflect information on the growth trajectory but is due to a decrease (increase) in the discount rate or driven by exuberance (pessimism) about long-run growth. This points to using the slope of valuation term structure as the return predictor rather than the level (i.e., the traditional price-dividend ratio of the equity index).

Our paper challenges the traditional approach to return and cash-flow prediction. For return prediction, an asset's own valuation ratio—the level of the asset's valuation term structure—is not the optimal predictor; instead, one should identify the cash-flow information cliff and compute the slope around the cutoff. For cash-flow prediction, the level and slope perform well together.

Our method can be applied to any asset that can be sliced along payout horizons. Recent studies show that strip prices can be computed for individual stocks and certain assets that are not publicly traded.⁴⁴ When applying our method, the first step is to examine the dimensionality of state space by analyzing strip valuation ratios that proxy for state variables. Next, state-variable proxies, such as the level and slope of valuation term structure, can be used for forecasting. For return prediction, one can further examine the cash-flow information cliff either directly from evidence on cash-flow predictability at different horizons or via sharper tests based on empirically motivated state-space models.⁴⁵ After identifying the information cliff, the slope of valuation term structure is computed around this cutoff and then used to predict the return of the asset.

⁴⁴Strip prices of a stock can be calculated from options or single-stock dividend futures and swaps (Gormsen and Lazarus, 2023). Strip values can also be computed for private equity (Gupta and Van Nieuwerburgh, 2021).

⁴⁵For the S&P 500 index, it suffices to consider two state variables that follow AR(1) processes and correspond to expectations of return and dividend growth rate, respectively. For assets with higher state-space dimensionality, AR(p) processes can be considered, and the test for information cliff involves estimating the AP(p) coefficients.

References

- Adam, Klaus, and Stefan Nagel, 2023, Chapter 16 - expectations data in asset pricing, in Rüdiger Bachmann, Giorgio Topa, and Wilbert van der Klaauw, eds., *Handbook of Economic Expectations*, 477–506 (Academic Press).
- Adrian, Tobias, Erkki Etula, and Tyler Muir, 2014, Financial Intermediaries and the Cross-Section of Asset Returns, *Journal of Finance* LXIX, 1–56.
- Ai, Hengjie, Mariano Max Croce, Anthony M Diercks, and Kai Li, 2018, News Shocks and the Production-Based Term Structure of Equity Returns, *Review of Financial Studies* 31, 2423–2467.
- Amromin, Gene, and Steven A. Sharpe, 2014, From the horse’s mouth: Economic conditions and investor expectations of risk and return, *Management Science* 60, 845–866.
- Andersen, Torben, and Luca Benzoni, 2010, Do bonds span volatility risk in the u.s. treasury market? a specification test for affine term structure models, *Journal of Finance* 65, 603–653.
- Ang, Andrew, and Geert Bekaert, 2007, Stock Return Predictability: Is it There?, *Review of Financial Studies* 20, 651–707.
- Ang, Andrew, and Monika Piazzesi, 2003, A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables, *Journal of Monetary Economics* 50, 745–787.
- Backus, David, Nina Boyarchenko, and Mikhail Chernov, 2018, Term structures of asset prices and returns, *Journal of Financial Economics* 129, 1–23.
- Baker, Malcolm, and Jeffrey Wurgler, 2000, The Equity Share in New Issues and Aggregate Stock Returns, *Journal of Finance* 55, 2219–2257.
- Baker, Malcolm, and Jeffrey Wurgler, 2006, Investor Sentiment and the Cross-Section of Stock Returns, *Journal of Finance* LXI, 1645–1680.
- Baker, Scott R., Nicholas Bloom, and Steven J. Davis, 2016, Measuring Economic Policy Uncertainty, *Quarterly Journal of Economics* 131, 1593–1636.
- 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.
- Bauer, Michael D., and Glenn D. Rudebusch, 2016, Resolving the Spanning Puzzle in Macro-Finance Term Structure Models*, *Review of Finance* 21, 511–553.
- Beeler, Jason, and John Y. Campbell, 2012, The Long-Run Risks Model and Aggregate Asset Prices: An Empirical Assessment, *Critical Finance Review* 1, 141–182.
- Bekaert, Geert, and Steven R. Grenadier, 1999, Stock and Bond Pricing in an Affine Economy, NBER Working Papers 7346, National Bureau of Economic Research, Inc.
- Belo, Frederico, Pierre Collin-Dufresne, and Robert Goldstein, 2015, Dividend Dynamics and the Term Structure of Dividend Strips, *Journal of Finance* LXX, 1115–1160.
- Bernanke, Ben S., and Kenneth N. Kuttner, 2005, What Explains the Stock Market’s Reaction to Federal Reserve Policy?, *Journal of Finance* 60, 1221–1257.

- Bikbov, Ruslan, and Mikhail Chernov, 2009, Unspanned stochastic volatility in affine models: Evidence from eurodollar futures and options, *Management Science* 55, 1292–1305.
- Bikbov, Ruslan, and Mikhail Chernov, 2010, No-arbitrage macroeconomic determinants of the yield curve, *Journal of Econometrics* 159, 166–182.
- Binsbergen, Jules H. van, 2021, Duration-Based Stock Valuation: Reassessing Stock Market Performance and Volatility, Working paper.
- Binsbergen, Jules H. Van, Michael W. Brandt, and Ralph S. J. Koijen, 2012, On the Timing and Pricing of Dividends, *American Economic Review* 102, 1596–1618.
- Binsbergen, Jules H van, Xiao Han, and Alejandro Lopez-Lira, 2022, Man versus Machine Learning: The Term Structure of Earnings Expectations and Conditional Biases, *Review of Financial Studies* 36, 2361–2396.
- Binsbergen, Jules H. Van, Wouter Hueskes, Ralph S. J. Koijen, and Evert B. Vrugt, 2013, Equity yields, *Journal of Financial Economics* 110, 503–519.
- Binsbergen, Jules H. Van, and Ralph S. J. Koijen, 2010, Predictive Regressions: A Present-Value Approach, *Journal of Finance* 65, 1439–1471.
- Binsbergen, Jules H. Van, and Ralph S. J. Koijen, 2017, The term structure of returns: Facts and theory, *Journal of Financial Economics* 124, 1–21.
- Boguth, Oliver, Murray Carlson, Adlai Fisher, and Mikhail Simutin, 2022, The Term Structure of Equity Risk Premia: Levered Noise and New Estimates, *Review of Finance* 27, 1155–1182.
- Bordalo, Pedro, Nicola Gennaioli, Rafael La Porta, and Andrei Shleifer, 2019, Diagnostic Expectations and Stock Returns, *Journal of Finance* 74, 2839–2874.
- Bordalo, Pedro, Nicola Gennaioli, Yueran Ma, and Andrei Shleifer, 2020, Overreaction in macroeconomic expectations, *American Economic Review* 110, 2748–82.
- Bordalo, Pedro, Nicola Gennaioli, Rafael La Porta, and Andrei Shleifer, 2024, Belief overreaction and stock market puzzles, *Journal of Political Economy* 132, 1450–1484.
- Bossaerts, Peter, and Pierre Hillion, 1999, Implementing Statistical Criteria to Select Return Forecasting Models: What Do We Learn?, *Review of Financial Studies* 12, 405–428.
- Bouchaud, Jean-Philippe, Philipp Krüger, Augustin Landier, and David Thesmar, 2019, Sticky Expectations and the Profitability Anomaly, *Journal of Finance* 74, 639–674.
- Bradshaw, Mark T., 2004, How Do Analysts Use Their Earnings Forecasts in Generating Stock Recommendations?, *The Accounting Review* 79, 25–50.
- Brennan, Michael J., Ashley W. Wang, and Yihong Xia, 2004, Estimation and test of a simple model of intertemporal capital asset pricing, *Journal of Finance* 59, 1743–1775.
- Campbell, John Y., and John Ammer, 1993, What Moves the Stock and Bond Markets? A Variance Decomposition for Long-Term Asset Returns, *Journal of Finance* 48, 3–37.
- Campbell, John Y., and Robert J. Shiller, 1988, The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors, *Review of Financial Studies* 1, 195–228.
- Campbell, John Y., and Robert J. Shiller, 1991, Yield Spreads and Interest Rate Movements: A Bird’s Eye View, *The Review of Economic Studies* 58, 495–514.

- Campbell, John Y., and Samuel B. Thompson, 2008, Predicting Excess Stock Returns Out of Sample: Can Anything Beat the Historical Average?, *Review of Financial Studies* 21, 1509–1531.
- Cejnek, Georg, and Otto Randl, 2016, Risk and return of short-duration equity investments, *Journal of Empirical Finance* 36, 181–198.
- Cejnek, Georg, and Otto Randl, 2020, Dividend Risk Premia, *Journal of Financial and Quantitative Analysis* 55, 1199–1242.
- Cejnek, Georg, Otto Randl, and Josef Zechner, 2021, The covid-19 pandemic and corporate dividend policy, *Journal of Financial and Quantitative Analysis* 56, 2389–2410.
- Charles, Constantin, Cary Frydman, and Mete Kilic, 2023, Insensitive Investors, Working paper, University of Southern California.
- Chauvet, Marcelle, and Jeremy Piger, 2008, A comparison of the real-time performance of business cycle dating methods, *Journal of Business & Economic Statistics* 26, 42–49.
- Chen, Long, Zhi Da, and Xinlei Zhao, 2013, What Drives Stock Price Movements?, *Review of Financial Studies* 26, 841–876.
- Chen, Yong, Zhi Da, and Dayong Huang, 2022, Short selling efficiency, *Journal of Financial Economics* 145, 387–408.
- Cieslak, Anna, and Hao Pang, 2021, Common Shocks in Stocks and Bonds, *Journal of Financial Economics* 142, 880–904.
- Cieslak, Anna, and Pavol Povala, 2015, Expected returns in treasury bonds, *Review of Financial Studies* 28, 2859–2901.
- Cieslak, Anna, and Pavol Povala, 2016, Information in the term structure of yield curve volatility, *Journal of Finance* 71, 1393–1436.
- Clark, Todd E., and Michael W. McCracken, 2001, Tests of equal forecast accuracy and encompassing for nested models, *Journal of Econometrics* 105, 85 – 110.
- Clark, Todd E., and Kenneth D. West, 2007, Approximately normal tests for equal predictive accuracy in nested models, *Journal of Econometrics* 138, 291 – 311.
- Cochrane, John H., 2008a, State-space vs. VAR models for stock returns, Technical report.
- Cochrane, John H., 2008b, The Dog That Did Not Bark: A Defense of Return Predictability, *Review of Financial Studies* 21, 1533–1575.
- Cochrane, John H., 2011, Presidential Address: Discount Rates, *Journal of Finance* 66, 1047–1108.
- Cochrane, John H., and Monika Piazzesi, 2005, Bond Risk Premia, *American Economic Review* 95, 138–160.
- Coibion, Olivier, and Yuriy Gorodnichenko, 2015, Information rigidity and the expectations formation process: A simple framework and new facts, *American Economic Review* 105, 2644–78.
- Collin-Dufresne, Pierre, and Robert S. Goldstein, 2002, Do bonds span the fixed income markets? theory and evidence for unspanned stochastic volatility, *Journal of Finance* 57, 1685–1730.

- Collin-Dufresne, Pierre, Robert S. Goldstein, and Christopher S. Jones, 2009, Can interest rate volatility be extracted from the cross section of bond yields?, *Journal of Financial Economics* 94, 47–66.
- Collin-Dufresne, Pierre, Michael Johannes, and Lars A. Lochstoer, 2016, Parameter Learning in General Equilibrium: The Asset Pricing Implications, *American Economic Review* 106, 664–98.
- Collin-Dufresne, Pierre, Robert S. Goldstein, and Christopher S. Jones, 2008, Identification of Maximal Affine Term Structure Models, *Journal of Finance* 63, 743–795.
- Cooper, Ilan, and Richard Priestley, 2008, Time-Varying Risk Premiums and the Output Gap, *Review of Financial Studies* 22, 2801–2833.
- Cooper, Rick A., Theodore E. Day, and Craig M. Lewis, 2001, Following the leader: A study of individual analysts' earnings forecasts, *Journal of Financial Economics* 61, 383 – 416.
- Copeland, Tom, Aaron Dolgoff, and Alberto Moel, 2004, The Role of Expectations in Explaining the Cross-Section of Stock Returns, *Review of Accounting Studies* 9, 149–188.
- Crump, Richard K., Stefano Eusepi, and Emanuel Moench, 2016, The term structure of expectations and bond yields, Staff Reports 775, Federal Reserve Bank of New York.
- Da, Zhi, and Mitch Warachka, 2011, The disparity between long-term and short-term forecasted earnings growth, *Journal of Financial Economics* 100, 424–442.
- D'Acunto, Francesco, and Michael Weber, 2024, Why survey-based subjective expectations are meaningful and important, Working Paper 32199, National Bureau of Economic Research.
- Dai, Qiang, and Kenneth J. Singleton, 2000, Specification analysis of affine term structure models, *Journal of Finance* 55, 1943–1978.
- De La O, Ricardo, and Sean Myers, 2021, Subjective Cash Flow and Discount Rate Expectations, *Journal of Finance* 76, 1339–1387.
- De la O, Ricardo, and Sean Myers, 2024, Which Subjective Expectations Explain Asset Prices?, *Review of Financial Studies* 37, 1929–1978.
- Dechow, Patricia M., and Richard G. Sloan, 1997, Returns to contrarian investment strategies: Tests of naive expectations hypotheses, *Journal of Financial Economics* 43, 3–27.
- Duffee, Gregory R., 2002, Term premia and interest rate forecasts in affine models, *Journal of Finance* 57, 405–443.
- Duffee, Gregory R., 2011, Information in (and not in) the Term Structure, *Review of Financial Studies* 24, 2895–2934.
- Duffee, Gregory R., 2013, Chapter 13 - bond pricing and the macroeconomy, volume 2 of *Handbook of the Economics of Finance*, 907–967 (Elsevier).
- Duffie, Darrell, 2001, *Dynamic Asset Pricing Theory*, Princeton Series in Finance, third edition (Princeton University Press, Princeton, N.J.).
- Duffie, Darrell, and Rui Kan, 1996, A Yield-Factor Model of Interest Rates, *Mathematical Finance* 6, 379–406.
- Duffie, Darrell, Jun Pan, and Kenneth Singleton, 2000, Transform analysis and asset pricing for affine jump-diffusions, *Econometrica* 68, 1343–1376.

- Eraker, Bjørn, 2008, Affine general equilibrium models, *Management Science* 54, 2068–2080.
- Fama, Eugene F., and Kenneth R. French, 1988a, Dividend yields and expected stock returns, *Journal of Financial Economics* 22, 3–25.
- Fama, Eugene F., and Kenneth R. French, 1988b, Permanent and temporary components of stock prices, *Journal of Political Economy* 96, 246–273.
- Farmer, Leland, Emi Nakamura, and Jon Steinsson, forthcoming, Learning about the long run, *Journal of Political Economy* .
- Feldhütter, Peter, Christian Heyerdahl-Larsen, and Philipp Illeditsch, 2016, Risk Premia and Volatilities in a Nonlinear Term Structure Model*, *Review of Finance* 22, 337–380.
- Feng, Guanhao, Stefano Giglio, and Dacheng Xiu, 2020, Taming the factor zoo: A test of new factors, *Journal of Finance* 75, 1327–1370.
- Ferson, Wayne E., Sergei Sarkissian, and Timothy T. Simin, 2003, Spurious Regressions in Financial Economics?, *Journal of Finance* 58, 1393–1413.
- Foster, F. Douglas, Tom Smith, and Robert E. Whaley, 1997, Assessing Goodness-Of-Fit of Asset Pricing Models: The Distribution of the Maximal R^2 , *Journal of Finance* 52, 591–607.
- Gabaix, Xavier, 2019, A Behavioral New Keynesian Model, Working Paper.
- 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, Bryan Kelly, and Serhiy Kozak, forthcoming, Equity Term Structures without Dividend Strips Data, *Journal of Finance* .
- Giglio, Stefano, Matteo Maggiori, Johannes Stroebel, and Stephen Utkus, 2021, Five facts about beliefs and portfolios, *American Economic Review* 111, 1481–1522.
- Gilchrist, Simon, and Egon Zakrajšek, 2012, Credit Spreads and Business Cycle Fluctuations, *American Economic Review* 102, 1692–1720.
- Golez, Benjamin, 2014, Expected Returns and Dividend Growth Rates Implied by Derivative Markets, *Review of Financial Studies* 27, 790–822.
- Golez, Benjamin, and Jens Jackwerth, 2024, Holding Period Effects in Dividend Strip Returns, *Review of Financial Studies* hhae002.
- Golez, Benjamin, and Peter Koudijs, 2018, Four centuries of return predictability, *Journal of Financial Economics* 127, 248 – 263.
- Golez, Benjamin, and Peter Koudijs, 2023, Equity Duration and Predictability, Technical report, University of Notre Dame.
- Gonçalves, Andrei S., 2019, What Moves Equity Markets? A Term Structure Decomposition for Stock Returns, Kenan institute of private enterprise research paper, University of North Carolina at Chapel Hill.
- Gonçalves, Andrei S., 2021, Reinvestment risk and the equity term structure, *Journal of Finance* 76, 2153–2197.
- Gormsen, Niels Joachim, 2021, Time Variation of the Equity Term Structure, *Journal of Finance* 76, 1959–1999.

- Gormsen, Niels Joachim, and Ralph S J Koijen, 2020, Coronavirus: Impact on Stock Prices and Growth Expectations, *The Review of Asset Pricing Studies* 10, 574–597.
- Gormsen, Niels Joachim, and Eben Lazarus, 2023, Duration-Driven Returns, *Journal of Finance* 78, 1393–1447.
- Goyal, Amit, and Ivo Welch, 2007, A Comprehensive Look at The Empirical Performance of Equity Premium Prediction, *Review of Financial Studies* 21, 1455–1508.
- Gu, Zhaoyang, and Joanna Shuang Wu, 2003, Earnings skewness and analyst forecast bias, *Journal of Accounting and Economics* 35, 5–29.
- Gupta, Arpit, and Stijn Van Nieuwerburgh, 2021, Valuing private equity investments strip by strip, *Journal of Finance* 76, 3255–3307.
- Hansen, Lars Peter, John C. Heaton, and Nan Li, 2008, Consumption Strikes Back? Measuring Long-Run Risk, *Journal of Political Economy* 116, 260–302.
- Hasler, Michael, and Mariana Khapko, 2023, Correlated cashflow shocks, asset prices, and the term structure of equity, *Management Science* 69, 5560–5577.
- Hasler, Michael, and Roberto Marfè, 2016, Disaster recovery and the term structure of dividend strips, *Journal of Financial Economics* 122, 116–134.
- He, Zhiguo, Bryan Kelly, and Asaf Manela, 2017, Intermediary asset pricing: New evidence from many asset classes, *Journal of Financial Economics* .
- Hillenbrand, Sebastian, and Odhrain McCarthy, 2022, The Optimal Stock Valuation Ratio, working paper, New York University.
- Hodrick, Robert J., 1992, Dividend Yields and Expected Stock Returns: Alternative Procedures for Inference and Measurement, *Review of Economic Dynamics* 5, 357–386.
- Hutton, Amy P., Gregory S. Miller, and Douglas J. Skinner, 2003, The role of supplementary statements with management earnings forecasts, *Journal of Accounting Research* 41, 867–890.
- Jagannathan, Ravi, and Binying Liu, 2018, Dividend Dynamics, Learning, and Expected Stock Index Returns, *Journal of Finance* 74, 401–448.
- Johnson, Travis L, 2019, A Fresh Look at Return Predictability Using a More Efficient Estimator, *Review of Asset Pricing Studies* 9, 1–46.
- Joslin, Scott, Marcel Pribsch, and Kenneth J. Singleton, 2014, Risk premiums in dynamic term structure models with unspanned macro risks, *Journal of Finance* 69, 1197–1233.
- Kelly, Bryan, Semyon Malamud, and Lasse Heje Pedersen, 2023, Principal portfolios, *Journal of Finance* 78, 347–387.
- Kelly, Bryan, Semyon Malamud, and Kangying Zhou, 2024, The virtue of complexity in return prediction, *Journal of Finance* 79, 459–503.
- Kelly, Bryan, and Seth Pruitt, 2013, Market Expectations in the Cross-Section of Present Values, *Journal of Finance* 68, 1721–1756.
- Kelly, Bryan, and Seth Pruitt, 2015, The three-pass regression filter: A new approach to forecasting using many predictors, *Journal of Econometrics* 186, 294–316.

- Knox, Benjamin, and Annette Vissing-Jørgensen, 2022, A Stock Return Decomposition Using Observables, Finance and economics discussion series, Federal Reserve Board of Governors.
- Koijen, Ralph S. J., Hanno Lustig, and Stijn Van Nieuwerburgh, 2015, The cross-section and time-series of stock and bond returns.
- Koijen, Ralph S. J., and Stijn Van Nieuwerburgh, 2011, Predictability of Returns and Cash Flows, *Annual Review of Financial Economics* 3, 467–491.
- Kostakis, Alexandros, Tassos Magdalinos, and Michalis P. Stamatogiannis, 2014, Robust econometric inference for stock return predictability, *Review of Financial Studies* 28, 1506–1553.
- Kothari, S.P., Jonathan Lewellen, and Jerold B. Warner, 2006, Stock returns, aggregate earnings surprises, and behavioral finance, *Journal of Financial Economics* 79, 537–568.
- 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* .
- Kragt, Jac, Frank de Jong, and Joost Driessen, 2020, The dividend term structure, *Journal of Financial and Quantitative Analysis* 55, 829–867.
- La Porta, Rafael, 1996, Expectations and the Cross-Section Returns of Stock, *Journal of Finance* 51, 1715–1742.
- Lacerda, Filipe, and Pedro Santa-Clara, 2010, Forecasting Dividend Growth to Better Predict Returns, Working paper, Universidade Nova de Lisboa.
- Larrain, Borja, and Motohiro Yogo, 2008, Does firm value move too much to be justified by subsequent changes in cash flow?, *Journal of Financial Economics* 87, 200 – 226.
- Lettau, Martin, and Sydney Ludvigson, 2001, Consumption, Aggregate Wealth, and Expected Stock Returns, *Journal of Finance* 56, 815–849.
- Lettau, Martin, and Sydney C. Ludvigson, 2005, Expected returns and expected dividend growth, *Journal of Financial Economics* 76, 583 – 626.
- Lettau, Martin, and Stijn Van Nieuwerburgh, 2007, Reconciling the return predictability evidence, *Review of Financial Studies* 21, 1607–1652.
- Lettau, Martin, and Jessica A. Wachter, 2007, Why Is Long-Horizon Equity Less Risky? A Duration-Based Explanation of the Value Premium, *Journal of Finance* 62, 55–92.
- Lettau, Martin, and Jessica A. Wachter, 2011, The term structures of equity and interest rates, *Journal of Financial Economics* 101, 90–113.
- Lewellen, Jonathan, 2004, Predicting returns with financial ratios, *Journal of Financial Economics* 74, 209–235.
- Li, Haitao, and Feng Zhao, 2006, Unspanned stochastic volatility: Evidence from hedging interest rate derivatives, *Journal of Finance* 61, 341–378.

- Litterman, Robert B., and José Scheinkman, 1991, Common factors affecting bond returns, *Journal of Fixed Income* 0, 54–61.
- Ludvigson, Sydney C., and Serena Ng, 2009, Macro Factors in Bond Risk Premia, *Review of Financial Studies* 22, 5027–5067.
- Malmendier, Ulrike, and Devin Shanthikumar, 2007, Are small investors naive about incentives?, *Journal of Financial Economics* 85, 457–489.
- Malmendier, Ulrike, and Devin Shanthikumar, 2014, Do security analysts speak in two tongues?, *Review of Financial Studies* 27, 1287–1322.
- Martin, Ian, 2017, What is the Expected Return on the Market?, *Quarterly Journal of Economics* 132, 367–433.
- McCarthy, Odhrain, and Sebastian Hillenbrand, 2021, Heterogeneous investors and stock market fluctuations, Working paper, Harvard University.
- Mikhail, Michael B., Beverly R. Walther, and Richard H. Willis, 1999, Does Forecast Accuracy Matter to Security Analysts?, *The Accounting Review* 185–200, 2.
- Miller, Gregory S., 2002, Earnings performance and discretionary disclosure, *Journal of Accounting Research* 40, 173–204.
- Miller, Shane, 2018, The Term Structures of Equity Risk Premia in the Cross-Section of Equities, Working paper, Duke University Fuqua School of Business.
- Nagel, Stefan, and Zhengyang Xu, 2022, Asset Pricing with Fading Memory, *Review of Financial Studies* 35, 2190–2245.
- Nelson, Charles R., and Myung J. Kim, 1993, Predictable Stock Returns: The Role of Small Sample Bias, *Journal of Finance* 48, 641–661.
- Newey, Whitney K., and Kenneth D. West, 1987, A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix, *Econometrica* 55, 703–708.
- Pan, Jun, 2002, The jump-risk premia implicit in options: evidence from an integrated time-series study, *Journal of Financial Economics* 63, 3–50.
- Pang, Hao, 2023, The term structure of inflation expectations and treasury yields, Working paper, Duke University.
- Pástor, Ľuboš, Meenakshi Sinha, and Bhaskaran Swaminathan, 2008, Estimating the Intertemporal Risk – Return Tradeoff Using the Implied Cost of Capital, *Journal of Finance* 63, 2859–2897.
- Patelis, Alex D., 1997, Stock Return Predictability and The Role of Monetary Policy, *Journal of Finance* 52, 1951–1972.
- Penman, Stephen H., 1980, An empirical investigation of the voluntary disclosure of corporate earnings forecasts, *Journal of Accounting Research* 18, 132–160.
- Pettenuzzo, Davide, Riccardo Sabbatucci, and Allan Timmermann, 2020, Cash flow news and stock price dynamics, *Journal of Finance* 75, 2221–2270.
- Piazzesi, Monika, Juliana Salomao, and Martin Schneider, 2015, Trend and Cycle in Bond Premia, working paper, Stanford University and University of Minnesota.

- Piotroski, Joseph D., and Eric C. So, 2012, Identifying Expectation Errors in Value/Glamour Strategies: A Fundamental Analysis Approach, *Review of Financial Studies* 25, 2841–2875.
- Presidential Task Force on Market Mechanisms (Chairman: Nicholas Brady), 1988, *Report of the Presidential Task Force on Market Mechanisms: submitted to The President of the United States, The Secretary of the Treasury, and The Chairman of the Federal Reserve Board* (U.S. Government Printing Office, Washington, D.C.).
- Pruitt, Seth, 2023, Dogs and Cats Living Together: A Defense of Cash-Flow Predictability, Working paper.
- Rapach, David E., Matthew C. Ringgenberg, and Guofu Zhou, 2016, Short interest and aggregate stock returns, *Journal of Financial Economics* 121, 46 – 65.
- Rapach, David E., Jack K. Strauss, and Guofu Zhou, 2010, Out-of-Sample Equity Premium Prediction: Combination Forecasts and Links to the Real Economy, *Review of Financial Studies* 23, 821–862.
- Rapach, David E., Jack K. Strauss, and Guofu Zhou, 2013, International stock return predictability: What is the role of the united states?, *Journal of Finance* 68, 1633–1662.
- Rytchkov, Oleg, 2012, Filtering Out Expected Dividends and Expected Returns, *Quarterly Journal of Finance* 02, 1250012.
- Sabbatucci, Riccardo, 2022, Are Dividends and Stock Returns Predictable? New Evidence Using M&A Cash Flows, Technical report, Stockholm School of Economics.
- Schmidt-Engelbertz, Paul, and Kaushik Vasudevan, 2023, Speculating on Higher Order Beliefs, Working paper, Yale University.
- Schulz, Florian, 2016, On the timing and pricing of dividends: Comment, *American Economic Review* 106, 3185–3223.
- Sims, Christopher A., 1980, Macroeconomics and Reality, *Econometrica* 48, 1–48.
- Song, Yang, 2016, Dealer funding costs: Implications for the term structure of dividend risk premia, working paper, University of Washington.
- Stambaugh, Robert F., 1999, Predictive Regressions, *Journal of Financial Economics* 54, 375–421.
- Vuolteenaho, Tuomo, 2002, What drives firm-level stock returns, *Journal of Finance* LVII, 233–264.
- Wang, Chen, 2020, Under- and Over-Reaction in Yield Curve Expectations, Working Paper.
- Wang, George H. K., Raphael J. Michalski, James V. Jordan, and Eugene J. Moriarty, 1994, An intraday analysis of Bid-Ask spreads and price volatility in the S&P 500 index futures market, *Journal of Futures Markets* 14, 837–859.

Internet Appendix

Appendix I: Derivation

I.1 Solving the valuation ratios

The price-dividend ratio of the dividend strip with maturity n , $P_{n,t}/D_t$, satisfies the following recursive equation

$$\frac{P_{n,t}}{D_t} = \mathbb{E}_t \left[M_{t+1} \frac{D_{t+1}}{D_t} \frac{P_{n-1,t+1}}{D_{t+1}} \right]. \quad (\text{A.1})$$

We conjecture that

$$\ln \left(\frac{P_{n,t}}{D_t} \right) = A(n) + B(n)^T X_t. \quad (\text{A.2})$$

Substituting this expression and expressions of stochastic discount factor and dividend growth into the recursive equation, we have

$$\begin{aligned} & \exp \{A(n) + B(n)^T X_t\} \\ = & \mathbb{E}_t \left[\exp \left\{ -r_f - \frac{1}{2} \lambda_t^T \Sigma \lambda_t - \lambda_t^T \epsilon_{t+1} + g_t + \sigma_D^T \epsilon_{t+1} + A(n-1) + B(n-1)^T X_{t+1} \right\} \right] \\ = & \mathbb{E}_t \left[\exp \left\{ g_t - r_f - \frac{1}{2} \lambda_t^T \Sigma \lambda_t + A(n-1) + B(n-1)^T \Pi X_t + (\sigma_D - \lambda_t + \sigma_X B(n-1))^T \epsilon_{t+1} \right\} \right] \\ = & \exp \left\{ g_t - r_f - \frac{1}{2} \lambda_t^T \Sigma \lambda_t + A(n-1) + B(n-1)^T \Pi X_t \right. \\ & \left. + \frac{1}{2} (\sigma_D - \lambda_t + \sigma_X B(n-1))^T \Sigma (\sigma_D - \lambda_t + \sigma_X B(n-1)) \right\} \\ = & \exp \left\{ g_t - r_f + A(n-1) + B(n-1)^T \Pi X_t - (\sigma_D + \sigma_X B(n-1))^T \Sigma \lambda_t \right. \\ & \left. + \frac{1}{2} (\sigma_D + \sigma_X B(n-1))^T \Sigma (\sigma_D + \sigma_X B(n-1)) \right\} \end{aligned} \quad (\text{A.3})$$

The coefficients on X_t should match $B(n)$ on the left hand side, so we have

$$B(n) = (\Pi^T - \theta \Sigma \sigma_X) B(n-1) + \phi - \theta \Sigma \sigma_D. \quad (\text{A.4})$$

The constants must sum up to $A(n)$ on the left hand side, so we have

$$\begin{aligned} A(n) = & A(n-1) + \bar{g} - \bar{r} - (\sigma_D + \sigma_X B(n-1))^T \Sigma \bar{\lambda} + \\ & \frac{1}{2} (\sigma_D + \sigma_X B(n-1))^T \Sigma (\sigma_D + \sigma_X B(n-1)). \end{aligned} \quad (\text{A.5})$$

The fact that $P_t^0 = D_t$ implies the boundary conditions, $A(0) = B(0) = 0$, which pins down a solution of $A(n)$ and $B(n)$.

Finally, we solve the log price-dividend ratio of the aggregate stock market. We conjecture

$$pd_t = \ln(P_t/D_t) = A + B^T X_t, \quad (\text{A.6})$$

and proceed to solve A and B . Following [Campbell and Shiller \(1988\)](#), we log-linearize the stock market return

$$\begin{aligned} r_{t+1}^{mkt} &= \kappa_0 + \kappa_1 pd_{t+1} - pd_t + \Delta d_{t+1} \\ &= \kappa_0 - (1 - \kappa_1) A - B^T (\mathbf{I} - \kappa_1 \Pi) X_t + g_t + (\kappa_1 \sigma_X B + \sigma_D)^T \epsilon_{t+1} \end{aligned} \quad (\text{A.7})$$

Under the no-arbitrage condition, we have

$$1 = \mathbb{E}_t [M_{t+1} \exp(r_{t+1}^{mkt})]. \quad (\text{A.8})$$

We follow the same method of matching undetermined coefficients in the analysis of dividend strip valuation ratios and solve

$$A = \frac{1}{1 - \kappa_1} \left[\bar{g} - \bar{r} + \kappa_0 - (\kappa_1 \sigma_X B + \sigma_D)^T \Sigma \bar{\lambda} + \frac{1}{2} (\kappa_1 \sigma_X B)^T \Sigma (\kappa_1 \sigma_X B) + (\kappa_1 \sigma_X B)^T \Sigma \sigma_D \right] \quad (\text{A.9})$$

$$B = (\mathbf{I} - \kappa_1 \Pi^T - \kappa_1 \theta \Sigma \sigma_X)^{-1} (\phi - \theta \Sigma \sigma_D - \gamma). \quad (\text{A.10})$$

I.2 Valuation ratios from the cross section

Consider an individual stock i . The dividend dynamics of firm i depend not only on the aggregate state variables, X_t , but also on the firm i -specific state variables, $Z_{i,t}$, that is K_i -dimensional and independent from X_t . Without loss of generality, we assume that $Z_{i,t}$ evolves as a first-order vector autoregression

$$Z_{i,t+1} = \Omega Z_{i,t} + \sigma_{i,Z}^T v_{i,t+1}, \quad (\text{A.11})$$

where $v_{i,t+1}$ is a N_i -by-1 vector of i -specific news that has a normal distribution $N(\mathbf{0}, \Sigma_i)$ and is independent over time and independent from the aggregate shocks ϵ_{t+1} . We use subscript i to differentiate firm i from the aggregate variables (without subscript i) and other firms (with subscript $j \neq i$).

The dividend growth rate of firm i loads on the aggregate and idiosyncratic shocks

$$\ln \left(\frac{D_{i,t+1}}{D_{i,t}} \right) = g_{i,t} + \sigma_{i,D}^T \epsilon_{t+1} + \sigma_{i,v}^T v_{i,t+1}, \quad (\text{A.12})$$

where the expected dividend growth rate is given by

$$g_{i,t} = \phi_i^T X_t + \delta_i^T Z_{i,t} + \bar{g}_i - \frac{1}{2} \sigma_{i,D}^T \Sigma \sigma_{i,D} - \frac{1}{2} \sigma_{i,v}^T \Sigma_i \sigma_{i,v}, \quad (\text{A.13})$$

which loads on the aggregate state variables, X_t , and firm i -specific state variables, $Z_{i,t}$.

The ratio of firm i 's dividend strip price, $P_{i,t}^n$, to firm i 's current dividend is

$$\frac{P_{i,t}^n}{D_{i,t}} = \exp \left\{ A_i(n) + B_i(n)^\top X_t + C_i(n)^\top Z_{i,t} \right\}, \quad (\text{A.14})$$

where $A_i(n)$, $B_i(n)$, and $C_i(n)$ are firm i -specific, deterministic functions of n given by the recursive equations

$$B_i(n) = (\Pi^\top - \theta \Sigma \sigma_X) B_i(n-1) + \phi_i - \gamma - \theta \Sigma \sigma_{i,D}. \quad (\text{A.15})$$

$$C_i(n) = \Omega^\top C_i(n-1) + \delta_i \quad (\text{A.16})$$

$$A_i(n) = A_i(n-1) + \bar{g}_i - \bar{r} - (\sigma_{i,D} + \sigma_X B_i(n-1))^\top \Sigma \bar{\lambda} + \frac{1}{2} (\sigma_{i,D} + \sigma_X B_i(n-1))^\top \Sigma (\sigma_{i,D} + \sigma_X B_i(n-1)) + \frac{1}{2} (\sigma_{i,v} + \sigma_{i,Z} C_i(n-1))^\top \Sigma_i (\sigma_{i,v} + \sigma_{i,Z} C_i(n-1)). \quad (\text{A.17})$$

with the initial conditions

$$A_i(0) = 0, \quad B_i(0) = 0, \quad \text{and} \quad C_i(0) = 0. \quad (\text{A.18})$$

The price of firm i 's stock, $P_{i,t}$, is the sum of all its dividend strips

$$\frac{P_{i,t}}{D_{i,t}} = \sum_{n=1}^{+\infty} \frac{P_{i,t}^n}{D_{i,t}} = \sum_{n=1}^{+\infty} \exp \left\{ A_i(n) + B_i(n)^\top X_t + C_i(n)^\top Z_{i,t} \right\}. \quad (\text{A.19})$$

In Appendix I, we use the log-linearization method of [Campbell and Shiller \(1988\)](#) to solve an approximate exponential-affine form, so the log price-dividend ratio of stock i is

$$\ln \left(\frac{P_{i,t}}{D_{i,t}} \right) \approx A_i + B_i^\top X_t + C_i^\top Z_{i,t}. \quad (\text{A.20})$$

Because $Z_{i,t}$ is independent from X_t , recovering the state space X_t using individual stocks' price-dividend ratio brings in noise. In a forecasting context, [Kelly and Pruitt \(2013\)](#) deal with this issue using partial least squares, which is a method to compress the cross-section of valuation ratios into signals (about the state variables) that are most relevant for the forecasting targets.

I.3 Solving the two-dimensional state space model

We conjecture that the market price-dividend ratio is exponential-affine in the state variables, so the log ratio is

$$pd_t = \ln(S_t/D_t) = A + B y_t + C z_t.$$

Next, we use the log-linearization of [Campbell and Shiller \(1988\)](#), i.e.,

$$r_{t+1} = \kappa_0 + \kappa_1 pd_{t+1} - pd_t + \Delta d_{t+1},$$

and substitute this log market return into the no-arbitrage condition

$$\mathbb{E}_t [M_{t+1} \exp\{r_{t+1}\}] = 1.$$

to obtain

$$\mathbb{E}_t \left[\exp \left\{ -r_f - \frac{1}{2} \lambda_t^2 (\sigma_\lambda^\top \Sigma \sigma_\lambda)^2 - \lambda_t \sigma_\lambda^\top \epsilon_{t+1} + \kappa_0 + \kappa_1 p d_{t+1} - p d_t + \Delta d_{t+1} \right\} \right] = 1 \quad (\text{A.21})$$

Using the conjecture of $p d_t$ and $p d_{t+1}$ and the specification of g_t and Δd_{t+1} , we obtain

$$\mathbb{E}_t \left[\exp \left\{ -r_f - \frac{1}{2} \lambda_t^2 (\sigma_\lambda^\top \Sigma \sigma_\lambda)^2 - \lambda_t \sigma_\lambda^\top \epsilon_{t+1} + \kappa_0 - A - B y_t - C z_t + z_t + \bar{g} - \frac{1}{2} \sigma_D^\top \Sigma \sigma_D + \sigma_D^\top \epsilon_{t+1} + \kappa_1 A + \kappa_1 B (\rho_y y_t + \sigma_y^\top \epsilon_{t+1}) + \kappa_1 C (\rho_z z_t + \sigma_z^\top \epsilon_{t+1}) \right\} \right] = 1 \quad (\text{A.22})$$

For the conjecture of $p d_t$ functional form to hold, the coefficient on z_t is zero, so we obtain

$$C = \frac{1}{1 - \kappa_1 \rho_z} \quad (\text{A.23})$$

Collecting all terms with shocks at $t + 1$ and using the moment-generating function, we obtain

$$\mathbb{E}_t \left[\exp \left\{ -\lambda_t \sigma_\lambda^\top \epsilon_{t+1} + \sigma_D^\top \epsilon_{t+1} + \kappa_1 B \sigma_y^\top \epsilon_{t+1} + \kappa_1 C \sigma_z^\top \epsilon_{t+1} \right\} \right] = \exp \left\{ \frac{1}{2} \lambda_t^2 (\sigma_\lambda^\top \Sigma \sigma_\lambda)^2 - (\sigma_D + \kappa_1 B \sigma_y + \kappa_1 C \sigma_z)^\top \Sigma \sigma_\lambda \lambda_t + \frac{1}{2} (\sigma_D + \kappa_1 B \sigma_y + \kappa_1 C \sigma_z)^\top \Sigma (\sigma_D + \kappa_1 B \sigma_y + \kappa_1 C \sigma_z) \right\} \quad (\text{A.24})$$

Substituting this expression into the no-arbitrage condition, we obtain

$$\exp \left\{ -r_f + \kappa_0 - A - B y_t - C z_t + z_t + \bar{g} - \frac{1}{2} \sigma_D^\top \Sigma \sigma_D - (\sigma_D + \kappa_1 B \sigma_y + \kappa_1 C \sigma_z)^\top \Sigma \sigma_\lambda (\bar{\lambda} + y_t) + \kappa_1 A + \kappa_1 B \rho_y y_t + \kappa_1 C \rho_z z_t + \frac{1}{2} (\sigma_D + \kappa_1 B \sigma_y + \kappa_1 C \sigma_z)^\top \Sigma (\sigma_D + \kappa_1 B \sigma_y + \kappa_1 C \sigma_z) \right\} = 1 \quad (\text{A.25})$$

For the conjecture of $p d_t$ functional form to hold, the coefficient on y_t is zero, so we obtain

$$B = -\frac{(\sigma_D + \kappa_1 C \sigma_z)^\top \Sigma \sigma_\lambda}{1 + \kappa_1 \sigma_y^\top \Sigma \sigma_\lambda - \kappa_1 \rho_y} \quad (\text{A.26})$$

Finally, all the constant terms should add up to zero, so we obtain

$$A = \frac{\bar{g} - r_f + \kappa_0 - \frac{1}{2} \sigma_D^\top \Sigma \sigma_D + \frac{1}{2} (\sigma_D + \kappa_1 B \sigma_y + \kappa_1 C \sigma_z)^\top \Sigma (\sigma_D + \kappa_1 B \sigma_y + \kappa_1 C \sigma_z - 2 \sigma_\lambda \bar{\lambda})}{1 - \kappa_1} \quad (\text{A.27})$$

In the main text, to clarify the notations, we use A_{pd} , B_{pd} , and C_{pd} to denote A , B , and C above, respectively.

Next, we solve the time- t log price-dividend ratio of the dividend strip that matures at $t + 1$. The no-arbitrage condition dictates

$$\mathbb{E}_t \left[M_{t+1} \frac{D_{t+1}}{P_t^1} \right] = 1, \quad (\text{A.28})$$

or equivalently

$$\mathbb{E}_t \left[M_{t+1} \frac{D_{t+1}}{D_t} \frac{D_t}{P_t^1} \right] = \mathbb{E}_t \left[M_{t+1} \exp \{ g_t + \sigma_D^\top \epsilon_{t+1} - s_t^1 \} \right] = 1, \quad (\text{A.29})$$

so we obtain

$$\mathbb{E}_t \left[\exp \left\{ -r_f - \frac{1}{2} \lambda_t^2 (\sigma_\lambda^\top \Sigma \sigma_\lambda)^2 - \lambda_t \sigma_\lambda^\top \epsilon_{t+1} + g_t + \sigma_D^\top \epsilon_{t+1} - s_t^1 \right\} \right] = 1. \quad (\text{A.30})$$

We conjecture

$$s_t^1 = A_1 + B_1 y_t + C_1 z_t.$$

Substituting this conjecture, the specification of g_t , and the specification of λ_t into the no-arbitrage condition, we obtain

$$\begin{aligned} \mathbb{E}_t \left[\exp \left\{ -r_f - \frac{1}{2} (\bar{\lambda} + y_t)^2 (\sigma_\lambda^\top \Sigma \sigma_\lambda)^2 - (\bar{\lambda} + y_t) \sigma_\lambda^\top \epsilon_{t+1} \right. \right. \\ \left. \left. + z_t + \bar{g} - \frac{1}{2} \sigma_D^\top \Sigma \sigma_D + \sigma_D^\top \epsilon_{t+1} - A_1 - B_1 y_t - C_1 z_t \right\} \right] = 1. \end{aligned}$$

Using the moment-generating function to simplify the expression, we obtain

$$\exp \left\{ -r_f + z_t + \bar{g} - A_1 - B_1 y_t - C_1 z_t - \sigma_\lambda^\top \Sigma \sigma_D (\bar{\lambda} + y_t) \right\} = 1. \quad (\text{A.31})$$

For the conjecture of s_t^1 functional form to hold, the coefficient of z_t and the coefficient of y_t must be zero, so we obtain

$$C_1 = 1, \quad (\text{A.32})$$

and

$$B_1 = -\sigma_\lambda^\top \Sigma \sigma_D. \quad (\text{A.33})$$

Finally, the constant terms add up to zero, so we obtain

$$A_1 = \bar{g} - r_f - \sigma_\lambda^\top \Sigma \sigma_D \bar{\lambda} \quad (\text{A.34})$$

Finally, we solve the conditional expected market return. First, we start with $\mathbb{E}_t[r_{t+1}] = \kappa_0 + \kappa_1 \mathbb{E}_t[pd_{t+1}] - pd_t + g_t$. Using the expression of pd_{t+1} , pd_t , and g_t , and the specifications of law of motion of z_t and y_t , we obtain

$$\mathbb{E}_t[r_{t+1}] = \kappa_0 - (1 - \kappa_1)A + \bar{g} - \frac{1}{2} \sigma_D^\top \Sigma \sigma_D - (1 - \kappa_1 \rho_y) B y_t. \quad (\text{A.35})$$

We collect the constant terms into A_{er} and define the coefficient of y_t to be B_{er} .

I.4 Proof of Proposition 3 on ρ_z and return forecasting errors

Proof. We know that the expected return is a function of the price of risk y_t :

$$\mathbb{E}_t[r_{t+1}] = A_{er} + B_{er}y_t,$$

and that

$$dr_t = A_{pd} - A_1 + (B_{pd} - B_1)y_t + (C_{pd} - C_1)z_t.$$

Combining the two equations, we have

$$\mathbb{E}_t[r_{t+1}] = A_{er} + \frac{B_{er}}{B_1 - B_{pd}} [dr_t - A_{pd} + A_1 - (C_{pd} - C_1)z_t] \quad (\text{A.36})$$

$$= \text{const.} + \frac{B_{er}}{B_1 - B_{pd}} [dr_t - (C_{pd} - C_1)z_t] \quad (\text{A.37})$$

If $\rho_z = 0$, $\mathbb{E}_t[r_{t+1}] = \text{const.} + \frac{B_{er}}{B_1 - B_{pd}} dr_t$. The forecast error is a white noise independent of time- t variables:

$$v_{t+1} = r_{t+1} - \mathbb{E}_t[r_{t+1}] = \epsilon_{t+1}.$$

However, if $\rho_z \neq 0$ but the investor still uses equation (A.37) to forecast $t + 1$ return, the forecast error is then

$$\begin{aligned} v_{t+1} &= r_{t+1} - \left[\text{const.} + \frac{B_{er}}{B_1 - B_{pd}} dr_t \right] = r_{t+1} - \left[\mathbb{E}_t[r_{t+1}] + \frac{B_{er}(C_{pd} - C_1)}{B_1 - B_{pd}} z_t \right] \\ &= \epsilon_{t+1} - \frac{B_{er}(C_{pd} - C_1)}{B_1 - B_{pd}} z_t = \epsilon_{t+1} - \frac{B_{er}}{B_1 - B_{pd}} \left(\frac{1}{1 - \kappa_1 \rho_z} - 1 \right) z_t. \end{aligned}$$

The correlation between $\hat{\rho}_{z,t}$ and v_{t+1} is therefore

$$\text{Corr}(\rho_{z,t}, v_{t+1}) = -\frac{B_{er}}{B_1 - B_{pd}} \text{Corr} \left(\rho_{z,t}, \left(\frac{1}{1 - \kappa_1 \rho_{z,t}} - 1 \right) z_t \right)$$

Based on our findings on return predictability, dr_t negatively predicts future returns. Therefore, the coefficient of dr_t in equation (A.37), $\frac{B_{er}}{B_1 - B_{pd}}$, is negative. Under this condition, we obtain

$$\begin{aligned} \text{sgn}(\text{Corr}(\rho_{z,t}, v_{t+1})) &= \text{sgn} \left(\text{Cov} \left(\rho_{z,t}, \left(\frac{1}{1 - \kappa_1 \rho_{z,t}} - 1 \right) z_t \right) \right) \\ &= \text{sgn} \left(\mathbb{E} \left(\frac{\kappa_1 \rho_{z,t}^2 z_t}{1 - \kappa_1 \rho_{z,t}} \right) - \mathbb{E}(\rho_{z,t}) \mathbb{E} \left(\frac{\kappa_1 \rho_{z,t} z_t}{1 - \kappa_1 \rho_{z,t}} \right) \right) \end{aligned}$$

As demonstrated by the rolling estimation results in Table 7, $\rho_{z,t}$ on average is close to zero (see also Table 7, we have $\mathbb{E}(\hat{\rho}_{z,t}) \approx 0$). Using 1-year earnings growth forecasts from IBES Global Aggregate (IGA) as a proxy for z_t and $\kappa_1 = 0.98$, we calculate the estimate of $\mathbb{E} \left(\frac{\kappa_1 \hat{\rho}_{z,t}^2 z_t}{1 - \kappa_1 \hat{\rho}_{z,t}} \right)$ in our

sample to be 0.005626 with p -value < 0.01 , which implies

$$\text{sgn}(\text{Corr}(\rho_{z,t}, v_{t+1})) = \text{sgn}\left(\mathbb{E}\left(\frac{\kappa_1 \rho_{z,t}^2 z_t}{1 - \kappa_1 \rho_{z,t}}\right)\right) > 0.$$

■

I.5 Deriving the Sharpe ratio of market-timing strategy

Following [Campbell and Thompson \(2008\)](#), we assume that the excess return can be decomposed as follows:

$$r_{t+1} = \mu + x_t + \varepsilon_{t+1}$$

where μ is the unconditional mean. The predictor x_t has mean 0 and variance σ_x^2 , independent from the error term ε_{t+1} . For simplicity, we assume that the mean-variance investor has a relative risk aversion coefficient $\gamma = 1$. When using x_t to time the market, the investor allocates

$$\alpha_t = \frac{\mu + x_t}{\sigma_\varepsilon^2}$$

to the risky asset and on average earns an excess return of

$$\mathbb{E}(\alpha_t r_{t+1}) = \mathbb{E}\left(\frac{(\mu + x_t)(\mu + x_t + \varepsilon_{t+1})}{\sigma_\varepsilon^2}\right) = \frac{\mu^2 + \sigma_x^2}{\sigma_\varepsilon^2}$$

The variance of the market-timing strategy is

$$\text{Var}(\alpha_t r_{t+1}) = \text{Var}\left[\frac{(\mu + x_t)(\mu + x_t + \varepsilon_{t+1})}{\sigma_\varepsilon^2}\right]$$

The (squared) market-timing Sharpe ratio s_1^2 can be written as

$$s_1^2 = \frac{[\mathbb{E}(\alpha_t r_{t+1})]^2}{\text{Var}(\alpha_t r_{t+1})} = A \cdot \frac{\mu^2 + \sigma_x^2}{\sigma_\varepsilon^2}$$

where A is a constant that depends on $\text{Var}[(\mu + x_t)(\mu + x_t + \varepsilon_{t+1})]$ and $(\mu^2 + \sigma_x^2)/\sigma_\varepsilon^2$.

Given the buy-and-hold Sharpe ratio s_0 ,

$$s_0^2 = \frac{\mu^2}{\sigma_x^2 + \sigma_\varepsilon^2}$$

and the predictive regression R^2 ,

$$R^2 = \frac{\sigma_x^2}{\sigma_x^2 + \sigma_\varepsilon^2},$$

we obtain the relationship between the buy-and-hold and market-timing Sharpe ratios as

$$s_1^2 = A \cdot \frac{\mu^2 + \sigma_x^2}{\sigma_\varepsilon^2} = A \cdot \frac{\mu^2 + \sigma_x^2}{(\sigma_x^2 + \sigma_\varepsilon^2)(1 - R^2)} = A \cdot \frac{s_0^2 + R^2}{1 - R^2}$$

When the predictor has no predictive power, we know that $R^2 = 0$ and $s_0 = s_1$. We therefore pin down the constant $A = 1$ and obtain

$$s_1 = \sqrt{\frac{s_0^2 + R^2}{1 - R^2}}. \tag{A.38}$$

Using data back to 1871, [Campbell and Thompson \(2008\)](#) obtain a long-term estimate of the market buy-and-hold Sharpe ratio (“ s_0 ”) of 0.37 (annualized). If a mean-variance investor uses the information from dr to construct a market-timing strategy, with an out-of-sample R^2 of 14.6%, she would obtain a Sharpe ratio (“ s_1 ”) of 0.58, representing a 54.7% improvement over the Sharpe ratio achieved by the buy-and-hold approach.

Appendix II: Filtering the Persistence of Expected Growth

Table A.1 Estimating the Persistence of Expected Cash-Flow Growth (State Space Model)

This table presents the estimation results for four models of dividend growth rates: (1) the unrestricted state-space model as specified in equations (18) and (19) in Section 2; (2) the restricted state-space model with the constraint $\rho_z = 0$; (3) the MA(1) model ($\Delta d_{t+1} = g + \sigma_D \varepsilon_{t+1} + \chi \sigma_D \varepsilon_t$); and (4) the AR(1) model ($\Delta d_{t+1} = g + \gamma \Delta d_t + \sigma_D \varepsilon_{t+1}$). Panel A reports results using the annual (non-overlapping) dividend growth of the S&P 500 index, while Panel B reports results using the annual (non-overlapping) dividend growth of the Fama-French market portfolio. For each model, the log-likelihood (“LogL”), AIC, and BIC are provided. t -statistics are presented in squared brackets.

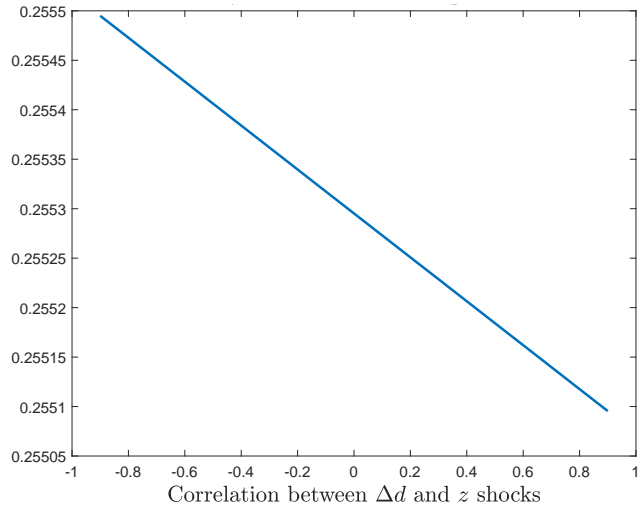
	$\hat{\rho}_z$	\hat{g}	$\hat{\sigma}_d$	$\hat{\sigma}_z$	$\hat{\chi}$	$\hat{\gamma}$	LogL	AIC	BIC
Panel A: S&P 500									
Unrestricted	0.26 [0.94]	0.06 [3.01]	0.00 [0.00]	0.11 [1.70]			74.44	-140.88	-128.97
Restricted		0.06 [4.68]	0.08 [0.00]	0.08 [0.00]			71.36	-136.72	-127.79
MA(1)		0.06 [3.38]	0.10 [13.45]		0.41 [6.11]		76.41	-146.82	-137.89
AR(1)		0.04 [3.64]	0.11 [14.90]			0.26 [3.51]	74.50	-142.99	-134.06
Panel B: MKT									
Unrestricted	-0.08 [-0.06]	0.06 [3.86]	0.00 [0.00]	0.15 [0.12]			43.96	-79.92	-69.8
Restricted		0.06 [3.62]	0.11 [0.10]	0.11 [0.10]			43.67	-81.34	-73.8
MA(1)		0.06 [3.94]	0.15 [6.99]		-0.09 [-1.02]		44.00	-82.00	-74.4
AR(1)		0.06 [3.89]	0.15 [6.98]			-0.08 [-0.87]	43.96	-81.93	-74.39

An alternative method to estimate ρ_z is to directly estimate the state-space model given by equations (18) and (19) with the realized dividend data. Using the standard Kalman filter, we obtain estimates of ρ_z . For comparison, we report results for both the S&P 500 index and the Fama-French market portfolio (“MKT”).⁴⁶ Since the model is set up at annual frequency, we use annual (non-overlapping) dividend growth data. The sample spans 1926 to 2019.⁴⁷ The results are reported in Table A.1, where Panel A and B are for S&P 500 and MKT, respectively. In the row “Unrestricted” of Panel A and B of Table A.1, the estimates of $\hat{\rho}_z$ are statistically indistinguishable from zero.⁴⁸ The restricted model with $\rho_z = 0$ generates similar likelihood and information criteria, indicating that allowing ρ_z to be a free parameter does not significantly improve the model fitness. We also

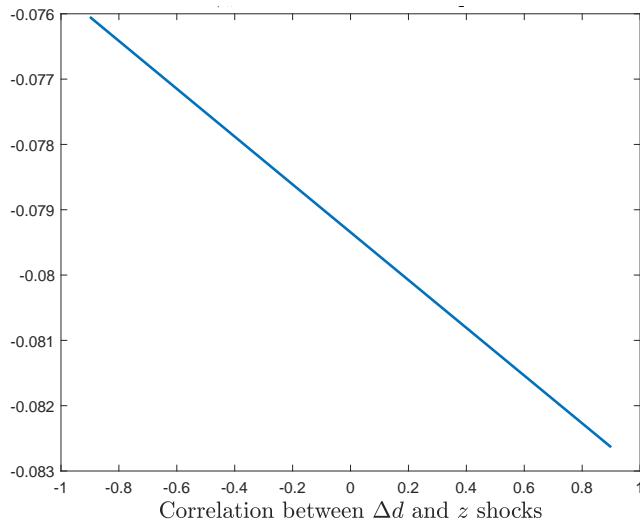
⁴⁶We obtain dividend data for the Fama-French market portfolio (the CRSP NYSE/NYSEMKT/Nasdaq Value-Weighted Market Index.)

⁴⁷We also used the longest available S&P 500 dividend series starting from 1872 and obtained similar results. The results are available upon request.

⁴⁸The Kalman filter assumes that the shocks to realized and expected dividend growth are uncorrelated. In the appendix, we demonstrate the robustness of our estimate of ρ_z by considering different values of the correlation, from -0.9 to 0.9, while fixing the volatility of realized-dividend shock at the estimate in Panel A. The estimated ρ_z barely moves with the value of shock correlations in $[-0.9, 0.9]$ as shown in Figure A.1.



A. S&P 500 dividend



B. MKT dividend

Figure A.1 ρ_z Estimates from the State-Space Model with Correlated Shocks

This figure presents the estimated values of the expected dividend growth autoregressive coefficient (ρ_z) in unrestricted state-space models, as discussed in Section 2, with varying correlations between the Δd and z shocks. The correlations range from -0.9 to 0.9, and the volatility of the Δd shock is adjusted to match the estimated $\hat{\sigma}_D$ from the state-space model with uncorrelated shocks. Panel A uses the annual (non-overlapping) dividend growth of the S&P 500 index, and Panel B uses the annual (non-overlapping) dividend growth of the Fama-French market portfolio.

estimate MA(1) and AR(1) models for comparison and find that the estimates of the autoregressive coefficient, i.e., χ and γ for MA(1) and AR(1), respectively, are statistically indistinguishable from zero. In sum, the state-space approach delivers a similar message as the estimation based on analyst forecasts: The persistence of growth expectation is close to zero.

Appendix III: Additional Tables and Figures

Table A.2 Predicting Dividend Growth Using Different Combinations of Valuation Ratios

This table reports regression results for predicting one-year S&P 500 Index dividend growth using various predictors and sets of valuation ratios. t -statistics based on Newey-West standard errors with autocorrelation of up to 18 lags are reported in parentheses. Data sample: 1988:01–2019:12. See Figure 3 for a detailed definition of each variable.

	Δd_{t+1}															
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
dr_t	-0.18*** (0.05)	-0.04** (0.02)														
pd_t	0.29*** (0.11)		0.02 (0.05)						0.07 (0.06)	0.10* (0.06)	7.93** (3.65)		0.10* (0.06)	4.24 (3.34)	2.26 (2.86)	2.15 (2.90)
g_t^F				0.68*** (0.25)												
KP_t^{CF}					0.22 (0.14)											
$s_t^{0.5}$						0.03 (0.03)	0.14*** (0.05)		0.14*** (0.05)			0.02 (0.03)	0.02 (0.03)	0.10** (0.04)		0.01 (0.03)
s_t^1						0.10*** (0.03)		0.18*** (0.06)		0.18*** (0.05)		0.16*** (0.06)	0.16*** (0.06)		0.15*** (0.04)	0.15*** (0.04)
s_t^{1+}							0.07 (0.05)	0.10* (0.06)			-7.69** (3.53)	0.10* (0.05)		-4.07 (3.23)	-2.11 (2.76)	-2.00 (2.80)
N	372	372	372	372	372	372	372	372	372	372	372	372	372	372	372	372
R^2	0.395	0.062	0.004	0.324	0.040	0.255	0.270	0.383	0.271	0.384	0.231	0.385	0.386	0.310	0.394	0.394

Table A.3 Predicting Returns Using Different Combinations of Valuation Ratios

This table reports regression results for predicting one-year S&P 500 Index returns using various predictors and sets of valuation ratios. t -statistics based on Newey-West standard errors with autocorrelation of up to 18 lags are reported in parentheses. Data sample: 1988:01–2019:12. See Figure 4 for a detailed definition of each variable.

	r_{t+1}														
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
dr_t	-0.16*** (0.03)														
pd_t		-0.20*** (0.07)						-0.13** (0.06)	-0.09 (0.07)	7.54 (4.58)		-0.09 (0.07)	1.18 (4.38)	-1.41 (4.03)	-1.87 (3.70)
μ_t^F			2.58*** (0.92)												
KP_t				0.90*** (0.31)											
$s_t^{0.5}$					0.03 (0.09)	0.18*** (0.06)		0.18*** (0.06)			0.03 (0.09)	0.03 (0.09)	0.17** (0.07)		0.04 (0.09)
s_t^1					0.25*** (0.09)		0.23*** (0.06)		0.23*** (0.06)		0.20** (0.10)	0.20** (0.10)		0.24*** (0.06)	0.21** (0.11)
s_t^{1+}						-0.13** (0.06)	-0.09 (0.07)			-7.53* (4.46)	-0.09 (0.07)		-1.28 (4.25)	1.29 (3.92)	1.74 (3.59)
N	372	372	372	372	372	372	372	372	372	372	372	372	372	372	372
R^2	0.248	0.138	0.156	0.149	0.245	0.230	0.265	0.230	0.265	0.183	0.266	0.266	0.231	0.266	0.268

Table A.4 Forecasting Macroeconomic Variables with dr and Valuation Ratios

This table presents the R^2 values for predicting one-year-ahead macroeconomic variables using predictors (dr , pd , and s^1) that contain information about the underlying state variables. The macroeconomic variables are categorized as follows: 1) Macroeconomic: nominal GDP Growth, Industrial Production Growth (“IP Growth”), Chicago Fed National Activity Index (“CFNAI”), Unemployment Rate, Real Consumption Growth, Total Business Inventories, Nonresidential Fixed Investment (nominal), Residential Fixed Investment (nominal), and GDP Deflator are all from FRED database. 2) Financial: Term Spread and Default Spread (“Baa-Aaa”) are from FRED; Gilchrist-Zakrajšek credit spread (GZ Credit Spread) is from [Gilchrist and Zakrajšek \(2012\)](#); CAPE is the cyclically adjusted price-earnings ratio from Robert Shiller’s website; cay is from [Lettau and Ludvigson \(2001\)](#). 3) Intermediary: Broker/Dealer leverage (“B/D Leverage”) is from [Adrian, Etula, and Muir \(2014\)](#); Broker/Dealer 1(5) year average CDS spreads (“B/D 1(5) Year Avg. CDS”) is from [Gilchrist and Zakrajšek \(2012\)](#); ROA of banks (“ROA Banks”) is from FRED. 4) Uncertainties: CBOE 1-month VIX index (“VIX”) and [Chauvet and Piger \(2008\)](#)’s smoothed U.S. recession probabilities estimates for given month (“CP Recession”) are from FRED; Economics policy uncertainties (“EPU”) is from [Baker, Bloom, and Davis \(2016\)](#); Survey of Professional Forecasters recession probability estimates (“SPF Recession”) is from the Philadelphia Fed. 5) Sentiments: Sentiment Index (both raw and orthogonalized against several macro variables), Number of IPOs (“IPO #”), and close-end fund NAV discount (“Close-end Discount”) are all from [Baker and Wurgler \(2006\)](#).

	$dr + pd$	dr	pd	$dr + s^1$	s^1
<i>Macroeconomic:</i>					
GDP Growth	0.222	0.061	0.000	0.224	0.178
IP Growth	0.202	0.062	0.001	0.203	0.167
Unemployment Growth	0.335	0.062	0.001	0.325	0.229
Real Consumption Growth	0.241	0.019	0.121	0.241	0.208
Business Inventories Growth	0.383	0.114	0.000	0.377	0.304
Nonres. Fixed Investment Growth	0.366	0.055	0.005	0.361	0.240
CPI Growth	0.311	0.227	0.311	0.302	0.071
<i>Financial:</i>					
Baa-Aaa	0.081	0.044	0.008	0.081	0.078
GZ Credit Spread	0.321	0.318	0.218	0.321	0.259
Term Spread	0.122	0.001	0.039	0.120	0.023
CAPE	0.474	0.300	0.465	0.478	0.062
cay	0.072	0.042	0.069	0.071	0.008
<i>Intermediary:</i>					
B/D Leverage	0.313	0.290	0.296	0.308	0.151
B/D 1 Year Avg. CDS	0.284	0.052	0.106	0.284	0.283
B/D 5 Year Avg. CDS	0.393	0.020	0.231	0.393	0.383
ROA Banks	0.403	0.080	0.275	0.388	0.002
<i>Uncertainties:</i>					
VIX	0.145	0.134	0.068	0.144	0.131
EPU	0.049	0.002	0.005	0.047	0.022
CP Recession	0.063	0.030	0.004	0.063	0.059
SPF Recession	0.194	0.017	0.008	0.191	0.109
<i>Sentiments:</i>					
Sentiment Index	0.131	0.131	0.099	0.131	0.094
Sentiment Index (orth.)	0.113	0.109	0.102	0.113	0.062
IPO #	0.144	0.142	0.093	0.144	0.120
Close-end Discount	0.123	0.109	0.119	0.118	0.054

Table A.5 Kostakis, Magdalinos, and Stamatogiannis (2014) IVX-Wald Test

This table presents results of the IVX-Wald test proposed by [Kostakis, Magdalinos, and Stamatogiannis \(2014\)](#) on the predictive coefficient β in [Table \(2\)](#). IVX-Wald is the Wald statistic to test $H_0 : \beta = 0$ against $H_1 : \beta \neq 0$. The test is designed to be robust to the persistence of the predictor. p -value of the IVX-Wald test is provided in the parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	dr_t	pd_t	μ_t^F	KP_t
IVX-Wald	9.29***	1.56	2.77*	5.74**
p -value	(0.002)	(0.212)	(0.096)	(0.017)

Table A.6 Predicting Annual Excess Return

This table presents the results of the predictive regression specified in equation (17). The dependent variable is the log excess return of the S&P 500 index over the next twelve months, r_{t+1}^e . The predictors include: the ‘slope’ of the term structure of valuation ratios dr_t , the price-dividend ratio pd_t , the filtered series for expected returns following Binsbergen and Koijen (2010) μ_t^F , and the single predictive factor extracted from 100 book-to-market and size portfolios as per Kelly and Pruitt (2013) KP_t . For each predictor, the table reports the raw β estimate followed by the Hodrick (1992) t -statistic, the Newey and West (1987) t -statistic (with 18 lags), and the β coefficient adjusted for Stambaugh (1999) bias. Starting from January 1998, we generate out-of-sample forecasts of the return for the next twelve months by estimating the regression using data up to the current month. These forecasts are then used to compute the out-of-sample R^2 , the ENC test (Clark and McCracken, 2001), and the p -value of the CW test (Clark and West, 2007). Data sample: 1988:01–2019:12.

	r_{t+1}^e				
	(1)	(2)	(3)	(4)	(5)
dr_t	-0.146				-0.228
<i>Hodrick t</i>	[-3.178]				[-2.945]
<i>Newey-West t</i>	(-3.867)				(-3.571)
<i>Stambaugh bias adjusted β</i>	-0.136				
pd_t		-0.180			0.161
		[-2.168]			[1.820]
		(-2.262)			(1.286)
		-0.170			
μ_t^F			2.293		
			[2.033]		
			(2.205)		
			2.303		
KP_t				0.827	
				[2.715]	
				(2.429)	
				0.837	
N	372	372	372	372	372
R^2	0.219	0.114	0.124	0.128	0.241
OOS R^2	0.098	-0.040	-0.096	0.005	0.138
ENC	1.924	0.296	0.021	2.175	4.539
$p(ENC)$	<0.10	>0.10	>0.10	<0.05	<0.05
$p(CW)$	0.058	0.379	0.493	0.072	0.028

Table A.7 Predicting Annual Return: Fama-French Market Return

This table presents the results of the predictive regression specified in equation (17). The dependent variable is the log market return from Fama-French in the next twelve months, r_{t+1}^{MKT} . The predictors include: the ‘slope’ of the term structure of valuation ratios dr_t , the price-dividend ratio pd_t , the filtered series for expected returns following Binsbergen and Kojen (2010) μ_t^F , and the single predictive factor extracted from 100 book-to-market and size portfolios as per Kelly and Pruitt (2013) KP_t . For each predictor, the table reports the raw β estimate followed by the Hodrick (1992) t -statistic, the Newey and West (1987) t -statistic (with 18 lags), and the β coefficient adjusted for Stambaugh (1999) bias. Starting from January 1998, we generate out-of-sample forecasts of the return for the next twelve months by estimating the regression using data up to the current month. These forecasts are then used to compute the out-of-sample R^2 , the ENC test (Clark and McCracken, 2001), and the p -value of the CW test (Clark and West, 2007). Data sample: 1988:01–2019:12.

	r_{t+1}^{MKT}				
	(1)	(2)	(3)	(4)	(5)
dr_t	-0.154				-0.222
Hodrick t					[-2.772]
Newey-West t					(-3.511)
Stambaugh bias adjusted β	-0.144				
pd_t		-0.198			0.133
		[-2.302]			[1.608]
		(-2.706)			(1.129)
		-0.188			
μ_t^F			2.486		
			[2.327]		
			(2.656)		
			2.496		
KP_t				0.794	
				[2.223]	
				(2.689)	
				0.805	
N	372	372	372	372	372
R^2	0.236	0.134	0.141	0.128	0.251
OOS R^2	0.144	0.022	-0.023	-0.001	0.181
ENC	3.083	0.963	0.598	2.483	6.163
$p(ENC)$	<0.05	>0.10	>0.10	<0.05	<0.01
$p(CW)$	0.017	0.166	0.321	0.048	0.019

Table A.8 Predicting Annual Return: Fama-French Market Excess Return

This table presents the results of the predictive regression specified in equation (17). The dependent variable is the log market excess return from Fama-French in the next twelve months, $r_{t+1}^{MKT,e}$. The predictors include: the ‘slope’ of the term structure of valuation ratios dr_t , the price-dividend ratio pd_t , the filtered series for expected returns following Binsbergen and Koijen (2010) μ_t^F , and the single predictive factor extracted from 100 book-to-market and size portfolios as per Kelly and Pruitt (2013) KP_t . For each predictor, the table reports the raw β estimate followed by the Hodrick (1992) t -statistic, the Newey and West (1987) t -statistic (with 18 lags), and the β coefficient adjusted for Stambaugh (1999) bias. Starting from January 1998, we generate out-of-sample forecasts of the return for the next twelve months by estimating the regression using data up to the current month. These forecasts are then used to compute the out-of-sample R^2 , the ENC test (Clark and McCracken, 2001), and the p -value of the CW test (Clark and West, 2007). Data sample: 1988:01–2019:12.

	$r_{t+1}^{MKT,e}$				
	(1)	(2)	(3)	(4)	(5)
dr_t	-0.144				-0.222
<i>Hodrick t</i>	[-3.060]				[-2.791]
<i>Newey-West t</i>	(-3.745)				(-3.503)
<i>Stambaugh bias adjusted β</i>	-0.134				
pd_t		-0.179			0.153
		[-2.108]			[1.704]
		(-2.199)			(1.202)
		-0.169			
μ_t^F			2.192		
			[2.075]		
			(2.057)		
			2.203		
KP_t				0.725	
				[2.044]	
				(2.251)	
				0.735	
N	372	372	372	372	372
R^2	0.206	0.108	0.109	0.105	0.225
OOS R^2	0.099	-0.018	-0.081	-0.037	0.140
ENC	2.000	0.376	-0.047	1.700	4.656
$p(ENC)$	<0.10	>0.10	>0.10	<0.10	<0.05
$p(CW)$	0.047	0.349	0.485	0.120	0.027

Table A.9 Monthly Return Prediction

This table presents the results of the predictive regression specified in equation (17). The dependent variable is the log return of the S&P 500 index over the next months, $r_{t+1/12}$. The predictors include: the ‘slope’ of the term structure of valuation ratios dr_t , the price-dividend ratio pd_t , the filtered series for expected returns following [Binsbergen and Kojen \(2010\)](#) μ_t^F , and the single predictive factor extracted from 100 book-to-market and size portfolios as per [Kelly and Pruitt \(2013\)](#) KP_t . For each predictor, the table reports the raw β estimate followed by the [Hodrick \(1992\)](#) t -statistic, the [Newey and West \(1987\)](#) t -statistic (with 7 lags), and the β coefficient adjusted for [Stambaugh \(1999\)](#) bias. Starting from January 1998, we generate out-of-sample forecasts of the return for the next month by estimating the regression using data up to the current month. These forecasts are then used to compute the out-of-sample R^2 , the ENC test ([Clark and McCracken, 2001](#)), and the p -value of the CW test ([Clark and West, 2007](#)). Data sample: 1988:01–2019:12.

	$r_{t+1/12}$				
	(1)	(2)	(3)	(4)	(5)
dr_t	-0.012				-0.017
<i>Hodrick t</i>	[-2.529]				[-1.427]
<i>Newey-West t</i>	(-2.826)				(-2.034)
pd_t		-0.015			0.011
		[-1.891]			[0.530]
		(-2.090)			(0.751)
μ_t^F			0.211		
			[2.224]		
			(2.401)		
KP_t				0.019	
				[0.656]	
				(0.680)	
N	383	383	383	383	383
R^2	0.021	0.011	0.015	0.001	0.022
OOS R^2	0.015	0.004	0.007	-0.012	0.005
ENC	2.678	1.122	1.673	-0.676	2.384
$p(ENC)$	<0.05	>0.10	<0.10	>0.10	<0.10
$p(CW)$	0.018	0.179	0.122	0.325	0.129

Table A.10 Monthly Excess Return Prediction

This table presents the results of the predictive regression specified in equation (17). The dependent variable is the log excess return of the S&P 500 index over the next months, $r_{t+1/12}^e$. The predictors include: the ‘slope’ of the term structure of valuation ratios dr_t , the price-dividend ratio pd_t , the filtered series for expected returns following [Binsbergen and Kojen \(2010\)](#) μ_t^F , and the single predictive factor extracted from 100 book-to-market and size portfolios as per [Kelly and Pruitt \(2013\)](#) KP_t . For each predictor, the table reports the raw β estimate followed by the [Hodrick \(1992\)](#) t -statistic, the [Newey and West \(1987\)](#) t -statistic (with 7 lags), and the β coefficient adjusted for [Stambaugh \(1999\)](#) bias. Starting from January 1998, we generate out-of-sample forecasts of the return for the next month by estimating the regression using data up to the current month. These forecasts are then used to compute the out-of-sample R^2 , the ENC test ([Clark and McCracken, 2001](#)), and the p -value of the CW test ([Clark and West, 2007](#)). Data sample: 1988:01–2019:12.

	$r_{t+1/12}^e$				
	(1)	(2)	(3)	(4)	(5)
dr_t	-0.011				-0.018
<i>Hodrick t</i>	[-2.394]				[-1.504]
<i>Newey-West t</i>	(-2.684)				(-2.134)
pd_t		-0.014			0.014
		[-1.687]			[0.670]
		(-1.873)			(0.944)
μ_t^F			0.188		
			[1.967]		
			(2.137)		
KP_t				0.015	
				[0.514]	
				(0.535)	
N	383	383	383	383	383
R^2	0.019	0.009	0.012	0.001	0.021
OOS R^2	0.012	0.001	0.003	-0.013	0.003
ENC	2.338	0.670	1.060	-0.766	2.233
$p(ENC)$	<0.05	>0.10	>0.10	>0.10	<0.10
$p(CW)$	0.038	0.283	0.228	0.302	0.159

Table A.11 Monthly Return Prediction: Fama-French MKT Return

This table presents the results of the predictive regression specified in equation (17). The dependent variable is the log market return from Fama-French in the next month, $r_{t+1/12}^{MKT}$. The predictors include: the ‘slope’ of the term structure of valuation ratios dr_t , the price-dividend ratio pd_t , the filtered series for expected returns following [Binsbergen and Koijen \(2010\)](#) μ_t^F , and the single predictive factor extracted from 100 book-to-market and size portfolios as per [Kelly and Pruitt \(2013\)](#) KP_t . For each predictor, the table reports the raw β estimate followed by the [Hodrick \(1992\)](#) t -statistic, the [Newey and West \(1987\)](#) t -statistic (with 7 lags), and the β coefficient adjusted for [Stambaugh \(1999\)](#) bias. Starting from January 1998, we generate out-of-sample forecasts of the return for the next month by estimating the regression using data up to the current month. These forecasts are then used to compute the out-of-sample R^2 , the ENC test ([Clark and McCracken, 2001](#)), and the p -value of the CW test ([Clark and West, 2007](#)). Data sample: 1988:01–2019:12.

	$r_{t+1/12}^{MKT}$				
	(1)	(2)	(3)	(4)	(5)
dr_t	-0.012				-0.017
<i>Hodrick t</i>	[-2.354]				[-1.330]
<i>Newey-West t</i>	(-2.742)				(-1.945)
pd_t		-0.015			0.011
		[-1.819]			[0.488]
		(-2.044)			(0.697)
μ_t^F			0.208		
			[2.091]		
			(2.306)		
KP_t				0.018	
				[0.588]	
				(0.626)	
N	383	383	383	383	383
R^2	0.019	0.011	0.014	0.001	0.021
OOS R^2	0.012	0.003	0.005	-0.014	0.003
ENC	2.227	0.876	1.275	-1.009	1.869
$p(ENC)$	<0.05	>0.10	<0.10	>0.10	>0.10
$p(CW)$	0.034	0.220	0.171	0.220	0.176

Table A.12 Monthly Return Prediction: Fama-French MKT excess Return

This table presents the results of the predictive regression specified in equation (17). The dependent variable is the log excess market return from Fama-French in the next month, $r_{t+1/12}^{MKT,e}$. The predictors include: the ‘slope’ of the term structure of valuation ratios dr_t , the price-dividend ratio pd_t , the filtered series for expected returns following Binsbergen and Koijen (2010) μ_t^F , and the single predictive factor extracted from 100 book-to-market and size portfolios as per Kelly and Pruitt (2013) KP_t . For each predictor, the table reports the raw β estimate followed by the Hodrick (1992) t -statistic, the Newey and West (1987) t -statistic (with 7 lags), and the β coefficient adjusted for Stambaugh (1999) bias. Starting from January 1998, we generate out-of-sample forecasts of the return for the next month by estimating the regression using data up to the current month. These forecasts are then used to compute the out-of-sample R^2 , the ENC test (Clark and McCracken, 2001), and the p -value of the CW test (Clark and West, 2007). Data sample: 1988:01–2019:12.

	$r_{t+1/12}^{MKT,e}$				
	(1)	(2)	(3)	(4)	(5)
dr_t	-0.011				-0.018
Hodrick t	[-2.228]				[-1.402]
Newey-West t	(-2.602)				(-2.042)
pd_t		-0.014			0.013
		[-1.621]			[0.623]
		(-1.830)			(0.885)
μ_t^F			0.185		
			[1.844]		
			(2.049)		
KP_t				0.014	
				[0.453]	
				(0.484)	
N	383	383	383	383	383
R^2	0.017	0.009	0.011	0.001	0.019
OOS R^2	0.010	0.000	0.001	-0.014	0.001
ENC	1.903	0.443	0.699	-1.082	1.718
$p(ENC)$	<0.10	>0.10	>0.10	>0.10	>0.10
$p(CW)$	0.065	0.340	0.297	0.205	0.209

Table A.13 Return Spanning Test: dr

The table presents the results of the following return spanning test:

$$r_{t+1} = \alpha + \beta dr_t + \gamma x_t + \epsilon_{t+1}.$$

The dependent variable is the log return of the S&P 500 Index over the next twelve months, r_{t+1} . x_t denotes an alternative return predictor. Detailed definitions and the sample period for each variable can be found in Figure 4. t -statistics, based on Newey-West standard errors with autocorrelation adjustments up to 18 lags, are provided in parentheses. Constant terms are omitted for brevity.

$x =$	r_{t+1}										
	<i>pd</i> (1)	<i>KP</i> (2)	μ^F (3)	<i>bm</i> (4)	<i>dy</i> (5)	<i>tbl</i> (6)	<i>lty</i> (7)	<i>ntis</i> (8)	<i>infl</i> (9)	<i>ltr</i> (10)	<i>svar</i> (11)
dr_t	-0.228*** (0.065)	-0.130*** (0.035)	-0.223*** (0.076)	-0.185*** (0.049)	-0.228*** (0.064)	-0.159*** (0.035)	-0.169*** (0.037)	-0.151*** (0.032)	-0.160*** (0.035)	-0.156*** (0.035)	-0.160*** (0.035)
x_t	0.141 (0.117)	0.307 (0.247)	-1.571 (1.709)	-0.207 (0.357)	-0.140 (0.114)	-0.387 (0.616)	-0.903 (0.730)	1.524 (1.356)	-7.034** (3.351)	0.202 (0.166)	2.885** (1.463)
N	372	372	372	372	372	372	372	372	372	372	372
R^2	0.264	0.259	0.259	0.253	0.263	0.251	0.259	0.289	0.269	0.249	0.254

$x =$	r_{t+1}										
	<i>csp</i> (1)	<i>ep</i> (2)	<i>de</i> (3)	<i>dfy</i> (4)	<i>dfr</i> (5)	<i>tms</i> (6)	<i>cay</i> (7)	<i>ik</i> (8)	<i>SII</i> (9)	<i>SVIX</i> (10)	dp^{Corr} (11)
dr_t	-0.149*** (0.033)	-0.159*** (0.040)	-0.157*** (0.035)	-0.156*** (0.035)	-0.155*** (0.034)	-0.160*** (0.036)	-0.169*** (0.047)	-0.178*** (0.050)	-0.150*** (0.038)	-0.221*** (0.045)	-0.209*** (0.072)
x_t	37.581* (21.626)	-0.192 (1.316)	-0.007 (0.046)	-0.282 (4.930)	0.421 (0.545)	-0.578 (1.352)	-0.351 (1.034)	2.796 (10.797)	-0.062*** (0.023)	1.373** (0.644)	0.011 (0.090)
N	180	372	372	372	372	372	124	124	324	193	210
R^2	0.380	0.248	0.248	0.248	0.249	0.250	0.248	0.248	0.408	0.293	0.304

Table A.14 Return Spanning Test: dr and pd

The table presents the results of the following return spanning test:

$$r_{t+1} = \alpha + \beta_1 dr_t + \beta_2 pd_t + \gamma x_t + \epsilon_{t+1}.$$

The dependent variable is the log return of the S&P 500 Index over the next twelve months, r_{t+1} . x_t denotes an alternative return predictor. Detailed definitions and the sample period for each variable can be found in Figure 4. t -statistics, based on Newey-West standard errors with autocorrelation adjustments up to 18 lags, are provided in parentheses. Constant terms are omitted for brevity.

$x =$	r_{t+1}										
	<i>KP</i> (1)	μ^F (2)	bm (3)	dy (4)	tbl (5)	lty (6)	ntis (7)	infl (8)	ltr (9)	svar (10)	csp (11)
pd_t	0.139 (0.113)	0.146 (0.273)	0.139 (0.130)	0.127 (0.195)	0.134 (0.127)	0.121 (0.130)	0.082 (0.130)	0.146 (0.114)	0.148 (0.119)	0.204 (0.125)	0.127 (0.146)
dr_t	-0.202*** (0.063)	-0.227*** (0.074)	-0.228*** (0.068)	-0.228*** (0.064)	-0.226*** (0.068)	-0.228*** (0.067)	-0.194*** (0.058)	-0.235*** (0.067)	-0.231*** (0.066)	-0.267*** (0.076)	-0.211*** (0.056)
x_t	0.299 (0.261)	0.072 (3.804)	-0.013 (0.428)	-0.015 (0.159)	-0.241 (0.639)	-0.699 (0.824)	1.355 (1.368)	-7.215* (3.737)	0.289* (0.175)	5.377*** (1.831)	31.269 (21.220)
N	372	372	372	372	372	372	372	372	372	372	180
R^2	0.275	0.264	0.264	0.264	0.266	0.271	0.294	0.286	0.267	0.284	0.393

$x =$	r_{t+1}									
	ep (1)	de (2)	dfy (3)	dfr (4)	tms (5)	cay (6)	ik (7)	SII (8)	SVIX (9)	dp^{Corr} (10)
pd_t	0.144 (0.117)	0.182 (0.151)	0.169 (0.110)	0.140 (0.116)	0.143 (0.118)	0.188 (0.131)	0.187 (0.129)	0.097 (0.139)	0.175 (0.266)	0.031 (0.154)
dr_t	-0.235*** (0.073)	-0.243*** (0.076)	-0.240*** (0.063)	-0.227*** (0.064)	-0.233*** (0.066)	-0.270*** (0.076)	-0.272*** (0.074)	-0.199*** (0.066)	-0.286** (0.112)	-0.215** (0.084)
x_t	-0.343 (1.492)	0.029 (0.060)	2.412 (5.439)	0.358 (0.489)	-0.631 (1.283)	-0.261 (0.964)	1.191 (9.805)	-0.060*** (0.021)	2.042 (1.436)	0.023 (0.084)
N	372	372	372	372	372	124	124	324	193	210
R^2	0.265	0.268	0.267	0.266	0.267	0.271	0.271	0.416	0.307	0.305

Table A.15 Correlations with Other Return Predictors

This table presents the correlations of various alternative return predictors with both dr_t and pd_t from 1988 to 2019. μ^F is the filtered demeaned expected return following [Binsbergen and Kojen \(2010\)](#). KP is a predictive factor extracted from 100 book-to-market and size portfolios from [Kelly and Pruitt \(2013\)](#). Most alternative predictors are from [Goyal and Welch \(2007\)](#) that include the default yield spread (dfy), the inflation rate (infl), stock variance (svar), the cross-section premium (csp, available in 1988-2002), the dividend payout ratio (de), the long-term yield (lty), the term spread (tms), the T-bill rate (tbl), the default return spread (dfr), the dividend yield (dy, log difference between current-period dividend and lagged S&P 500 index price), the long-term rate of return (ltr), the earnings-to-price ratio (ep), the book to market ratio (bm), the investment-to-capital ratio (ik), the net equity expansion ratio (ntis), the percent equity issuing ratio (eqis), and the consumption-wealth-income ratio (cay). SII is the short interests index from [Rapach, Ringgenberg, and Zhou \(2016\)](#) (1988-2014). SVIX is an option-implied lower bound of 1-year equity premium from [Martin \(2017\)](#) (1996-2012). dp^{Corr} is the dividend-price ratio corrected for option-implied expected dividend growth following [Golez \(2014\)](#) (1994-2011).

	dr	pd
bm	-0.788	-0.851
cay	-0.262	-0.272
csp	0.345	0.428
de	-0.245	-0.463
dfr	-0.024	0.005
dfy	-0.078	-0.273
dr	1.000	0.873
dy	-0.879	-0.990
ep	-0.558	-0.453
dp^{Corr}	-0.784	-0.834
ik	0.656	0.637
infl	-0.100	-0.074
KP	-0.565	-0.496
ltr	-0.000	-0.056
lty	-0.368	-0.425
μ^F	-0.892	-0.967
ntis	-0.074	0.075
pd	0.873	1.000
SII	0.047	-0.015
svar	0.149	-0.053
SVIX	0.042	-0.359
tbl	-0.175	-0.243
tms	-0.255	-0.217

Table A.16 Time-varying ρ_z and Return Predictability

This table presents the results from regressions examining the relationship between return predictability from dr_t and the time-varying persistence of the expected cash-flow growth. The dependent variables are the in-sample residuals from return predictive regressions ($\hat{\varepsilon}_t$) and the out-of-sample return forecast errors (v_t). The independent variable is the time-varying expected cash-flow growth persistence ($\hat{\rho}_{z,t}$), estimated using three-year rolling windows. t -statistics based on Newey-West standard errors with 18 lags are reported in parentheses. The sample period begins in January 1998 when the first out-of-sample forecast is made, and concludes in December 2019.

	$\hat{\varepsilon}_t$	v_t
	(1)	(2)
Intercept	-0.011 (-1.127)	-0.046 (-4.076)
$\hat{\rho}_{z,t}$	0.556 (5.143)	0.469 (4.599)
N	252	252
R^2	0.173	0.094

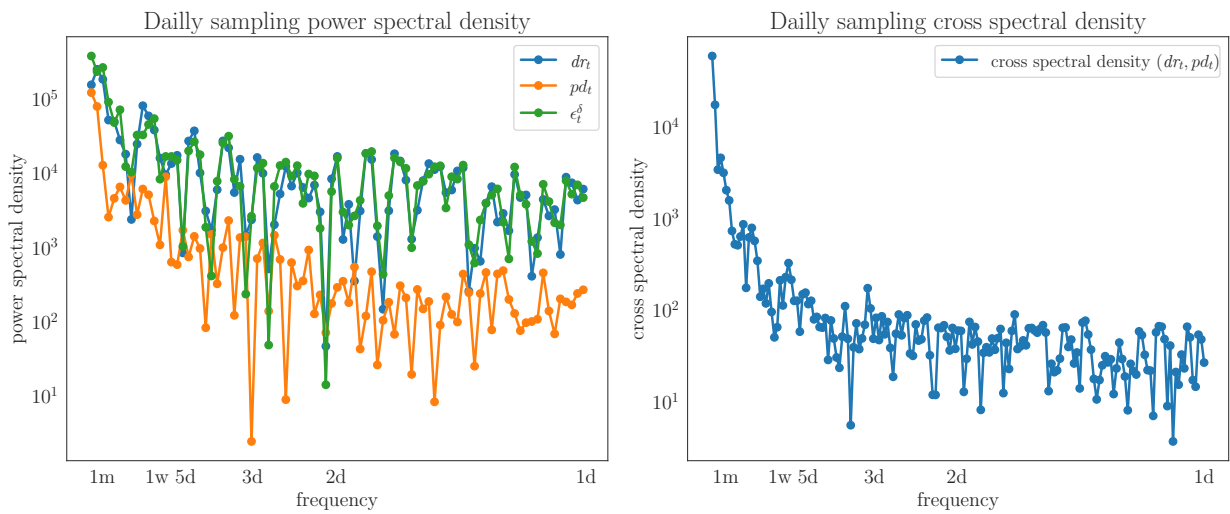


Figure A.2 Spectrum and Cross-spectrum of dr and pd (Daily Frequency)

The left panel displays the estimated spectral densities of dr , pd , and the residuals of dr from projection onto pd (ϵ^{Pr}). The integral of the spectral density represents the variance. The horizontal axis ranges from zero to π and is labeled with the corresponding cycle lengths. The right panel illustrates the cross-spectral density between dr and pd , with the integral representing the covariance.

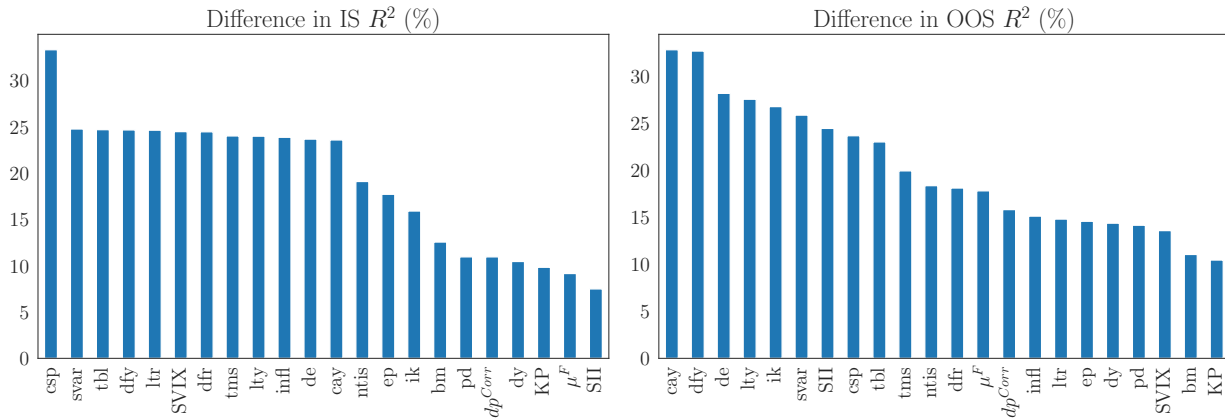


Figure A.3 In-Sample and Out-of-Sample R^2 Wedge between dr and Other Return Predictors: Excess Return

This figure compares annual return predictive R^2 between dr_t and other commonly studied predictors. The forecast target is the annual log excess return of the S&P 500 Index. Panels A and B report, respectively, the differences in in-sample (IS) and out-of-sample (OOS) R^2 between dr and an alternative predictor. A positive value signifies that dr has a stronger predictive power than the alternative within the same sample period. Most predictors are from [Goyal and Welch \(2007\)](#) and include the price-dividend ratio (pd), the default yield spread (dfy), the inflation rate (infl), stock variance (svar), the cross-section premium (csp, available in 1988-2002), the dividend payout ratio (de), the long-term yield (lty), the term spread (tms), the T-bill rate (tbl), the default return spread (dfr), the dividend yield (dy), the long-term rate of return (ltr), the earnings-to-price ratio (ep), the book to market ratio (bm), the investment-to-capital ratio (ik), the net equity expansion ratio (ntis), the percent equity issuing ratio (eqis), and the consumption-wealth-income ratio (cay). KP is the predictive factor extracted from 100 book-to-market and size portfolios from [Kelly and Pruitt \(2013\)](#). dp^{Corr} is the dividend-price ratio corrected for option-implied dividend growth in [Golez \(2014\)](#) (available in 1994-2011). μ^F is the filtered series for expected returns following [Binsbergen and Koijen \(2010\)](#). SII is the short interests index from [Rapach, Ringgenberg, and Zhou \(2016\)](#) (available in 1988-2014). SVIX is an option-implied lower bound of annual equity premium in [Martin \(2017\)](#) (available in 1996-2012).

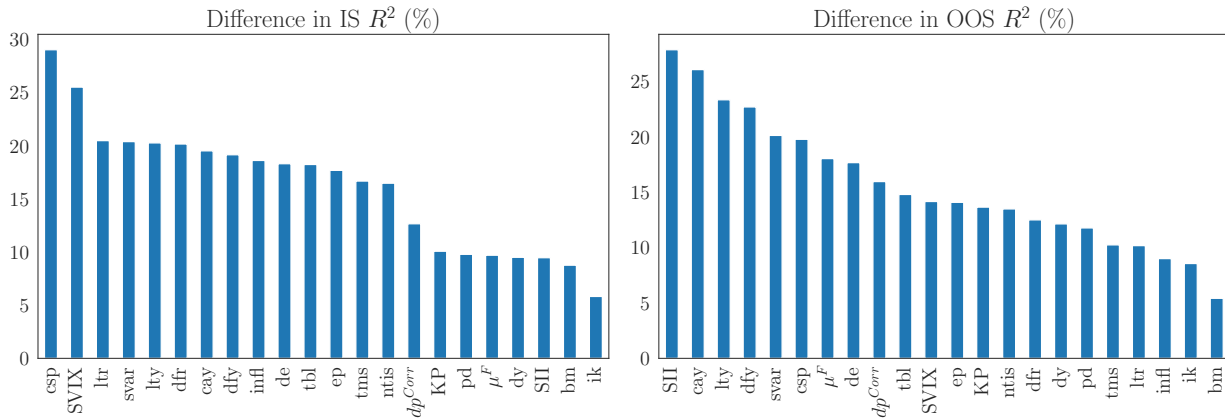


Figure A.4 In-Sample and Out-of-Sample R^2 Wedge between dr and Other Return Predictors: Fama-French Market Return

This figure compares annual return predictive R^2 between dr_t and other commonly studied predictors. The forecast target is the annual log market return from Fama-French. Panels A and B report, respectively, the differences in in-sample (IS) and out-of-sample (OOS) R^2 between dr and an alternative predictor. A positive value signifies that dr has a stronger predictive power than the alternative within the same sample period. Most predictors are from [Goyal and Welch \(2007\)](#) and include the price-dividend ratio (pd), the default yield spread (dfy), the inflation rate (infl), stock variance (svar), the cross-section premium (csp, available in 1988-2002), the dividend payout ratio (de), the long-term yield (lty), the term spread (tms), the T-bill rate (tbl), the default return spread (dfr), the dividend yield (dy), the long-term rate of return (ltr), the earnings-to-price ratio (ep), the book to market ratio (bm), the investment-to-capital ratio (ik), the net equity expansion ratio (ntis), the percent equity issuing ratio (eqis), and the consumption-wealth-income ratio (cay). KP is the predictive factor extracted from 100 book-to-market and size portfolios from [Kelly and Pruitt \(2013\)](#). dp^{Corr} is the dividend-price ratio corrected for option-implied dividend growth in [Golez \(2014\)](#) (available in 1994-2011). μ^F is the filtered series for expected returns following [Binsbergen and Koijen \(2010\)](#). SII is the short interests index from [Rapach, Ringgenberg, and Zhou \(2016\)](#) (available in 1988-2014). SVIX is an option-implied lower bound of annual equity premium in [Martin \(2017\)](#) (available in 1996-2012).

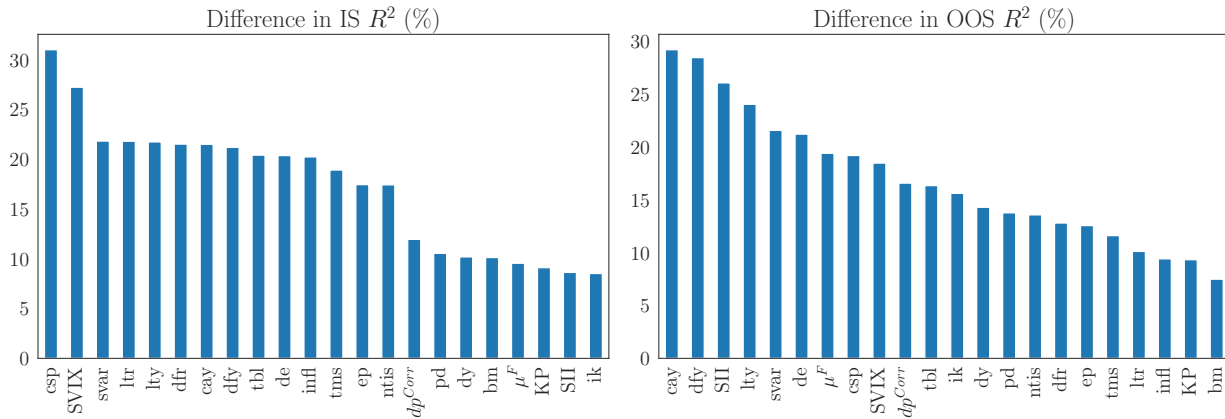


Figure A.5 In-Sample and Out-of-Sample R^2 Wedge between dr and Other Return Predictors: Fama-French Market Excess Return

This figure compares annual return predictive R^2 between dr_t and other commonly studied predictors. The forecast target is the annual log market excess return from Fama-French. Panels A and B report, respectively, the differences in in-sample (IS) and out-of-sample (OOS) R^2 between dr and an alternative predictor. A positive value signifies that dr has a stronger predictive power than the alternative within the same sample period. Most predictors are from [Goyal and Welch \(2007\)](#) and include the price-dividend ratio (pd), the default yield spread (dfy), the inflation rate (infl), stock variance (svar), the cross-section premium (csp, available in 1988-2002), the dividend payout ratio (de), the long-term yield (lty), the term spread (tms), the T-bill rate (tbl), the default return spread (dfr), the dividend yield (dy), the long-term rate of return (ltr), the earnings-to-price ratio (ep), the book to market ratio (bm), the investment-to-capital ratio (ik), the net equity expansion ratio (ntis), the percent equity issuing ratio (eqis), and the consumption-wealth-income ratio (cay). KP is the predictive factor extracted from 100 book-to-market and size portfolios from [Kelly and Pruitt \(2013\)](#). dp^{Corr} is the dividend-price ratio corrected for option-implied dividend growth in [Golez \(2014\)](#) (available in 1994-2011). μ^F is the filtered series for expected returns following [Binsbergen and Koijen \(2010\)](#). SII is the short interests index from [Rapach, Ringgenberg, and Zhou \(2016\)](#) (available in 1988-2014). SVIX is an option-implied lower bound of annual equity premium in [Martin \(2017\)](#) (available in 1996-2012).

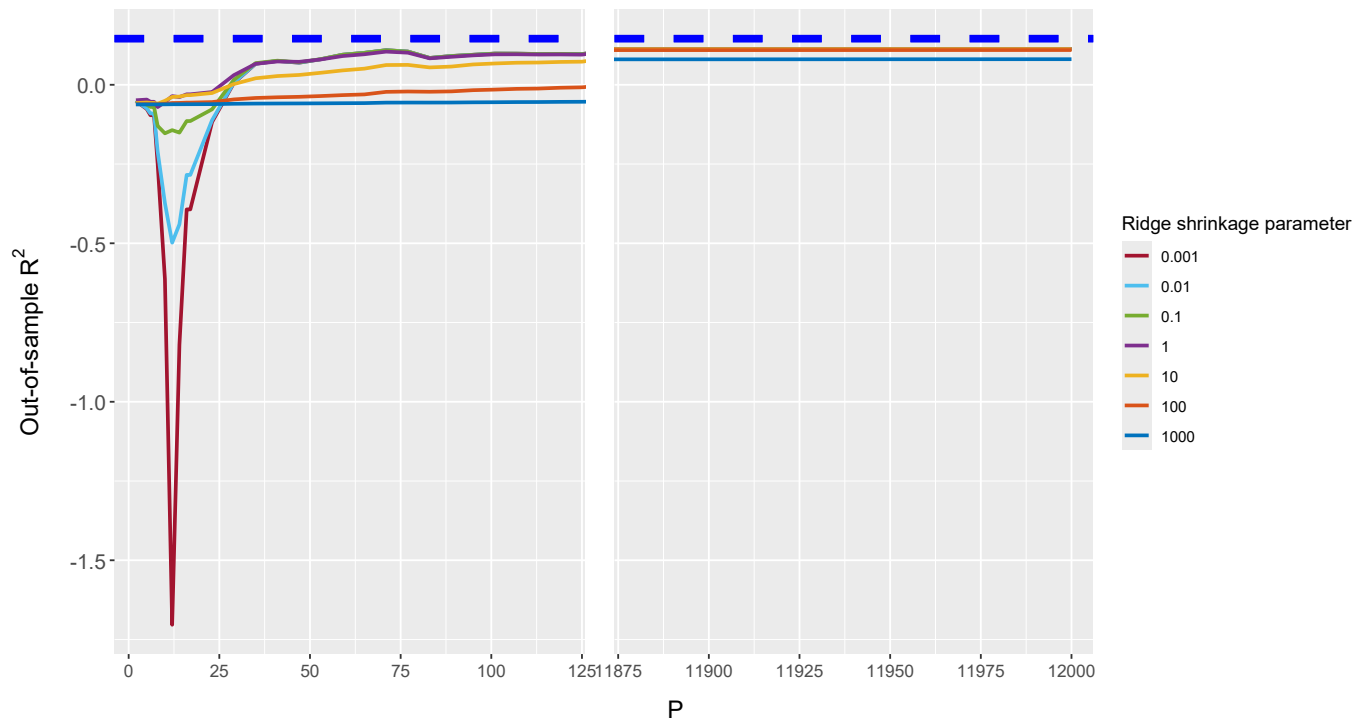


Figure A.6 Out-of-sample R^2 and model complexity

This figure presents the out-of-sample (OOS) R^2 against the degree of model complexity for various values of the ridge shrinkage parameter, using the machine learning method developed by Kelly, Malamud, and Zhou (2024). The analysis is based on ridge regressions and forecasts the annual log return of the S&P index. The initial OOS prediction starts in January 1998, and the OOS R^2 is calculated following Goyal and Welch (2007). The machine learning models employ a 12-month training window, $\gamma = 2$, a Random Fourier Features (RFF) count P ranging from 2 to 12,000, and the shrinkage parameter ranging from 0.001 to 1000. The blue dashed line indicates the OOS R^2 obtained from the standard univariate predictive regression using dr as the predictor.

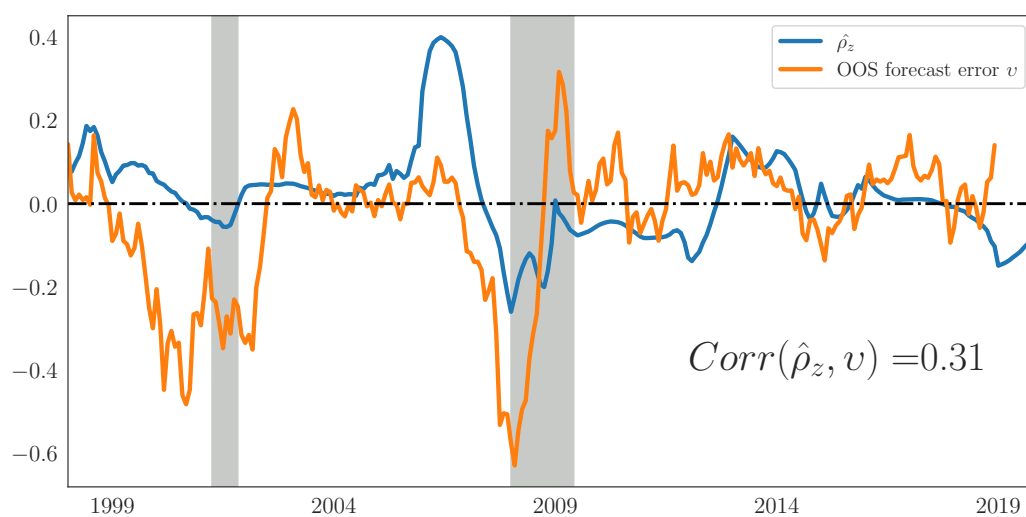


Figure A.7 Rolling Estimates of Expected Growth Persistence and Out-of-sample Return Prediction Errors

This figure plots the rolling estimates of the autoregressive coefficient of expected cash flow growth, $\hat{\rho}_{z,t}$, and the out-of-sample return prediction errors using the slope of S&P 500 valuation term structure (dr_t) as the predictor. $\hat{\rho}_{z,t}$ is estimated using analyst forecasts of S&P 500 aggregate earnings in three-year rolling windows. The initial out-of-sample forecast begins in 1998. The graph also reports the correlation between these two time series. The monthly sample period spans from January 1998 to December 2019.

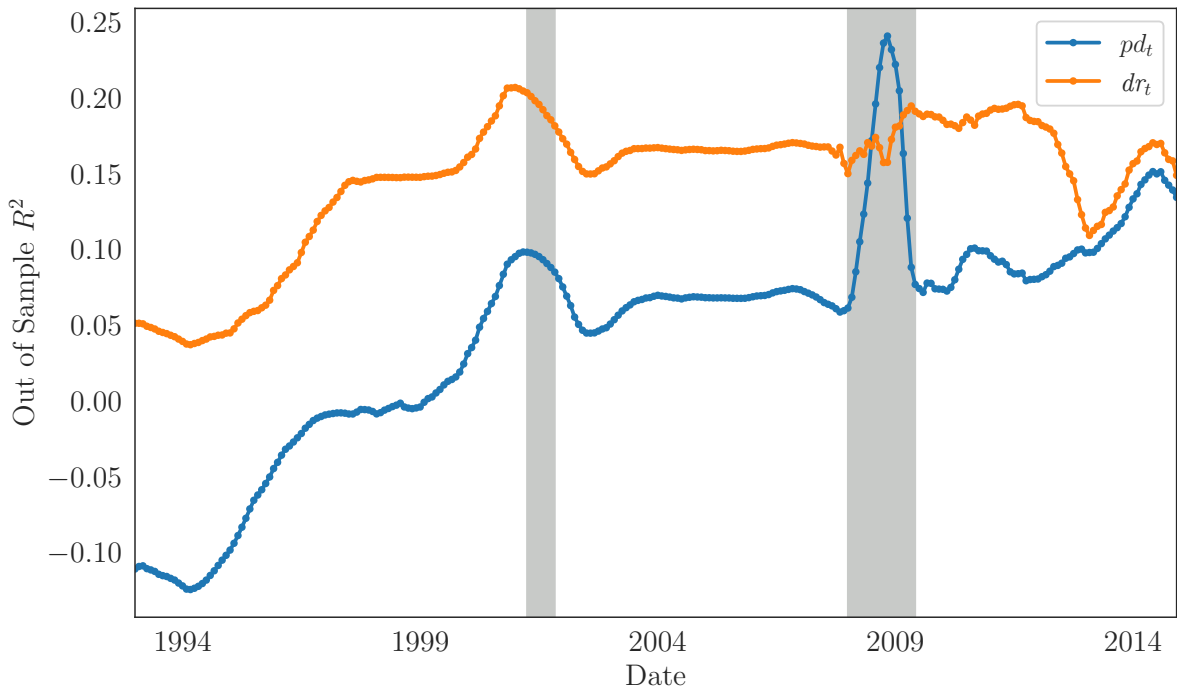
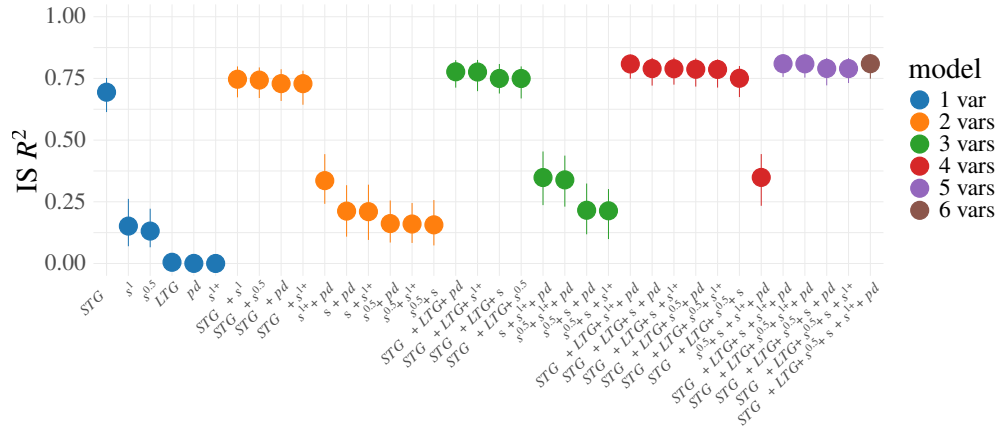
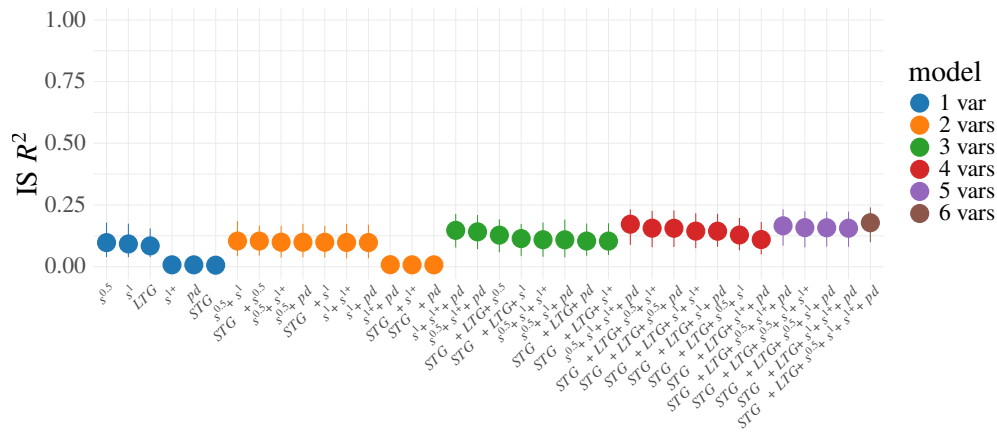


Figure A.8 Out-of-sample R^2 by Sample Split Dates

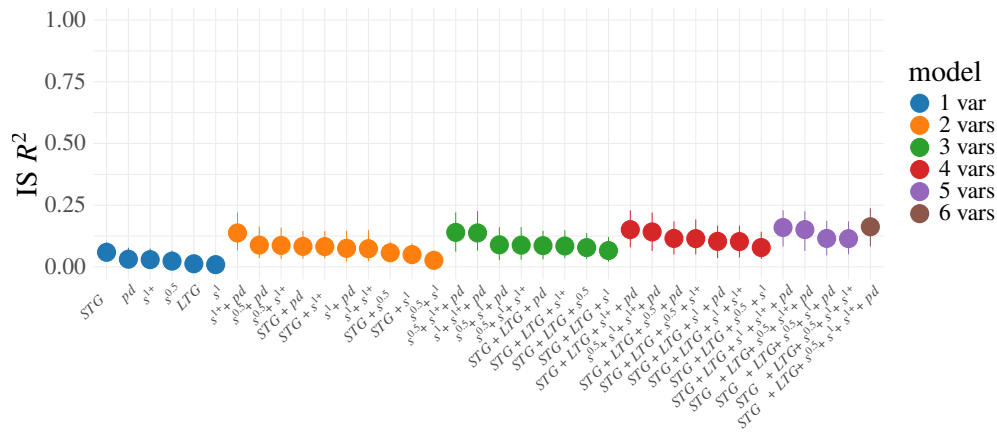
This figure presents the out-of-sample R^2 for annual return predictions as specified in equation (17), using various sample split dates. The first out-of-sample split date is January 1993, and the last date is June 2015. We calculate the out-of-sample R^2 from univariate regressions with dr and pd as predictors, respectively. ⁱ



A. IGA one-year earnings growth: Δe_{t+1}



B. IGA earnings growth between years 1 and 2: Δe_{t+2}



C. IGA earnings growth between years 2 and 3: Δe_{t+3}

Figure A.9 R^2 from Earnings Growth Predictive Regressions at Various Horizons with Bootstrapped Confidence Interval

This figure presents the in-sample R^2 for predicting annual S&P 500 Index earnings growth across different horizons using various predictors. The dependent variables are S&P earnings growth, sourced from IGA, for Year 1, Years 1 to 2, and Years 2 to 3, displayed in Panels A, B, and C, respectively. The predictors used include the price-dividend ratio (pd), one-year earnings growth forecast ($STG \equiv \mathbb{E}_t^A(\Delta e_{t+1})$), long-term earnings growth forecast (LTG), and various combinations of pd , $s^{0.5}$ (price-dividend ratio of a six-month strip), s^1 (price-dividend ratio of a one-year strip), and s^{1+} (price-dividend ratio of dividends beyond one year). Each round dot represents the in-sample R^2 with a 95% bootstrapped confidence interval.