The Information Cliff

Ye Li* Chen Wang†

March 3, 2026

Abstract

We characterize an information cliff in the stock market: the supply of information on aggregate cash flows drops precipitously beyond a one-year horizon, and so does analysts' forecast accuracy. We use a generalized state-space model to explore the implications for expected cash-flow growth and expected returns. Identifying the state-space dimensionality is the only necessary step for sharpening the model structure. Once done, the information cliff has a direct mathematical representation: the expected cash-flow component of the state space must be non-persistent. Furthermore, the expected market returns only depend on the valuation wedge between the total market and one-year dividend strip.

*We are grateful to Nick Barberis, Jules van Binsbergen, Zhi Da, Kent Daniel, Carter Davis, Joost Driessen, Stefano Giglio, Will Goetzmann, Benjamin Golez, Bryan Kelly, Ralph Koijen, Kai Li, Lars Lochstoer, Shane Miller, Alan Moreira, Riccardo Sabbatucci, Tano Santos, José Scheinkman, Jessica Wachter, Michael Weber, and conference and seminar participants at the Columbia Business School, European Finance Association, Federal Reserve Board, Florida State University Truist Beach Conference, GMO, INQUIRE UK, Johns Hopkins University, Midwest Finance Association, the Office of Financial Research 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

Understanding financial market participants' information about the aggregate economy is crucial. Asset prices reflect their expectations of future macroeconomic conditions, directly influencing resource allocation and policymaking. Market participants' expectations depend on the availability of information and how the information is processed. A large literature examines the latter, often exploring how biases in information processing distort expectation formation. Instead, we begin by characterizing the supply of information in this paper, revealing an information cliff—a sharp decline in available information on economic fundamentals beyond a one-year horizon. We then show that this critical feature sheds new light on several central objects in a canonical asset pricing framework and generates broader implications for the dynamics of expected cash flows and returns.

A critical source of information in the stock market is corporate disclosure, particularly the quarterly guidance firms release with their earnings announcements. We find that 96% of guidance pertains to performance within the next year.[1] Importantly, this information cliff is reflected in the information content of analysts' forecasts. Analyst forecasts of aggregate earnings growth for the upcoming year yield an R^2 of 73% when predicting realized earnings growth, but their accuracy plummets for subsequent periods. Specifically, the R^2 drops to around 20% when analyst forecasts for the second (months 13-24) or third year are used to predict earnings growth for the corresponding periods. The predictive power of long-term growth (LTG) forecasts is similarly weak.

Corporate announcements in the first month of a quarter are more informative for revealing aggregate economic conditions than those in other months (Guo, 2025; Guo and Wachter, 2025a). Since the information supplied by firms is predominantly about performance within the next year, we would expect the wedge between the predictive power of analysts' next-year forecasts and that of their forecasts for subsequent years to be most pronounced in the first month of a quarter. Our results confirm this hypothesis. The alignment between the timing of information supply and predictive power wedge across the horizon lends further support to the one-year cliff of available information.

[1] Beyond corporate announcements, we tabulate in the Appendix III the horizons of major professional forecasts of macroeconomic variables. The majority have horizons within four to six quarters.

To explore the asset pricing implications of the information cliff, we extend the model in Lettau and Wachter (2007) by allowing the dynamics of cash flows and risk prices to be driven by an arbitrary set of latent state variables and shocks.[2] The model provides a necessary analytical structure without imposing restrictive assumptions. Our analysis takes three steps. The first step is critical for sharpening the model structure: we develop a method to identify the dimensionality of the state space. Second, we show that once this dimensionality is pinned down, market participants' information set has a simple mathematical representation, and the information cliff has a direct implication for the (lack of) persistence of expected cash-flow growth—a key object in the asset pricing literature. In the last step, we connect the information cliff and expected market return.

To examine the state-space dimensionality, we show that logarithm of dividend strip prices scaled by realized dividends ("valuation ratios") are linear functions of the state variables. As strips across maturities differ in their state-variable loadings, their valuation ratios are linearly independent combinations of state variables and empirically span the state space. We compute the valuation ratios of S&P 500 dividend strips and find that the first two principal components account for 96% of total variance. In addition, when forecasting returns and dividend growth, the best-performing pairs of valuation ratios perform as well as three or more valuation ratios. Therefore, to span the dividend valuations, expected returns, and expected dividend growth, a two-dimensional state space suffices, despite the many economic forces that affect the aggregate stock market.

The two state variables can be rotated to represent, respectively, the conditional expected return and conditional expected dividend growth rate, and their laws of motion are AR(1) processes.[3] While prior studies have assumed such a two-dimensional structure in different contexts (e.g., Lettau and Wachter, 2007; Binsbergen and Koijen, 2010), they have not tested this assumption. For our purposes, formally identifying the state-space dimensionality is crucial; otherwise, as will become clear next, our first theoretical result on the parametric representation of information cliff would follow directly from an assumed (rather than empirically grounded) two-dimensional state space.

The conditional expectation of dividend growth rate over the next year—one of the two state

[2]The model does not impose rational expectation. The stochastic discount factor may reflect belief distortions.

[3]Note that adding more lags in AR processes means having more than two state variables.

variables—encapsulates all the information that market participants have about future cash flows. Since it is an AR(1) process, a non-zero autoregressive coefficient would imply that the current information set contains persistent signals, allowing market participants to forecast growth for the next two, three, and subsequent years. Thus, the information cliff—market participants do not have information about growth beyond the next year—implies a zero autoregressive coefficient. This autoregressive coefficient has been one of the central objects in the asset pricing literature.[4]

This direct connection between the information cliff and the persistence (autoregressive coefficient) of expected cash-flow growth rates can be tested, providing further support to the information cliff. Using analyst forecasts to proxy for cash-flow expectations, we estimate the autoregressive coefficient of expected cash-flow growth rate and find it to be consistently around zero across different specifications. For robustness, we also fit a state-space model to dividend data, using a latent variable to represent the expected growth rate. This approach yields similar results.

After exploring the connection with the information cliff and expected cash-flow growth rate, we turn to the expected return and derive another set of results from our model: if the expected cash-flow growth rate has a zero autoregressive coefficient, the expected return is a univariate function of the slope of valuation term structure; otherwise, the sign of the slope's return forecasting error aligns with that of the autoregressive coefficient. Here valuation term structure refers to the collection of valuation ratios of strips with varying maturities. The market price-dividend ratio (pd) reflects the overall valuation level, and the slope is given by the wedge between pd and the valuation ratio of one-year dividend strip. When this wedge widens, the term structure steepens: a larger (smaller) fraction of market value comes from dividends beyond the next year (within the next year).

The return predictive power of the slope is quite intuitive under the information cliff. Since market participants have limited information on cash flows beyond the very next year, the rise of valuation of cash flows beyond one year relative to valuation of next-year cash flows is not driven by improved long-term growth but due to a lower discount rate that benefits the valuation of long-term

[4]There is a large literature on the dynamics of expected cash-flow growth rate and its asset pricing implications (e.g., Bansal and Yaron, 2004; Beeler and Campbell, 2012; Belo, Collin-Dufresne, and Goldstein, 2015; Collin-Dufresne, Johannes, and Lochstoer, 2016). In an endowment economy, the cash flow is both firms' payout and aggregate consumption, but in reality, they differ. Our theoretical and empirical analysis are about on the former.

cash flows more than that of near-term cash flows. Thus, a steepening of valuation term structure predicts lower market returns. The slope delivers an in-sample R^2 of 25% and an out-of-sample R^2 of 15% and subsumes the predictive power of pd. Augmenting the slope with other predictors from the literature does not improve the performance, in line with our result on the expected return being a univariate function of the slope.[5] Finally, our rolling-window estimation shows that when the autoregressive coefficient of expected cash-flow growth deviates from zero, its sign aligns with that of the slope's return forecasting error, consistent with our model prediction.

Our exponential-affine model based on Lettau and Wachter (2007) implies a linear mapping from the slope of valuation term structure to the expected return under the information cliff, which corresponds to the standard predictive regression. To address the concern of nonlinearity, we show that the slope in a linear regression outperforms machine learning algorithms in (Kelly, Malamud, and Zhou, 2024) that nonlinearly aggregate a large set of predictors (including the slope itself).

Beyond discount-rate variation, our analytical framework and return predictability results have a mispricing interpretation. Traditional market timing bets against the overall valuation level, pd. Our findings indicate that market participants are well informed of the near term but face an information cliff beyond one year, so mispricing is likely in the long-term cash flows, and the focus should be on the slope rather than the level of valuation term structure. To time the market is to reduce exposure when the valuation term structure steepens and increase exposure when it flattens.[6] Betting against the slope is betting against exuberance or pessimism about the long term.

Literature. We characterize an information cliff that market participants face, stemming from the prevailing corporate practice of providing guidance on performance within the next year. In contrast, prior studies do not discuss information supply but instead focus on distortions in market participants' information usage and how to formalize various biases in their long-term expectations, such as over- or under-reaction (Afrouzi et al., 2023; Bordalo et al., 2024a,b; Enke and Graeber,

[5]The slope outperforms other predictors, including those summarized in Goyal and Welch (2007) and from more recent papers, across evaluation metrics, such as Hodrick (1992) adjustment for standard errors and out-of-sample tests (e.g., encompassing (ENC) and Clark-West (CW) tests).

[6]Its Sharpe ratio of 0.58 is 55% higher than that of buy-and-hold strategy (Campbell and Thompson, 2008).

2023; Wang, 2020), false pattern recognition (Barberis et al., 1998; Guo and Wachter, 2025b), optimism (Cassella et al., 2023), bounded rationality (De Silva and Thesmar, 2024), and failure to distinguish old versus new information (Guo, 2025; Guo and Wachter, 2025a).

We do not aim to model or pin down a specific bias in how agents utilize information like those listed above; instead, we emphasize the supply of information and focus on a sharp decline of information at the one-year horizon and its asset pricing implications.[7] Leveraging the analytical framework of Lettau and Wachter (2007) and our novel method of identifying the state-space dimensionality, we provide findings that have not been previously discussed, such as the lack of persistence of expected annual growth rate of aggregate cash flows and the connection between the expected market return and the slope of the valuation term structure. This focus on a specific cutoff horizon (one year) and our asset pricing results on the aggregate market distinguish our paper from the existing literature that study how specific belief biases distort agents' long-term versus short-term expectations of firm-level growth (e.g., Da and Warachka, 2011; Bordalo et al., 2019; Cassella et al., 2023; De Silva and Thesmar, 2024; Guo and Wachter, 2025a).

Given the information cliff at the one-year horizon, we characterize how it affects the information content of analyst forecasts, contributing to 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).[8] Our paper is particularly related to studies on how distortions in long-term forecasts lead to mispricing and return predictability (e.g., La Porta, 1996; Nagel and Xu, 2022; Bordalo et al., 2024b). Our unique focus on the one-year information cliff leads to the discovery of the slope of valuation term structure as a powerful return predictor that outperforms other predictors across various metrics. These results contribute to the literature on return predictability.[9]

[7]We do not study short-termism, a related topic in the accounting literature on managerial incentives, disclosure, and firms' performance (e.g., Bushee, 2001; Bhojraj and Libby, 2005; Call et al., 2014; Brochet et al., 2015).

[8]This literature includes studies on expectations of firm-level performance (e.g., 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; Guo and Wachter, 2025a), aggregate market returns and cash flows (e.g., Chen et al., 2013; De La O and Myers, 2021; Gao and Martin, 2021; Hillenbrand and McCarthy, 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 the macroeconomy (e.g., 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).

[9]There is a vast and growing literature on return predictability (e.g., Fama and French, 1988; Campbell and Shiller,

The timing of information supply plays an important role in our findings: while analyst forecasts predict next-year earnings better than subsequent-year earnings, this difference is most pronounced in the first month of a quarter when firms supply more information about the aggregate economy (Guo, 2025; Guo and Wachter, 2025a). Our focus on the impact of information-supply timing on market participants' cash-flow expectations differs sharply from prior work, which predominantly examines its effect on returns. This literature has grown significantly since Savor and Wilson (2013) and Lucca and Moench (2015) established that stock market returns are significantly higher on announcement days than non-announcement days (with theoretical explanations provided by, e.g., Ai and Bansal (2018) and Wachter and Zhu (2022)). Further studies have shown that information-supply timing affects equity-market anomalies (Engelberg et al., 2018), influences the performance of CAPM and factor models (Savor and Wilson, 2014; Gilbert et al., 2018), and drives cross-firm spillover effects (Savor and Wilson, 2016; Ben-Rephael et al., 2021).

2 Direct Evidence on the Information Cliff

The overall data structure. This paper draws on three main categories of data, primarily from 1988 to 2019, to document the information cliff in agents' cash-flow expectations, embed it in a state-space model, identify the state-space dimensionality, and explore the asset pricing implications. First, to provide direct, model-free evidence of the information cliff in Section 2, we use data on corporate guidance of earnings and sales growth from IBES Guidance; subjective expectations (analyst forecasts) and realized earnings growth rates from IBES Global Aggregates (IGA); and self-aggregated long-term growth (LTG) forecasts from firm-level IBES Unadjusted Summary Files. Second, to construct our state variable proxies in Section 3, we compute valuation ratios of dividend strips primarily using prices of S&P 500 index futures from LSEG Datastream and Fama-Bliss zero-coupon bonds from the CRSP US Treasury Database and alternatively using options

1988; Baker and Wurgler, 2000; Lettau and Ludvigson, 2001; Lewellen, 2004; Cochrane, 2008; 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; Kelly et al., 2024; Bordalo et al., 2024b).

Table 1 Summary Statistics

This table reports summary statistics, including the number of observations, mean, standard deviation, minimum, maximum, and quartiles, for the main variables. The sample is monthly from 1988:01 to 2019:12. Panel A summarizes $\mathbb{E}_t^A[\Delta e_{t+1}]$, $\mathbb{E}_t^A[\Delta e_{t+2}]$, and $\mathbb{E}_t^A[\Delta e_{t+3}]$, analysts' forecasts of earnings growth for the first, second, and third year into the future from IBES Global Aggregate (IGA). Δe_t is the realized annual earnings growth from IGA. LTG_t is forecasts of long-term earnings growth, which we self-aggregate from the IBES Unadjusted US Summary Statistics File. Panel B summarizes the rest of variables. These include our main return predictor, dr ("slope"); the price-dividend ratio pd of the S&P 500 index; the filtered series for demeaned expected returns (μ^F) and dividend growth (g^F) from Binsbergen and Koijen (2010); the predictive factors for return (KP) and dividend growth (KP^{CF}) from Kelly and Pruitt (2013); dividend strip price-to-dividend ratios for 0.5 and 1 years ($s^{0.5}$, s^1); the long-term dividend strip price-to-dividend ratio for horizons beyond 1 year (s^{1+}); one-month and one-year log returns of the S&P 500 index ($r_{t+1/12}$, r_{t+1}); one-month and one-year log returns of the Fama-French market portfolio ($r^{MKT}_{t+1/12}$, r^{MKT}_{t+1}); and the 1-year dividend growth rate of the S&P 500 index and the Fama-French market portfolio (Δd_{t+1}, Δd^{MKT}_{t+1}).

	mean	std	min	25%	50%	75%	max
Panel A: Analyst earnings growth forecasts							
$\mathbb{E}_t^A[\Delta e_{t+1}]$	0.103	0.096	-0.167	0.056	0.103	0.154	0.425
$\mathbb{E}_t^A[\Delta e_{t+2}]$	0.134	0.043	-0.069	0.104	0.127	0.157	0.269
$\mathbb{E}_t^A[\Delta e_{t+3}]$	0.130	0.036	0.052	0.100	0.122	0.159	0.217
Δe_t	0.072	0.135	-0.380	-0.008	0.092	0.148	0.425
LTG_t	0.125	0.018	0.093	0.115	0.120	0.129	0.187
Panel B: Predictors, returns and dividend growth							
dr_t	4.027	0.494	2.952	3.727	4.044	4.208	6.632
pd_t	3.883	0.289	3.239	3.656	3.930	4.047	4.524
μ_t^F	-0.039	0.024	-0.091	-0.051	-0.041	-0.024	0.010
g_t^F	0.019	0.059	-0.233	-0.002	0.031	0.056	0.132
KP_t	-0.504	0.073	-0.725	-0.562	-0.482	-0.450	-0.378
KP_t^{CF}	-0.385	0.068	-0.605	-0.422	-0.389	-0.338	-0.220
$s_t^{0.5}$	-0.819	0.281	-2.629	-0.883	-0.768	-0.666	-0.280
s_t^1	-0.142	0.280	-2.241	-0.210	-0.098	0.016	0.393
s_t^{1+}	3.863	0.297	3.204	3.629	3.913	4.030	4.521
$r_{t+1/12}$	0.009	0.041	-0.184	-0.015	0.013	0.034	0.108
r_{t+1}	0.095	0.157	-0.568	0.046	0.126	0.187	0.429
$r^{MKT}_{t+1/12}$	0.009	0.042	-0.187	-0.016	0.014	0.036	0.108
r^{MKT}_{t+1}	0.096	0.159	-0.554	0.036	0.128	0.194	0.440
Δd_{t+1}	0.059	0.070	-0.237	0.025	0.068	0.112	0.168
Δd^{MKT}_{t+1}	0.058	0.081	-0.207	0.018	0.051	0.107	0.262

(from OptionMetrics) and dividend futures (from Goldman Sachs and Bloomberg) of the S&P 500 index for robustness. These ratios allow us to map out the latent state variables, which is key for identifying the state-space dimensionality. Finally, to test the model's implications on return predictability in Section 4, we benchmark our return predictor against a wide array of predictors

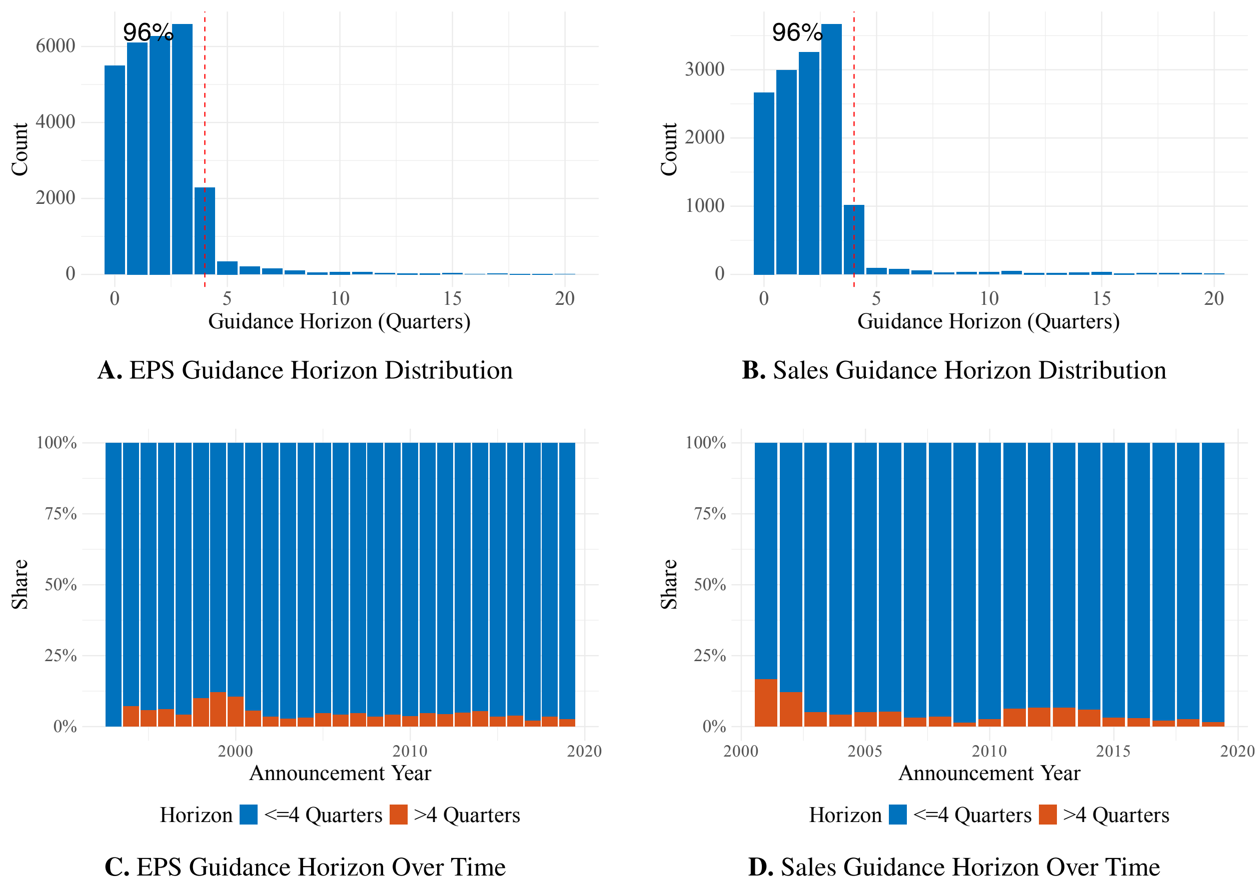


Figure 1 The One-Year Cliff of Corporate Earnings and Sales Guidance

This figure shows the horizon of management guidance announcements for S&P 500 firms from 1992 to 2019 (IBES Guidance). Panels A and B show the distribution of guidance by quarterly horizons for EPS and sales, respectively. The annotations highlight the percentage of guidance for the next four quarters or less. Panels C and D plot the time series of the proportion of guidance for a horizon of four quarters or less for EPS and sales, respectively.

from the literature and machine learning models in a series of forecasting and spanning exercises. Table 1 provides the summary statistics for the main variables. We will introduce the variable names and specific details of data sources and variable construction in each relevant section.

Information supply and the one-year cliff. Information releases from publicly listed firms are an important source of signals about aggregate economic conditions. We characterize a sharp decline of information supply at the one-year horizon, using the IBES Guidance dataset that spans the period from 1992 to 2019 and includes 28,000 management guidance announcements from S&P

500 firms.[10] In Panel A of Figure 1, we plot the frequency distribution of firms' earnings guidance. The vertical axis reports the number of earnings guidance, and the horizontal axis shows the guidance horizon. Firms provide guidance towards the next fiscal year end and very rarely towards the next two or three fiscal year ends. The guidance horizon is the difference between a fiscal year end and the announcement date. In the frequency distribution, there is a sharp discontinuity around the one-year horizon, beyond which firms seldom provide earnings guidance. An overwhelming 96% of all guidance announcements are for horizons of one year or less.

This scarcity of long-term corporate guidance creates an information cliff for the market participants, leaving analysts and investors with little direct information from firms to anchor their forecasts beyond the one-year mark. As shown in Panel C of Figure 1—where we compare the fraction of guidance within four quarters to the fraction of guidance beyond four quarters, this short-term focus in corporate guidance is not a recent phenomenon but has been a consistent feature of corporate disclosure practices throughout our sample period.

One natural concern regarding earnings guidance is that corporate earnings may be subject to management manipulation, which may limit their informativeness. However, we find a similar pattern in sales guidance (see Panel B and D of Figure 1). Sales are less susceptible to manipulation and is thus a more objective measure of corporate performance. The consistency across both earnings and sales guidance suggests that the information cliff is not about firms' earnings management in conjunction with strategic disclosure but reflects standard corporate practices.

Since the aggregate stock market's cash flow is closely tied to the macroeconomy, investors may also rely on forecasts of macroeconomic variables from professionals and policymakers. Table A.4 in the Internet Appendix summarizes the maximum forecast horizons for U.S. economic growth from several leading sources, including the Blue Chip Economic Indicators, the Survey of Professional Forecasters, the Livingston Survey, Consensus Economics, the Wall Street Journal Economic Survey, and the Federal Reserve's Summary of Economic Projections (SEP). While some surveys, like the SEP, provide three-year forecasts for annual GDP growth, the majority

[10] Earnings guidance has become increasingly prevalent (Penman, 1980; Hutton et al., 2003; Miller, 2002).

of professional forecasts cluster around the annual horizon. Some longer-term forecasts exist (e.g., Blue Chip's semiannual 5- to 10-year projections) but are typically multi-year averages. Such averaging smooths out fluctuations and diminishes their informativeness about economic environments at specific horizons beyond one year. This scarcity of detailed, long-term economic forecasts reinforces the information cliff that the stock-market participants face.

The information cliff and cash-flow expectations. We examine how the scarcity of information on long-term growth affects market participants' expectations of future cash flows. IBES Global Aggregates (IGA) provides analysts' forecast of earnings growth for the S&P 500 index based on firm-level forecasts. The aggregation procedure weighs individual companies by their market capitalization.[11] 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 of one, two, and three years (i.e., $k = 1, 2, 3$), and the growth rate forecasts are denoted by $\mathbb{E}_t^A[\Delta e_{t+1}]$, $\mathbb{E}_t^A[\Delta e_{t+2}]$, and $\mathbb{E}_t^A[\Delta e_{t+3}]$, respectively.[12] The data is available at a weekly frequency. We will consider both weekly and monthly frequencies in our regression analysis. For estimation at monthly frequency, we take the last weekly observation of each month. Our sample is from January 1988 to December 2019. We also utilize analyst forecasts for long-term growth (LTG). We aggregate firm-level LTG from IBES to the index level.[13] This data is available at the monthly frequency. It is aggregated to the index level via the same aggregation procedure described above.

Using analysts' forecasts, we predict earnings growth over the next 12 months, from the 13th to 24th month, and from the 25th to 36th month. The information cliff manifests as a sharp deterioration in forecasting performance once the horizon extends beyond one year. Table 2 report the results. In column (1), we regress the realized one-year growth rate of aggregate earnings from firms covered by IGA on the analysts' forecast of the same one-year horizon. The R^2 is 73%,

[11]To 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.

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

[13]IBES 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., 2024b, and De la O and Myers, 2024).

Table 2 Predicting Earnings Growth Across Horizons with Analyst Forecasts

This table reports the results of regressions that predict earnings growth at various horizons with analyst forecasts. The dependent variables are realized earnings growth from IGA of next year, the year after, and the third year in the future, and the average earnings growth between years 3 to 5. The independent variables are analysts' forecasts of one-year earnings growth between $t+\tau$ to $t+\tau+1$ across horizons ($\mathbb{E}_t^A[\Delta e_{t+\tau+1}]$, for $\tau = 0, 1, 2$) from IGA and the self-aggregated long-term earnings growth forecasts (LTG_t) of the S&P 500 Index. The t-statistics are calculated based on Newey-West standard errors with 18 lags are reported in parentheses.

	Full Sample				Feb–Jun Subsample			
	Δe_{t+1}	Δe_{t+2}	Δe_{t+3}	$\overline{\Delta e_{t+2,t+5}}$	Δe_{t+1}	Δe_{t+2}	Δe_{t+3}	$\overline{\Delta e_{t+2,t+5}}$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Intercept	-0.056	-0.097	-0.149	0.004	-0.071	-0.163	-0.111	-0.031
	(-4.127)	(-2.063)	(-2.130)	(0.047)	(-2.805)	(-2.520)	(-1.570)	(-0.357)
$\mathbb{E}_t^A[\Delta e_{t+1}]$	1.204				1.217			
	(20.101)				(8.562)			
$\mathbb{E}_t^A[\Delta e_{t+2}]$		1.164				1.615		
		(3.683)				(4.216)		
$\mathbb{E}_t^A[\Delta e_{t+3}]$			1.598				1.363	
			(3.342)				(2.596)	
LTG_t				0.486				0.770
				(0.717)				(1.107)
N	372	360	348	324	155	150	145	135
R^2	0.73	0.15	0.21	0.01	0.55	0.21	0.17	0.03

indicating that analysts in general are able to forecast near-term growth quite well.

Next, we change the forecasting target while maintaining the setup of right-side variables in the regressions. In column (2) of Table 2, the analysts' forecast for earnings growth from the 13th to 24th month only has a R^2 of 15% for predicting earnings growth at the same horizon, which stands in sharp contrast to the forecasting performance at one-year horizon in Column (1). We obtain similar results when forecasting growth from the 25th to 36th month in column (3). In the last column, we replicate the exercise for the average growth for years 3, 4 and 5, $\overline{\Delta e_{t+2,t+5}} = (\Delta e_{t+3} + \Delta e_{t+4} + \Delta e_{t+5})/3$, which corresponds to the forecast horizon of LTG (a period of three to five years as per IBES's definition).

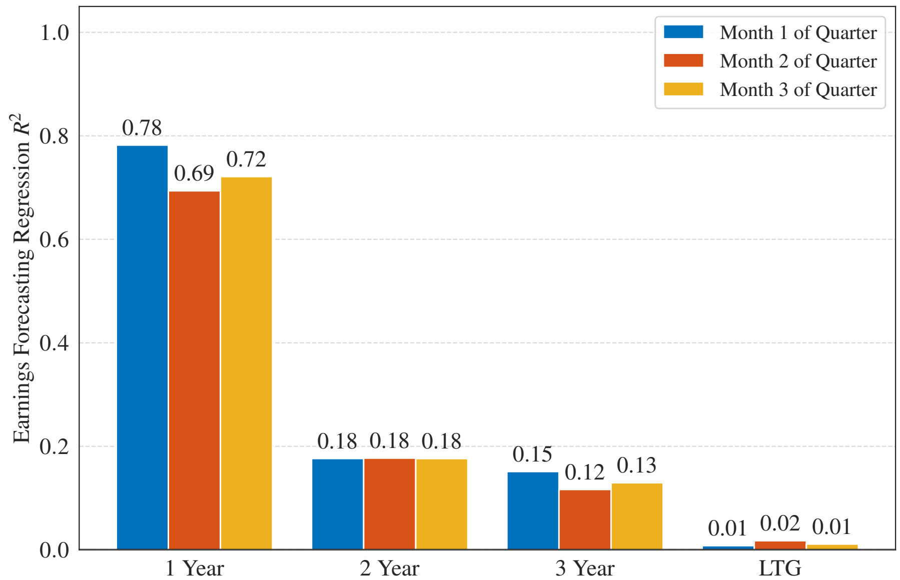
In the IBES Global Aggregates database, a fiscal year (FY) for the U.S. aggregate stock market is defined as the period from June to May next year. For example, in October 2000, the earnings growth rate over the next FY is the growth rate from FY2000 (June 1999–May 2000) to FY2001 (June 2000–May 2001). Consequently, the next FY's data in the numerator of this growth rate incorporates realized earnings from firms whose fiscal years end in June, July, August, or September 2000.[14] Similarly, in July 2000, the next FY contains realized earnings from firms with fiscal years ending in June 2000, while in January 2001, the next FY contains realized earnings from firms with fiscal years ending between June and December 2000. IBES shifts the FY designation forward in February. As a result, the earnings growth rate in February 2001 does not contain any realized earnings, since the next FY becomes FY2002, spanning June 2001 to May 2002. The same holds for March, April, May, and June 2001: none of these months contain realized earnings in the earnings growth rate over the next FY.

When the earnings growth rate does contain realized earnings, so do analyst forecasts, which tend to inflate the R^2 in regressions that use analyst forecasts to explain the next year's realized earnings growth, as both the forecasts and the forecast targets incorporate already-realized earnings data. Therefore, in Columns (5) through (8), we re-estimate the same regression but restrict the sample to the subset of months—February through June—that are free from this issue, as the next FY in these months does not yet contain any realized earnings.

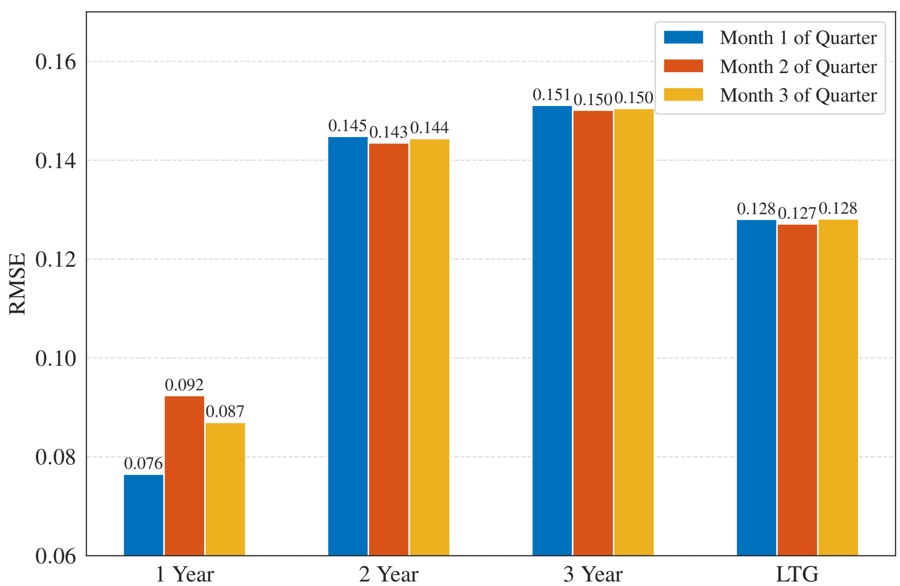
Exploring information seasonality. Next, we provide more evidence on the information cliff by exploring the seasonality of information release. As pointed out by Guo (2025) and Guo and Wachter (2025a), corporate announcements in the first month of a quarter contain more information about the aggregate economic conditions than the second and third months.

Given that the information supplied is mainly about firms' performance within the coming year rather than in the subsequent years, we have three hypotheses. First, when predicting earnings growth within the next year, analyst forecasts should exhibit stronger predictive power in the first month of a quarter as firms supply information. Second, when predicting earnings growth beyond one year, the predictive power of analyst forecasts should not vary a lot across months of a quarter, because even though firms supply more information in the first month, the information is not about

[14]Note that firms rarely report earnings in the same month as their fiscal period ends, typically requiring at least a one-month lag. It is therefore unlikely that, as of October 2000, any firm has reported earnings for a fiscal period ending in October 2000. For this reason, the next FY's data in October 2000 is unlikely to contain realized earnings from firms with fiscal years ending in October.



A. In-sample R^2 by month of the quarter



B. RMSE by month of the quarter

Figure 2 Predicting Earnings Growth Across Horizons by Month of a Quarter

This figure reports, respectively in Pane A and B, the R^2 of predictive regression and the Root Mean Squared Error (RMSE) of predicting earnings at various horizons with analyst forecasts (as done in Table 2), with the regression samples separated by month of a quarter to highlight the effect of information supply. The RMSE for LTG (3-year average growth) is annualized by multiplying by $\sqrt{3}$ for comparability.

growth beyond the very next year. Third, the wedge between one-year and long-term predictive power of analyst forecasts should be wider in the first month of a quarter with information released on the upcoming year but not for the subsequent years.

To examine the three hypotheses, we use analyst forecasts to predict earnings growth as in Table 2 but separate the sample by month of a quarter. In Figure 2, we plot the predictive regression R^2 in the top panel and the root mean squared error (RMSE) in the bottom panel, with regression results in Table A.2 of Appendix III. Note that since the forecasting target in the last specification involving the LTG is an average of growth rates over three years, we multiply its RMSE by $\sqrt{3}$ so that its magnitude is comparable to that of RMSE from other forecasting regressions.

Our findings in Figure 2 support the three hypotheses. First, analyst forecasts of aggregate earnings show greater accuracy measured in both R^2 and RMSE in the first month of a quarter than in the second and third months, a result attributable to the fresh information from firms' performance guidance in the first month. Second, once the forecasting horizon extends to 2 or 3 years, the accuracy of analyst forecasts does not exhibit as large a variation across months of a quarter. Third, comparing R^2 and RMSE across forecasting horizons (i.e., within the same color), we can see that the accuracy wedge between analyst forecast of one-year earnings and analyst forecast for the subsequent years is wider in the first month of a quarter. For example, the R^2 drops from 78% to 18% (a wedge of 60%) in the first month when the forecasting period changes from the first to the second year, but in the second month, the R^2 wedge is 51% (= 69% - 18%).

Overall, our empirical exercises in this section demonstrate that the supply of information from corporate announcements has a one-year cliff, which corresponds to the significantly higher accuracy of analyst earnings forecasts at the one-year horizon compared to longer horizons. The importance of information supply as a driver of this accuracy wedge is further supported by the finding that the wedge is widest during the first month of the quarter, coinciding with peak information release.[15] The following sections embed the information cliff into a canonical model and explore its implications on several central objects in the asset pricing literature.

We reproduce the within-quarter predictability patterns from Figure 2 using only data from

[15] A potential concern is whether the sharp decline of forecasting accuracy is due to agents' extrapolating past fundamentals (e.g., Barberis et al., 1998; Barberis, 2018). We test this possibility by including lagged earnings growth as a control in our forecasting regressions. Table A.3 in the Internet Appendix shows that our main results are robust.

February to April. Data from Feb to June does not have the issue of next FY's earnings containing realized data. We exclude May and June so that we have one first month of a quarter (April), one second month (February), and one third month (March), together constituting a sample with balanced exposure to information seasonality (rather than having more observations to the second and third months). Figure A.2 in the Internet Appendix conveys similar messages to those in Figure 2. For one-year earnings growth, the predictive R^2 is 0.55 in the first month of the quarter (April), declining to 0.47 in the second and third months (February and March). The RMSE exhibits a similar pattern as well, remaining lower in the first month than in the second and third months.[16]

3 A Parametric Representation of Information Cliff

We provide a parametric representation of the information cliff through a state-space model (e.g., Lettau and Wachter, 2007). In the model, cash-flow expectations are driven by the state variables. To sharpen the characterization of information cliff, we develop a method to identify the dimensionality of state space. We find that the state space is two-dimensional, and therefore, the model can be reduced to having one state variable driving the expected return (via the price of risk) and the other driving the conditional expectation of cash-flow growth rate. The information cliff is equivalently represented as a condition on the autocorrelation of the expected cash-flow growth rate.

3.1 A generic state-space model

We consider a dynamic economy where the information filtration is given by a Markov process. 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

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

[16]Figure A.3 in the Internet Appendix reports similar results for the Feb.–Jul. sample. While July may suffer from the problem of next FY's earnings containing realized data, we include July to have a balanced seasonality exposure in the sample, with two first months (Apr. and Jul.), two second months (Feb. and May), and two third months (Mar. and Jun.).

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,

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

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

$$g_t = \overline{g} + \phi^\top X_t - \frac{1}{2}\sigma_D^\top \Sigma \sigma_D. \tag{3}$$

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

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

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

$$\lambda_t = \overline{\lambda} + \theta^\top X_t. \tag{5}$$

We do not impose restrictions on θ, the state-variable loadings of the prices of risks, λ_t. The prices of risks (or associated the change of measure) may also be interpreted as reflecting belief distortions, in which case agents' subjective expectation may deviate from the rational expectation.

The information cliff is about whether agents' current information set contains useful signal about cash-flow growth beyond the next year. In this generic setup, characterizing this property seems rather complicated. Next, we investigate the dimensionality of X_t, and then we show that once the dimensionality is pinned down, characterizing the information cliff becomes straightforward.

3.2 State space dimensionality

The measurement framework. To determine the dimensionality of X_t, we develop a method based the mapping between dividend strip prices and the state variables, X_t. Let P_t^n denote the time-t price of dividend at $t+n$. The no-arbitrage condition implies a recursive equation: for $n \geq 1$,

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

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

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

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

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

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

Thus, a collection of log price-dividend ratios of dividend strips can span the state space. The revealed (rotated) state variables may represent various underlying economic forces.[17] Next, we empirically identify the state-space dimensionality by analyzing the strip valuation ratios.

Mapping out the state variables. To obtain P_t^n, we first calculate the value of dividends paid beyond the first n years, denoted by P_t^{n+}. Under the risk-neutral measure ("RN"),

$$P_t^{n+} = e^{-nr_f} \mathbb{E}_t^{RN} \left[\sum_{\tau=1}^{+\infty} e^{-\tau r_f} D_{t+n+\tau} \right] = e^{-nr_f} \mathbb{E}_t^{RN} \left[\mathbb{E}_{t+n}^{RN} \left[\sum_{\tau=1}^{+\infty} e^{-\tau r_f} D_{t+n+\tau} \right] \right], \tag{9}$$

[17] 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 (Lettau and Wachter, 2007). 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; Lettau and Wachter, 2007, 2011; Koijen, Lustig, and Van Nieuwerburgh, 2015; Backus, Boyarchenko, and Chernov, 2018; Kragt, de Jong, and Driessen, 2020). Our analysis relies on the prices of dividend strips to map out state variables. Giglio, Kelly, and Kozak (2024) analyze the dual problem—that is, they compute strip prices from empirically specified and observed dynamics of state variables.

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

$$P_{t+n} = \mathbb{E}_{t+n}^{RN}\left[\sum_{\tau=1}^{+\infty} e^{-\tau r_f} D_{t+n+\tau}\right], \tag{10}$$

so we have

$$P_t^{n+} = e^{-nr_f}\mathbb{E}_t^{RN}[P_{t+n}]. \tag{11}$$

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).

Once we obtain, for example, P^{1+}, the price of one-year dividend, P_t^1, is given by

$$P_t^1 = P_t - P_t^{1+}, \tag{12}$$

which is the difference between the price of dividends across all horizons, P_t (i.e., the equity price), and the price of dividends paid after the first 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, $s_t^{1+} = \ln(P^{1+}/D_t)$.[18] We consider these equity strips as the futures data at 0.5 and 1 year maturities are the most liquid.

For futures prices, we use S&P 500 index futures, which are the most actively traded equity futures. The futures prices are from Datastream.[19] The zero-coupon bond prices are from the Fama-Bliss database. The return and market capitalization 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.[20] After the market crash of October

[18] There 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.

[19] We 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.

[20] 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.

1987, regulators overhauled several trade-clearing protocols.[21] Our sample ends in December 2019. Lastly, Fama-French factors at the monthly frequency are obtained from Ken French's website.

For robustness, we construct dividend strip prices with two alternative methods.[22] First, using S&P 500 index options data from OptionMetrics (1996-2019), we derive the implied dividend strip prices for 0.5 to 2 years maturity via the put-call parity, following Binsbergen et al. (2012). Second, we use S&P 500 dividend futures data from Goldman Sachs and Bloomberg (2005-2019) and interpolate the futures curve for dividend strip prices with maturities ranging from 1 to 7 years, following Binsbergen et al. (2013).[23] We choose the futures-implied dividend strips as our main sample due to the longer time series and greater liquidity.

Measuring state space dimensionality. As shown in Section 3.1, valuation ratios of dividend strips are linear combinations of the underlying state variables. In Panel A of Figure 3, we report the results from principal component analysis (PCA) of strip valuation ratios computed from the index futures data (our baseline sample). The first two components account for 96% of total variance. We show the PCA results based on the options data in Panel B of Figure 3 and the results based on dividend futures data in Panel C. These results indicate that the state space is two-dimensional.[24]

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, such as the expected return and cash-flow growth. Next, we take a predictive regression approach as the expected return and expected dividend growth rate are driven by the state variables.

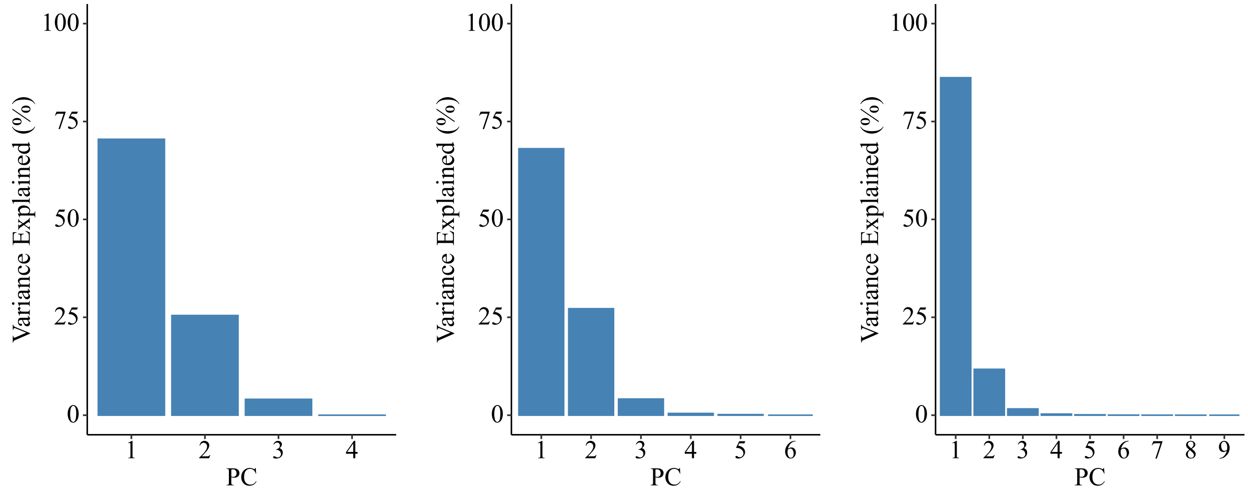
In Figure 4, we report R^2 of predicting S&P 500 dividend growth over the next year using

[21] The 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."

[22] A 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).

[23] We are grateful to Christian Mueller-Glissmann at Goldman Sachs for sharing the dividend futures data

[24] We formally test this dimensionality in Panel C of Figure 3, where we have the largest universe of (9 distinct) valuation ratios and thus potentially the richest set of state-variable proxies, using Horn's Parallel Analysis (see Figure A.8 in the Internet Appendix). The eigenvalues of the first two factors significantly exceed those from randomly resampled data, while the third does not, confirming the two-dimensional specification.



A. Futures (1988-2019) **B.** Options (1996-2019) **C.** Dividend Futures (2005-2019)

Figure 3 Principal Component Analysis of Dividend Strip Valuation Ratios

This figure presents three plots from a principal component analysis of valuation ratios computed from three data sources. Panel A is based on dividend strip prices from the S&P 500 futures from 1988 to 2019. The valuation ratios include pd_t, $s_t^{0.5}$, s_t^1, and s_t^{1+}. Panel B is based on dividend strip prices from the S&P 500 options from 1996 to 2019. The valuation ratios include pd_t, $s_t^{0.5}$, s_t^1, $s_t^{1.5}$, s_t^2, and s_t^{2+}. Panel C is based on dividend strip prices from the S&P 500 dividend futures from 2005 to 2019. The valuation ratios include pd_t, s_t^i (for $i \in \{1, 2, \ldots, 7\}$), and s_t^{7+}.

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.5 in the Internet Appendix. Our predictive regression is run on monthly observations. For comparison, we include predictors from prior studies. 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.

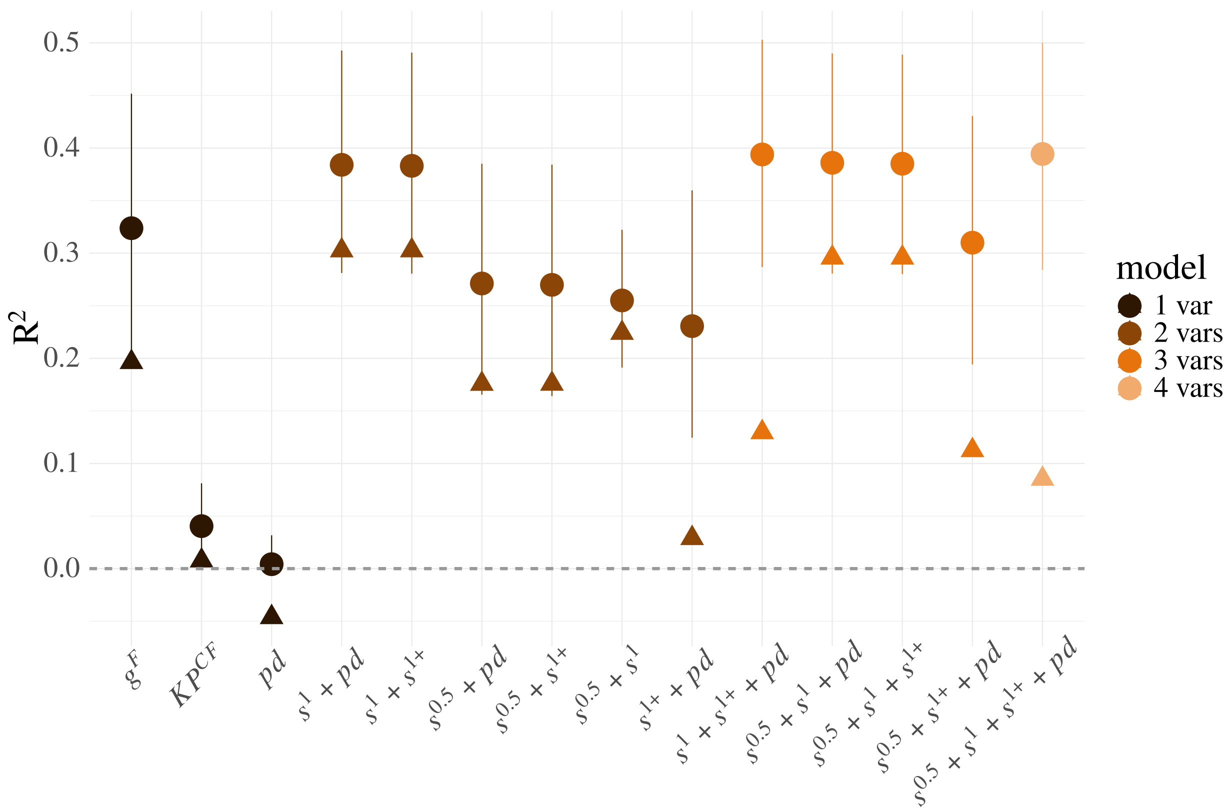


Figure 4 In-Sample and Out-of-Sample R^2 from Dividend Growth Predictive Regressions

This figure reports in- and out-of-sample R^2 for predicting annual S&P 500 Index dividend growth. The predictors include the predictor from Binsbergen and Koijen (2010) (g^F), the 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 dividend growth beginning in 1998:01.

Our results show that there exist two valuation ratios, for example, s^1 and s^{1+}, sufficient for forecasting dividend growth, indicating the state space being two-dimensional. Note that the cash flow predictive power varies across different sets of valuation ratios, suggesting that it is important to take a predictive regression approach rather than simply relying on the PCA of valuation ratios. Any given pair of valuation ratios can almost fully span the two principal components, but they may contain distinct information about return and cash flow dynamics.

In Figure 5, we report the R^2 of predicting annual returns of the S&P 500 with different sets of valuation ratios.[25] Our regression is run monthly. We report the detailed regression results in Table

[25] Binsbergen et al. (2013) use valuation ratios of dividend strips to forecast strip returns and dividends. Different from Binsbergen et al. (2013), our focus is on the asset that strips aggregate up to (equity index) rather than strips.

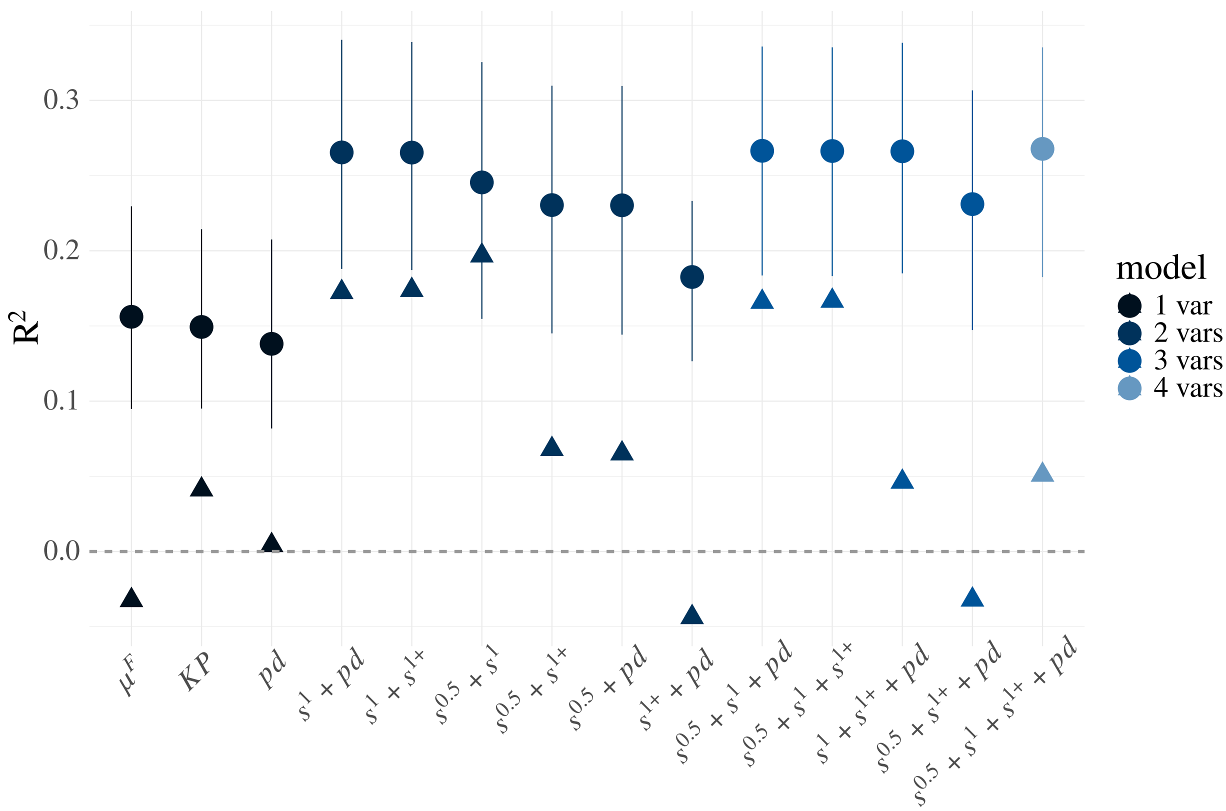


Figure 5 In-Sample and Out-of-Sample R^2 from Return Predictive Regressions

This figure reports in- and out-of-sample R^2 for predicting annual S&P 500 Index returns. The predictors include the predictor from Binsbergen and Koijen (2010) (μ^F), the 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.

A.6 in the Internet Appendix. For comparison, we include predictors motivated by a state-space approach, such as μ^F from Binsbergen and Koijen (2010) and KP from Kelly and Pruitt (2013), and we add the price-dividend ratio. 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, for example, s^1 and s^{1+}, indicating that the state space is two-dimensional.

3.3 The information cliff: a parametric presentation

We have shown that two state variables suffice to span the spaces of dividend valuations across maturities, expected returns, and expected cash-flow growth rates. Next, we set the dimension of state variables, X_t, to two in the model from Section 3.1. This will allow us to derive a sharp implication of the information cliff on expected cash-flow growth and develop an empirical test.

As in Lettau and Wachter (2007), we rotate the state variables so that one drives the price of risk, while the other drives the expected dividend growth rate. Let $X_t = [y_t, z_t]^\top$. The state variable, y_t, with a law of motion

$$y_{t+1} = \rho_y y_t + \sigma_y^\top \epsilon_{t+1}\,, \tag{13}$$

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

$$\lambda_t = \overline{\lambda} + y_t, \tag{14}$$

and the stochastic discount factor (SDF) is given by

$$M_{t+1} = \exp\left\{-r_f - \frac{1}{2}\lambda_t^2(\sigma_\lambda^\top \Sigma \sigma_\lambda) - \lambda_t \sigma_\lambda^\top \epsilon_{t+1}\right\}, \tag{15}$$

where, as in Section 3.1, ϵ_{t+1} is a N-by-1 vector of shocks, independent over time with normal distribution $N(\mathbf{0}, \Sigma)$. 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 expected dividend growth rate, $g_t = \mathbb{E}[\ln(D_{t+1}/D_t)]$, is given by

$$g_t = \overline{g} + z_t - \frac{1}{2}\sigma_D^\top \Sigma \sigma_D, \tag{16}$$

where the state variable, z_t, has the following law of motion

$$z_{t+1} = \rho_z z_t + \sigma_z^\top \epsilon_{t+1}\,. \tag{17}$$

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.

Throughout our analysis, the expectation operator represents the econometricians' belief or rational expectation, and g_t, is rational expectation of cash-flow growth over the next year that is a univariate function of z_t. We allow agents' subjective expectation to deviate from rational expectation. In this two-dimensional setup, a distorted belief, denoted by $\hat{\mathcal{P}}$ can be defined through

Table 3 Predicting Dividend Growth Across Horizons.

This table presents the results of predictive regressions for dividend growth across multiple horizons. The dependent variables are the realized dividend growth rates for the subsequent first, second, and third years, sourced from Bloomberg. The independent variables are the slope of the valuation term structure, dr_t, and the price-dividend ratio, pd_t. t-statistics calculated based on Newey-West standard errors with 18 lags are reported in parentheses. Data sample: 1988:01–2019:12.

	Δd_{t+1}	Δd_{t+2}	Δd_{t+3}
	(1)	(2)	(3)
Intercept	-0.312	-0.105	0.122
	(-1.526)	(-0.681)	(0.684)
dr_t	-0.181	-0.116	0.018
	(-3.535)	(-2.789)	(0.561)
pd_t	0.285	0.163	-0.035
	(2.782)	(2.139)	(-0.464)
N	372	360	348
R^2	0.39	0.16	0.01

a Radon-Nikodym derivative with respect to the physical measure $\mathcal{P}$ given by

$$\frac{d\hat{\mathcal{P}}}{d\mathcal{P}} = \exp\left\{-\frac{1}{2}\hat{\lambda}_t^2(\hat{\sigma}_\lambda^\top \Sigma \hat{\sigma}_\lambda)^2 - \hat{\lambda}_t \hat{\sigma}_\lambda^\top \epsilon_{t+1}\right\}, \tag{18}$$

where $\hat{\lambda}_t$, in its most general form, can be a linear function of both of the state variables, y_t and z_t:

$$\hat{\lambda}_t = \hat{\alpha}_0 + \hat{\alpha}_y y_t + \hat{\alpha}_z z_t\,. \tag{19}$$

Under the distorted belief, agents' expected dividend growth over the next year is given by

$$\begin{aligned}\hat{g}_t =& \bar{g} + z_t - \sigma_z^\top \Sigma \hat{\sigma}_\lambda\left(\hat{\alpha}_0 + \hat{\alpha}_y y_t + \hat{\alpha}_z z_t\right) - \frac{1}{2}\sigma_D^\top \Sigma \sigma_D \\ =& \bar{g} + \hat{\beta}_y y_z + \hat{\beta}_z z_t - \frac{1}{2}\sigma_D^\top \Sigma \sigma_D,\end{aligned} \tag{20}$$

where, to simplify notations, we introduce the linear coefficients, $\hat{\beta}_z$ and $\hat{\beta}_y$. The law of motion (17) under rational expectation has z_t's coefficient equal to one. Belief distortion may cause $\hat{\beta}_z \neq 1$. Moreover, y_t does not enter g_t in (17), but may enter the distorted expectation, $\hat{g}_t$ if $\hat{\beta}_y \neq 0$.[26]

No matter whether the agents have rational expectation or distorted beliefs, should ρ_z—the autoregressive coefficient of g_t—be not zero, the information cliff cannot exist. To see this, consider the cash-flow growth from $t+1$ to $t+2$, which is the sum of g_{t+1} and a shock to the realized growth

[26] Since λ_t linearly drives the expected return (as will be shown in Section 4.1), belief distortion implies the expected return being correlated with agents' growth expectation errors, in line with prior findings (see our literature review).

at $t+2$ with a zero mean, so the time-t expectation (or best forecast) of growth from $t+1$ to $t+2$ is

$$\begin{aligned}\mathbb{E}_t[\ln(D_{t+2}/D_{t+1})] =& \mathbb{E}_t[g_{t+1}] = \overline{g} + \mathbb{E}_t[z_{t+1}] - \frac{1}{2}\sigma_D^\top \Sigma \sigma_D \\ =& \overline{g} + \rho_z z_t - \frac{1}{2}\sigma_D^\top \Sigma \sigma_D . \end{aligned} \tag{21}$$

Here, as in our empirical exercise in Section 2, we are evaluating whether agents' information set has signal about future growth, so the expectation operator, $\mathbb{E}_t[\cdot]$, reflects the rational expectation or econometricians' belief, and accordingly, $\mathbb{E}_t[z_{t+1}] = \rho_z z_t$ under the physical-measure law of motion (17). If agents have rational expectation, their growth expectation is g_t, which is linear in z_t; otherwise, under belief distortions, agents' growth expectation can be a linear function of z_t and y_t. In either case, agents' time-t information set contains z_t. If $\rho_z \neq 0$, agents have signal about growth from $t+1$ to $t+2$, which implies that the information cliff does not exist. Therefore, a necessary condition for the information cliff is $\rho_z = 0$. The next proposition summarizes this result.

Proposition 1 (Information cliff: a parametric presentation) *If agents do not have information about cash-flow growth beyond the next year, we have $\rho_z = 0$.*

Discussion: misperception of the persistence. Several recent studies highlight the importance of agents' perceived persistence of state variables (e.g., Gabaix, 2019; Wang, 2020). Our theoretical framework can accommodate subjective persistence of expected cash-flow growth rate. As shown in (20), even though z_t is not persistent under $\rho_z = 0$, y_t may still be persistent. In this case, agents' taking (20) as the model for the expected growth rate would commit a similar mistake as in Guo and Wachter (2025b): they have the wrong mental model of a persistent expected growth rate. As emphasized in our literature review, our paper focuses on the implication of information cliff on z_t (the actual signal about future cash flows), i.e., $\rho_z = 0$, rather than how bias in agents' belief behaves and propagates over the forecasting horizon. That said, our model has a flexible setup that accommodates potential belief distortions like the one in Guo and Wachter (2025b).

3.4 Testing the parametric representation

We examine whether the expected cash-flow growth lacks persistence, i.e., $\rho_z = 0$. As the coverage of dividend forecasts started in 2003, we follow the literature and consider analysts' earnings forecasts as proxy for cash-flow expectations (available in a longer sample starting 1976).[27] An 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 earnings growth rates are close to those of dividends.

Analyst forecasts may not perfectly reflect the rational expectations of growth. Therefore, we add a noise term between the analyst expectations and expectations in our model:

$$\mathbb{E}_t^A \left[\Delta e_{t+k}\right] = \mathbb{E}_t \left[\Delta e_{t+k}\right] + \varepsilon_{t,k}^A, \tag{22}$$

where we consider $k = 1, 2, 3$, and $\mathbb{E}_t\left(\cdot\right)$ represents the rational expectation as in the model. From equation (16) in the model, we obtain

$$\begin{aligned}
\mathbb{E}_t^A \left[\Delta e_{t+1}\right] &= c + z_t + \varepsilon_{t,1}^A \\
\mathbb{E}_t^A \left[\Delta e_{t+2}\right] &= c + \mathbb{E}_t\left[z_{t+1}\right] + \varepsilon_{t,2}^A = c + \rho_z z_t + \varepsilon_{t,2}^A \\
\mathbb{E}_t^A \left[\Delta e_{t+3}\right] &= c + \mathbb{E}_t\left[z_{t+2}\right] + \varepsilon_{t,3}^A = c + \rho_z^2 z_t + \varepsilon_{t,3}^A,
\end{aligned}$$

where c is a constant and $\varepsilon_{t,k}^A$ has a zero mean, $k = 1, 2, 3$. Using the first equation to substitute out z_t in the second and third equations, we obtain a system:

$$\underbrace{\begin{bmatrix} \mathbb{E}_t^A \left[\Delta e_{t+2}\right] \\ \mathbb{E}_t^A \left[\Delta e_{t+3}\right] \end{bmatrix}}_{\equiv \mathbf{y}_t^A} = \left(1 - \rho_z\right) g + \rho_z \underbrace{\begin{bmatrix} \mathbb{E}_t^A \left[\Delta e_{t+1}\right] \\ \mathbb{E}_t^A \left[\Delta e_{t+2}\right] \end{bmatrix}}_{\equiv \mathbf{x}_t^A} + \underbrace{\begin{bmatrix} \varepsilon_{t,1}^A - \rho_z \varepsilon_{t,0}^A \\ \varepsilon_{t,2}^A - \rho_z \varepsilon_{t,1}^A \end{bmatrix}}_{\equiv \epsilon_t^A}. \tag{23}$$

We estimate ρ_z by regressing $\mathbf{y}_t^A$ on $\mathbf{x}_t^A$. The system of equations becomes a linear regression system of earnings growth estimates, $\mathbb{E}_t^A\left[\Delta e_{t+1}\right]$, $\mathbb{E}_t^A\left[\Delta e_{t+2}\right]$, and $\mathbb{E}_t^A\left[\Delta e_{t+3}\right]$. Note that ϵ_t^A is allowed to

[27] Analyst 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; Hillenbrand and McCarthy, 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.

Table 4 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 rate, based on equation (23). 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 (23), while columns (3) and (4) only use earnings growth forecasts for 1 and 2 years ahead ("Y1:Y2") to estimate the first equation in (23). 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	0.122	0.141	0.133
	(13.995)	(16.906)	(15.536)	(16.745)
ρ_z	0.028	0.015	-0.071	-0.073
	(0.690)	(0.381)	(-1.379)	(-1.295)
N	768	1887	384	943
R^2	0.003	0.001	0.025	0.028
Sample	Monthly	Weekly	Monthly	Weekly
Periods	Y1:Y3	Y1:Y3	Y1:Y2	Y1:Y2

be serially correlated across the starting dates of growth periods, i.e., t, $t+1$, and $t+2$.

The results are reported in Panel A of Table 4. We estimate equation (23) 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 (23), 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 (17), 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 5, our estimate is statistically indistinguishable from zero, which implies $\rho_z = 0$, consistent with our findings in Table 4.

For robustness, we consider a method to estimate ρ_z without using analyst forecasts (Appendix II). We fit the latent state model given by (16) and (17) to dividend data to filter out the expected

Table 5 Estimating the Persistence of Expected cash-flow growth (LTG Forecasts)

This table reports estimates of ρ_z^{LT} in the regression, $\log(1 + LTG_t) = \text{const} + \rho_z^{LT}\mathbb{E}_t^A[\Delta e_{t+1}] + \varepsilon_t$, 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. t-statistics based on Newey-West standard errors with autocorrelation of up to 18 lags are reported in parentheses.

	(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

cash-flow growth rate. The results corroborate our findings on ρ_z being close to zero.

4 The Information Cliff and Expected Returns

The previous section presents a critical implication of the information cliff on the expected cash-flow growth rate. Next, we draw the connection between the information cliff and the expected return. We not only provide additional asset-pricing results but also corroborate our findings on the information cliff. So far, we have provided direct evidence in Section 2 and tested its parametric implication in Section 3. While these tests directly target the information cliff, they are not standard. In this section, we adopt a well-established set of asset-pricing tools in our empirical tests.

4.1 Return predictability: the other side of information cliff

The theoretical framework. We utilize the two-dimensional state-space model in Section 3.3 to drive a necessary and sufficient condition of $\rho_z = 0$ and formalize the tests. Note that our focus in this section is the expected return (i.e., the expectation of future returns under the econometricians' belief or rational expectation), not agents' subjective expectations of future returns.

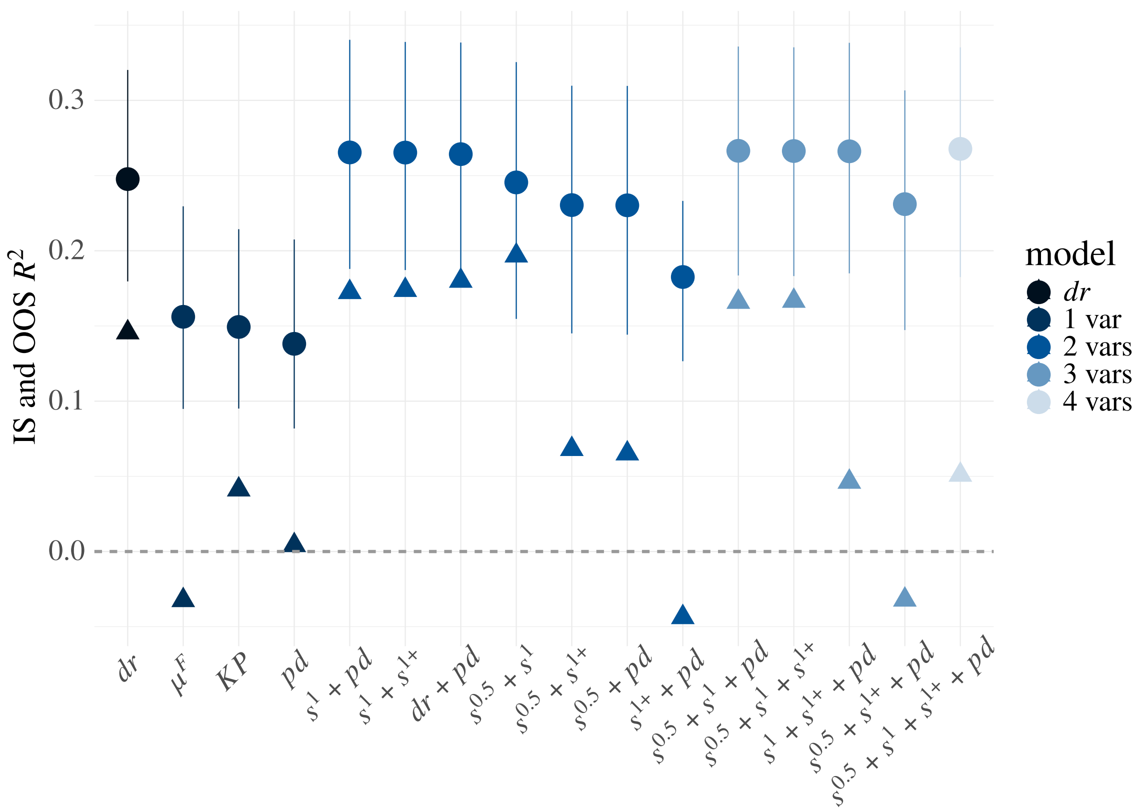


Figure 6 In-Sample and Out-of-Sample R^2 from Return Predictive Regressions: Including dr

This figure reports in- and out-of-sample R^2 for predicting annual S&P 500 Index returns. The predictors include slope of valuation term structure (dr), the predictor from Binsbergen and Koijen (2010) (μ^F), the 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.

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

$$pd_t = A_{pd} + B_{pd}y_t + C_{pd}z_t, \tag{24}$$

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

$$s_t^1 = A_1 + B_1 y_t + C_1 z_t. \tag{25}$$

We define the *slope* of valuation term structure as the difference between the price-dividend ratio—valuation level of the whole market—and the valuation ratio of one-year dividend strip:

$$dr_t = pd_t - s_t^1 = A_{pd} - A_1 + (B_{pd} - B_1)y_t + (C_{pd} - C_1)z_t. \tag{26}$$

An increase in dr_t reflects a greater fraction market value coming from cash flows beyond the very next year, a steepening of the valuation term structure. 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. Hence, the notation, dr, represents "discount rate".

Proposition 2 (Return predictability under the information cliff) *The expected return at time t is a linear function of y_t: $\mathbb{E}_t[r_{t+1}] = A_{er} + B_{er} y_t$, where A_{er} and B_{er} are constant. A necessary and sufficient condition for $\rho_z = 0$ is that dr_t's loading on z_t is zero, i.e., $C_{pd} - C_1 = 0$, or equivalently, that dr_t is a function of only y_t, so $\mathbb{E}_t[r_{t+1}]$ is a univariate linear function of dr_t, and vice versa.*

Therefore, to test $\rho_z = 0$, we can test the necessary and sufficient condition—that is, dr_t drives the expected return. This link between dr_t and the expected return has an intuitive explanation. If z_t lacks persistence ($\rho_z = 0$), market participants are not informed about growth beyond $t + 1$.[28] Therefore, when the valuation term structure steepens (i.e., dr_t increases), a greater fraction of market value is from $t + 1$ onward not due to objectively improved long-run growth but due to a lower discount rate (a lower y_t) that benefits the valuation of long-term cash flows more than that of near-term cash flows; 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 that brings down the value of dividends beyond one year more than valuation of dividends within one year.

Note that we do not study the term structure of equity risk premium (the difference in average returns between short- and long-horizon strips), which has attracted arguably most attention among studies on dividend strips.[29] We only use dividend strip prices for information on state variables, and our focus is on the expected market return. 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.

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

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

Predictive regression. Next, we provide a thorough analysis of the return predictive power of dr_t. We start with standard predictive regression for annual returns of S&P 500 index:

$$r_{t+1} = \alpha + \beta dr_t + \epsilon_{t+1}, \tag{27}$$

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.[30] In the Internet Appendix (Table A.7), we also report the IVX-Wald test (Kostakis, Magdalinos, and Stamatogiannis, 2014) that explicitly accounts for predictor persistence.

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

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

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

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

Table 6 Return Prediction

This table reports the return prediction results. The dependent variable of the predictive regression is the log annual return of the S&P 500 index, r_{t+1}. We include the following predictors: the slope of valuation term structure dr_t, price-dividend ratio pd_t, the return predictor from Binsbergen and Koijen (2010) μ^F, and the predictor from Kelly and Pruitt (2013) KP. We report t-statistics for each coefficient based Hodrick (1992) standard error (in squared brackets) and Newey and West (1987) standard error with 18 lags (in parentheses). Starting from January 1998, we construct out-of-sample forecasts using rolling regressions estimated with data up to the forecast date. We report in- and out-of-sample R^2s and p-values from the ENC (Clark and McCracken, 2001) and CW test (Clark and West, 2007).

	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)
pd_t		-0.199			0.141
		[-2.367]			[1.721]
		(-2.747)			(1.209)
μ_t^F			2.584		
			[2.313]		
			(2.804)		
KP_t				0.895	
				[2.960]	
				(2.857)	
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
$p(ENC)$	<0.05	>0.10	>0.10	<0.05	<0.01
$p(CW)$	0.022	0.200	0.303	0.031	0.021

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 6 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%.[31] Out-of-sample forecasts deliver an R^2 of 14.6%, significantly outperforming the historical mean as shown by the p-values of ENC and CW. The predictive coefficient is also large in magnitude, indicating high volatility of the conditional expected return.

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

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 (1) of Table 6 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).[32]

Column (2) of Table 6 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. The IVX-Wald test of Kostakis, Magdalinos, and Stamatogiannis (2014) in Table A.7 in the Internet 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 and pd. In Column (3) of Table 6, we follow the procedure in Binsbergen and Koijen (2010) to construct their return predictor, μ^F. While μ_t^F outperforms pd, its predictive power is significantly 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 shown in Internet Appendix I.2, they contain firm-level noise that is orthogonal to the expected market return. Kelly and Pruitt (2013) use partial least squares to reduce noise. Our approach differs as dr does not contain firm-level noise under the information cliff. Following the procedure in Kelly and Pruitt (2013), we construct their return predictor, denoted by KP. In column (4) of Table 6, we report the results. KP significantly outperforms pd but underperforms

[32] In the Internet Appendix, we show how to calculate the Sharpe ratio based on the out-of-sample R^2.

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

This table reports summary statistics (mean, median, and percentiles) of rolling-window estimates of ρ_z, the autoregressive coefficient of expected cash-flow growth rate. Each estimate is obtained from a three-year window of weekly observations using equation (23) applied to IGA aggregate earnings growth forecasts for the S&P 500 Index.

	count	mean	10%	15%	20%	25%	50%	75%	80%	90%
$\hat{\rho}_{z,t}$	384	0.025	-0.130	-0.105	-0.082	-0.066	0.001	0.074	0.094	0.196

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 that is in Table A.7 in the Internet Appendix.

In the Internet Appendix, we demonstrate the robustness of our results by repeating the analysis for alternative forecasting targets, even though our theoretical framework requires the forecasting target to be S&P 500 return as dr_t is about the valuation term structure of S&P 500. In Table A.8, we replace the S&P 500 annual return with the excess annual return. In Table A.9 and A.10, we consider the Fama-French market portfolio return and excess return, respectively.[33]

4.2 Rolling-window estimation

Our full-sample estimation in Section 3.4 has shown that ρ_z is close to zero, and our full-sample estimation of return predictive regressions has also demonstrated that dr has very strong predictive power, lending further support to $\rho_z = 0$ as the return predictive power of dr corresponds to $\rho_z = 0$ (see Proposition 2 of Section 4.1). In the following, we empirically examine validity of this necessary and sufficient condition through rolling-window estimation. In each rolling window, we estimate ρ_z as in Section 3.4 (Table 4) and predict returns using dr. We show that in around 20% of the rolling windows where ρ_z deviates from zero, the predictive power of dr is indeed weaker.

Each rolling window contains three years of weekly observations of analyst forecasts.[34] We summarize the statistics of the rolling-window estimates of ρ_z in Table 7. Naturally, how agents'

[33]As we have made clear, our goal is to predict annual returns (one period-ahead in our model). However, we also show that dr demonstrates superior return predictive power at a monthly horizon. Our baseline results are reported in Table A.11, and see Table A.12 for results on predicting monthly S&P 500 excess return. Table A.13 and Table A.14 report the results on predicting the monthly Fama-French market portfolio return and excess return, respectively.

[34]The results are similar if we use alternative window lengths from one to five years (available upon request). Our sample period is 1988–2019. The first estimate of ρ_z uses three years of IGA data starting in 1985.

belief formation model may vary over time, so the estimate, $\hat{\rho}_z$, fluctuates. However, its mean and median across rolling windows are close to zero, in line with the full-sample estimate in Table 4. In the next proposition based on the model in Section 3.3, 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 Appendix I.4.

Proposition 3 (ρ_z and return forecast errors) *Let ν_{t+1} denote the forecast error from predicting r_{t+1} with dr_t, and let ρ_z denote the autoregressive coefficient of expected cash-flow growth z_t in equation (17). If $\rho_z > 0$, then ν_{t+1} is positive. If $\rho_z < 0$, then ν_{t+1} is negative.*

The theoretical result in Proposition 3 shows the tight connection between dr's return predictive power and $\rho_z = 0$, a parametric manifestation of the information cliff, from a new perspective. When ρ_z deviates from zero, the return forecasting error of dr exhibits systematic patterns. This result receives direct empirical support. in Figure 7, we plot the average in-sample and out-of-sample forecasting errors for four categories of rolling windows: 1) the estimate of ρ_z is significantly negative; 2) the estimate of ρ_z is negative but statistically insignificant; 3) the estimate of ρ_z is positive but statistically insignificant; 4) the estimate of ρ_z is significantly positive. As predicted by Proposition 3, the sign of return forecasting errors aligns well with that of the estimate of ρ_z.

In Table 8, we report results of regressing dr's in- and out-of-sample return forecasting errors on the estimate of ρ_z or the sign of the estimate of ρ_z. The number of observations for this regression is the number of rolling windows. As predicted by Proposition 3, we obtain a positive regression coefficient, in line with the message in Figure 7.

In summary, Proposition 2 states a necessary and sufficient condition of $\rho_z = 0$, connecting it to the return predictive power of dr, the slope of valuation term structure. Proposition 3 characterizes dr's return forecasting error varies with ρ_z's deviation from zero, strengthening the link between dr as a return predictor and the information cliff.

Therefore, the other side of cash-flow information cliff is return predictability. This result echoes the findings that the absence of cash flow predictability indicates return predictability (e.g., Cochrane, 2008) but differs in meaningful ways. First, as we have shown in Section 2, cash-flow growth within one year horizon is in fact highly predictable. Second, the return predictive

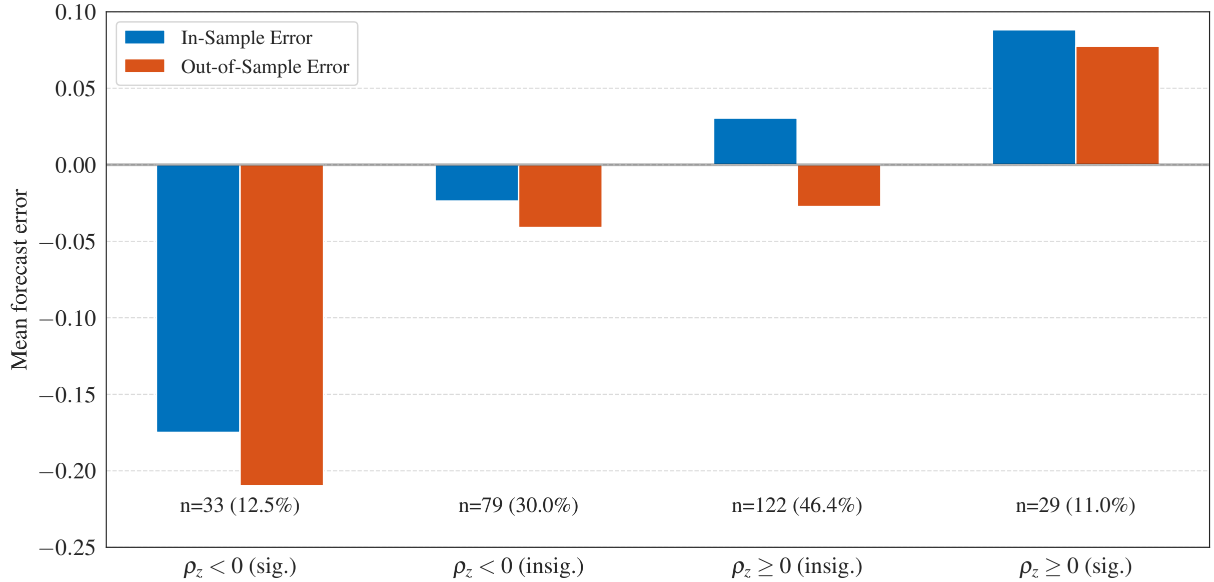


Figure 7 Average Return Prediction Errors Across Subsamples by ρ_z Estimate

This figure plots mean return forecast errors computed in-sample and out-of-sample using dr_t as the predictor. The sample is divided into four categories based on three-year rolling-window estimates of ρ_z: significantly negative, insignificantly negative, insignificantly positive, and significantly positive (significance determined at the 5% level) The number of observations and their percentage of the total sample are indicated below each category.

Table 8 Time-Varying ρ_z and Return Predictability

This table reports the regression results examining the relationship between return prediction errors and rolling-window estimate of the autoregressive coefficient of expected cash-flow growth rate. The dependent variables are the in-sample residuals (ε_{t+1}) and out-of-sample forecast errors (υ_{t+1}) from return predictive regressions based on dr_t. Independent variables include the expected cash-flow growth persistence parameter $\hat{\rho}_{z,t}$ (estimated using three-year rolling windows of analyst forecasts) and an indicator variable $\mathbb{I}_{\hat{\rho}_{z,t}>0}$ that equals one for $\hat{\rho}_{z,t} > 0$. Newey-West t-statistics (18 lags) are reported in parentheses. The sample period is 1998:01–2019:12, beginning with the first out-of-sample forecast.

	ε_{t+1}		υ_{t+1}	
	(1)	(2)	(3)	(4)
Intercept	-0.011	-0.057	-0.046	-0.078
	(-1.127)	(-3.217)	(-4.076)	(-3.982)
$\hat{\rho}_{z,t}$	0.556		0.469	
	(5.143)		(4.599)	
$\mathbb{1}\left(\hat{\rho}_{z,t} > 0\right)$		0.098		0.071
		(5.039)		(3.141)
N	252	252	252	252
R^2	0.173	0.110	0.094	0.043

power does not come from the traditional price-dividend ratio but from the slope of valuation term structure, dr. In the last part of our paper, we compare dr against the other return predictors from the literature and nonlinear (machine learning) methods. Under our result from Proposition 2, the strong return predictive power of dr lends support to the existence of information cliff.

4.3 Benchmarking the return prediction results

Before we embark on a battery of tests of dr's return predictive power against that of other predictors, we want to highlight the intuitive nature of using the slope of valuation term structure to predict returns under the cash flow information cliff. 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 (under the information cliff), 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. Likewise, when dr decreases, the flattening of valuation term structure is driven by a higher discount rate. In summary, under the cash flow information cliff, the slope reflects the discount rate.

Comparing the slope and other predictors. We have compared the return predictive power of dr with that of pd, μ^F, and KP. Figure 8 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), and 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).[35] We also include pd, μ^F, and KP. In Figure 8, we report in- and out-of-sample R^2 of dr minus those of other predictors. All columns

[35] Note 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).

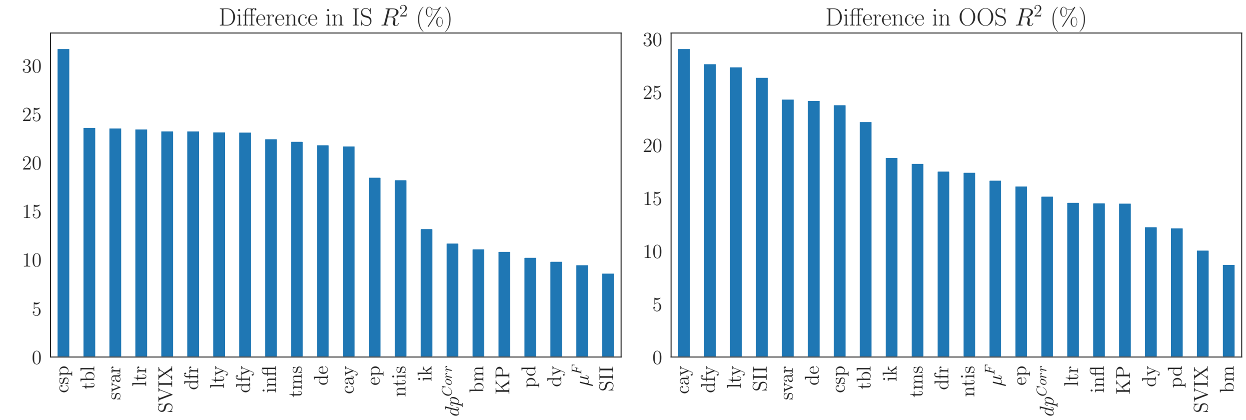


Figure 8 In-Sample and Out-of-Sample R^2 Wedge between dr and Other Return Predictors

This figure compares annual return prediction R^2 between dr_t and other predictors from prior studies. 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), 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).

are in the positive region, indicating dr performs better.

Finally, in the Internet Appendix, we repeat the exercise in Figure 8 for alternative forecasting targets, such as S&P 500 excess annual return (Figure A.4), Fama-French market portfolio annual return (Figure A.5), and Fama-French market portfolio excess annual return (Figure A.6).

The role of nonlinearity. Following Lettau and Wachter (2007), our theoretical framework is an exponential-affine model. 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, our forecasting exercises may have mistakenly ignored important 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

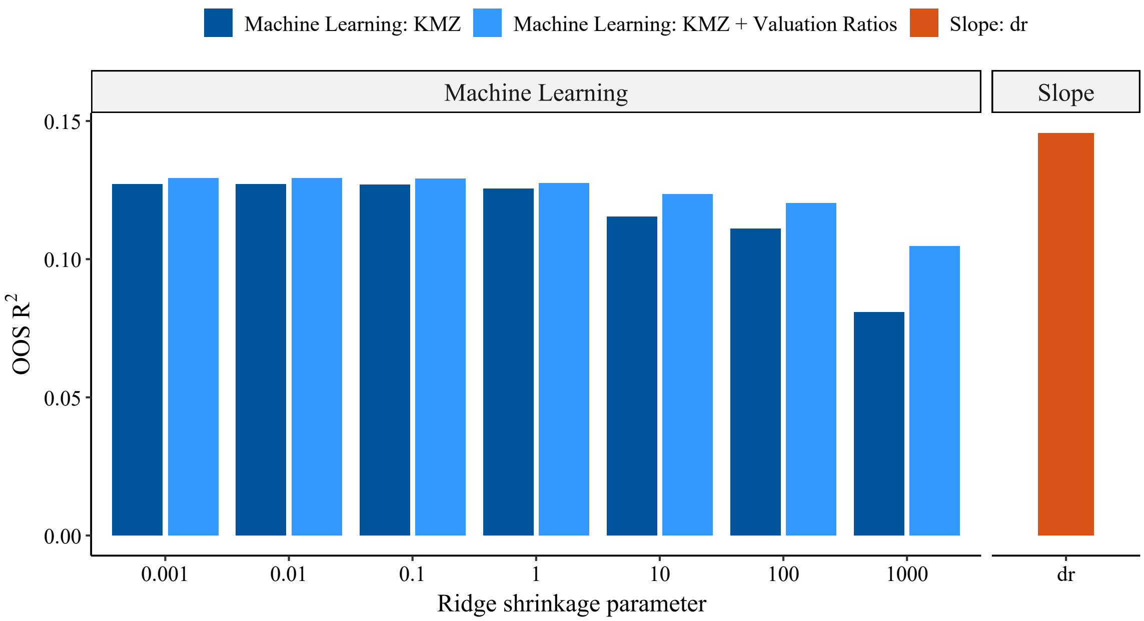


Figure 9 Return Prediction: The Role of Nonlinearity

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).

progressively to incorporate nonlinear terms ("model complexity"). 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.7 in the Internet Appendix). In Figure 9, 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. We consider two cases, one with a signal base including all of our strip valuation ratios (state variable proxies) and other predictors and the second signal set 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.[36] Overall, our analysis has two implications. First,

[36]Note that due to estimation errors, the machine learning model may underperform our simple OLS with dr as the

in terms of raw signals, *dr* contains sufficient information, so one may not seek "big data" (i.e., alternative signals) for forecasting market returns. Second, combining all signals (including our state variable proxies) nonlinearly does not improve forecasting performance, validating our exponential-affine framework that implies a linear relationship between state variables and the expected return.

Discussion: Spanning tests. The bond literature highlights the critical issue of unspanned state variables (i.e., state variables that are not spanned by bond yields, the equivalent of dividend strip valuation ratios in the bond markets).[37] We find that this is not the case in our analysis of expected stock-market return. We have performed the following spanning tests in the Internet Appendix. In Table A.15, 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 8, 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.15 but also an out-of-sample R^2 close to zero. Table A.16 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.

5 Conclusion

This paper characterizes an "information cliff"—a structural feature of the stock market where the supply of information about aggregate cash flows drops precipitously beyond a one-year horizon. In contrast to a large literature that emphasizes biases in how agents process information, we

predictor even when *dr* and other valuation ratios are included as signals.

[37] Since Duffie and Kan (1996), using bond yields to map out the state variables has been key to the estimation of term structure models (Duffee, 2013). Unspanned state variables include stochastic volatility and macro factors (e.g., Collin-Dufresne and Goldstein, 2002; Cooper and Priestley, 2008; Bikbov and Chernov, 2009; Ludvigson and Ng, 2009; Andersen and Benzoni, 2010; Duffee, 2011; Joslin et al., 2014; Cieslak and Povala, 2015).

begin with the source of the information itself. We provide strong evidence for this cliff using the horizons of corporate guidance. By exploring the timing of information supply (Guo, 2025; Guo and Wachter, 2025a), we demonstrate the connection between information supply and the sharp decline of analysts' forecast accuracy at the one-year horizon, contributing to the growing literature on market participants' short- and long-term cash-flow expectations.

To understand the asset pricing implications of this information structure, we develop a flexible framework that extends the model of Lettau and Wachter (2007). A key methodological innovation of our paper is to empirically discipline the model's structure. By analyzing the valuation ratios of aggregate dividend strips that map out the (latent) state variables, we show that two state variables, which can be rotated to represent expected returns and expected dividend growth, are sufficient to capture the dynamics of the aggregate stock market.

This empirically-grounded two-dimensional structure allows us to draw a sharp link between the information cliff and canonical objects in asset pricing. The cliff implies that expected cash-flow growth must lack persistence; any persistent component would provide a basis for forecasting growth beyond one year, contradicting the cliff. We test and confirm this prediction: the autoregressive coefficient of expected aggregate cash-flow growth is indistinguishable from zero.

Our model also shows that under the information cliff, the expected market return is a univariate function of the slope of the valuation term structure. Intuitively, with no new information about long-term growth, a steepening of the slope—where the valuation of distant cash flows rises relative to near-term cash flows—must be driven by a lower discount rate. Empirically, we find this slope is a remarkably robust return predictor, outperforming a large set of established predictors in both in-sample and out-of-sample tests and subsuming their forecasting power. A linear regression with the slope as a predictor outperforms non-linear aggregator of a large set of predictors.

In summary, our paper documents the information cliff and its implications on the expected cash-flow growth and expected return. Our findings also shed light on potential market mispricing. The success of the valuation slope as a return predictor suggests that while market participants are well-informed about the near term, mispricing is concentrated in the valuation of long-term cash flows. This shifts the focus of market timing from betting against the overall valuation level (e.g., the price-dividend ratio) to betting against the valuation term structure. The information cliff provides a rational, information supply-side foundation for why long-horizon expectations can appear untethered from fundamentals (Bordalo et al., 2024b) and why the relative pricing of long- and short-term cash flows holds the key to excess volatility and time-varying expected returns.

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).

Afrouzi, Hassan, Spencer Y Kwon, Augustin Landier, Yueran Ma, and David Thesmar, 2023, Overreaction in expectations: Evidence and theory, *Quarterly Journal of Economics* 138, 1713–1764.

Ai, Hengjie, and Ravi Bansal, 2018, Risk preferences and the macroeconomic announcement premium, *Econometrica* 86, 1383–1430.

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.

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.

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.

Barberis, Nicholas, 2018, Psychology-based models of asset prices and trading volume, in *Handbook of behavioral economics: applications and foundations 1*, volume 1, 79–175 (Elsevier).

Barberis, Nicholas, Andrei Shleifer, and Robert W. Vishny, 1998, A model of investor sentiment, *Journal of Financial Economics* 49, 307–343.

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.

Ben-Rephael, Azi, Bruce I. Carlin, Zhi Da, and Ryan D Israelsen, 2021, Information consumption and asset pricing, *Journal of Finance* 76, 357–394.

Bhojraj, Sanjeev, and Robert Libby, 2005, Capital market pressure, disclosure frequency-induced earnings/cash flow conflict, and managerial myopia (retracted), *The Accounting Review* 80, 1–20.

Bikbov, Ruslan, and Mikhail Chernov, 2009, Unspanned stochastic volatility in affine models: Evidence from eurodollar futures and options, *Management Science* 55, 1292–1305.

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, Matthew OBrien, and Andrei Shleifer, 2024a, Long-term expectations and aggregate fluctuations, *NBER Macroeconomics Annual* 38, 311–347.

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, Rafael La Porta, and Andrei Shleifer, 2024b, Belief overreaction and stock market puzzles, *Journal of Political Economy* 132, 1450–1484.

Bordalo, Pedro, Nicola Gennaioli, Yueran Ma, and Andrei Shleifer, 2020, Overreaction in macroeconomic expectations, *American Economic Review* 110, 2748–82.

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.

Brochet, Francois, Maria Loumioti, and George Serafeim, 2015, Speaking of the short-term: Disclosure horizon and managerial myopia, *Review of Accounting Studies* 20, 1122–1163.

Bushee, Brian J, 2001, Do institutional investors prefer near-term earnings over long-run value?, *Contemporary Accounting Research* 18, 207–246.

Call, Andrew C, Shuping Chen, Bin Miao, and Yen H Tong, 2014, Short-term earnings guidance and accrual-based earnings management, *Review of accounting studies* 19, 955–987.

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 Samuel B. Thompson, 2008, Predicting Excess Stock Returns Out of Sample: Can Anything Beat the Historical Average?, *Review of Financial Studies* 21, 1509–1531.

Cassella, Stefano, Benjamin Golez, Huseyin Gulen, and Peter Kelly, 2023, Horizon bias and the term structure of equity returns, *Review of Financial Studies* 36, 1253–1288.

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.

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 Pavol Povala, 2015, Expected returns in treasury bonds, *Review of Financial Studies* 28, 2859–2901.

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., 2008, The Dog That Did Not Bark: A Defense of Return Predictability, *Review of Financial Studies* 21, 1533–1575.

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, Michael Johannes, and Lars A. Lochstoer, 2016, Parameter Learning in General Equilibrium: The Asset Pricing Implications, *American Economic Review* 106, 664–98.

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 sudy 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.

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.

De Silva, Tim, and David Thesmar, 2024, Noise in expectations: Evidence from analyst forecasts, *Review of Financial Studies* 37, 1494–1537.

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., 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.

Engelberg, Joseph, R David McLean, and Jeffrey Pontiff, 2018, Anomalies and news, *Journal of Finance* 73, 1971–2001.

Enke, Benjamin, and Thomas Graeber, 2023, Cognitive uncertainty, *The Quarterly Journal of Economics* 138, 2021–2067.

Fama, Eugene F., and Kenneth R. French, 1988, 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* .

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, 2024, Equity term structures without dividend strips data, *Journal of Finance* 79, 4143–4196.

Giglio, Stefano, Matteo Maggiori, Johannes Stroebel, and Stephen Utkus, 2021, Five facts about beliefs and portfolios, *American Economic Review* 111, 1481–1522.

Gilbert, Thomas, Christopher Hrdlicka, and Avraham Kamara, 2018, The structure of information release and the factor structure of returns, *Journal of Financial Economics* 127, 546–566.

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.

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, *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.

Guo, Hongye, 2025, Earnings extrapolation and predictable stock market returns, *Review of Financial Studies* 38, 1730–1782.

Guo, Hongye, and Jessica A Wachter, 2025a, Correlation neglect in asset prices, working paper, The Wharton School of the University of Pennsylvania.

Guo, Hongye, and Jessica A Wachter, 2025b, "Superstitious" investors, *Review of Asset Pricing Studies* 15, 1–45.

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.

Hillenbrand, Sebastian, and Odhrain McCarthy, 2021, Heterogeneous investors and stock market fluctuations, Working paper, Harvard 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.

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 Priebsch, 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 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.

Kostakis, Alexandros, Tassos Magdalinos, and Michalis P. Stamatogiannis, 2014, Robust econometric inference for stock return predictability, *Review of Financial Studies* 28, 1506–1553.

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.

Lettau, Martin, and Sydney Ludvigson, 2001, Consumption, Aggregate Wealth, and Expected Stock Returns, *Journal of Finance* 56, 815–849.

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.

Lucca, David O., and Emanuel Moench, 2015, The Pre-FOMC Announcement Drift, *Journal of Finance* 70, 329–371.

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.

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.

Penman, Stephen H., 1980, An empirical investigation of the voluntary disclosure of corporate earnings forecasts, *Journal of Accounting Research* 18, 132–160.

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.).

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.

Savor, Pavel, and Mungo Wilson, 2013, How much do investors care about macroeconomic risk? evidence from scheduled economic announcements, *Journal of Financial and Quantitative Analysis* 48, 343–375.

Savor, Pavel, and Mungo Wilson, 2014, Asset pricing: A tale of two days, *Journal of Financial Economics* 113, 171–201.

Savor, Pavel, and Mungo Wilson, 2016, Earnings announcements and systematic risk, *Journal of Finance* 71, 83–138.

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.

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.

Wachter, Jessica A, and Yicheng Zhu, 2022, A model of two days: Discrete news and asset prices, *Review of Financial Studies* 35, 2246–2307.

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
"The Information Cliff"

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

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

We conjecture that

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

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

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

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

$$B\left(n\right) = \left(\Pi^\top - \theta\Sigma\sigma_X\right)B\left(n-1\right) + \phi - \theta\Sigma\sigma_D. \tag{A.4}$$

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

$$\begin{aligned}
A\left(n\right) =& A\left(n-1\right) + \overline{g} - \overline{r} - \left(\sigma_D + \sigma_X B\left(n-1\right)\right)^\top \Sigma\overline{\lambda} + \\
&\frac{1}{2}\left(\sigma_D + \sigma_X B\left(n-1\right)\right)^\top \Sigma\left(\sigma_D + \sigma_X B\left(n-1\right)\right).
\end{aligned} \tag{A.5}$$

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

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

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

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

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

Under the no-arbitrage condition, we have

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

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

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

$$B = \left(\mathbf{I} - \kappa_1\Pi^\top + \kappa_1\theta\Sigma\sigma_X\right)^{-1}\left(\phi - \theta\Sigma\sigma_D - \gamma\right). \tag{A.10}$$

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

$$Z_{i,t+1} = \Omega Z_{i,t} + \sigma_{i,Z}^\top \upsilon_{i,t+1}, \tag{A.11}$$

where $\upsilon_{i,t+1}$ is a N_i-by-1 vector of i-specific news that has a normal distribution $N\left(\mathbf{0}, \Sigma_i\right)$ 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

$$\ln\left(\frac{D_{i,t+1}}{D_{i,t}}\right) = g_{i,t} + \sigma_{i,D}^\top \epsilon_{t+1} + \sigma_{i,\upsilon}^\top \upsilon_{i,t+1}, \tag{A.12}$$

where the expected dividend growth rate is given by

$$g_{i,t} = \phi_i^\top X_t + \delta_i^\top Z_{i,t} + \overline{g}_i - \frac{1}{2}\sigma_{i,D}^\top \Sigma\sigma_{i,D} - \frac{1}{2}\sigma_{i,\upsilon}^\top \Sigma_i \sigma_{i,\upsilon}, \tag{A.13}$$

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

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

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

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

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

$$\begin{aligned} A_i(n) = & A_i(n-1) + \overline{g}_i - \overline{r} - \left(\sigma_{i,D} + \sigma_X B_i(n-1)\right)^\top \Sigma\overline{\lambda} + \frac{1}{2}\left(\sigma_{i,D} + \sigma_X B_i(n-1)\right)^\top \Sigma \\ & \left(\sigma_{i,D} + \sigma_X B_i(n-1)\right) + \frac{1}{2}\left(\sigma_{i,v} + \sigma_{i,Z} C_i(n-1)\right)^\top \Sigma_i \left(\sigma_{i,v} + \sigma_{i,Z} C_i(n-1)\right). \end{aligned} \tag{A.17}$$

with the initial conditions

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

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

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

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

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

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 and Proposition 2

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

$$pd_t = \ln(S_t/D_t) = A + By_t + Cz_t.$$

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

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

to obtain

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

Using the conjecture of pd_t and pd_{t+1} and the specification of g_t and Δd_{t+1}, we obtain

$$\begin{aligned}\mathbb{E}_t\Bigg[\exp\Bigg\{ & -r_f - \frac{1}{2}\lambda_t^2(\sigma_\lambda^\top\Sigma\sigma_\lambda) - \lambda_t\sigma_\lambda^\top\epsilon_{t+1} + \kappa_0 - A - By_t - Cz_t + z_t + \overline{g} - \frac{1}{2}\sigma_D^\top\Sigma\sigma_D + \sigma_D^\top\epsilon_{t+1} \\ & +\kappa_1 A + \kappa_1 B(\rho_y y_t + \sigma_y^\top\epsilon_{t+1}) + \kappa_1 C(\rho_z z_t + \sigma_z^\top\epsilon_{t+1}) \Bigg\}\Bigg] = 1 \end{aligned} \tag{A.22}$$

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

$$C = \frac{1}{1-\kappa_1\rho_z} \tag{A.23}$$

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

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

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

$$\begin{aligned}\exp\Bigg\{&-r_f + \kappa_0 - A - By_t - Cz_t + z_t + \overline{g} - \frac{1}{2}\sigma_D^\top\Sigma\sigma_D - (\sigma_D + \kappa_1 B\sigma_y + \kappa_1 C\sigma_z)^\top\Sigma\sigma_\lambda(\overline{\lambda} + y_t) \\ &+\kappa_1 A + \kappa_1 B\rho_y y_t + \kappa_1 C\rho_z z_t + \frac{1}{2}(\sigma_D + \kappa_1 B\sigma_y + \kappa_1 C\sigma_z)^\top\Sigma(\sigma_D + \kappa_1 B\sigma_y + \kappa_1 C\sigma_z)\Bigg\} = 1\end{aligned} \tag{A.25}$$

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

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

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

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

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

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

or equivalently

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

so we obtain

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

We conjecture

$$s_t^1=A_1+B_1y_t+C_1z_t.$$

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

$$\begin{aligned}\mathbb{E}_t\Bigg[\exp\Bigg\{&-r_f-\frac{1}{2}(\overline{\lambda}+y_t)^2(\sigma_\lambda^\top\Sigma\sigma_\lambda)-(\overline{\lambda}+y_t)\sigma_\lambda^\top\epsilon_{t+1}\\&+z_t+\overline{g}-\frac{1}{2}\sigma_D^\top\Sigma\sigma_D+\sigma_D^\top\epsilon_{t+1}-A_1-B_1y_t-C_1z_t\Bigg\}\Bigg]=1.\end{aligned}$$

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

$$\exp\left\{-r_f+z_t+\overline{g}-A_1-B_1y_t-C_1z_t-\sigma_\lambda^\top\Sigma\sigma_D(\overline{\lambda}+y_t)\right\}=1. \tag{A.31}$$

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

$$C_1=1, \tag{A.32}$$

and

$$B_1=-\sigma_\lambda^\top\Sigma\sigma_D. \tag{A.33}$$

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

$$A_1=\overline{g}-r_f-\sigma_\lambda^\top\Sigma\sigma_D\overline{\lambda} \tag{A.34}$$

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

$$\mathbb{E}_t[r_{t+1}]=\kappa_0-(1-\kappa_1)A+\overline{g}-\frac{1}{2}\sigma_D^\top\Sigma\sigma_D-(1-\kappa_1\rho_y)By_t. \tag{A.35}$$

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

I.4 Proposition 3: 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

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

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

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

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

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

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

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

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

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

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

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 $\mathbb{E}\left(\hat{\rho}_{z,t}\right) \approx 0$. 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

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

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

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

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

The variance of the market-timing strategy is

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

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

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

where A is a constant that depends on $Var\left[\left(\mu + x_t\right)\left(\mu + x_t + \varepsilon_{t+1}\right)\right]$ 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

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

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: Estimating A State Space Model of Cash Flows

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

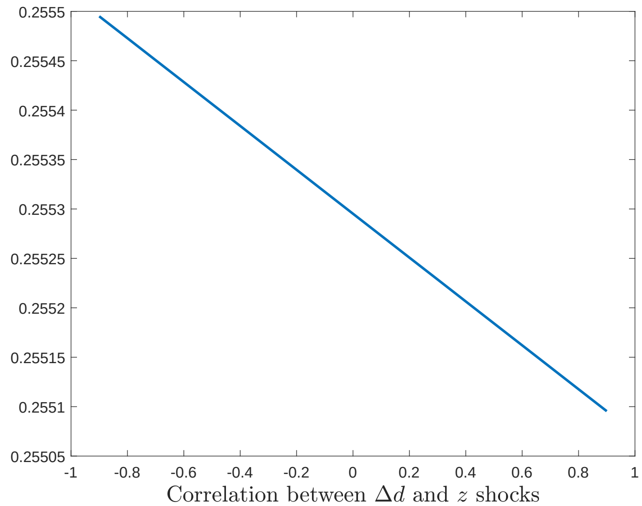
This table presents the estimation results for four models of dividend growth rates: (1) the unrestricted state-space model as specified in equations (16) and (17) in Section 3.3; (2) the restricted state-space model with the constraint $\rho_z = 0$; (3) the MA(1) model ($\Delta d_{t+1} = g + \sigma_D \varepsilon_{t+1} + \chi \sigma_D \varepsilon_t$); 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	0.06	0.00	0.11			74.44	-140.88	-128.97
	[0.94]	[3.01]	[0.00]	[1.70]					
Restricted		0.06	0.08	0.08			71.36	-136.72	-127.79
		[4.68]	[0.00]	[0.00]					
MA(1)		0.06	0.10		0.41		76.41	-146.82	-137.89
		[3.38]	[13.45]		[6.11]				
AR(1)		0.04	0.11			0.26	74.50	-142.99	-134.06
		[3.64]	[14.90]			[3.51]			
Panel B: MKT									
Unrestricted	-0.08	0.06	0.00	0.15			43.96	-79.92	-69.8
	[-0.06]	[3.86]	[0.00]	[0.12]					
Restricted		0.06	0.11	0.11			43.67	-81.34	-73.8
		[3.62]	[0.10]	[0.10]					
MA(1)		0.06	0.15		-0.09		44.00	-82.00	-74.4
		[3.94]	[6.99]		[-1.02]				
AR(1)		0.06	0.15			-0.08	43.96	-81.93	-74.39
		[3.89]	[6.98]			[-0.87]			

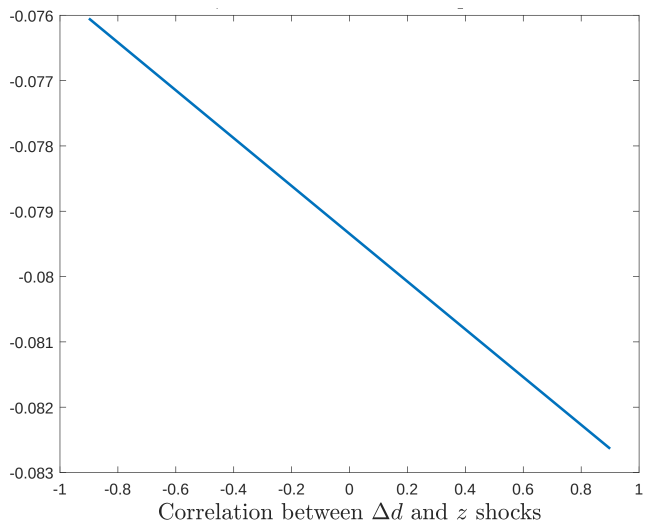
An alternative method to estimate ρ_z is to directly estimate the state-space model given by equations (16) and (17) with the realized dividend data. Using the standard Kalman filter, we obtain estimates of ρ_z. For comparison, we report results for both the S&P 500 index and the Fama-French market portfolio ("MKT"). We obtain dividend data for the Fama-French market portfolio (the CRSP NYSE/NYSEMKT/Nasdaq Value-Weighted Market Index). Since the model is set up at annual frequency, we use annual (non-overlapping) dividend growth data. The sample spans 1926 to 2019.[38] 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.[39] 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 estimate MA(1) and AR(1) models for comparison

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

[39] The Kalman filter assumes that the shocks to realized and expected dividend growth are uncorrelated. In Figure A.1, 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 of Table A.1. The estimated ρ_z barely moves with the value of shock correlations in $[-0.9, 0.9]$.



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 3.3, 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.

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 autoregressive coefficient of expected cash-flow growth rate is close to zero.

Appendix III: Additional Tables and Figures

Table A.2 Predicting Earnings Growth with Analyst Forecasts: By Month of the Quarter

This table reports the results of regressions that predict earnings growth at various horizons with analyst forecasts, with the regression samples separated by month of a quarter to highlight the effect of information supply. The dependent variables are realized earnings growth from IGA of next year, year after next, the third year into the future, and the average earnings growth between years 3 to 5. The independent variables are analysts' forecasts of one-year earnings growth between $t + \tau$ and $t + \tau + 1$ across horizons ($\mathbb{E}_t^A [\Delta e_{t+\tau+1}]$, for $\tau = 0, 1, 2$) from IGA and the self-aggregated long-term earnings growth forecasts (LTG_t) of the S&P 500 Index. The t-statistics are calculated based on Newey-West standard errors with 18 lags are reported in parentheses.

	Δe_{t+1}	Δe_{t+2}	Δe_{t+3}	$\overline{\Delta e_{t+2,t+5}}$
	(1)	(2)	(3)	(4)
Panel A: First month of the quarter				
Intercept	-0.048	-0.099	-0.134	0.019
	(-4.056)	(-2.491)	(-1.883)	(0.236)
$\mathbb{E}_t^A [\Delta e_{t+1}]$	1.184			
	(21.935)			
$\mathbb{E}_t^A [\Delta e_{t+2}]$		1.168		
		(4.078)		
$\mathbb{E}_t^A [\Delta e_{t+3}]$			1.503	
			(3.045)	
LTG_t				0.373
				(0.587)
N	128	128	128	122
R^2	0.78	0.18	0.15	0.01
Panel B: Second month of the quarter				
Intercept	-0.059	-0.128	-0.107	-0.001
	(-4.020)	(-2.430)	(-1.408)	(-0.015)
$\mathbb{E}_t^A [\Delta e_{t+1}]$	1.202			
	(17.752)			
$\mathbb{E}_t^A [\Delta e_{t+2}]$		1.346		
		(3.801)		
$\mathbb{E}_t^A [\Delta e_{t+3}]$			1.340	
			(2.588)	
LTG_t				0.541
				(0.855)
N	128	128	128	122
R^2	0.69	0.18	0.12	0.02
Panel C: Third month of the quarter				
Intercept	-0.056	-0.125	-0.119	0.011
	(-4.024)	(-2.485)	(-1.532)	(0.147)
$\mathbb{E}_t^A [\Delta e_{t+1}]$	1.209			
	(19.611)			
$\mathbb{E}_t^A [\Delta e_{t+2}]$		1.318		
		(3.852)		
$\mathbb{E}_t^A [\Delta e_{t+3}]$			1.428	
			(2.694)	
LTG_t				0.439
				(0.696)
N	128	128	128	122
R^2	0.72	0.18	0.13	0.01

Table A.3 Forecasting Earnings Growth Across Horizons: Controlling for Extrapolation

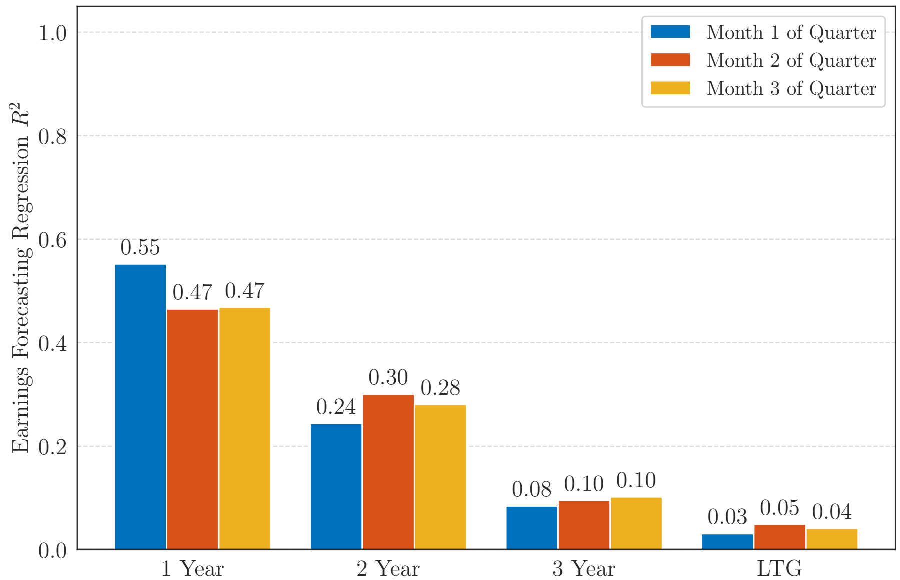
This table reports the results of regressions that predict earnings growth at various horizons. The dependent variables are realized earnings growth from IGA of next year, the year after, and the third year in the future, and the average earnings growth rate between years 3 to 5. The independent variables are analysts' forecasts of one-year earnings growth between $t + \tau$ and $t + \tau + 1$ across horizons ($\mathbb{E}_t^A [\Delta e_{t+\tau+1}]$, for $\tau = 0, 1, 2$) from IGA, the self-aggregated long-term earnings growth forecasts (LTG_t) of the S&P 500 Index, and the lagged realized earnings growth ($\Delta e_{t-1,t}$). t-statistics calculated based on Newey-West standard errors with 18 lags are reported in parentheses.

	Δe_{t+1}		Δe_{t+2}		Δe_{t+3}		$\overline{\Delta e_{t+2,t+5}}$	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Intercept	-0.056	-0.059	-0.097	-0.033	-0.149	-0.072	0.004	0.014
	(-4.127)	(-3.936)	(-2.063)	(-0.643)	(-2.130)	(-0.779)	(0.047)	(0.163)
$\mathbb{E}_t^A [\Delta e_{t+1}]$	1.204	1.189						
	(20.101)	(19.936)						
$\mathbb{E}_t^A [\Delta e_{t+2}]$			1.164	0.798				
			(3.683)	(2.483)				
$\mathbb{E}_t^A [\Delta e_{t+3}]$					1.598	1.119		
					(3.342)	(1.894)		
LTG_t							0.486	0.541
							(0.717)	(0.782)
Δe_t		0.075		-0.210		-0.199		-0.228
		(0.992)		(-1.671)		(-1.683)		(-3.389)
N	372	372	360	360	348	348	324	324
$R^2$2	0.73	0.74	0.15	0.19	0.21	0.23	0.01	0.20

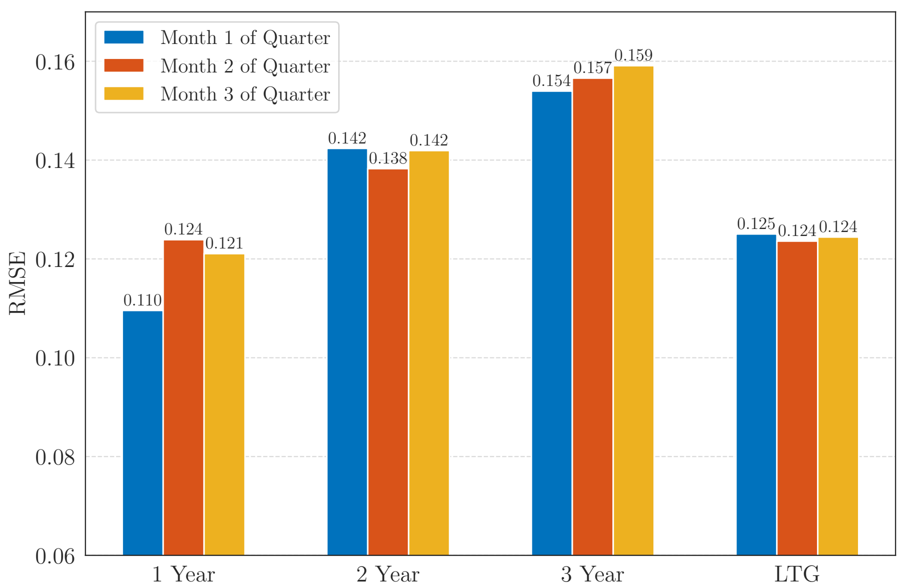
Table A.4 Forecast Horizons of Popular Macroeconomic Surveys

This table summarizes the forecast horizons (in quarters) of several major surveys that cover macroeconomic variables such as GDP growth, unemployment rate and industrial production.

Survey	Survey Start Year	Macro Variable	Survey Frequency	Maximal forecast horizon
Blue Chip Economic Indicators	1976	Real GDP growth, Unemployment rate, Industrial production	Monthly	~4-5 quarters
Survey of Professional Forecasters	1968	Real GDP growth, Unemployment rate	Quarterly	Quarterly: 4 quarters Annual: 3 years
Livingston Survey	1946	Real GDP growth, Unemployment rate	Semi-annually	~4-5 quarters
Consensus Economics	1989	Real GDP growth, Unemployment rate	Monthly	6 quarters
Wall Street Journal Economic Survey	2008	Real GDP growth (annual and quarterly), Unemployment rate	Monthly	Quarterly GDP: 6 quarters Annual: 3 years
Federal Reserve Summary of Economic Projections	2007	Real GDP growth, Unemployment rate	Quarterly	3 years



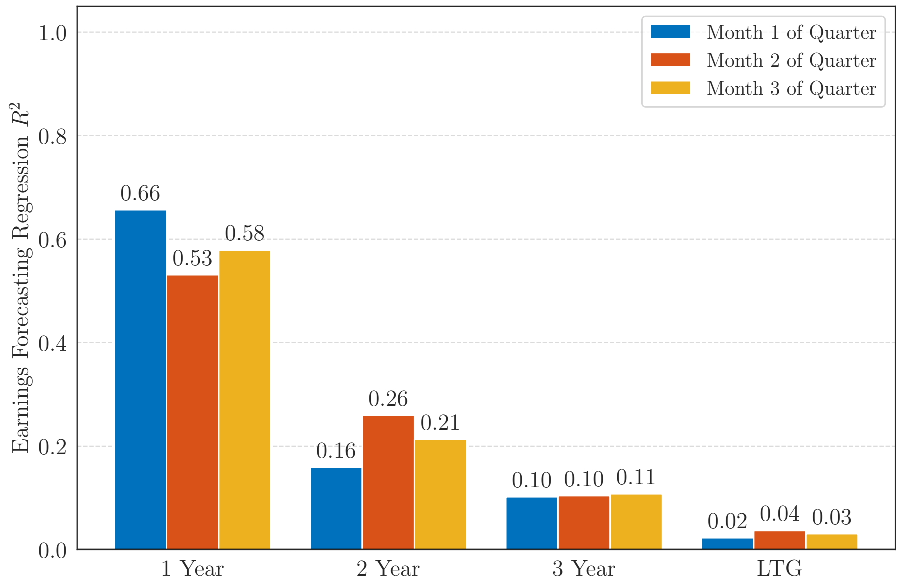
A. In-sample R^2 by month of the quarter



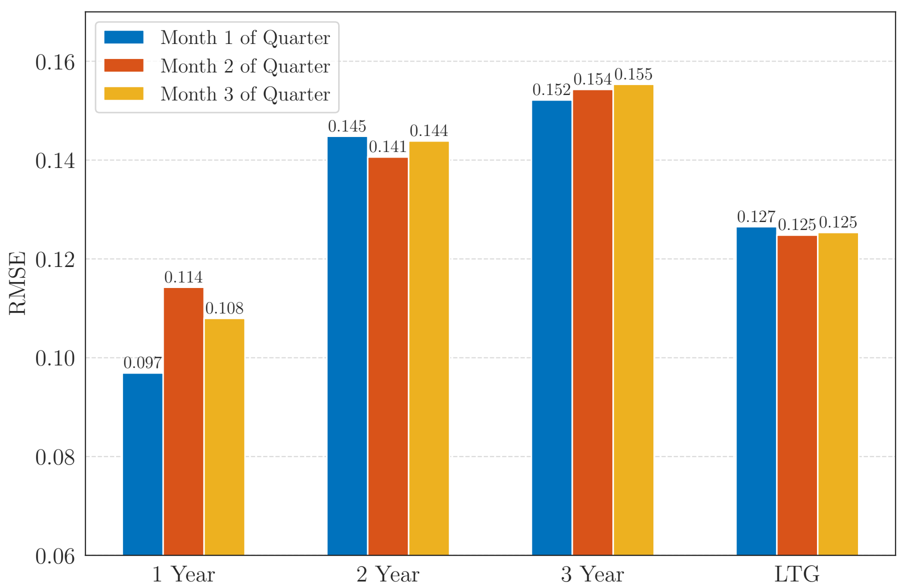
B. RMSE by month of the quarter

Figure A.2 Predicting Earnings Growth Across Horizons by Month of a Quarter: February–April

This figure reports, respectively in Pane A and B, the R^2 of predictive regression and the Root Mean Squared Error (RMSE) of predicting earnings at various horizons with analyst forecasts (as done in Table 2), with the regression samples separated by month of a quarter to highlight the effect of information supply. The RMSE for LTG (3-year average growth) is annualized by multiplying by $\sqrt{3}$ for comparability. We restrict the sample to February–April observations, a window during which the aggregate year-1 earnings forecast do not contain partially realized annual earnings.



A. In-sample R^2 by month of the quarter



B. RMSE by month of the quarter

Figure A.3 Predicting Earnings Growth Across Horizons by Month of a Quarter: February–July

This figure reports, respectively in Pane A and B, the R^2 of predictive regression and the Root Mean Squared Error (RMSE) of predicting earnings at various horizons with analyst forecasts (as done in Table 2), with the regression samples separated by month of a quarter to highlight the effect of information supply. The RMSE for LTG (3-year average growth) is annualized by multiplying by $\sqrt{3}$ for comparability. We restrict the sample to February–July observations, a window during which the aggregate year-1 earnings forecast do not contain partially realized annual earnings.

Table A.5 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. 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.

	Δd_{t+1}															
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
dr_t	-0.18***	-0.04**														
	(0.05)	(0.02)														
pd_t	0.29***		0.02						0.07	0.10*	7.93**		0.10*	4.24	2.26	2.15
	(0.11)		(0.05)						(0.06)	(0.06)	(3.65)		(0.06)	(3.34)	(2.86)	(2.90)
g_t^F				0.68***												
				(0.25)												
KP_t^{CF}					0.22											
					(0.14)											
$s_t^{0.5}$						0.03	0.14***		0.14***			0.02	0.02	0.10**		0.01
						(0.03)	(0.05)		(0.05)			(0.03)	(0.03)	(0.04)		(0.03)
s_t^1						0.10***		0.18***		0.18***		0.16***	0.16***		0.15***	0.15***
						(0.03)		(0.06)		(0.05)		(0.06)	(0.06)		(0.04)	(0.04)
s_t^{1+}							0.07	0.10*			-7.69**	0.10*		-4.07	-2.11	-2.00
							(0.05)	(0.06)			(3.53)	(0.05)		(3.23)	(2.76)	(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.6 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. Newey-West standard errors with autocorrelation of up to 18 lags are reported in parentheses. Data sample: 1988:01–2019:12. See Figure 5 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***						-0.13**	-0.09	7.54		-0.09	1.18	-1.41	-1.87
		(0.07)						(0.06)	(0.07)	(4.58)		(0.07)	(4.38)	(4.03)	(3.70)
μ_t^F			2.58***												
			(0.92)												
KP_t				0.90***											
				(0.31)											
$s_t^{0.5}$					0.03	0.18***		0.18***			0.03	0.03	0.17**		0.04
					(0.09)	(0.06)		(0.06)			(0.09)	(0.09)	(0.07)		(0.09)
s_t^1					0.25***		0.23***		0.23***		0.20**	0.20**		0.24***	0.21**
					(0.09)		(0.06)		(0.06)		(0.10)	(0.10)		(0.06)	(0.11)
s_t^{1+}						-0.13**	-0.09			-7.53*	-0.09		-1.28	1.29	1.74
						(0.06)	(0.07)			(4.46)	(0.07)		(4.25)	(3.92)	(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.7 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 (6). 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.8 Predicting Annual Excess Return

This table presents the results of the predictive regression specified in equation (27). The dependent variable is the log excess return of the S&P 500 index over the next twelve months, r^e_{t+1}. 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^F_t, 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, and the Newey and West (1987) t-statistic (with 18 lags). Starting from January 1998, we generate out-of-sample forecasts by estimating the regression using data up to the current month. These forecasts are then used to compute the out-of-sample R^2, and the p-values of ENC test (Clark and McCracken, 2001) and CW test (Clark and West, 2007).

	r^e_{t+1}				
	(1)	(2)	(3)	(4)	(5)
dr_t	-0.146				-0.228
Hodrick t	[-3.178]				[-2.945]
Newey-West t	(-3.867)				(-3.571)
pd_t		-0.180			0.161
		[-2.168]			[1.820]
		(-2.262)			(1.286)
μ^F_t			2.293		
			[2.033]		
			(2.205)		
KP_t				0.827	
				[2.715]	
				(2.429)	
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.9 Predicting Annual Return: Fama-French Market Return

This table presents the results of the predictive regression specified in equation (27). 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 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, and the Newey and West (1987) t-statistic (with 18 lags). Starting from January 1998, we generate out-of-sample forecasts by estimating the regression using data up to the current month. These forecasts are then used to compute the out-of-sample R^2, and the p-values of ENC test (Clark and McCracken, 2001) and CW test (Clark and West, 2007).

	r_{t+1}^{MKT}				
	(1)	(2)	(3)	(4)	(5)
dr_t	-0.154				-0.222
Hodrick t	[-3.233]				[-2.772]
Newey-West t	(-4.464)				(-3.511)
pd_t		-0.198			0.133
		[-2.302]			[1.608]
		(-2.706)			(1.129)
μ_t^F			2.486		
			[2.327]		
			(2.656)		
KP_t				0.794	
				[2.223]	
				(2.689)	
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.10 Predicting Annual Return: Fama-French Market Excess Return

This table presents the results of the predictive regression specified in equation (27). 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, and the Newey and West (1987) t-statistic (with 18 lags). Starting from January 1998, we generate out-of-sample forecasts by estimating the regression using data up to the current month. These forecasts are then used to compute the out-of-sample R^2, and the p-values of ENC test (Clark and McCracken, 2001) and CW test (Clark and West, 2007).

	$r_{t+1}^{MKT,e}$				
	(1)	(2)	(3)	(4)	(5)
dr_t	-0.144				-0.222
Hodrick t	[-3.060]				[-2.791]
Newey-West t	(-3.745)				(-3.503)
pd_t		-0.179			0.153
		[-2.108]			[1.704]
		(-2.199)			(1.202)
μ_t^F			2.192		
			[2.075]		
			(2.057)		
KP_t				0.725	
				[2.044]	
				(2.251)	
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.11 Monthly Return Prediction

This table presents the results of the predictive regression specified in equation (27). 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 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, and the Newey and West (1987) t-statistic (with 7 lags). Starting from January 1998, we generate out-of-sample forecasts by estimating the regression using data up to the current month. These forecasts are then used to compute the out-of-sample R^2, and the p-values of ENC test (Clark and McCracken, 2001) and CW test (Clark and West, 2007).

	$r_{t+1/12}$				
	(1)	(2)	(3)	(4)	(5)
dr_t	-0.012				-0.017
Hodrick t	[-2.529]				[-1.427]
Newey-West t	(-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.12 Monthly Excess Return Prediction

This table presents the results of the predictive regression specified in equation (27). The dependent variable is the log excess return of the S&P 500 index over the next months, $r^e_{t+1/12}$. 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 Koijen (2010) μ^F_t, 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, and the Newey and West (1987) t-statistic (with 7 lags). Starting from January 1998, we generate out-of-sample forecasts by estimating the regression using data up to the current month. These forecasts are then used to compute the out-of-sample R^2, and the p-values of ENC test (Clark and McCracken, 2001) and CW test (Clark and West, 2007).

	$r^e_{t+1/12}$				
	(1)	(2)	(3)	(4)	(5)
dr_t	-0.011				-0.018
Hodrick t	[-2.394]				[-1.504]
Newey-West t	(-2.684)				(-2.134)
pd_t		-0.014			0.014
		[-1.687]			[0.670]
		(-1.873)			(0.944)
μ^F_t			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.13 Monthly Return Prediction: Fama-French MKT Return

This table presents the results of the predictive regression specified in equation (27). The dependent variable is the log market return from Fama-French in the next month, $r^{MKT}_{t+1/12}$. 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 Koijen (2010) μ^F_t, 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, and the Newey and West (1987) t-statistic (with 7 lags). Starting from January 1998, we generate out-of-sample forecasts by estimating the regression using data up to the current month. These forecasts are then used to compute the out-of-sample R^2, and the p-values of ENC test (Clark and McCracken, 2001) and CW test (Clark and West, 2007).

	$r^{MKT}_{t+1/12}$				
	(1)	(2)	(3)	(4)	(5)
dr_t	-0.012				-0.017
Hodrick t	[-2.354]				[-1.330]
Newey-West t	(-2.742)				(-1.945)
pd_t		-0.015			0.011
		[-1.819]			[0.488]
		(-2.044)			(0.697)
μ^F_t			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.14 Monthly Return Prediction: Fama-French MKT excess Return

This table presents the results of the predictive regression specified in equation (27). The dependent variable is the log excess market return from Fama-French in the next month, $r^{MKT,e}_{t+1/12}$. 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 Koijen (2010) μ^F_t, 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, and the Newey and West (1987) t-statistic (with 7 lags). Starting from January 1998, we generate out-of-sample forecasts by estimating the regression using data up to the current month. These forecasts are then used to compute the out-of-sample R^2, and the p-values of ENC test (Clark and McCracken, 2001) and CW test (Clark and West, 2007).

	$r^{MKT,e}_{t+1/12}$				
	(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^F_t			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.15 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 8. Newey-West standard errors with autocorrelation adjustments up to 18 lags are provided in parentheses. Constant terms are omitted for brevity.

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

$x =$	csp	ep	de	dfy	dfr	tms	cay	ik	SII	SVIX	dp^{Corr}
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
dr_t	-0.149***	-0.159***	-0.157***	-0.156***	-0.155***	-0.160***	-0.169***	-0.178***	-0.150***	-0.221***	-0.209***
	(0.033)	(0.040)	(0.035)	(0.035)	(0.034)	(0.036)	(0.047)	(0.050)	(0.038)	(0.045)	(0.072)
x_t	37.581*	-0.192	-0.007	-0.282	0.421	-0.578	-0.351	2.796	-0.062***	1.373**	0.011
	(21.626)	(1.316)	(0.046)	(4.930)	(0.545)	(1.352)	(1.034)	(10.797)	(0.023)	(0.644)	(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.16 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 8. Newey-West standard errors with autocorrelation adjustments up to 18 lags are provided in parentheses. Constant terms are omitted for brevity.

	r_{t+1}										
$x =$	KP (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	0.146	0.139	0.127	0.134	0.121	0.082	0.146	0.148	0.204	0.127
	(0.113)	(0.273)	(0.130)	(0.195)	(0.127)	(0.130)	(0.130)	(0.114)	(0.119)	(0.125)	(0.146)
dr_t	-0.202***	-0.227***	-0.228***	-0.228***	-0.226***	-0.228***	-0.194***	-0.235***	-0.231***	-0.267***	-0.211***
	(0.063)	(0.074)	(0.068)	(0.064)	(0.068)	(0.067)	(0.058)	(0.067)	(0.066)	(0.076)	(0.056)
x_t	0.299	0.072	-0.013	-0.015	-0.241	-0.699	1.355	-7.215*	0.289*	5.377***	31.269
	(0.261)	(3.804)	(0.428)	(0.159)	(0.639)	(0.824)	(1.368)	(3.737)	(0.175)	(1.831)	(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

$x =$	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	0.182	0.169	0.140	0.143	0.188	0.187	0.097	0.175	0.031
	(0.117)	(0.151)	(0.110)	(0.116)	(0.118)	(0.131)	(0.129)	(0.139)	(0.266)	(0.154)
dr_t	-0.235***	-0.243***	-0.240***	-0.227***	-0.233***	-0.270***	-0.272***	-0.199***	-0.286**	-0.215**
	(0.073)	(0.076)	(0.063)	(0.064)	(0.066)	(0.076)	(0.074)	(0.066)	(0.112)	(0.084)
x_t	-0.343	0.029	2.412	0.358	-0.631	-0.261	1.191	-0.060***	2.042	0.023
	(1.492)	(0.060)	(5.439)	(0.489)	(1.283)	(0.964)	(9.805)	(0.021)	(1.436)	(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

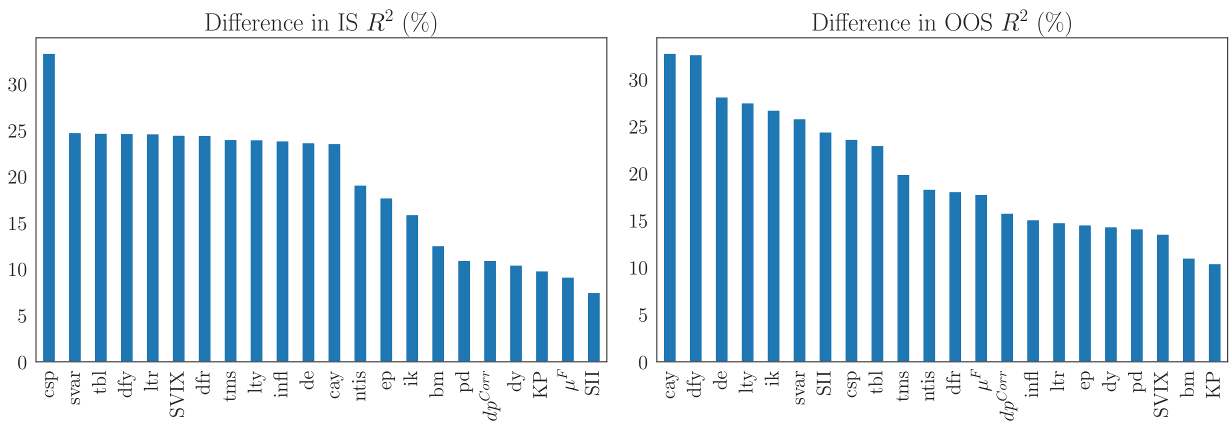


Figure A.4 In-Sample and Out-of-Sample R^2 Wedge between dr and Other Return Predictors: Excess Return

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), 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).

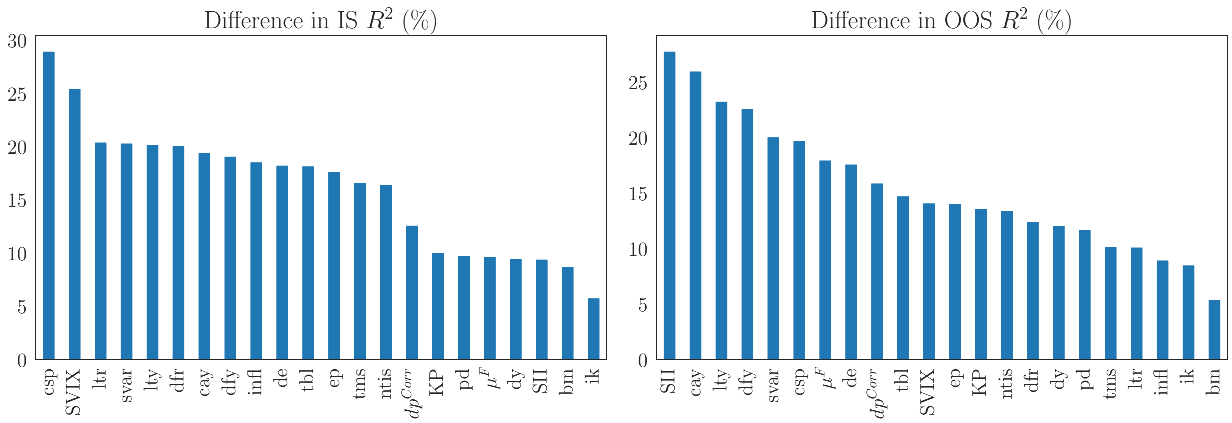


Figure A.5 In-Sample and Out-of-Sample R^2 Wedge between dr and Other Return Predictors: Fama-French Market Return

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), 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).

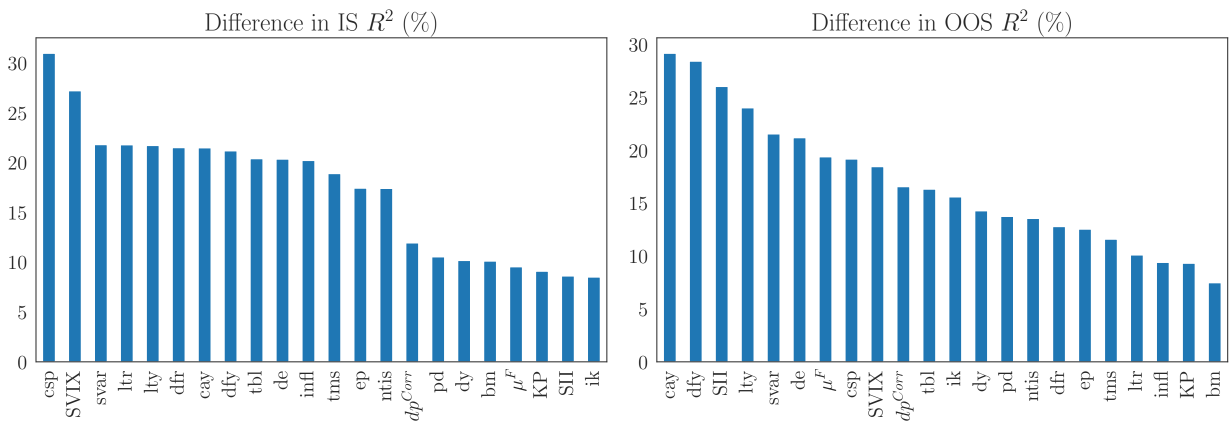


Figure A.6 In-Sample and Out-of-Sample R^2 Wedge between dr and Other Return Predictors: Fama-French Market Excess Return

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), 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).

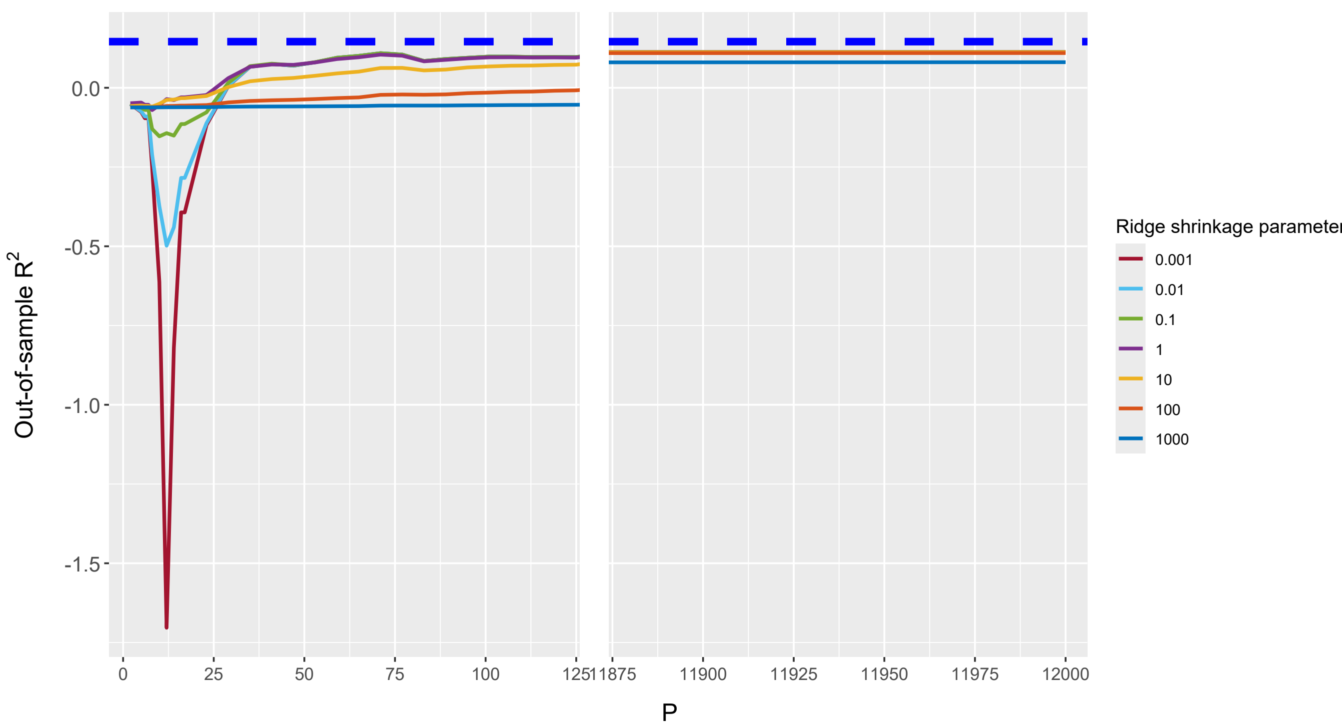


Figure A.7 Out-of-Sample R^2 and Model Complexity

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.

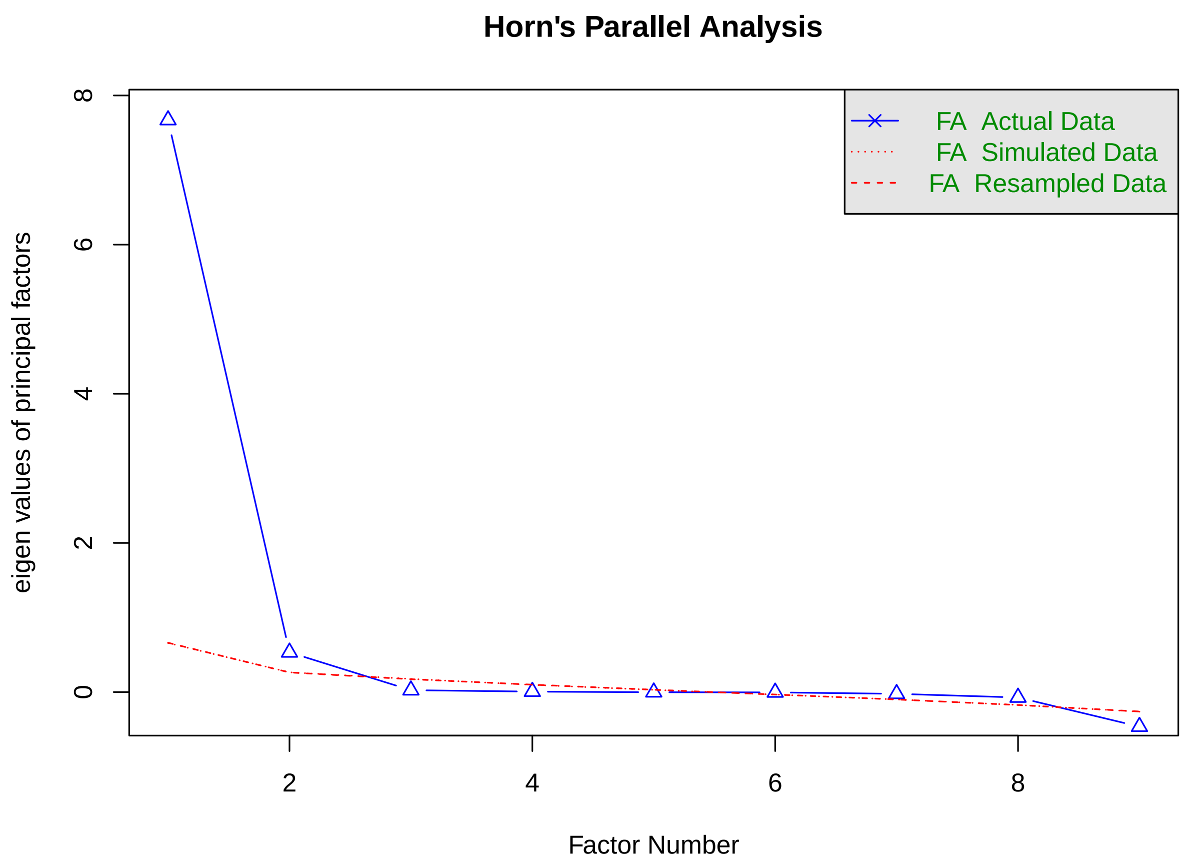


Figure A.8 Horn's Parallel Analysis to Determine Number of Factors.

This scree plot illustrates the results of Horn's Parallel Analysis applied to dividend futures data. The solid blue line displays the eigenvalues of the principal factors extracted from the data, while the red lines represent the thresholds for statistical significance based on random noise (dotted for simulated data, dashed for resampled data). We run the simulation 1,000 times to generate these noise benchmarks.