The Hidden Trend in Stock-Market Cash Flows*

Sebastian Hillenbrand†[1] and Odhrain McCarthy‡[2]

[1]Harvard Business School
[2]NYU Stern School of Business at NYU Abu Dhabi

August 17, 2026

Abstract

We identify a persistent growth component in stock-market cash flows. A present-value state-space model of dividends and the price-dividend ratio over 1871–2022 decomposes valuations into a persistent cash flow trend, a cash flow cycle, and an expected return. Trend growth rises from 2% in the early 1900s to 6% by the 2000s and explains 80% of price-dividend variation. Stripped of the trend, expected returns forecast returns out of sample, at an R^2 of 9% and 22% over one and five years. The trend unifies four facts: valuations' secular increase, near-unit-root persistence, weak in-sample predictability, and out-of-sample failure. The results suggest a reframing of the excess-volatility puzzle.

JEL Classification: G4, G11, G12

Keywords: Hidden Trend, Return predictability, Persistence, Excess-Volatility Puzzle

*This paper supersedes an earlier working paper called "The Optimal Stock Market Valuation Ratio". We are particularly grateful to Sam Hanson and John Campbell for their feedback. We also thank Ben Knox, Martin Lettau, Sydney Ludvigson, Andrea Tamoni, Riccardo Sabbatucci, Shrihari Santosh, Alexi Savov, Adi Sunderam, Erik Stafford, Jeffrey Wurgler and Luis Viceira as well as seminar and conference participants at the Annual Valuation Workshop, the SGF Conference, the EFA, the Macro-Finance Junior Conference, University of Maryland Smith School of Business, Berkeley Haas School of Business, and Harvard Business School for helpful comments.

†shillenbrand@hbs.edu
‡otm210@nyu.edu

1 Introduction

Why do stock-market valuations move so much? Since Shiller (1981), the dominant answer has been that valuations vary primarily because expected returns vary, with a limited role for cash flows. Cochrane (2011) went further, attributing essentially *all* of the variation to expected returns. Under this view, a high valuation ratio signals low expected returns. And because canonical valuation ratios, such as the price-dividend ratio and CAPE, are close to a random walk over 150 years, expected returns must carry a persistent component of their own.

The alternative, posited by Barsky and De Long (1993) and Bansal and Yaron (2004), is that the persistent component sits in expected cash flow growth. A high valuation ratio then need not signal low expected returns—it may signal high persistent growth. And expected returns need not carry a very persistent component to reconcile with near-unit-root valuation ratios.

In this paper we identify a persistent component in stock-market cash flows, and reconcile the two views. We estimate a present-value state-space model of dividends and the price-dividend ratio for the S&P 500 from 1871 to 2022, decomposing the ratio into a persistent growth trend, a transitory cash flow cycle, and an autoregressive expected return. Trend-growth shocks are small—the one-year change in trend growth has a volatility of 18 basis points—but highly persistent, so they accumulate. Because they accumulate, the trend eventually dominates price-dividend variation, its share converging to 73%. But convergence is slow. In a typical 150-year history the trend reaches only 35%. The last 150 years were not a typical path. Trend-growth shocks ran predominantly in one direction, raising S&P500 trend growth from roughly 2% per year in the early 1900s to 6% by the 2000s. Along this path the trend accounts for 80% of price-dividend variation.

Where does the secular trend come from? An accounting decomposition of realized dividend-per-share growth over the post-1926 sample offers suggestive evidence. Inflation contributes to the mid-century acceleration, but over the full sample the rise in per-share growth is overwhelmingly real. In the first third of the sample, dividends grew 0.9 percentage points slower than GDP; in the last third, 1.7 points faster. This swing in the dividend share of GDP accounts for the bulk of the rise in dividend-per-share growth between the two periods, with rising buybacks offsetting slower real GDP growth in more recent decades. The pattern is consistent with the literature on factor shares, market power, investment, and superstar firms (Karabarbounis and Neiman, 2014; Autor, Dorn, Katz, Patterson, and Van Reenen, 2020; Barkai, 2020; De Loecker, Eeckhout, and Unger, 2020; Greenwald, Lettau, and Ludvigson, 2025; Atkeson, Heathcote, and Perri, 2025).

We identify the trend by bringing the macroeconomist's trend-cycle decomposition (Beveridge and

Nelson, 1981; Nelson and Plosser, 1982) to stock-market cash flows, extending the present-value filter of Van Binsbergen and Koijen (2010). Cash flow growth is split into a persistent trend and a transitory cycle.[1] The price-dividend ratio is then log-linearized around a time-varying trend-growth state (Lettau and Van Nieuwerburgh, 2008), rather than around a fixed steady state (Campbell and Shiller, 1988). This departure matters because equity duration is not constant—it rises as trend growth drifts upward—and a fixed expansion point misses it. Identification rests on two facts. A growth shock raises both dividends and valuations while a discount-rate shock moves valuations alone. Additionally, valuations are highly sensitive to persistent growth so even a small trend shift leaves a large imprint on prices. The model combines the direct information in cash flows with the amplified information in prices to recover a persistent component of the kind long-run-risk models emphasize (Bansal and Yaron, 2004; Bansal, Kiku, and Yaron, 2012) but macroeconomic data cannot easily detect (Hansen, Heaton, and Li, 2008; Croce, Lettau, and Ludvigson, 2015).

This hidden trend is a single mechanism that reconciles four facts, usually studied separately. Start with the secular increase in valuation ratios. As trend growth rises, equity duration lengthens. At our estimates, a one-percentage-point increase in trend growth raised valuations by about 15% at the start of our sample and 45% by its end. The model thus implies a secular rise in valuations that is most pronounced in recent decades, explaining why valuations are now so high.

Because the trend dominates realized variation, the price-dividend ratio largely inherits its persistence. The model implies the persistence of the price-dividend ratio is a variance-weighted average of its components' persistences. The trend's dominant realized share thus pulls the ratio's persistence up to the near-unit-root level found in the data.

A valuation ratio dominated by cash flow variation is a poor return signal, which explains its weak in-sample predictive power (Lettau and Ludvigson, 2005; Ang and Bekaert, 2007).[2] Our model quantifies the noise. Expected returns account for only about a fifth of price-dividend variation over the full sample, so four-fifths are cash-flow noise. Accordingly, the univariate slope is attenuated roughly fivefold. The one-year estimate bears this out—it is -0.04, statistically indistinguishable from zero. However, since we have identified that noise in the form of the filtered trend and cycle, we can control for it. Doing so strengthens the slope to -0.24, about six times the univariate value. The filtered expected-return component of the price-dividend ratio offers a direct, alternative route. Free of cash

[1] In the model we treat trend growth as a unit-root process. Over a finite 150-year sample, however, a unit root is observationally equivalent to a highly persistent stationary process (Campbell and Perron, 1991). Consistent with this, replacing it with such a process delivers nearly identical filtered states (see Section 7).

[2] Modified predictors and improved inference restore some in-sample evidence (e.g., Lettau and Ludvigson, 2001; Lewellen, 2004; Boudoukh, Michaely, Richardson, and Roberts, 2007; Lettau and Van Nieuwerburgh, 2008).

flow variation by construction, it forecasts one-year returns on its own, at the same slope. Adding the cash flow states leaves it unchanged, and they enter insignificantly. The same pattern holds at ten years.

In sample, the noise attenuates. Out of sample, it misleads. As trend growth drifted upward, valuation ratios rose in tandem. Univariate regressions, unable to distinguish higher growth from lower discount rates, systematically underpredicted returns and failed to beat the historical mean (Goyal and Welch, 2008; Goyal, Welch, and Zafirov, 2024).[3] The filtered expected-return measure, estimated in real time, accounts for the rise in trend growth, delivering higher return forecasts and an out-of-sample R^2 of 9% at one year and 22% at five years.

Cash flows provide the mirror-image test. Expected-return variation is the noise when the price-dividend ratio forecasts dividend growth, and here the noise is small—about a fifth of price-dividend variation—so the raw ratio is mildly attenuated. Accordingly, the ratio significantly predicts long-horizon cash flows, with dividend growth over the next ten years accounting for a quarter of price-dividend variation, and growth over the next twenty for nearly half. The trend and cycle do better still, and once the cash flow states are included, expected returns carry no incremental information about future growth.

This hidden trend reconciles several long-standing views. Our decomposition operationalizes Cochrane (2008)'s "dog that did not bark" argument, since a small but persistent trend can make dividend growth nearly unpredictable at short horizons yet influential over longer-horizons. Expected returns are the reverse. They are volatile—roughly 6% a year—and mean-revert with a half-life of four years,[4] so their variance dominates over shorter-windows. But as trend variance accumulates, the cash flow share catches up, and at 50 years, over our realized path, cash flows and expected returns each explain roughly half of price-dividend variation. This helps reconcile our results with Cochrane (2011), who analyzes an effective window of roughly fifty years and finds that *realized returns* explain almost all of price-dividend variation. Why the gap? The answer lies in the Stambaugh (1999) bias, which inflates the return coefficient, especially when observations are overlapping (Boudoukh, Israel, and Richardson, 2022). Most of trend growth's rise, from 2% when valuations were low to 6% when they were high, fell inside this window. The model implies that rise was *unforecastable*, so nearly half of price-dividend variation came from cash flow news—all of it reflected in realized returns.[5]

[3]Imposing economic restrictions on forecasts or combining predictors restores some out-of-sample performance (Campbell and Thompson, 2008; Rapach, Strauss, and Zhou, 2010).

[4]These expected return dynamics are consistent with risk premia varying over the business cycle, through habit (Campbell and Cochrane, 1999), disasters (Barro, 2006; Gabaix, 2012), or intermediary capital (He and Krishnamurthy, 2013), or with beliefs, as when investors extrapolate recent returns (Greenwood and Shleifer, 2014; Barberis, Greenwood, Jin, and Shleifer, 2015) or overreact to cash flow news (Bordalo, Gennaioli, Porta, and Shleifer, 2024).

[5]This same pattern attenuates the share of price-dividend variation attributed to expected cash flow growth. Concretely, low valuations forecast low dividend growth, yet realized dividend growth came in high, and vice versa.

Because the trend is estimated from prices and dividends, a natural concern is that it reflects the filter rather than a distinct cash flow state. We therefore subject it to three further cross-examinations. First, the drift is visible in the raw data: trailing averages of realized dividend- and earnings-per-share growth rise alongside the trend. Second, it is recoverable in real time. Estimated recursively on expanding samples, the filtered trend tracks its full-sample counterpart closely. Third, shutting down trend-growth shocks strongly decreases the model's likelihood.

An alternative is that the low-frequency movements in valuations are driven by a persistent component in expected returns (Fama and French, 2002). We give that explanation a free hand, letting long-run expected returns drift as an unrestricted random walk. Trend growth is largely unchanged, and the long-run expected return rises slightly rather than falls over the sample–the opposite of what rising valuation ratios would imply. Allowing the two persistent components to comove, as consumption-based models imply (e.g., Bansal and Yaron, 2004), only strengthens the trend's role rather than weakening it. A separate concern is the secular shift toward repurchases (Grullon and Michaely, 2002), which mechanically inflates the price-dividend ratio. But re-estimating the system on earnings per share, which payout policy does not affect, leaves trend growth and expected returns largely unchanged.

Our analysis implies that today's high valuations are not solely a signal of low expected returns. They also reflect a rise in the persistent growth component of stock-market cash flows, suggesting that markets expect no near-term decline in the profit share. Stripping out the trend yields a 150-year expected-return series that is volatile and recoverable in real time over the post-war period, addressing what Cochrane (2011) identifies as the central question of asset pricing: how expected returns vary. Over short and medium windows, expected returns dominate price-dividend variation. Over the last 150 years, the hidden trend dominates. The results suggest a reframing of the excess-volatility puzzle.

The next section presents motivating evidence on persistence, return predictability, and cash flow trends and cycles. Sections 3 and 4 develop and estimate the state-space system. Section 5 tests the model's predictions. Section 6 explores the implications and economic origins of the trend. Section 7 presents model extensions. Section 8 concludes.

2 Motivating Evidence

Four features of the canonical valuation ratios motivate our analysis. First, valuation ratios have steadily drifted upward in recent decades. Second, they are persistent enough that standard tests struggle to distinguish them from a random walk. Third, over the full 1871–2022 sample, they predict returns only weakly. Fourth, they fail to beat the historical mean out of sample. Together they sit awkwardly with the benchmark view that valuation ratios primarily reflect time-varying expected returns (Campbell and

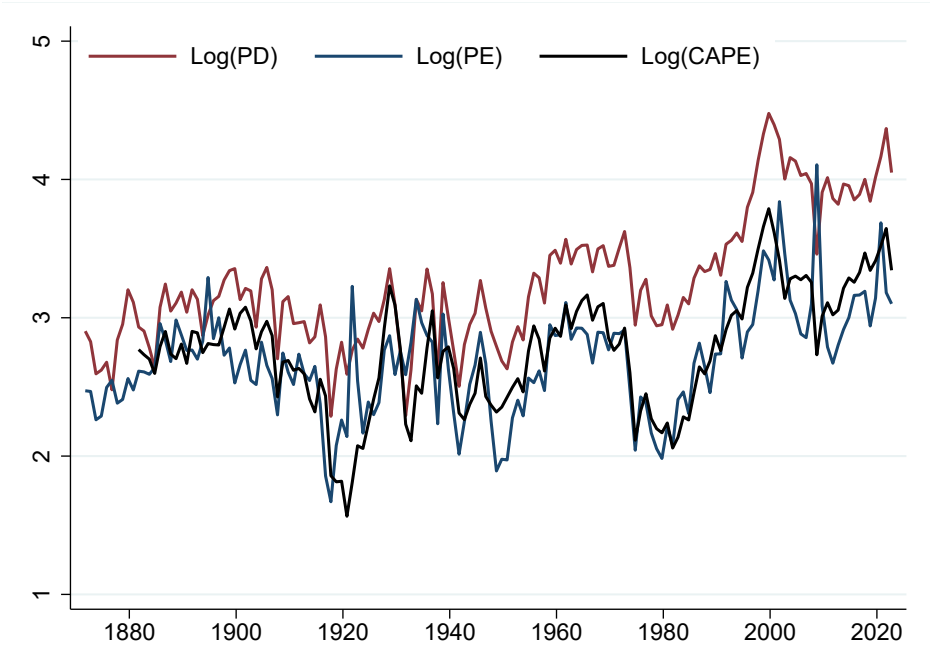
Shiller, 1988; Cochrane, 2011)—under that view, expected returns must contain a component persistent enough to be indistinguishable from a random walk, one that drifted lower as valuations rose. The alternative is that the persistent component sits in cash flow growth.

2.1 The Persistence and Drift in Valuation Ratios

Figure 1 plots three canonical valuation ratios: the log price-dividend ratio (pd), the log price-earnings ratio (pe), and the log cyclically adjusted price-earnings ratio ($cape$); see Appendix A for data details. All three drift upward over the sample: pd rises from about 3 in the late nineteenth century to about 4 by the 2000s, and $cape$ from about 2.7 to 3.3.

The pd and $cape$ ratios are also highly persistent. Table 1 quantifies the persistence. The bias-adjusted AR(1) coefficient is 0.93 for both pd and $cape$.[6] At this level of persistence, standard tests cannot distinguish the series from a random walk: augmented Dickey-Fuller tests fail to reject a unit root, and KPSS tests reject stationarity around a constant mean.

Figure 1: Valuation Ratios



Note: This figure plots the time series of the log price-dividend ratio, the log price-earnings ratio, and the log cyclically adjusted price-earnings ratio. The sample is from 1871–2022.

The price-earnings ratio is an instructive exception. Its AR(1) is 0.75, and ADF rejects the unit root. The reason, we argue—and consistent with the structure we model in Section 3—is not that pe lacks the low-frequency component, but that earnings are more volatile than dividends: firms smooth dividends

[6]The OLS autoregressive root is downward-biased in finite samples. We apply the Kendall (1954) correction, $\hat{\rho}_{adj} = \hat{\rho} + (1 + 3\hat{\rho})/T$.

Table 1: Persistence of Valuation Ratios

	Log(PD) (1)	Log(PE) (2)	Log(CAPE) (3)
Panel A: AR(1) persistence			
$\hat{\rho}$	0.90	0.73	0.91
$\hat{\rho}_{\text{adj}}$	0.93	0.75	0.93
Panel B: Unit-root and stationarity tests			
ADF	-1.54	-4.84***	-2.52
KPSS	3.10***	2.67***	3.33***
Sample	1871–2022	1871–2022	1881–2022

Note: This table reports persistence statistics for three valuation ratios: the log price-dividend ratio (*pd*), the log price-earnings ratio (*pe*), and the log cyclically adjusted price-earnings ratio (*cape*). Data are annual. Panel A reports the autoregressive root from an AR(1) regression of each ratio on its first lag, both unadjusted ($\hat{\rho}$) and adjusted for the downward small-sample bias of Kendall (1954), $\hat{\rho}_{adj} = \hat{\rho} + (1 + 3\hat{\rho})/T$. Panel B reports two unit-root and stationarity tests. The augmented Dickey-Fuller (ADF) test has a unit root as the null hypothesis. The Kwiatkowski, Phillips, Schmidt, and Shin (1992) (KPSS) test has stationarity as the null hypothesis. The ADF augmentation lag is selected by BIC and increased, if necessary, until the residuals satisfy a Ljung and Box (1978) test at the 10% level; the KPSS truncation lag is set to the same value. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

(Lintner, 1956; Brav, Graham, Harvey, and Michaely, 2005), and earnings include transitory accounting items that dividends do not (Hillenbrand and McCarthy, 2024). The *pe* ratio therefore inherits a sizable transitory cash flow cycle that lowers its overall persistence. The raw CAPE fits this interpretation. Averaging earnings over ten years removes much of the transitory variation, and its persistence rises to 0.93—the same as the price-dividend ratio.

Structural break tests point the same way. Following Lettau and Van Nieuwerburgh (2008), we apply the Bai and Perron (1998) procedure and find significant breaks in all three ratios, with *pd* shifting upward in 1954 and again in 1992 (Appendix Table A.5).

We do not interpret this evidence as establishing that valuation ratios contain a unit root. Following Campbell and Perron (1991), we read it as evidence that they are persistent enough that the two are nearly indistinguishable in finite samples.

2.2 Weak Return Predictability

The third and fourth facts concern forecasting. Table 2 reports predictive regressions of one-year-ahead log raw and excess returns on the three ratios. Panel A presents in-sample estimates for the full 1871–2022 sample, with slopes and standard errors adjusted for small-sample bias (Stambaugh, 1999) following Amihud and Hurvich (2004). The predictive power is limited. The slopes on *pd* and *pe* are statistically indistinguishable from zero. Only *cape* predicts raw returns, but it does not reliably predict excess returns. This weakness is well documented, and modified predictors or alternative methodolog-

ical choices can restore some evidence of predictability (e.g., Lettau and Ludvigson, 2001; Lewellen, 2004; Boudoukh, Michaely, Richardson, and Roberts, 2007). Our focus is on why the canonical ratios, used directly, contain at best a noisy signal.

Table 2: The Predictive Power of Valuation Ratios for Returns

	Raw returns			Excess returns		
	Log(PD) (1)	Log(PE) (2)	Log(CAPE) (3)	Log(PD) (4)	Log(PE) (5)	Log(CAPE) (6)
Panel A: In-sample						
b_1^{adj}	−0.019	−0.052	−0.072**	−0.012	−0.032	−0.058
	(0.032)	(0.037)	(0.036)	(0.032)	(0.038)	(0.037)
	[0.545]	[0.164]	[0.046]	[0.700]	[0.398]	[0.116]
Sample	1871–2022	1871–2022	1881–2022	1871–2022	1871–2022	1881–2022
Panel B: Out-of-sample						
R^2_{oos}	−1.82%	−0.57%	−2.50%	−2.76%	−2.99%	−7.11%
CW p-val	[0.377]	[0.238]	[0.114]	[0.392]	[0.444]	[0.259]
Sample	1965–2022	1965–2022	1965–2022	1965–2022	1965–2022	1965–2022

Note: This table reports predictive regressions of log one-year-ahead raw returns and log one-year-ahead excess returns on three valuation ratios: log price–dividend (PD), log price–earnings (PE), and log cyclically adjusted price–earnings (CAPE). Data are annual. Panel A presents in-sample estimates, where the slope coefficient b_1 and its standard error are adjusted for the small-sample bias of Stambaugh (1999) following the correction in Amihud and Hurvich (2004). Standard errors and p-values are shown in round and square brackets, respectively. Panel B reports out-of-sample performance over the 1965–2022 period, based on recursive estimates of the predictive model $r_{t+1} = \alpha_t + \beta_t x_t + u_{t+1}$. The out-of-sample ($R^2_{\text{oos}}$) is defined in Equation (1). The row labeled "CW p-val" reports one-sided p-values for the Clark and West (2007) MSFE-adjusted test. Under the nested null, the predictor has no incremental forecasting value, $\beta_1 = 0$; the alternative is that the unrestricted predictive model, which allows $\beta_1 \neq 0$, has a lower MSFE than the historical-mean benchmark. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

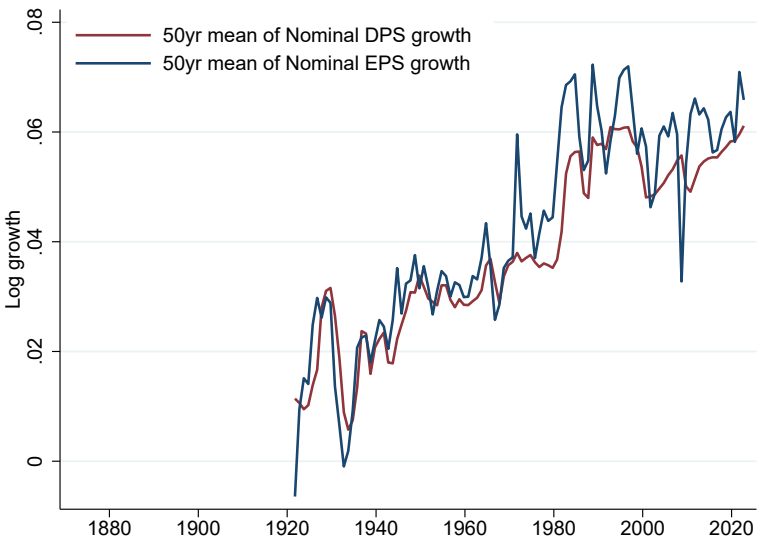
Panel B turns to out-of-sample forecasting. As Goyal and Welch (2008) emphasize, in-sample predictability often fails to translate into real-time forecasting gains. Following their methodology, we estimate each regression recursively, $\hat{r}^{OOS}_{t+1} = \hat{\alpha}_t + \hat{\beta}_t x_t$, using only data available at time t, and evaluate forecasts against the expanding historical mean $\bar{r}_t$ using the out-of-sample R^2:

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

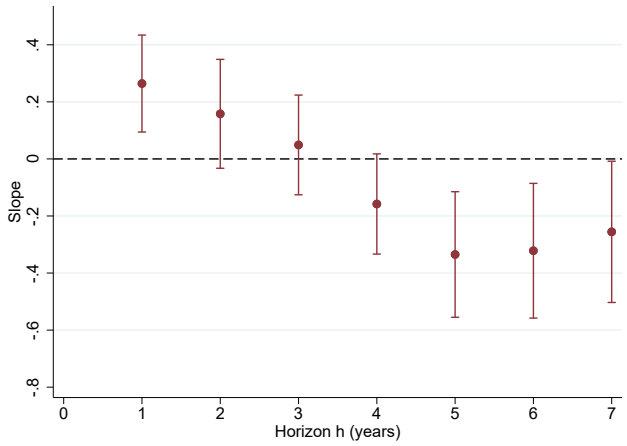
A negative R^2_{OOS} indicates that the predictive model performs worse than the historical-mean benchmark. The R^2_{OOS} is negative in every specification, ranging from −0.6% to −7.1%, and the Clark and West (2007) p-values provide no evidence that the predictive models outperform the historical-mean benchmark. In real time, the canonical valuation ratios have not beaten the historical mean.

Figure 2: Trends and Cycles in Cash Flows

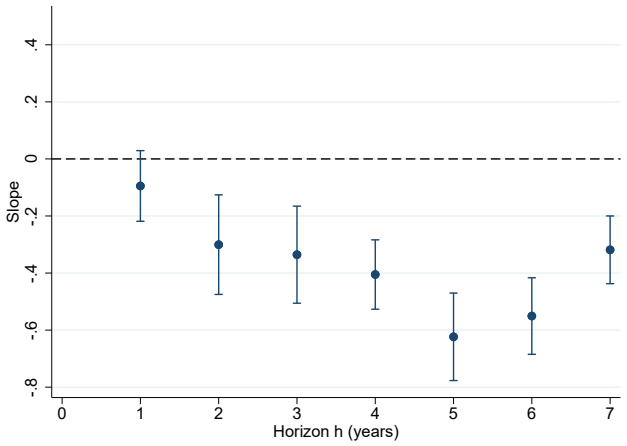
(A) Trend: 50-year rolling average growth



(B) Cycle: dividends



(C) Cycle: earnings



Note: This figure presents evidence on the trend and cycle components of cash flow growth. Panel (A) plots 50-year trailing averages of annual log dividend and earnings per share growth. Panels (B) and (C) plot slope coefficients β_h from regressions of cumulative h-year growth on current one-year growth, $\Delta x_{t,t+h} = \alpha_h + \beta_h \Delta x_t + \nu_{t,t+h}$, following eq. (2), for log dividends (Panel B) and log earnings (Panel C). Vertical bars represent 90% confidence intervals computed using Newey–West HAC standard errors. Data are annual. The sample is 1871–2022.

2.3 Trends and Cycles in Cash Flows

Why do valuation ratios struggle to pass unit-root tests, and why do they forecast returns so poorly? Our hypothesis is that cash flow growth contains both a persistent trend and a transitory cycle, and valuation ratios inherit both. The trend raises their persistence. Together, trend and cycle obscure the return-predictability signal (Lettau and Ludvigson, 2005). Figure 2 presents evidence on each component.

Panel A plots trailing 50-year averages of annual log dividend-per-share and earnings-per-share

growth. Both rise by roughly 4 percentage points over the sample. This drift is visible in the raw average. Long-run cash flow growth and valuation levels rise together, consistent with an upward drift in trend growth contributing to the secular rise in valuation ratios.

Panels B and C examine the cyclical component. We regress cumulative h-year growth on past one-year growth,

$$\Delta x_{t,t+h} = \alpha_h + \beta_h \Delta x_t + \nu_{t,t+h}, \qquad h = 1, \ldots, 7. \tag{2}$$

Under i.i.d. growth, $\beta_h = 0$. Mean reversion implies $\beta_h < 0$: a period of below-average growth is followed by above-average growth as cash flows revert to trend. Because the trend moves little at these horizons, the regressions primarily capture the cyclical component. The estimated slopes are negative for both dividends and earnings, consistent with a mean-reverting cyclical component. The pattern is somewhat stronger for earnings, consistent with dividend smoothing reducing the cycle in dividends.

3 A Present-Value Model with Trends and Cycles

To reconcile these facts, we develop a present-value model in which cash flows contain the trend and cycle documented above. The model builds on Van Binsbergen and Koijen (2010), who treat expected returns and expected dividend growth as latent processes identified jointly from prices and dividends, with two departures. First, instead of modeling cash flow growth as a single autoregressive process, we split it into a persistent trend and a transitory cycle, following the trend-cycle tradition in macroeconomics (Beveridge and Nelson, 1981; Nelson and Plosser, 1982; Harvey, 1985).[7] Second, we log-linearize the price-dividend ratio around the time-varying trend state rather than a fixed steady state. Both departures matter for identifying trend growth.

3.1 Model Setup

We formulate a multivariate unobserved components (UC) model of expected returns and cash flows, which jointly imply a process for the price-dividend ratio through a present-value identity.

We decompose dividends into a persistent trend and a transitory cycle. In the baseline, we treat persistent trend growth as a random walk. The same treatment is standard for other persistent latent trends in macroeconomics, including the natural rate of interest and inflation (Laubach and Williams, 2003; Stock and Watson, 2007; Bauer and Rudebusch, 2020). Campbell and Perron (1991) show that a very persistent process and its unit-root limit are nearly indistinguishable in a finite sample, and that

[7] See also, among others, Clark (1987); Watson (1986); King, Plosser, Stock, and Watson (1991); Harvey (1989); Durbin and Koopman (2001).

imposing the unit root often lowers mean squared error even when the restriction is false. We therefore treat the random walk as an estimation device rather than a claim about integration, while we continue to regard trend growth as a very persistent but stationary process, so that its first and second moments are well-defined. Section 7.3 estimates a stationary version and finds similar filtered states.

Cash Flows Formally, log dividends, d_t, are the sum of a trend, τ_t, and a cycle, c_t:

$$\begin{aligned} d_t &= \tau_t + c_t, \\ \tau_t &= \tau_{t-1} + g_{t-1}, \\ g_t &= g_{t-1} + \varepsilon_{g,t}, \quad \varepsilon_{g,t} \sim \mathcal{N}(0, \sigma_g^2) \\ c_t &= \rho_c \, c_{t-1} + \theta_c \, \varepsilon_{c,t-1} + \varepsilon_{c,t}, \quad \varepsilon_{c,t} \sim \mathcal{N}(0, \sigma_c^2) \end{aligned} \tag{3}$$

This is a "local linear" trend model since the trend τ_t is locally (from $t-1$ to t) linear in the growth rate g_{t-1}. Setting $\sigma_g^2 = 0$ recovers constant trend growth, as in classical present-value models (Shiller, 1981). The cycle c_t follows a parsimonious stationary ARMA(1,1) process ($|\rho_c| < 1$), capturing the mean-reverting dynamics documented in Figure 2.

Expected Returns We model expected returns, $\mathbb{E}_t[r_{t,t+1}] \equiv \mu_t$, as an AR(1) process around a constant steady-state level $\bar{\mu}$:

$$\mu_t = \bar{\mu} + \rho_\mu \, (\mu_{t-1} - \bar{\mu}) + \epsilon_{\mu,t}, \quad \varepsilon_{\mu,t} \sim \mathcal{N}(0, \sigma_\mu^2) \tag{4}$$

This is a frequent specification for expected returns (e.g., Pástor and Stambaugh, 2009; Koijen and Van Nieuwerburgh, 2011; Campbell, 2017). We assume that all shocks are mutually uncorrelated, an identification assumption which is both standard and convenient in UC models. We later relax this assumption in Section 7.

Present-Value Identity Present-value identities link stock prices to future cash flows and returns. The starting point is the definition of log returns:

$$r_{t,t+1} = \Delta d_{t+1} + \log(1 + \exp(pd_{t+1})) - pd_t \tag{5}$$

in which the nonlinear term $\log(1 + \exp(pd_{t+1}))$ is replaced with a first-order linear Taylor approximation. Campbell and Shiller (1988) famously linearize around the sample mean $\overline{pd}$. In our setting, that choice is problematic. A persistent trend-growth component can drive pd_t far from its sample mean for decades at a time, so a fixed expansion point produces large approximation errors.

We therefore follow Lettau and Van Nieuwerburgh (2008) and linearize around a time-varying

steady-state ratio $\overline{pd}_t$, defined as the price-dividend ratio that would prevail if expected returns and cash flow growth remained permanently at their current steady-state levels. For $g_t < \bar{\mu}$, it is defined by

$$\overline{pd}_t \equiv \overline{pd}(g_t) = g_t - \log(\exp(\bar{\mu}) - \exp(g_t)). \tag{6}$$

The constraint never binds in our sample—the filtered g_t stays well below $\bar{\mu}$ throughout—and it disappears in the stationary version of Section 7.3.

The sensitivity of $\overline{pd}_t$ to trend growth is itself not constant. It rises as growth approaches the steady-state discount rate. This sensitivity is the duration of equity. With a steady-state expected return of 8.5%, a one-percentage-point rise in trend growth raises the steady-state price-dividend ratio by roughly 15% when growth is 2%, and 40% when growth is 6%. A fixed steady-state expansion misses this nonlinearity.

Performing a first-order Taylor approximation and iterating forward on the deviation of pd_t from its trend counterpart $\overline{pd}_t$, the log-linearized present-value identity becomes

$$pd_t - \overline{pd}_t = \delta(\rho_t) + \sum_{j=1}^{\infty} \rho_t^{j-1} \, \mathbb{E}_t\left[\Delta d_{t+j} - g_t\right] - \sum_{j=1}^{\infty} \rho_t^{j-1} \, \mathbb{E}_t\left[r_{t+j-1,t+j} - \bar{\mu}\right], \tag{7}$$

where $\delta(\rho_t)$ is a Jensen's-inequality term close to zero (Appendix B gives the full derivation). Eq. (7) has the familiar Campbell and Shiller (1988) form, with two differences. It relates deviations of the price-dividend ratio from its time-varying steady-state value, $pd_t - \overline{pd}_t$, to deviations in cash flows and returns from theirs. And it discounts those deviations at a weight $\rho_t = \frac{\exp(\overline{pd}_t)}{1+\exp(\overline{pd}_t)} = \exp(g_t - \bar{\mu})$, which rises with trend growth. Intuitively, as trend growth rises, ρ_t rises, and future cyclical states are discounted less. However, since cyclical deviations are not long-lasting, the effect of this channel is quantitatively small. For simplicity, we assume ρ_t is constant in the baseline model. Section 7 relaxes this assumption, letting ρ_t vary, and verifies that the channel is small. Setting $\rho_t = \rho$ yields[8]

$$pd_t - \overline{pd}_t = \delta + \sum_{j=1}^{\infty} \rho^{j-1} \, \mathbb{E}_t\left[\Delta d_{t+j} - g_t\right] - \sum_{j=1}^{\infty} \rho^{j-1} \, \mathbb{E}_t\left[r_{t+j-1,t+j} - \bar{\mu}\right]. \tag{8}$$

Substituting the cashflow and expected return dynamics into (8) allows us to express the valuation ratio

[8] ρ_t stays between 0.98 and 0.995 at a quarterly frequency across our sample. We set $\rho = 0.99$—the standard choice at quarterly frequency (Campbell, 2017).

purely in terms of the state variables:

$$pd_t = \delta + \overline{pd}_t - \frac{1}{1-\rho\rho_\mu}(\mu_t - \bar{\mu}) - \frac{1-\rho_c}{1-\rho\rho_c}\, c_t \;+\; \frac{1-\rho}{1-\rho\rho_c}\,\theta_c\,\varepsilon_{c,t}. \tag{9}$$

or, more compactly,

$$pd_t = \delta + \overline{pd}_t + \mathcal{M}_t + \mathcal{C}_t \tag{10}$$

The price-dividend ratio is thus a three-component process. The time-varying steady-state component, $\overline{pd}_t$, captures cash flow movements driven by shifts in trend-growth. The term $\mathcal{M}_t = -\frac{1}{1-\rho\rho_\mu}(\mu_t - \bar{\mu})$ captures the expected-return cycle and $\mathcal{C}_t = -\frac{1-\rho_c}{1-\rho\rho_c}\, c_t \;+\; \frac{1-\rho}{1-\rho\rho_c}\,\theta_c\,\varepsilon_{c,t}$ captures the cash-flow cycle. When the expected returns are high relative to their steady-state values, the price-dividend ratio is low. Conversely, when trend growth is high, the price-dividend ratio is high. The cash flow cycle cuts both ways: a dividend level above trend depresses the ratio—mean reversion implies lower growth ahead—while a positive dividend shock raises it.

How does the model identify low-frequency trend growth? Two features of the joint system do the work. First, dividends and the price-dividend ratio respond differently to the underlying shocks: a rise in trend growth raises both dividends and valuations, while a fall in discount rates raises valuations only. Second, through the duration mechanism above, the price-dividend ratio is highly sensitive to persistent growth, so even small trend shocks have a large effect on prices. The system combines the direct information in cash flows with the amplified information in prices to separate trend growth from discount rates.

3.2 Valuation Ratios as Noisy Predictors

Why should predictive regressions with valuation ratios have limited forecasting ability for returns? The answer is classical measurement error.

Suppose we want to predict the j-period return $r_{t,t+j} = \sum_{i=1}^{j} \rho^{i-1} r_{t+i}$. The best possible predictor is the conditional expectation:

$$r_{t,t+j} = \mathbb{E}_t[r_{t,t+j}] + \psi_{t,t+j}, \qquad \mathbb{E}_t(\psi_{t,t+j}) = 0. \tag{11}$$

Regressing realized returns on $\mathbb{E}_t[r_{t,t+j}]$ yields a slope coefficient of one. In practice, this predictor is not

observed. Suppose instead, we observe a noisy proxy

$$x_t = \mathbb{E}_t[r_{t,t+j}] + u_t, \qquad \mathbb{E}(u_t \mid \mathbb{E}_t[r_{t,t+j}]) = 0. \tag{12}$$

The predictive slope becomes

$$\frac{\text{Cov}(r_{t,t+j}, x_t)}{\text{Var}(x_t)} = \frac{\text{Var}(\mathbb{E}_t[r_{t,t+j}])}{\text{Var}(x_t)} = 1 - \frac{\text{Var}(u_t)}{\text{Var}(x_t)}. \tag{13}$$

Thus, the slope is no longer one but is attenuated toward zero by the presence of noise – the classical measurement error result.

The price-dividend ratio is such a proxy. From equation (9), the price-dividend ratio can be written as

$$pd_t = \delta + \mathcal{M}_t + \eta_t, \qquad \eta_t \equiv \overline{pd}(g_t) + \mathcal{C}_t \tag{14}$$

The valuation ratio thus carries the expected-return signal plus cash flow noise. The coefficient linking j-period conditional expected returns to the price-dividend ratio over a finite window of length T is therefore

$$b_{j,T} = \frac{\text{Cov}_T(\mathbb{E}_t[r_{t,t+j}], pd_t)}{\text{Var}_T(pd_t)} = -[1 - (\rho\rho_\mu)^j] \frac{\text{Var}_T(\mathcal{M}_t)}{\text{Var}_T(pd_t)} = -[1 - (\rho\rho_\mu)^j] \left[1 - \frac{\text{Var}_T(\eta_t)}{\text{Var}_T(pd_t)}\right]. \tag{15}$$

where throughout we write $\text{Var}_T(X) \equiv \mathbb{E}\left[\widehat{\text{Var}}_T(X)\right]$ and $\text{Cov}_T(X,Y) \equiv \mathbb{E}\left[\widehat{\text{Cov}}_T(X,Y)\right]$. These moments are well defined once $\overline{pd}(g_t)$ is approximated by its first-order expansion in g_t, as throughout the paper.[9]

The two terms have distinct interpretations. The first, $-[1 - (\rho\rho_\mu)^j]$, reflects the fact that the valuation ratio contains the expected-returns signal for the entire term structure of returns. As $j \to \infty$, its magnitude converges to one. The second term captures attenuation from cash flow noise. With no noise, it equals one, and the entire variance of pd_t would be driven by $\mathcal{M}_t$. As the noise variance increases, the slope shrinks toward zero. Solving for $\text{Var}_T(\eta_t)$, we can write

$$b_{j,T} = -[1 - (\rho\rho_\mu)^j] \left[1 - \frac{\text{Var}_T(\overline{pd}(g_t)) + \text{Var}_T(\mathcal{C}_t)}{\text{Var}_T(pd_t)}\right]. \tag{16}$$

[9]Under the stationary trend growth model (derived in Appendix G), no expansion of the trend component is needed, and the variance shares $\text{Var}_T(\overline{pd}_t)/\text{Var}_T(pd_t)$ under the two specifications agree within two percentage points over our sample.

The appendix provides analytical formulas for $\mathrm{Var}_T(\overline{pd}(g_t))$, $\mathrm{Var}_T(\mathcal{C}_t)$, $\mathrm{Var}_T(\mathcal{M}_t)$, and $\mathrm{Var}_T(pd_t)$. In the data we predict realized returns, not conditional expectations, so that $\mathrm{Cov}_T(r_{t,t+j}, pd_t)/\mathrm{Var}_T(pd_t) = b_{j,T} + \mathrm{Cov}_T(\psi_{t,t+j}, pd_t)/\mathrm{Var}_T(pd_t)$. The second term is the finite-window analogue of the familiar Stambaugh (1999) bias mechanism. It is zero in the population, but non-zero in the finite sample, and the bootstrap of Section 5 accounts for this term. The model leads to the following proposition.

Proposition 1. *When predicting returns with the pd ratio, variation in the cash flow cycle and trend growth attenuate the slope coefficient toward zero.*

Univariate predictive regressions may therefore recover little of the expected-return signal whenever valuation ratios are substantially driven by cash flow trends and cycles—a potential explanation for the weak evidence in the return-predictability literature. The proposition also delivers quantitative predictions, which we take to the data in Section 5.

Prediction 1. *Predicting returns with the pd ratio:*

- *The slope coefficient linking j-period conditional expected returns $\mathbb{E}_t[r_{t,t+j}]$ to pd_t is given by eq.* (16) *and is attenuated toward zero.*
- *A multivariate regression that controls for the cycle components c_t and $\epsilon_{c,t}$ and the trend growth component $\overline{pd}_t$ recovers the unattenuated coefficient $-[1-(\rho_\mu\rho)^j]$.*

Both claims follow directly from eq. (16). The model therefore delivers a sharp testable prediction: if cash flow trends and cycles attenuate return predictability, a multivariate regression controlling for these components should yield a slope coefficient substantially larger in magnitude than its univariate counterpart.

If the expected-return component $\mathcal{M}_t$ were directly observable, it could serve as a predictor in its own right:

$$b_j \;=\; \frac{\mathrm{Cov}(r_{t,t+j},\, \mathcal{M}_t)}{\mathrm{Var}(\mathcal{M}_t)} \;=\; -[1-(\rho_\mu\rho)^j] \tag{17}$$

Prediction 2. *Predicting returns with the expected-return signal $\mathcal{M}_t$:*

- *Regressing j-period realized returns $r_{t,t+j}$ on the expected-return signal $\mathcal{M}_t$ yields the unattenuated coefficient $-[1-(\rho_\mu\rho)^j]$.*
- *Augmenting this regression with c_t, $\epsilon_{c,t}$, and $\overline{pd}_t$ leaves the coefficient unchanged.*

Because $\mathcal{M}_t$ is free of cash flow variation by construction, its slope is not attenuated, and cash flow controls should have no effect.

The same logic gives a cash flow restriction. Predictions 1 and 2 concern returns. There, the expected-return component is signal, while the cash flow trend and cycle are noise. For cash flows, the roles reverse. The trend and cycle are signal. The expected-return component is noise. It is straightforward to show (see Appendix F)

$$\mathbb{E}_t\big[\Delta d_{t,t+j}\big] \equiv \mathbb{E}_t\big[\sum_{k=1}^{j} \rho^{k-1}\,\Delta d_{t+k}\big] \;=\; a_j\, g_t \;+\; b_j\, c_t \;+\; c_j \varepsilon_{c,t}, \tag{18}$$

where $a_j > 0$, $b_j < 0$ and c_j takes the sign of θ_c. The expected-return component M_t does not enter.

Prediction 3. *Predicting cash flow growth:*

- *When predicting future cash flow growth with the pd ratio, variation in expected returns attenuates the slope coefficient.*
- *Replacing pd_t with the filtered cash flow states raises predictive power. The states span the model's forecast of future dividend growth, while pd_t is a single noisy combination of them. Trend growth, g_t, enters positively; the cash flow cycle, c_t, enters negatively, forecasting lower growth as dividends revert to trend.*
- *Once the cash flow states are included, the expected-return component M_t carries no incremental information about future dividend growth.*

4 Estimating the Model

We take the model to the data. Estimation filters the cash flow trend, cycle and expected returns contained in the price-dividend ratio. Section 5 uses the filtered states to test the model's forecasting restrictions. Section 6 uses the parameter estimates to decompose price-dividend variation into the variance shares of the three components, and to characterize the implied persistence of the price-dividend ratio.

4.1 State-Space Representation

We can write our unobserved components as a state-space system with the state vector

$$s_t = (\tau_t,\; g_t,\; c_t,\; \varepsilon_{c,t},\; \mu_t,\; 1)^\top \tag{19}$$

where the transition equation $s_{t+1} = A\, s_t + C\, \tilde{\epsilon}_{t+1}$ is such that

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & \rho_c & \theta_c & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \rho_\mu & (1-\rho_\mu)\bar{\mu} \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \qquad C = \begin{bmatrix} 0 & 0 & 0 \\ \sigma_g & 0 & 0 \\ 0 & \sigma_c & 0 \\ 0 & \sigma_c & 0 \\ 0 & 0 & \sigma_\mu \\ 0 & 0 & 0 \end{bmatrix}, \tag{20}$$

where C is a 6×3 "shock-loading" matrix governing how each structural innovation enters and scales the state variables, the state innovations $\tilde{\epsilon}_{t+1}^{\top} = \left(\tilde{\epsilon}_{g,t+1},\ \tilde{\epsilon}_{c,t+1},\ \tilde{\epsilon}_{\mu,t+1}\right) \sim \mathcal{N}(0, I_3)$ are mutually independent, and A determines the loading on past states.

The measurement equation stacks log dividends and the log price-dividend ratio:

$$x_t = (d_t, pd_t)^{\top} \tag{21}$$

The standard Kalman filter requires a linear measurement equation, but the valuation-ratio equation contains the steady-state log price–dividend ratio, $\overline{pd}_t$, which is a non-linear function of the trend-growth state g_t. We therefore employ an Extended Kalman Filter (EKF), which replaces the nonlinear measurement function with its first-order Taylor approximation, updated at each time step. Doing so, the linearized measurement equation takes the form $x_t = G_t\, s_t$, where (see the appendix for a detailed derivation)

$$G_t = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & \lambda_t & -\dfrac{1-\rho_c}{1-\rho\rho_c} & \dfrac{1-\rho}{1-\rho\rho_c}\,\theta_c & -\dfrac{1}{1-\rho_\mu\rho} & \kappa_t \end{bmatrix}, \tag{22}$$

with

$$\lambda_t \equiv \left.\frac{\partial \overline{pd}_t}{\partial g_t}\right|_{g_t = \hat{g}_{t|t-1}} = 1 + \frac{\exp(\hat{g}_{t|t-1})}{\exp(\bar{\mu}) - \exp(\hat{g}_{t|t-1})}, \tag{23}$$

$$\kappa_t = \delta + \frac{\bar{\mu}}{1-\rho\rho_\mu} + \hat{g}_{t|t-1} - \log\!\big(\exp(\bar{\mu}) - \exp(\hat{g}_{t|t-1})\big) - \lambda_t\, \hat{g}_{t|t-1}. \tag{24}$$

The first row of G_t maps dividends to the trend and cycle. The second row encodes the structural present-value constraint. At each step of the filter, G_t is re-evaluated at the current prior estimate $\hat{g}_{t|t-1}$,

so that the linearization tracks the evolving steady-state of the pd ratio.

4.2 Filtered Cash Flows

We estimate the model at quarterly frequency using maximum likelihood. For each candidate parameter vector, the EKF delivers filtered state estimates and an approximate log-likelihood; we select the parameters that maximize it. Appendix C.2 provides further details. Table 3 reports the parameter estimates and bootstrapped standard errors.

Table 3: Trend-Cycle Model: Parameter Estimates

Parameter	ρ_c	θ_c	$\sigma_g(\cdot 100)$	σ_c	ρ_μ	σ_μ	$\bar{\mu}$
Estimate	0.9711	0.4927	0.0225	0.0273	0.9564	0.0048	0.0214
	(0.0135)	(0.0375)	(0.0146)	(0.0008)	(0.0193)	(0.0016)	(0.0018)

Note: This table reports parameter estimates for the trend-cycle model, estimated using quarterly frequency data on log dividends and the log price-dividend ratio. The sample period is 1871–2022. Bootstrapped standard errors are shown in parentheses.

Figure 3 plots the filtered states. Panel A shows the key finding: trend growth roughly triples over the sample, drifting from about 2% per year before 1900 to about 6% in recent decades. One might worry that the filter manufactures this drift. The raw data suggest otherwise: 50-year trailing averages of realized dividend-per-share and earnings-per-share growth track the filtered trend (Panel A). Section 6.2 examines the economic forces behind it.

The trend-level (Panel B) tracks the dividend level closely but is extremely smooth, reflecting the removal of the cyclical component. The cycle (Panel C) is more volatile, spiking around major historical episodes but consistently reverting toward zero. In annual terms, the cycle is moderately persistent in *levels*—the AR(1) of cyclical dividends is 0.90—and its growth is volatile, with an annualized volatility of 7.5%.[10]

Innovations to trend growth are individually small—the one-year change in annualized trend growth has a standard deviation of just 18bps.[11] However, because trend growth is extremely persistent, they accumulate over long horizons.[12] Is σ_g statistically distinguishable from zero? Because σ_g is a standard deviation, the null $\sigma_g = 0$ lies on the boundary of the parameter space, so standard Wald tests—including those based on the bootstrap standard errors in Table 3—do not apply. We instead compare

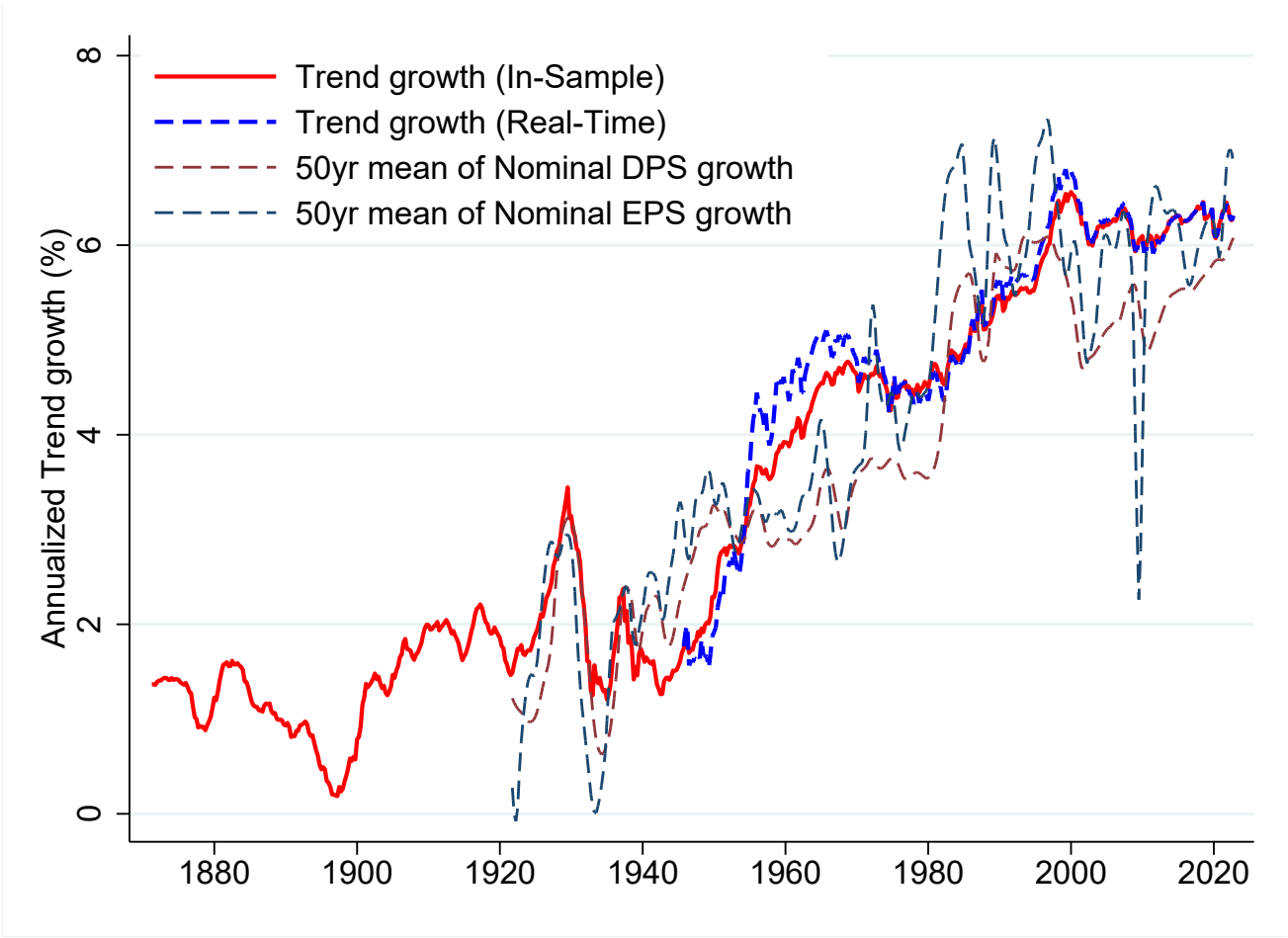
[10] Annual cycle growth $c_t - c_{t-4}$ has variance $2[\gamma(0) - \gamma(4)]$, where $\gamma(0) = (1 + \theta_c^2 + 2\rho_c\theta_c)\sigma_c^2/(1-\rho_c^2)$, $\gamma(1) = \rho_c\,\gamma(0) + \theta_c\,\sigma_c^2$, and $\gamma(h) = \rho_c\,\gamma(h-1)$ for $h \geq 2$. Substituting our estimates yields an annualized cycle growth volatility of 7.5%.

[11] Annualized trend growth is $4g_t$, so its one-year change is $4(g_t - g_{t-4}) = 4\sum_{i=0}^{3}\varepsilon_{g,t-i}$, with standard deviation $4\sqrt{4}\,\hat{\sigma}_g =$ 18bps.

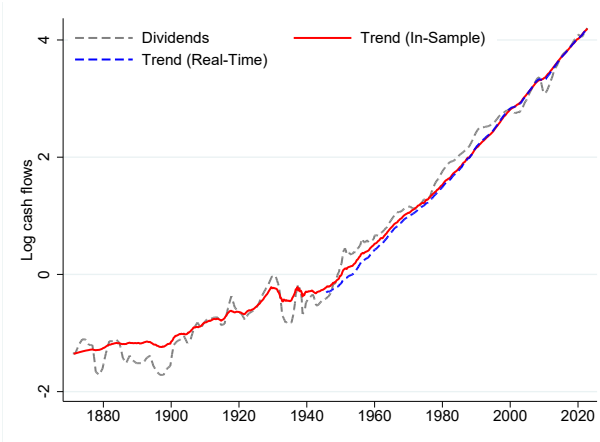
[12] Under the unit-root limit, the standard deviation of the cumulative change in annualized trend growth over 150 years is $0.18\% \cdot \sqrt{150} \approx 2.2\%$.

Figure 3: Trend-Cycle Model: Filtered Cash Flows

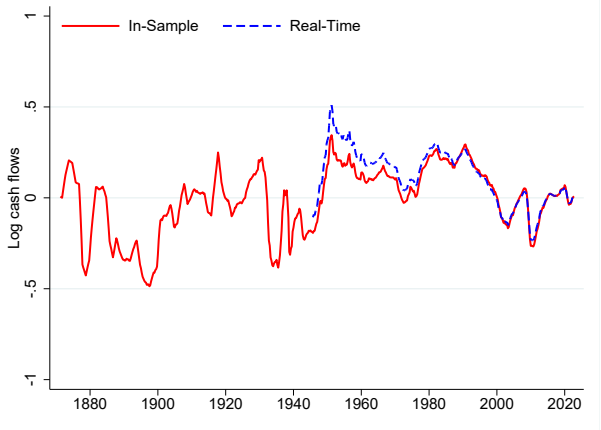
(A) Annualized Trend growth ($4\hat{g}_t$)



(B) Trend ($\hat{\tau}_t$)



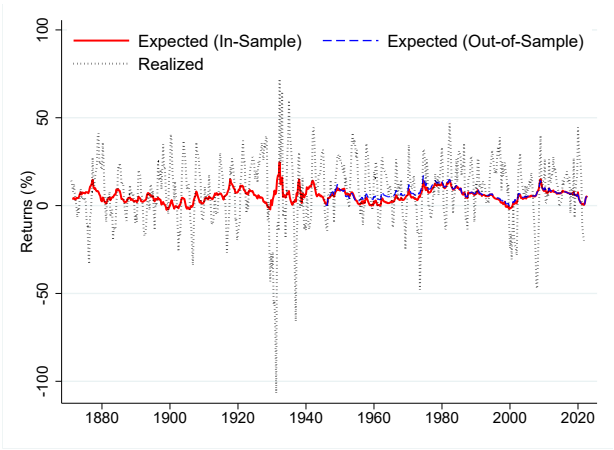
(C) Cycle ($\hat{c}_t$)



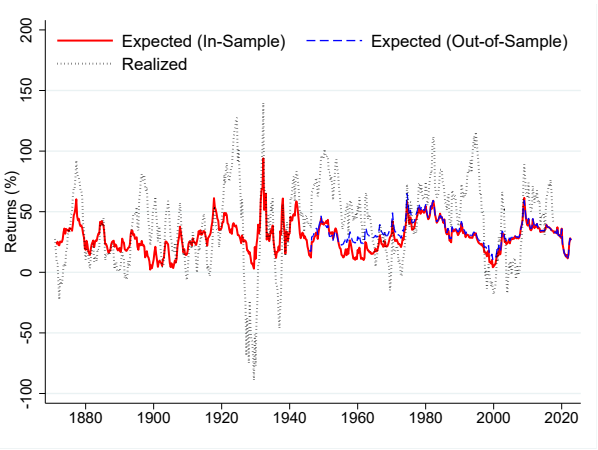
Note: This figure displays the filtered estimates for trend growth ($\hat{g}_t$) in Panel A, the trend ($\hat{\tau}_t$) in Panel B, and the cycle ($\hat{c}_t$) in Panel C. In Panel A, trend growth is annualized ($4 * \hat{g}_t$). The solid red line plots the in-sample estimates derived from the full sample (1871–2022), while the dashed blue line plots the real-time estimates derived from an expanding training window beginning in 1945. Panel A additionally plots the 50-year (dashed maroon) and 50-year (dotted navy) trailing means of annual dividend-per-share and earnings-per-share growth.

Figure 4: Trend-Cycle Model: Filtered Expected Returns

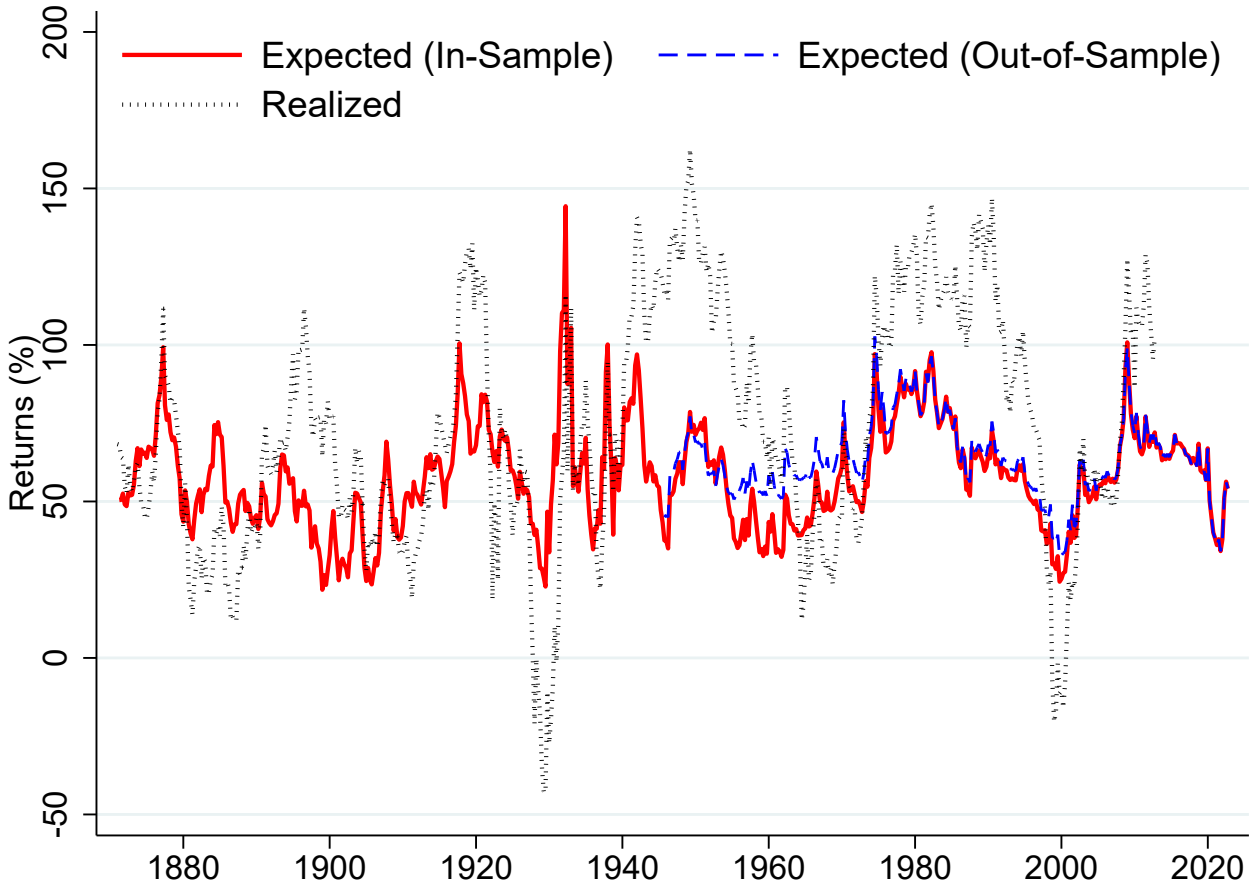
(A) 1-year



(B) 5-year



(C) 10-year



Note: This figure plots model-implied expected returns at the 1-year (Panel A), 5-year (Panel B), and 10-year (Panel C) horizons, constructed from Equation (4) using filtered states and parameter estimates. The solid red line uses full-sample estimates (1871–2022); the dashed blue line uses real-time expanding-window estimates beginning in 1945. Realized returns are shown for comparison (dotted black line).

likelihoods. Figure A.1 plots the profile log-likelihood for σ_g. The likelihood is maximized at $\hat{\sigma}_g$ and declines steeply toward zero. Re-estimating with σ_g fixed at zero yields a likelihood-ratio statistic of 46.6, rejecting the null at any conventional level. Section 7.1 provides the economic reasoning: forcing a deterministic trend pushes the cash flow cycle and expected-return process toward unit roots as they absorb the secular variation trend growth used to handle, at a large likelihood cost.

4.3 Filtered Expected Returns

Given the parameter estimates and filtered states, we construct expected returns over the next j years as the ρ-discounted sum of iterated AR(1) forecasts from Equation (4):

$$\widehat{\mathbb{E}}_t[r_{t,t+j}] = \sum_{i=1}^{j} \rho^{i-1}\, \widehat{\mathbb{E}}_t[r_{t+i-1,t+i}] = \frac{1-\rho^j}{1-\rho}\,\bar{\mu} + \frac{1-(\rho\rho_\mu)^j}{1-\rho\rho_\mu}\,(\hat{\mu}_t - \bar{\mu}). \tag{25}$$

As $j \to \infty$, the deviation term converges to $(\hat{\mu}_t - \bar{\mu})/(1-\rho\rho_\mu)$—the expected-return component of the price-dividend ratio with the sign flipped.

Figure 4 plots the expected-return series at the 1-, 5-, and 10-year horizons alongside realized returns. Three features stand out. First, with the persistent trend purged, expected returns retain substantial variation—an annualized volatility of 6.0%—but their persistence is now moderate, with an annual AR(1) of 0.84 and a half-life of roughly four years. Both tests in Table A.2 point to stationarity from opposite nulls: the ADF test rejects a unit root, and the KPSS test does not reject stationarity. Second, the series aligns with economic intuition: expected returns are countercyclical, spiking during the Great Depression and the stagflation of the 1970s–80s, and compressing to historical lows during the dot-com bubble. Third, the fit with realized returns improves with the horizon: at one year, expected-return shocks have largely yet to materialize in realized returns, whereas by ten years nearly 90% of the expected-return component has been captured.

Finally, we verify the filter in real time. Beginning in 1945, we re-estimate all parameters by maximum likelihood at each subsequent quarter, using only data available at that date, and retain the terminal filtered state from each vintage. The real-time states (dashed blue in Figure 3) track the full-sample estimates closely, confirming that an econometrician estimating the model in real time would have recovered essentially the same trend.

5 Predictability as Model Validation

The hidden trend is a cash flow state, and the expected-return signal is a return state. If the model is right, each should forecast its own object, at the slope the model pins down, and carry no incremental

information about the other's. This section puts that claim to the data three ways. The in-sample return regressions test the analytical restrictions of Predictions 1 and 2. The out-of-sample regressions ask whether the expected-return signal works in real time. The cash-flow regressions test Prediction 3 at long horizons, where a small but persistent trend is most influential.

5.1 In-Sample Return Predictability

Table 4 reports predictive regressions of realized returns at the one- and ten-year horizons. Return predictability regressions are subject to several well-known statistical issues: persistent predictors bias OLS coefficients, overlapping long-horizon returns inflate t-statistics, and the two interact (Stambaugh, 1999; Torous, Valkanov, and Yan, 2004; Boudoukh, Israel, and Richardson, 2022). Our approach mitigates these issues on two counts. First, expected returns, stripped of the trend, are less persistent than the raw price-dividend ratio. Second, we base p-values on a parametric bootstrap (square brackets) that repeats the full empirical procedure—simulation, re-estimation, re-filtering, regression—in every replication to account for parameter estimation error, generated regressors, predictor persistence, and overlapping multi-year returns (see Appendix C.3 for details). The bottom panel reports the model-implied coefficients and p-values for the null that the empirical and model-implied coefficients are equal.

We begin with Prediction 1. The expected-return share of price-dividend variation, estimated in Section 6, makes the prediction quantitative. The model implies that the expected-return share of pd_t variation over the full sample is roughly a fifth, and that the unattenuated one-year slope is -0.20, predicting a univariate slope of $0.20 \times (-0.20) \approx -0.04$. Column (1) delivers exactly this value. The slope is -0.04, statistically indistinguishable from zero. If the attenuation reflects cash flow noise, controlling for that noise should restore the model-implied slope of -0.20. Column (2) adds the filtered cash flow components, and the slope on pd_t is -0.24, six times the univariate value, statistically significant, and close to the analytical prediction. The remaining coefficients also align with the model.

Prediction 2 offers the alternative route. Rather than controlling for cash flow noise, we use the filtered expected-return signal $\widehat{\mathcal{M}}_t$ directly, which is free of cash flow variation by construction. In column (3), $\widehat{\mathcal{M}}_t$ predicts one-year returns with a slope of -0.22, close to the predicted -0.20, and raises the R^2 from 1% to 6%. Column (4) shows that the slope is stable when cash flow controls are added, and the controls themselves are insignificant—consistent with $\widehat{\mathcal{M}}_t$ fully absorbing the expected-return signal in pd_t.

Columns (5)–(8) repeat the analysis for ten-year returns and show the same pattern. The univariate slope on pd_t is -0.23 and insignificant (column 5); controlling for cash flow noise strengthens it five-fold (column 6); $\widehat{\mathcal{M}}_t$ alone delivers comparable predictive power (column 7); and cash flow controls leave

Table 4: Testing the Model: In-Sample Return Predictability

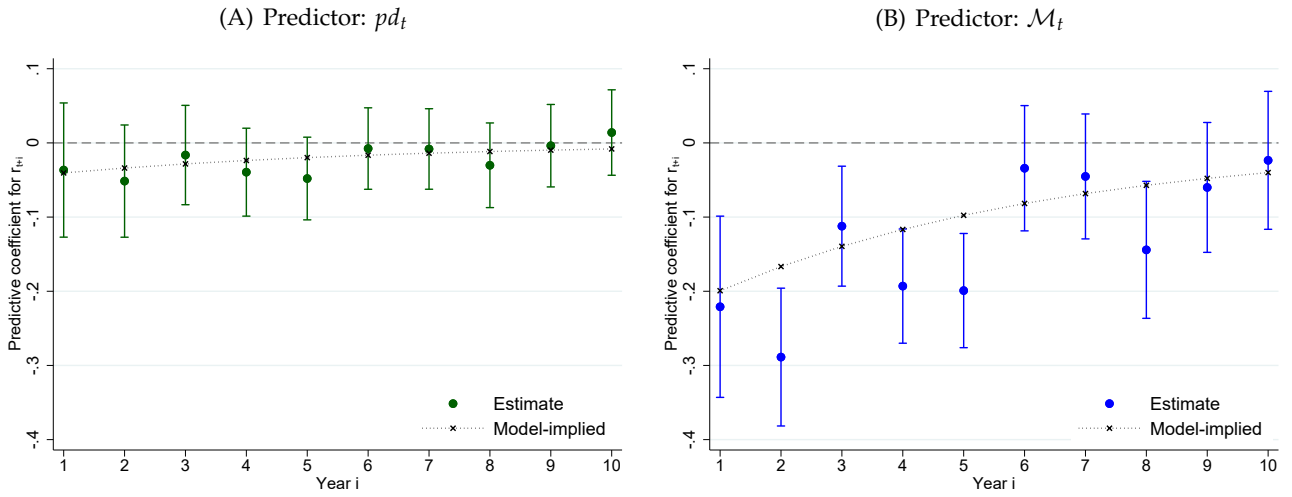
	$r_{t,t+j} = \sum_{i=1}^{j} \rho^{i-1} r_{t+i}$								
	1-year				10-year				20-year
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$\widehat{\mathcal{M}}_t$			-0.22***	-0.24***			-1.16***	-1.16***	
			(0.00)	(0.00)			(0.00)	(0.00)	
			[0.01]	[0.00]			[0.01]	[0.01]	
pd_t	-0.04	-0.24***			-0.23	-1.16***			-0.25
	(0.22)	(0.00)			(0.05)	(0.00)			(0.10)
	[0.68]	[0.00]			[0.48]	[0.01]			[0.61]
$\widehat{\overline{pd}}_t$		0.25***		0.01		1.21***		0.05	
		(0.00)		(0.75)		(0.00)		(0.75)	
		[0.00]		[0.53]		[0.01]		[0.45]	
$\hat{c}_t$		-0.26*		-0.08		-0.89*		-0.02	
		(0.01)		(0.46)		(0.01)		(0.94)	
		[0.06]		[0.35]		[0.08]		[0.81]	
$\hat{\varepsilon}_{c,t}$		-0.21		-0.24		0.42		0.28	
		(0.44)		(0.38)		(0.41)		(0.59)	
		[0.51]		[0.43]		[0.44]		[0.60]	
Observations	151	151	151	151	142	142	142	142	132
R^2	0.01	0.07	0.06	0.07	0.07	0.34	0.34	0.34	0.05
Adj. R^2	0.00	0.04	0.05	0.04	0.06	0.32	0.34	0.32	0.04
Model-implied coefficients:									
$\mathcal{M}_t$			-0.20	-0.20			-0.89	-0.89	
			{0.88}	{0.56}			{0.34}	{0.31}	
pd_t	-0.04	-0.20			-0.20	-0.89			-0.32
	{0.34}	{0.57}			{0.44}	{0.31}			{0.13}
$\overline{pd}_t$		0.20		0.00		0.89		0.00	
		{0.31}		{0.26}		{0.24}		{0.25}	
c_t		-0.15		0.00		-0.66		0.00	
		{0.35}		{0.37}		{0.70}		{0.88}	
$\varepsilon_{c,t}$		0.03		0.00		0.11		0.00	
		{0.64}		{0.64}		{0.65}		{0.69}	

Note: This table tests Predictions 1 and 2 and reports predictive regressions of realized returns, $r_{t,t+j} = \sum_{i=1}^{j} \rho^{i-1} r_{t+i}$, over 1-year, 10-year and 20-year horizons. Independent variables include the filtered expected-return signal ($\widehat{\mathcal{M}}_t$), log price-dividend ratio (pd_t), and filtered cash-flow noise components ($\widehat{\overline{pd}}_t, \hat{c}_t, \hat{\varepsilon}_{c,t}$), derived from full-sample parameter estimates (1871–2022). The top panel reports OLS regression coefficients, p-values based on Newey-West t-stats in round brackets, and p-values based on a parametric bootstrap procedure in square brackets (see Appendix C.3). The bottom panel reports the model-implied (analytical) coefficient and the p-value (in curly brackets) for the test $H_0: \beta = \beta^{\text{analytical}}$ versus $H_1: \beta \neq \beta^{\text{analytical}}$. Significance levels based on bootstrapped p-values: *(p<0.10), **(p<0.05), ***(p<0.01).

the slope on $\widehat{\mathcal{M}}_t$ largely unchanged (column 8). Across both horizons, we cannot reject equality of the empirical and analytical coefficients at conventional levels. Column (9) predicts twenty-year returns with the pd ratio and finds only a small increase in the coefficient compared to the ten-year horizon.

A natural question is whether these results depend on early dividend data, which is reconstructed from pre-CRSP sources and less reliable than the post-1926 CRSP record. Appendix Table A.6 repeats the regression analysis for the post-1926 CRSP subsample. The patterns are robust and, if anything, sharper.

Figure 5: Term Structure of Expected Returns



Note: This figure reports predictive regressions of realized future one-year returns, r_{t+i} on the log price-dividend ratio (pd_t) in Panel (A) and the filtered expected-return signal ($\widehat{\mathcal{M}}_t$) in Panel (B). 90% confidence intervals are constructed using bootstrapped standard errors.

Finally, Figure 5 examines the term structure of expected returns. Because expected returns are autoregressive, the model predicts that the slope coefficient in regressions of future one-year returns r_{t+j} on current information decays geometrically with the horizon j, at rate ρ_μ^{j-1}. Panel A predicts with pd_t, Panel B with $\widehat{\mathcal{M}}_t$. In both, the empirical slopes decline with horizon and track the model-implied decay.

5.2 Out-of-Sample Return Predictability

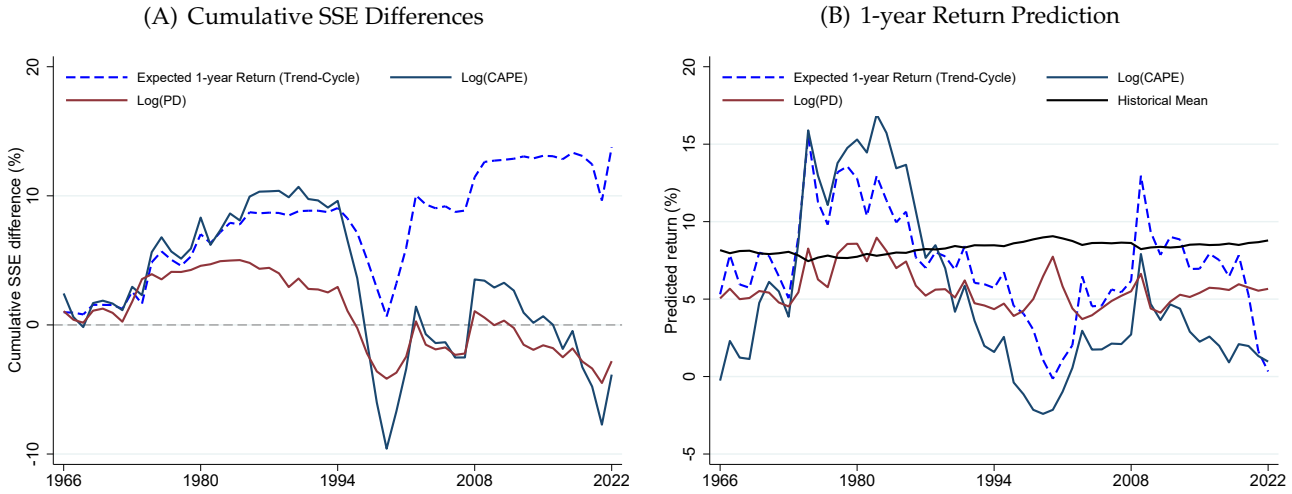
We now ask whether the trend-cycle model predicts returns in real time. We follow Goyal and Welch (2008), re-estimating the model on an expanding window and comparing its forecasts to the expanding historical mean. The test is demanding. Raw valuation ratios that look informative in sample routinely fail it (Goyal, Welch, and Zafirov, 2024).

Our model explains why. The *pd* ratio mixes two components—high-frequency mean reversion in expected returns, which forecasts returns, and persistent shifts in cash flow growth, which do not—

and a regression estimated in real time cannot tell them apart. Campbell (2008) takes the persistent side seriously. He shows that if $\log(D_{t+1}/P_t)$ follows a random-walk, regardless of its source, then $\mathbb{E}_t\, r_{t+1} \approx D_{t+1}/P_t + \mathbb{E}_t\, g_{t+1}$. Campbell and Thompson (2008) implement this idea, proxying $\mathbb{E}_t\, g_{t+1}$ with moving averages of past profitability and payouts. Our paper builds on this work, providing evidence that a persistent component lives in cash flow growth and using a filter to estimate it in real time.

Table 5 reports out-of-sample R^2_{OOS} over two evaluation periods. For the longer post-1945 sample, the model delivers an R^2_{OOS} of 7.29% at one year and 12.28% at five years for raw returns, and 8.64% and 17.14% for excess returns. Turning to Goyal and Welch (2008)'s prominent post-1965 evaluation period, there the model delivers 8.93% at one year and 22.20% at five years for raw returns, and 9.15% and 16.57% for excess returns. By contrast, canonical valuation ratios uniformly deliver negative out-of-sample R^2 (Table 2).

Figure 6: Out-of-Sample Return Predictability



Note: This figure shows the out-of-sample return prediction performance for various predictors. Panel (A) plots the cumulative difference in squared forecast errors between the historical mean benchmark and each predictor, $\sum_t \left[(r_{t+1} - \bar{r}_t)^2 - (r_{t+1} - \hat{r}^{\text{OOS}}_{t+1})^2\right]$, so that positive values indicate the predictor outperforms the benchmark. Panel (B) plots the out-of-sample return predictions, $\hat{r}^{\text{OOS}}_{t+1}$, alongside the expanding historical mean. The predictors are $\log(PD)$, $\log(CAPE)$, and the expected 1-year return from the trend-cycle model estimated using an expanding window. The evaluation sample is annual, spanning 1966–2022. All forecasts are plotted at the year to which the predicted return corresponds.

Figure 6 illustrates the difference in real time. Panel (A) plots the cumulative difference in squared forecast errors between the historical mean and each predictor, where an upward slope indicates outperformance. The trend-cycle model accumulates gains almost monotonically, whereas *pd* and *cape* begin underperforming the historical mean from the early 1980s and mid-1990s, respectively.

Panel (B) reveals the source of the divergence. The trend-cycle expected return is stationary by con-

Table 5: Return Predictability: Out-of-Sample R^2

	Raw Returns		Excess Returns	
	1-yr	5-yr	1-yr	5-yr
1945–2022	7.29%**	12.28%**	8.64%***	17.14%**
	[0.014]	[0.039]	[0.005]	[0.015]
1965–2022	8.93%**	22.20%**	9.15%**	16.57%**
	[0.020]	[0.022]	[0.013]	[0.033]

Note: This table reports out-of-sample R^2_{OOS} for expected returns implied by the trend-cycle model at 1- and 5-year horizons, over the post-1945 and post-1965 evaluation periods. Multi-year returns are discounted at $\rho = 0.96$. The out-of-sample estimates are obtained by re-estimating the trend-cycle model on an expanding window. Realized excess returns subtract cumulated future three-month Treasury bill rates over the forecast horizon. Model-implied excess returns subtract the current annualized three-month bill rate projected over the horizon assuming a flat term structure, discounted at ρ. Stars denote significance of the Clark and West (2007) MSFE-adjusted statistic with Newey–West standard errors ($h - 1$ lags); p-values in brackets. $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

struction: it fluctuates around its long-run mean, spiking during market stress—the 1970s stagflation, the 2008 financial crisis—and compressing during booms, consistent with countercyclical risk premia. By accounting for time variation in g_t, the model isolates this stationary signal even as valuation ratios drift far from historical norms. The *pd* prediction, by contrast, is nearly flat, its slope severely attenuated—a direct manifestation of Proposition 1. As trend growth drifted higher over the post-war period, valuation ratios rose in tandem, and the univariate regression, unable to distinguish higher growth from lower discount rates, systematically underpredicted returns.

5.3 Long-Horizon Cash Flow Predictability

If filtered cash flows indeed reflect cash flows, they should forecast them. The filtered expected-return state, by contrast, should carry no incremental forecasting power. Short-run predictability is well established (Van Binsbergen and Koijen, 2010; Koijen and Van Nieuwerburgh, 2011). The model's distinctive restriction is at low frequency—a small persistent component is hard to see at short horizons but should matter over long ones—so we test Prediction 3 at long horizons in Table 6.

The raw price-dividend ratio forecasts future dividend growth. At ten years the coefficient on pd_t is 0.26, matching the model-implied 0.26 (column 1). At twenty years it is 0.47, above the model-implied 0.37, though the two are statistically indistinguishable (column 4). In other words, dividend growth over the next decade accounts for a quarter of today's price-dividend variation, and growth over two decades for nearly half. Consistent with a low-frequency trend component, the coefficient rises with the horizon.

But pd_t forecasts cash flows more poorly than it could, for two reasons. First, pd_t contains the

Table 6: Testing the Model: Long-Horizon Cash Flow Predictability

	$\Delta d_{t,t+j} = \sum_{i=1}^{j} \rho^{i-1} \Delta d_{t+i}$						$\rho^j pd_{t+j}$	
	10-year			20-year			10-year	20-year
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
pd_t	0.26**			0.47***			0.46***	0.22**
	(0.00)			(0.00)			(0.00)	(0.00)
	[0.04]			[0.01]			[0.00]	[0.03]
$\hat{g}_t$		36.95**	37.75**		73.02**	75.23**		
		(0.00)	(0.00)		(0.00)	(0.00)		
		[0.03]	[0.03]		[0.03]	[0.02]		
$\hat{c}_t$		-0.68*	-0.74		-1.25	-1.33		
		(0.00)	(0.00)		(0.00)	(0.00)		
		[0.10]	[0.11]		[0.10]	[0.12]		
$\hat{\varepsilon}_{c,t}$		-0.13	-0.06		0.70	0.76		
		(0.80)	(0.90)		(0.29)	(0.33)		
		[0.83]	[0.91]		[0.25]	[0.26]		
$\widehat{\mathcal{M}}_t$			-0.19			-0.16		
			(0.22)			(0.40)		
			[0.41]			[0.58]		
Observations	142	142	142	132	132	132	142	132
R^2	0.15	0.29	0.30	0.26	0.46	0.47	0.42	0.17
Adj. R^2	0.15	0.27	0.28	0.26	0.45	0.45	0.41	0.16
Model-implied coefficients:								
pd_t	0.26			0.37			0.53	0.31
	{0.86}			{0.44}			{0.81}	{1.00}
$\hat{g}_t$		33.10	33.10		55.25	55.25		
		{0.31}	{0.35}		{0.15}	{0.17}		
$\hat{c}_t$		-0.59	-0.59		-0.72	-0.72		
		{0.78}	{0.95}		{0.29}	{0.26}		
$\hat{\varepsilon}_{c,t}$		0.21	0.21		0.14	0.14		
		{0.42}	{0.52}		{0.23}	{0.18}		
$\widehat{\mathcal{M}}_t$			0.00			0.00		
			{0.24}			{0.34}		

Note: This table reports predictive regressions of realized dividend growth, $\Delta d_{t,t+j} = \sum_{i=1}^{j} \rho^{i-1} \Delta d_{t+i}$, over 10-year and 20-year horizons. Independent variables include the log price-dividend ratio (pd_t), the filtered cash-flow states ($\hat{\bar{g}}_t, \hat{c}_t, \hat{\varepsilon}_{c,t}$) and the filtered expected-return signal ($\widehat{\mathcal{M}}_t$), derived from full-sample parameter estimates (1871–2022). The top panel reports OLS regression coefficients, p-values based on Newey-West t-statistics in round brackets, and p-values based on a parametric bootstrap procedure (see Appendix C.3). The bottom panel reports the model-implied (analytical) coefficient and the p-value (in curly brackets) for the test $H_0 : \beta = \beta^{\text{analytical}}$ versus $H_1 : \beta \neq \beta^{\text{analytical}}$. Significance levels based on bootstrapped p-values: *(p<0.10), **(p<0.05), ***(p<0.01).

expected-return component, which is noise for cash flows—though the noise is only about a fifth of price-dividend variation, so the univariate slope is attenuated mildly. Second, the cash flow content of pd_t is an infinite-horizon forecast, while the ten- and twenty-year regressions call for horizon-specific weights. Replacing pd_t with the filtered states removes the expected-return noise and lets the trend and cycle enter at their own weights, positive for trend growth and negative for the cycle as dividends revert to trend. The R^2 rises from 15% to 29% at ten years and from 26% to 46% at twenty. The coefficient on trend growth is large because the dependent variable cumulates trend growth quarter after quarter, and for the same reason it roughly doubles with the horizon, from 37 at ten years to 73 at twenty.

The expected-return state does not forecast cash flows. Adding $\widehat{\mathcal{M}}_t$ to the cash-flow-state regression leaves a coefficient that is small and insignificant, against a model-implied value of zero. This is the separation the filter is meant to deliver: $\widehat{\mathcal{M}}_t$ forecasts returns, while $\hat{g}_t$, $\hat{c}_t$, and $\widehat{\varepsilon}_{c,t}$ forecast cash flows. Across all specifications, we cannot reject equality of the OLS and model-implied coefficients.

The post-1926 CRSP sample carries the same message with less precision and smaller magnitudes on trend growth (Appendix Table A.7). This is expected. A small permanent component is harder to identify as the sample shortens.

In short, the filtered states have the economic content the model assigns them. The hidden trend forecasts 10-year and 20-year cash flow growth and, conditional on the cash flow states, expected returns do not.

6 Implications and Origins of the Hidden Trend

6.1 Asset-Pricing Implications

A single force accounts for three features of the price-dividend ratio. Over 1871–2022, trend growth explains 80 percent of its variation, the whole of its secular rise, and its near-unit-root persistence. This section derives these implications and then turns to their origins.

Sample-Length Dependence: Variance Shares For a window of length T, define the contribution of component $X_t \in \{\mathcal{M}_t, \mathcal{C}_t, \overline{pd}_t\}$ to variation in pd_t as

$$b_{X,T} = \frac{\mathrm{Cov}_T(X_t, pd_t)}{\mathrm{Var}_T(pd_t)}. \tag{26}$$

In the baseline model, the trend component $\overline{pd}(g_t)$ is a nonlinear function of a random walk. To approximate its variance, we take a first-order expansion in g_t around a fixed point g_0, yielding (Appendix D.1),

$$\mathrm{Var}_T\big(\overline{pd}(g_t)\big) \approx \frac{\mathrm{Var}_T\big(g_t\big)}{(1-\rho)^2}, \qquad \mathrm{Var}_T\big(g_t\big) = \sigma_g^2\,\frac{T^2-1}{6T}. \tag{27}$$

Because g_t is a random walk, $\mathrm{Var}_T(g_t)$ grows without bound, approximately linearly in T. Longer windows therefore attribute a growing share of pd_t to the trend component. The same unit root means the decomposition exists only at finite windows, and the unconditional variance is undefined. That unit root is our stand-in for a highly persistent stationary process, so we also solve the model under that process directly (Appendix G) with $g_t = \bar{g} + \rho_g\,(g_{t-1} - \bar{g}) + \varepsilon_{g,t}$. We set $\rho_g = 0.999$. Valuing the persistent trend component directly, we obtain:

$$\begin{aligned} pd_t &= \delta_{\rho_g} + F_{\rho_g}(g_t) + \mathcal{M}_t + \mathcal{C}_t, \\ F_{\rho_g}(g_t) &= \log \sum_{j=1}^{\infty} \exp\Big[\, j\,(\bar{g} - \bar{\mu}) + \Lambda_j\,(g_t - \bar{g})\,\Big], \qquad \Lambda_j = \frac{1-\rho_g^{\,j}}{1-\rho_g}. \end{aligned} \tag{28}$$

Ignoring constants, this is eq. (10) with $F_{\rho_g}(g_t)$ in place of $\overline{pd}(g_t)$. Because a high g_t decays toward $\bar{g}$, F_{ρ_g} is finite for every g_t and its unconditional population moments exist, requiring only $\bar{g} < \bar{\mu}$. The stationary model reproduces the finite-window shares and adds the unconditional limit the baseline cannot reach.

Panel A of Table 7 reports the decomposition for both approaches. Throughout, $\overline{pd}$ denotes the trend component of pd_t—$\overline{pd}(g_t)$ in the baseline, $F_{\rho_g}(g_t)$ in the stationary model. The first block computes the model-implied shares in closed form under the baseline approximation; the second simulates the stationary model using the baseline parameters shown in Table 3; the third computes the shares by regressing the filtered states on the price-dividend ratio, averaging across all windows of length T for our 1871-2022 sample.

The decomposition depends on the sample length. Expected returns are volatile—roughly 6 percent a year—but mean-revert within about four years, so their variance is bounded and accounts for most of the variation in short and medium samples. Trend growth barely moves year to year, but its shocks are persistent, so they accumulate. Accumulation is slow. Under the stationary model, where trend growth has a half-life of 173 years, the trend's share of pd variation eventually converges to 73 percent. Over a typical 150-year sample the trend explains 33–35 percent of the variation in pd_t, with expected returns accounting for most of the rest.

Table 7: Asset-Pricing Implications

Panel A: Variance Decomposition of pd_t

	Baseline			Stationary trend growth			Historical filtered path		
Window	$\mathcal{M}$	$\mathcal{C}$	$\overline{pd}$	$\mathcal{M}$	$\mathcal{C}$	$\overline{pd}$	$\mathcal{M}$	$\mathcal{C}$	$\overline{pd}$
T=1	0.87	0.07	0.06	0.86	0.07	0.07	0.76	0.14	0.11
T=10	0.83	0.10	0.07	0.82	0.10	0.08	0.68	0.17	0.15
T=50	0.73	0.11	0.16	0.71	0.11	0.18	0.52	0.11	0.37
T=150	0.57	0.09	0.33	0.56	0.09	0.35	0.19	0.01	0.80
T=∞	—	—	—	0.23	0.04	0.73	—	—	—

Panel B: Persistence of pd_t (annual)

	Baseline			Stationary trend growth			Implied ϕ_{pd}, by shares	
Window	$\phi_{\mathcal{M}}$	$\phi_{\mathcal{C}}$	$\phi_{\overline{pd}}$	$\phi_{\mathcal{M}}$	$\phi_{\mathcal{C}}$	$\phi_{\overline{pd}}$	Model	Historical
T=150	0.82	0.88	0.97	0.82	0.88	0.97	0.88	0.94
T=∞	0.84	0.90	—	0.84	0.90	0.99	0.95	—

Note: Panel A decomposes the variance of pd_t over windows of length T; Panel B reports the model-implied annual persistence of each component, and the persistence of pd_t implied by weighting the components by their variance shares. In the stationary trend growth columns, $\overline{pd}$ denotes $F_{\rho_g}(g_t)$. Both models use the estimated parameters (Table 3). Baseline entries are computed in closed form (Appendix D); stationary trend growth entries are simulated, with the $T = \infty$ row computed by Monte Carlo from the invariant distribution. The historical block applies the same variance-share calculation to the filtered states over 1871–2022, averaging sample moments across all windows of length T before taking ratios.

The last 150 years of stock-market history, however, were not a typical draw. Trend-growth shocks ran predominantly in one direction, lifting trend growth from about 2% a year in the early 1900s to 6% by the 2000s—a cumulative move of roughly two standard deviations relative to an average sample. Because these shocks reinforced one another, the trend's realized share of price-dividend variation far exceeds its 33–35% average.[13] Expected returns still dominate at short and medium windows, but over the full sample the trend is the predominant driver, explaining 80% of the variation against 19% for expected returns.

Persistence The high persistence of valuation ratios follows from combining these shares with the components' own persistences. Define $\phi_{x,T}$ as the annual autocorrelation of x computed from window moments of length T (Appendix D). The model implies the persistence of pd_t is the variance-weighted

[13] Appendix H reports the companion variance decomposition of returns. The return decomposition answers a different question—which news moves returns each period, not which accumulated components move valuations—yet in the model's average history the two deliver a broadly similar split over 150 years, with discount rates just under half and the trend at roughly 40 percent.

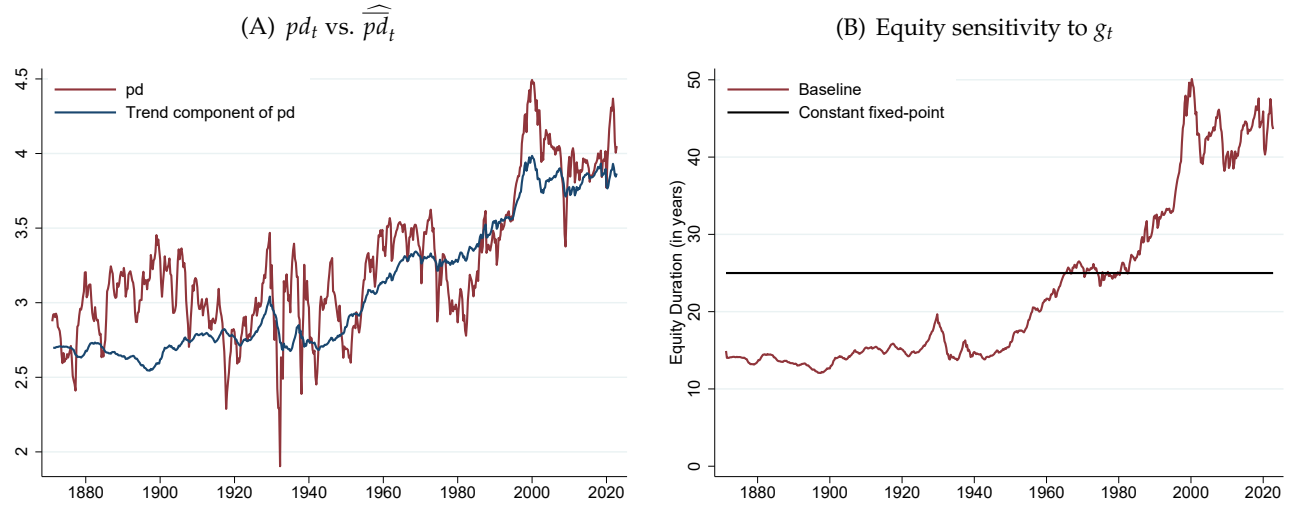
average of its components' persistences,

$$\phi_{pd,T} = \phi_{\widehat{pd},T}\, b_{\widehat{pd},T} + \phi_{\mathcal{M},T}\, b_{\mathcal{M},T} + \phi_{\mathcal{C},T}\, b_{\mathcal{C},T}. \tag{29}$$

At short windows every within-window autocorrelation is attenuated, so Panel B reports it for our 150-year sample and, under stationary trend growth, for the population.[14] At 150 years, expected returns carry an annual persistence of 0.82, the cycle 0.88, the trend 0.97. If expected returns drove all *pd* variation, its AR(1) would be 0.82. The trend pulls it up. At the model's average shares—the trend at one third—implied persistence is 0.88. At our sample's realized shares—the trend at 80 percent—it is 0.94. In the raw data, the comparable window objects are 0.90 for *pd* and 0.91 for *cape*. The bias-adjusted values are 0.93 for both (Table 1). Valuation ratios look like unit roots because, in a long window dominated by the trend, they nearly are.

Secular Increase in Valuations The rise in trend growth also explains the rise in stock-market valuations, as Panel A of Figure 7 shows. The trend component of the price-dividend ratio rose from 2.7 to 3.9 over the sample, accounting for the entire secular increase in pd_t, which rose from 2.9 to 4.0. The cyclical component was roughly flat.

Figure 7: The Secular Increase in the Price-Dividend Ratio



Note: Panel A plots pd_t versus the estimated trend component $\widehat{pd}_t$ under the baseline model. Panel B plots λ_t, the equity sensitivity (duration) to g_t, as defined by eq. (23) in annualized terms.

The increase is amplified by equity duration—the sensitivity of the price-dividend ratio to the growth and discount rate spread—which itself rises with trend growth. To see why, take the Gordon

[14] Appendix Table A.10 reports the shorter windows.

growth model $P/D = 1/(r-g)$. The derivative of the log price with respect to growth is the price-dividend ratio itself, $\partial \log P/\partial g = 1/(r-g) = P/D$. Intuitively, a higher P/D means value sits further in the future, so a growth shock has more periods to compound before the cash flows arrive and lifts the price by proportionally more. In our model this sensitivity is time-varying, characterized by λ_t.[15] Panel B of Figure 7 shows that duration roughly tripled over the sample, from about 15 to 45 years—a one-percentage-point growth shock that raised the log price-dividend ratio by roughly 15% early in the sample and 45% by the end. Duration therefore interacts with the trend shocks themselves—as trend growth rose, it raised the sensitivity of *pd* to the further growth shocks that happened also to be positive, amplifying their effect on valuations. This is exactly what the Campbell–Shiller fixed-point expansion misses. As the constant line in Panel B shows, a fixed duration is too high in the first half of the sample and too low in the second.

Relationship with Literature Our decomposition gives content to Cochrane (2008)'s "dog that did not bark" argument. He contends that the absence of strong short-run cash flow predictability indirectly strengthens the case for discount-rate dominance. But as he acknowledges, the argument loses force if the *pd* ratio contains a small but persistent component, which can explain little *pd* variation at short horizons yet dominate over long ones—precisely the conditions of our model.

The same trend reconciles our results with Cochrane (2011). Regressing 15-year returns and 15-year dividend growth on pd_t over 1947–2009—an effective window of roughly fifty years—he finds returns drive almost all price-dividend variation, and dividend growth almost none. Columns (1) and (5) of Table 8 reproduce that result for this sample.[16]

The key distinction is that these are regressions of *realized* outcomes, not conditional expectations. Over the same window, our filtered 15-year expected returns explain 42% of *pd* variation (Column 2) and 15-year expected dividend growth, as predicted by the filtered cash-flows, explains 31% of *pd* variation (Column 6). Why do the realized regressions produce such different coefficients? The answer lies in the mechanism behind the Stambaugh (1999) bias. In population, future realized forecast errors are orthogonal to time-t information. In short samples with a persistent predictor, however, their sample covariance with pd_t need not be close to zero. Fifteen-year overlapping observations magnify the problem (Boudoukh, Israel, and Richardson, 2022).

[15] Equity duration also operates through a second channel. Beyond raising the sensitivity of pd_t to trend-growth shocks (the λ_t effect above), higher trend growth changes how the *cyclical* components are discounted: the linearization weight $\rho_t = \exp(g_t - \bar{\mu})$ rises with g_t, so cyclical deviations further in the future are priced more heavily, raising the loadings on the expected-return cycle $\mathcal{M}_t$ and the cash-flow cycle $\mathcal{C}_t$. The baseline shuts this channel off by fixing $\rho_t = \rho$ (eq. (8)). Section 7.3 shows the effect is quantitatively small.

[16] Cochrane (2011) uses the dividend yield, $dp_t = -pd_t$, and reports direct fifteen-year coefficients of 1.01 for returns and -0.11 for dividend growth (his Table II). The difference between his and our coefficients is largely due to dividend annualization—his annual dividend reinvests intra-year distributions at the market return, ours at a fixed rate.

Table 8: Reconciliation with Cochrane (2011)

	15-year						
	$r_{t,t+15} = \sum_{i=1}^{15} \rho^{i-1} r_{t+i} = \widehat{E}_t[r_{t,t+15}] + \widehat{\psi}_{t,t+15}$,			$\widehat{\psi}_{t,t+15} = \widehat{N}^{CF}_{t,t+15} - \widehat{N}^{DR}_{t,t+15}$	$\Delta d_{t,t+15} = \widehat{E}_t[\Delta d_{t,t+15}] + \widehat{N}^{\Delta d}_{t,t+15}$		
	$r_{t,t+15}$	$\widehat{E}_t[r_{t,t+15}]$	$\widehat{N}^{CF}_{t,t+15}$	$-\widehat{N}^{DR}_{t,t+15}$	$\Delta d_{t,t+15}$	$\widehat{E}_t[\Delta d_{t,t+15}]$	$\widehat{N}^{\Delta d}_{t,t+15}$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
pd_t	-1.16	-0.42	-0.58	-0.16	-0.07	0.31	-0.38
	(0.23)	(0.14)	(0.06)	(0.08)	(0.07)	(0.07)	(0.06)
Observations	48	48	48	48	48	48	48
R^2	0.61	0.42	0.75	0.16	0.02	0.58	0.56

Note: This table reports OLS regressions of fifteen-year outcomes on pd_t over the sample of Cochrane (2011), namely 1947–2009. Expectations are model-implied from the filtered states, and news is realized minus expected. Columns (1)–(4) decompose the realized discounted fifteen-year return into its conditional expectation and return news, $\psi_{t,t+15} = N^{CF}_{t,t+15} - N^{DR}_{t,t+15}$. Columns (5)–(7) decompose realized discounted dividend growth into its conditional expectation and dividend-growth news. Newey-West standard errors in round brackets.

We decompose realized returns into an expected and unpredictable part, $r_{t,t+15} = E_t[r_{t,t+15}] + \psi_{t,t+15}$. Our model estimates imply that *unforecastable* news accounts for 74% of *pd* variation, with 58% from cash flow news (Column 3) and 16% from discount rate news (Column 4). Why did these unforecastable news line up with valuations? Most of trend growth's rise, from 2 to 6 percent a year, fell between 1947 and 1994—the effective window Cochrane (2011) analyzes. Early on, valuations sat below average and the trend was still rising; the next fifteen years brought positive cash flow news, and realized returns beat expectations. By the 1990s valuations were high and the climb was over; the news stopped, and returns fell short. Low pd_t drew good news, high pd_t none. Figure 8 Panel (A) plots cash flow news against pd_t, where we see a sharp negative relationship that the regression reads as predictable explained variation.

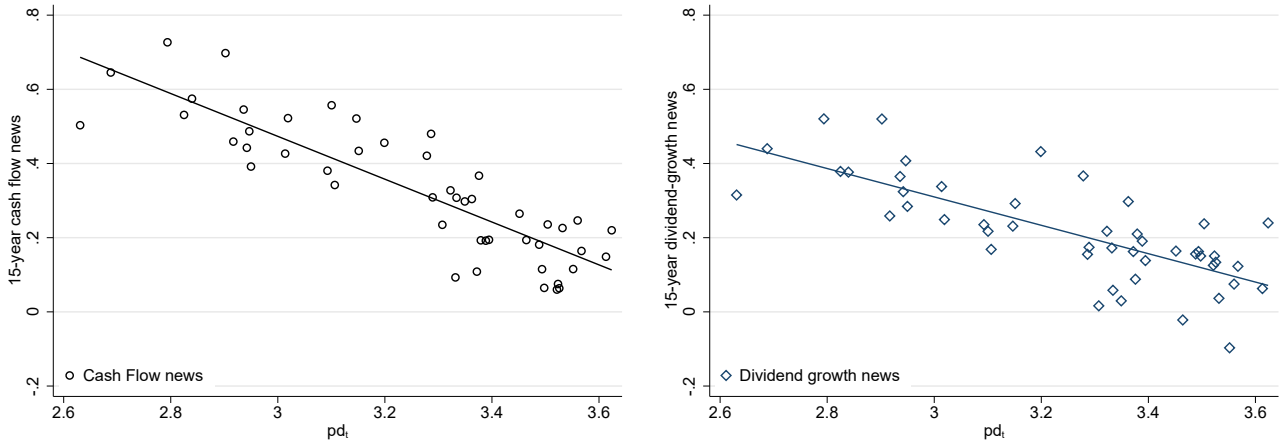
The trend's rise over this window also explains the gap between 15-year expected dividend growth, which explains 31% of variation in *pd*, and 15-year realized dividend growth, which explains none. When *pd* was low in 1947 it forecast low dividend growth, but subsequent dividend growth news was high. When *pd* was high at the end, trend growth shocks were closer to zero, and dividend growth news was low. Low *pd* drew good dividend news, high *pd* none. This generates a negative slope per Figure 8 Panel (B), meaning the realized slope is much lower than the model-implied expected slope.[17]

Over the 150-year sample we have more power, and the sample coefficient should align closer to the population coefficient in expectation. Our model implies 19% of *pd* variation is due to expected-returns and 80% due to trend growth. Consistent with this, Table 4 shows that 10-year returns, which accounts for most of the full return term-structure, explains 23% of the variation in *pd*. This 23% is only

[17] 15-year cash flow news differs from 15-year dividend growth news because cash flow news carries the whole term structure, including dividend growth news plus the revaluation of later dividends capitalized in the terminal price.

Figure 8: Cash Flow and Dividend Growth News over the Cochrane (2011) Sample

(A) Cash flow news, $\widehat{N}^{CF}_{t,t+15}$, vs. pd_t (B) Dividend growth news, $\widehat{N}^{\Delta d}_{t,t+15}$, vs. pd_t



Note: Panel (a) plots fifteen-year cash-flow news $\widehat{N}^{CF}_{t,t+15}$ as estimated by our model against pd_t. Panel (b) plots fifteen-year dividend-growth news $\widehat{N}^{\Delta d}_{t,t+15} = \Delta d_{t,t+15} - \widehat{E}_t[\Delta d_{t,t+15}]$. Expectations are model-implied from the filtered states. The sample is 1947-2009 annual.

slightly above the model-implied value suggesting the Stambaugh (1999) bias is small over the full-sample. Similarly, Table 6 shows that 10-year and 20-year dividend growth explain 26% and 47% of pd variation, which are not significantly different from the model-implied values of roughly a quarter and 40%, respectively. With trend growth, the term-structure is longer because shocks are persistent, and so even 20-year expected growth does not reflect the full variation in pd due to cash flows.

These results reframe the excess-volatility debate. Shiller (1981) argued that de-trended stock prices are too volatile relative to de-trended dividends discounted at a constant rate. Kleidon (1986), Marsh and Merton (1986), and Barsky and De Long (1993) objected that detrending dividends with a deterministic full-sample growth rate removes low-frequency fundamental variation by construction. Our method offers a way of estimating the trend in real time, reconciling the two views. Over short and medium sample windows, the majority of variation in pd_t is associated with expected returns, consistent with Shiller (1981)'s puzzle. In the long-run, however, trend growth compounds, eventually accounting for the majority of valuation ratio movements.

6.2 Economic Origins of the Trend

The previous sections show that stock-market cash flows contain a highly persistent trend growth component. We now give suggestive evidence on its economic origins by decomposing the low-frequency movements in realized dividend-per-share growth.

Let D_t denote aggregate nominal dividends paid by firms in the S&P 500 index, Y^N_t nominal GDP, P_t

the GDP deflator, and S_t an effective share count. Define $\phi_t \equiv D_t / Y_t^N$ as the dividend share of GDP and $Y_t^R \equiv Y_t^N / P_t$ as real GDP. We construct dividends per share on the S&P 500 index as D_t / S_t from CRSP, which matches Shiller's published series up to a residual, $\log D_t^{Sh} = \log(D_t / S_t) + \omega_t$. Shiller dividends per index share, D_t^{Sh}, can then be written as

$$\log D_t^{Sh} = \underbrace{\log\left(\frac{D_t}{Y_t^N}\right)}_{\text{dividend share of GDP}} + \underbrace{\log\left(\frac{Y_t^N}{P_t}\right)}_{\text{real GDP}} + \underbrace{\log P_t}_{\text{price level}} - \underbrace{\log S_t}_{\text{share count}} + \omega_t, \tag{30}$$

where ω_t collects index reconstitution, imperfect historical coverage, and remaining differences between Shiller's series and our CRSP construction (Appendix A.2). Taking annual log differences and averaging over a 30-year trailing window gives

$$\overline{\Delta \log D^{Sh}}_{t,30} = \overline{\Delta \log \phi}_{t,30} + \overline{\Delta \log Y^R}_{t,30} + \overline{\pi}_{t,30} - \overline{\Delta \log S}_{t,30} + \overline{\Delta \omega}_{t,30}. \tag{31}$$

We label the terms growth in the dividend share of GDP, real GDP growth, inflation and net dilution. Net dilution is positive when repurchases shrink the effective share count and negative when issuance dilutes continuing shareholders (Bernstein and Arnott, 2003; Boudoukh, Michaely, Richardson, and Roberts, 2007). All quantities in Figure 9 and Table 9 are in percent per year. Because aggregate dividends and the share count require CRSP data from year-end 1926, the first complete 30-year trailing window for those components starts in 1956.

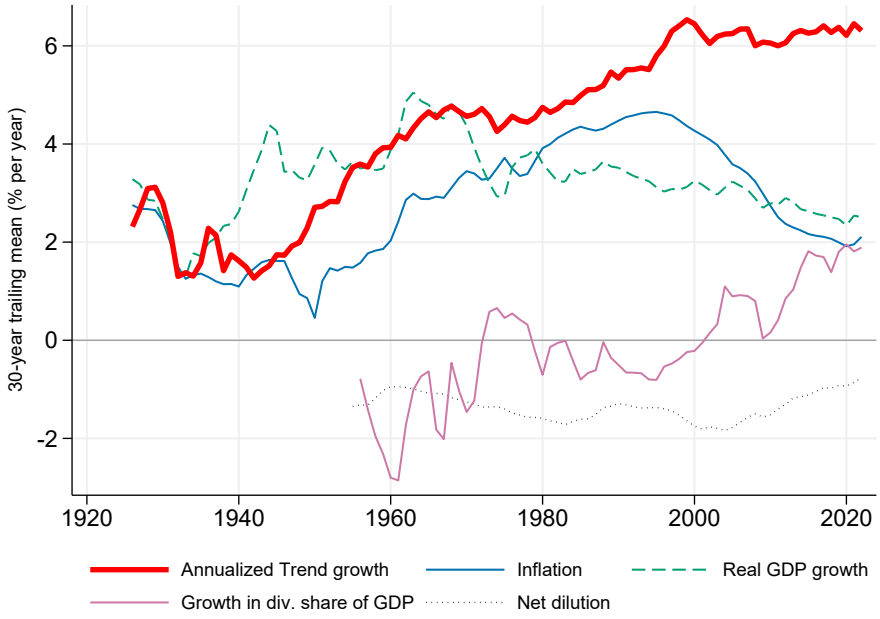
Table 9: Dividend-Per-Share Growth and Its Components

	Sample average over			
	1926–1958	1959–1990	1991–2022	Δ
Growth in div. share of GDP	-0.9	-0.5	1.7	2.6
Real GDP growth	3.4	3.6	2.5	-0.9
Inflation	1.6	4.3	2.2	0.6
Net dilution	-1.5	-1.3	-0.8	0.7
ω	0.3	-0.1	-0.1	-0.4
Dividend per share growth	2.9	6.0	5.5	2.6
Dividend per share growth (Real)	1.3	1.7	3.3	2.0
Trend growth	2.3	4.6	6.1	3.8

Note: This table reports sample averages of the components of realized Shiller dividend-per-share growth alongside trend growth. The five components sum to dividend-per-share growth through the differenced identity in equation (31). The trend-growth row reports the annualized filtered state $4\hat{g}_t$. The Δ column gives the change from the first window to the last. See Appendix A.2 for construction.

Figure 9 shows the raw patterns. Table 9 splits the post-CRSP period into three roughly equal sub-

Figure 9: Trend Growth and Components of Dividend-Per-Share Growth



Note: This figure plots annualized filtered trend dividend growth, $4\hat{g}_t$, together with 30-year trailing averages of the components of realized Shiller dividend-per-share growth, in percent per year, following equation (31). The inflation and real GDP series extend to a window ending in 1926, using macroeconomic data from the 1890s onward. The dividend-share and net-dilution series require CRSP data and so begin with the first complete window, ending in 1956. See Appendix A.2 for construction.

samples. Between 1926–1958 and 1959–1990, average trend growth rose by 2.3 points, and realized dividend-per-share growth rose with it, by 3.1. The decomposition splits that rise three ways. The real per-share margins—the dividend share of GDP, net dilution, and real GDP growth—contributed 0.8 points, inflation 2.7, and an accounting residual the rest. The inflation piece is consistent with the Great Inflation working its way into the persistent growth of nominal cash flows.

Between 1959–1990 and 1991–2022, average trend growth rises a further 1.5 points, to 6.1 percent. Yet nominal dividend-per-share growth falls by 0.5 points. The decomposition shows what moved. Real dividend-per-share growth rises by 1.6 points—a 2.2 point increase in the growth of the dividend share of GDP and a 0.5 point easing of the dilution drag, partly offset by a 1.1 point decline in real GDP growth. A 2.1 point fall in inflation more than offsets the real gain in nominal terms. A natural interpretation is that the real components of per-share growth are more persistent than inflation and therefore matter more for trend growth.

Over the full 1926–2022 period the message is plain. Inflation drives the mid-century acceleration but over the full-sample, the rise in dividend-per-share growth is overwhelmingly a real phenomenon. Nominal growth rises by 2.6 points between the first and last sub-samples, and 2.0 of that is real. The

swing in the growth of the dividend share of GDP—from -0.9 to $+1.7$ points per year, a move of 2.6 points—accounts for the entire rise.

The unit root in trend growth is thus best read as a parsimonious proxy for the persistent forces underlying dividend-per-share growth—factor shares, payout and tax policy, investment, and reallocation across firms.[18]

Our per-share evidence connects to recent corporate-sector evidence, though the objects are distinct. Greenwald, Lettau, and Ludvigson (2025) decompose the level of real corporate equity, while Atkeson, Heathcote, and Perri (2025) construct enterprise value and free cash flow for the U.S. corporate sector from the national accounts. These papers point to low-frequency movements in the cash flows accruing to owners—changes in factor shares, net investment, and taxes. We instead decompose the price-dividend ratio, and nominal dividends per share, from the perspective of a continuing shareholder in the S&P 500 index—the aggregate long studied in the return-predictability literature.

The 1959–1990 period illustrates how the approaches relate. Greenwald, Lettau, and Ludvigson (2025) attribute the post-war rise in real equity to economic growth and our accounting agrees that real GDP growth is the largest single source of real dividend-per-share growth in levels. Yet it is unlikely to explain why trend growth accelerated: average real GDP growth rose only about 0.2 percentage points between 1926–1958 and 1959–1990. Our decomposition instead assigns much of the acceleration to the dividend share of GDP, especially in recent decades, consistent with the factor-share channel in Greenwald, Lettau, and Ludvigson (2025), who assign 40% of the 1989–2017 rise in equity to reallocation toward shareholders. The same force appears in the broad corporate sector, where Atkeson, Heathcote, and Perri (2025) document a secular rise in free cash flow. Atkeson, Heathcote, and Perri (2026) push the point to valuation ratios: the falling labor share and weak net investment lift them mechanically, and the ratios should not be expected to return to their historical means—the accounting counterpart of our finding that the price-dividend ratio moved with trend growth.

Our exercise relates to Schorfheide, Song, and Yaron (2018), who recover a persistent component in real consumption growth—on which dividends load—from mixed-frequency cash flow data alone, and sharpen it by imposing the equilibrium pricing restrictions of an endowment economy with recursive preferences. The two exercises recover different objects. In their framework, dividends inherit persistent growth from the aggregate component; in ours, dividends can move when aggregate growth does not—in fact, we find that much of its low-frequency movement traces to the dividend share of GDP.[19]

[18] See, among others, Gutiérrez and Philippon (2017); Farhi and Gourio (2018); Eggertsson, Robbins, and Wold (2021); Kuvshinov and Zimmermann (2022); Crouzet and Eberly (2023).

[19] In related work, Liu and Matthies (2022) recover a persistent component of aggregate consumption growth from a clever third source—news coverage of economic prospects.

7 Extensions

This section confronts the baseline with four concerns. First, the secular rise in valuations may reflect downward drifting expected returns rather than rising trend growth. Second, the shift toward repurchases may distort dividends as a cash flow measure. Third, the results may ride on two conveniences of the baseline—the unit root for trend growth and the fixed linearization weight ρ. Fourth, the real-time forecasting gains may be specific to the baseline rather than a property of every specification that allows for trend growth.

7.1 Expected-Return Drift and No Trend-Growth Counterfactual

7.1.1 Allowing Expected Returns to Drift

A natural concern is that the secular rise in valuations reflects a fall in expected returns rather than a rise in trend growth. By holding the steady-state expected return fixed, the baseline may load onto trend growth what properly belongs to a declining discount rate. The cleanest way to confront this is to give the discount-rate explanation a free hand. We let the steady-state expected return drift as an unrestricted random walk $\bar{\mu}_t$, with expected returns mean-reverting toward it, and we leave its innovations independent of those to trend growth. Nothing ties the long-run discount rate to growth. Estimation details are in Appendix J.

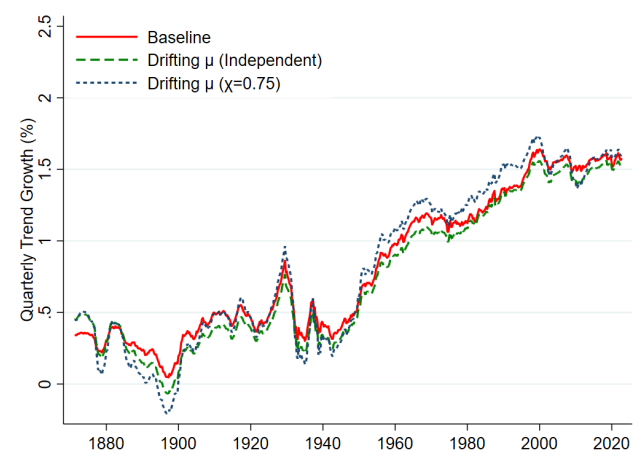
Two findings stand out. First, filtered trend growth barely moves from the baseline, lying slightly below it before 1980 and converging thereafter (Figure 10, Panel A). Second, the long-run expected return rises over the sample, from roughly 7% per year in the late 1800s to 8–8.5% by the 1980s, closely tracking the expanding-window mean of realized returns (Panel B). Given free rein, the discount rate climbs rather than falls—the opposite of what rising valuations would imply. The rearranged Gordon identity, $r = g + D/P$, gives the intuition. Trend growth rose by enough that expected returns and valuations could rise together. Model-implied one-year expected returns are correspondingly close across the two specifications (Panel C).[20]

Independence serves the expected-return explanation well, since it lets the long-run rate move however the data prefer. But theory suggests the two long-run components are positively linked, so that a rise in trend growth should pull the long-run discount rate up with it. In consumption-based models,

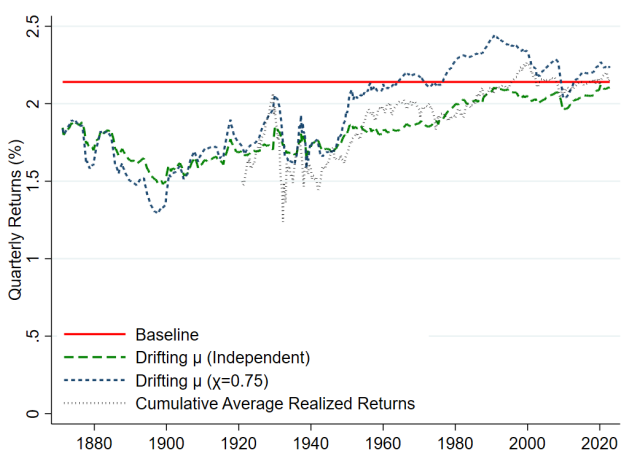
[20] The steady-state valuation gap $pd_t - \overline{pd}_t$ decomposes into an expected-return cycle, $-\frac{1}{1-\rho\rho_\mu}(\mu_t - \bar{\mu})$, and a cash-flow cycle (Appendix J). In the drifting specification, a lower $\bar{\mu}_t$ pre-1980 raises $\overline{pd}_t$ and closes the gap directly; in the baseline, the constant higher $\bar{\mu}$ forces a small negative expected-return cycle to absorb the residual. One-year expected returns are therefore slightly lower in the baseline early in the sample, with the difference diminishing as the cycle attenuates.

Figure 10: Drifting Steady-State Expected Returns

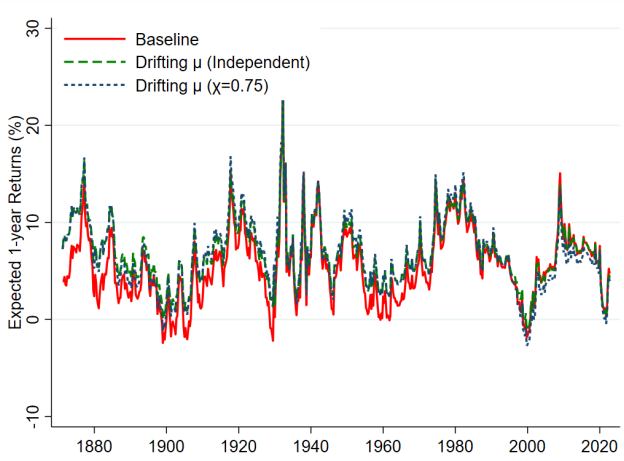
(A) Quarterly Trend Growth



(B) Quarterly Steady-State Expected Returns



(C) Expected 1-year Returns



Note: This figure compares the baseline model, in which the long-run mean of expected returns $\bar{\mu}$ is constant, to the extended models in which $\bar{\mu}_t$ follows a random walk where its shocks are independent of trend growth (Appendix J) and where they are correlated at $\chi = 0.75$ (Appendix K). Panel (A) plots filtered quarterly trend dividend growth g_t under each specification. Panel (B) plots the filtered long-run expected return $\bar{\mu}$ (baseline) and $\bar{\mu}_t$ (extended) alongside the expanding-window average quarterly realized returns. Panel (C) plots model-implied expected one-year returns under both specifications.

persistent shifts in long-run growth move the long-run discount rate directly (Bansal and Yaron, 2004).[21] Because our model is nominal, persistent inflation supplies a second channel, raising both nominal cash flow growth and nominal expected returns (Boudoukh and Richardson, 1993; Bauer and Rudebusch, 2020). The strength of the pass-through is debated. Gormsen and Lazarus (2026) find that only the pure discount-rate component of secular real-rate movements transmits strongly to valuations, while sticky-discount-rate and money-illusion accounts imply partial pass-through (Katz, Lustig, and Nielsen, 2017;

[21] See also Lucas Jr (1978); Breeden (1979); Hansen and Singleton (1982); Mehra and Prescott (1985); Weil (1989); Laubach and Williams (2003); Hansen, Heaton, and Li (2008). A related literature studies the discount rate and nominal-growth link at business-cycle frequencies (Fama and Schwert, 1977; Campbell, Pflueger, and Viceira, 2020; Cieslak and Pflueger, 2023).

Modigliani and Cohn, 1979).

We therefore re-estimate the model allowing the two permanent innovations to be correlated,

$$\varepsilon_{g,t} = \sigma_g \, \epsilon_{g,t}, \qquad \varepsilon_{\bar{\mu},t} = \sigma_{\bar{\mu}} \big(\chi \, \epsilon_{g,t} + \sqrt{1 - \chi^2} \, \epsilon_{v,t} \big), \tag{32}$$

nesting the independent specification at $\chi = 0$. Valuations primarily identify innovations to the long-run spread, $\bar{\mu}_t - g_t$, rather than the two permanent components separately. We therefore fix χ and re-estimate the remaining parameters, focusing on the positive-correlation nodes $\chi \in \{0.25, 0.50, 0.75\}$.[22]

The correlated specifications carry two messages. First, the valuation-relevant dynamics are robust. Valuations discipline one object above all—the volatility of the long-run spread shock, $\sigma(\bar{\mu}_t - g_t)$, and it barely moves across the grid, holding between 2.1 and 2.4 bps per quarter (Table A.4). The profile likelihood tilts mildly toward positive correlation—rising 2.8 log-likelihood units from $\chi = 0.25$ to $\chi = 0.75$—but filtered trend growth follows the independent path closely (Figure 10).

Second, where the specifications part ways, they strengthen the role of trend growth. When a growth shock also raises the discount rate, the two effects on valuations partly offset. To trace the same rise in valuations and dividends, the model raises the volatility of trend growth: σ_g climbs from 2.32 bps per quarter at $\chi = 0.25$ to 3.12 bps at $\chi = 0.75$. In short, positive correlation means trend growth rises even more sharply alongside a comoving long-run discount rate, leaving the role of trend growth at least as large as before.

7.1.2 A Sharper Test: Shutting Down Trend Growth

We now ask the sharper question: can drifting expected returns substitute for trend growth entirely? We re-estimate the drifting-$\bar{\mu}$ model with trend growth shut off, replacing the random-walk component with a constant growth rate $\bar{g}$. Estimation details are in Appendix L.

Appendix Figure A.2 traces why the substitution fails. Trend growth is pinned at a constant 3.1% per year—above the pre-war average of about 2% and well below the post-2000s 6% (Panel A)—so the deterministic trend overshoots dividends early in the sample and undershoots them late (Panel B). The model must absorb these long-lasting deviations somewhere. It pushes the cash flow cycle toward a unit root ($\rho_c = 0.988$), generating a persistently negative cycle for the first century and a persistently positive one in recent decades (Panel C). But a near-unit-root cycle innovation shifts the level of dividends, not their growth rate, and so moves prices too little. The model closes the remaining gap in the *pd* ratio through expected returns, pushing ρ_μ to 0.995 and leaving the random-walk channel essentially unused

[22] At the perfectly correlated boundary $\chi = 1$, the two permanent innovations are collinear. The likelihood is nearly flat along directions that raise both permanent volatilities while preserving the spread volatility, making the boundary poorly identified.

($\sigma_{\bar{\mu}} \approx 0$). The result, in Panel D, is persistently low expected returns through the post-war period that systematically underpredict realized returns.

The likelihood records the strain. Imposing $\sigma_g = 0$ on the drifting expected-return model yields a likelihood-ratio statistic of 21.8, well above conventional critical values for the boundary test. Even with steady-state expected returns free to drift, the data reject shutting down trend growth.

7.2 Trend-Cycle Model for Earnings

The stock price is the present value of all future per-share dividends, making dividends per share the relevant cash flow for a buy-and-hold investor. But the secular shift toward repurchases (Grullon and Michaely, 2002; Skinner, 2008) raises the concern that dividends understate the market's cash-flow-generating capacity, biasing the trend-growth estimate. A shift toward repurchases has two offsetting effects in the dividend model. It depresses realized dividend growth, which the model may read as lower trend growth, and it mechanically raises the price-dividend ratio, which the model may read as higher trend growth. The net effect is ambiguous.

We re-estimate the model using earnings per share, which are unaffected by payout policy, at the cost of being a noisier measure of the cash flow ultimately relevant to investors. Using $d_t = e_t + de_t$ with $de_t \equiv d_t - e_t$, and assuming a constant steady-state log payout ratio $\overline{de}$,

$$de_{t+j} = \overline{de} + v_{t+j}, \qquad v_{t+j} \overset{i.i.d.}{\sim} (0, \sigma_v^2), \tag{33}$$

earnings and dividends share the same non-stationary trend, and the present-value identity for the price–earnings ratio takes the same form as for the price–dividend ratio (Appendix I):

$$pe_t - \overline{pe}_t \approx \delta + \sum_{j=1}^{\infty} \rho^{j-1}\, \mathbb{E}_t\left[\Delta e_{t+j} - g_t\right] - \sum_{j=1}^{\infty} \rho^{j-1}\, \mathbb{E}_t\left[r_{t+j-1,t+j} - \bar{\mu}\right], \tag{34}$$

where $\overline{pe}_t \equiv \overline{pd}_t + \overline{de}$. The payout ratio enters only through $\overline{de}$, which shifts the steady-state valuation level. We fix $\bar{\mu}$ at the baseline estimate and $\overline{de}$ at the full-sample mean of $d_t - e_t$.

The constant-payout assumption is deliberate: it biases the test against finding rising trend growth. Panel (A) of Appendix Figure A.3 shows that the trailing 10-year average of de_t has drifted downward, lying above $\overline{de}$ in the first half of the sample and below it in the second. Holding $\overline{de}$ constant therefore overstates $\overline{pe}_t$ post-repurchases, and with $\bar{\mu}$ fixed the filter can only reconcile this by lowering estimated trend growth. If anything, the earnings specification understates trend growth in the repurchase era.

Panel C shows that filtered trend growth from the earnings specification is nearly identical to the dividend baseline, despite a payout assumption built to push it lower. Model-implied expected one-year

returns are also nearly identical (Panel D).

7.3 Very Persistent but Stationary Trend Growth

The baseline makes two simplifications. It treats trend growth as a unit root. And it fixes the linearization weight ρ at its sample value, though ρ_t is itself a duration weight that rises with trend growth. As it rises, cyclical deviations far in the future are priced more heavily. We relax the first assumption, then both, re-estimating all parameters under each specification—in Section 6, we evaluated the stationary model at the baseline parameter estimates.

The first specification is the stationary model of equation (28). Trend growth mean-reverts at $\rho_g = 0.999$, the trend component $F_{\rho_g}(g_t)$ replaces $\overline{pd}(g_t)$, and the cycle terms keep their constant-ρ loadings. The second specification relaxes the fixed ρ as well. Expanding the stationary cyclical part around its mean, gives (see Appendix G):

$$pd_t \approx \delta_{\rho_g} + F_{\rho_g}(g_t) + \mathcal{M}_t^F + \mathcal{C}_t^F, \tag{35}$$

$$\mathcal{M}_t^F = -\Big(\sum_{j\geq 1} \omega_j(g_t)\,\Lambda_j^{\mu}\Big)(\mu_t - \bar{\mu}), \qquad \omega_j(g_t) = \frac{\exp(a_j)}{\sum_{k\geq 1}\exp(a_k)}, \qquad \Lambda_j^{\mu} = \frac{1-\rho_\mu^j}{1-\rho_\mu}, \tag{36}$$

with $\mathcal{C}_t^F$ built the same way from the cash-flow profiles. Equity duration now also flows through the cyclical deviations of the valuation equation—higher trend growth tilts the strip shares, $\omega_j(g_t)$, toward long maturities, raising the cycle loadings. This is the channel Section 3 shut down by fixing ρ_t. As $\rho_g \to 1$, the shares become geometric at the state-dependent rate $\rho_t = \exp(g_t - \bar{\mu})$—the dynamic coefficient of Section 3.

Appendix Figure A.4 compares the three sets of filtered states. Trend growth is nearly indistinguishable from the baseline under both stationary specifications, correlating at 0.999 under constant ρ and 0.995 under dynamic ρ_t (Panel A). Expected returns are very similar under constant ρ, correlating at 0.98, and differ slightly under dynamic ρ_t, at 0.95. The unit-root, constant-ρ baseline is a close and simple proxy for the stationary solutions, and we adopt it.

7.4 Model Extensions: Return Predictability

The previous subsections show that, in every case that allows for trend growth, the model demands it and the filtered states stay close to the baseline. A harder test is out-of-sample return predictability—the test the canonical valuation ratios fail, beaten by the historical mean in real time. We ask whether our predictability survives it across specifications, again over the post-1965 evaluation period.

Appendix Table A.11 reports out-of-sample R^2 in Panel (B) and the in-sample counterparts in Panel (A).

The four specifications that let trend growth move—the baseline, the drifting-$\bar{\mu}$, earnings, and stationary trend growth models—differ in how they treat discount rates, payout policy, and the persistence of g_t, yet every one delivers positive and significant R^2_{OOS} at both horizons, for raw and excess returns.[23] The baseline forecasts 8.93% of one-year and 22.20% of five-year raw returns, and the other three land in the same range. In every case the Clark and West (2007) test rejects equal predictive accuracy against the historical mean.

The one specification that fails is the one that holds trend growth constant. It gives the discount rate a free hand—$\bar{\mu}_t$ drifts as an unrestricted random walk—but shuts the cash flow trend down. Its in-sample R^2 is already sharply negative, so we do not report the out-of-sample R^2, where it fares worse still. In short, once trend growth is allowed, the model's real-time forecasting power is robust.

8 Conclusion

A three-component process consisting of a persistent growth trend, a transitory cash flow cycle and a time-varying discount rate is a powerful description of aggregate stock-price variation for the 1871-2022 period. It accounts for the weak in-sample predictability of valuation ratios, their failure to forecast returns out of sample, their near-unit-root persistence, and their secular increase. The results suggest a reframing of the Shiller (1981) excess-volatility puzzle: over short and medium windows, expected returns dominate price-dividend variation; over the last 150 years, trend growth dominates.

Filtering out the hidden trend yields a 150-year trend-cycle expected-return series that speaks to what Cochrane (2011) identifies as the central question of asset pricing: how discount rates vary over time. The full-sample series provides a historical path of the equity discount rate, suited to measurement: assessing time-varying discount-rate theories, attributing historical performance, and evaluating long-horizon accounting estimates. Its real-time counterpart, estimated from 1945 onwards, forecasts returns out of sample where valuation ratios fail. If that performance reflects more than chance—and there are good theoretical reasons to think it might—the measure also offers a usable input for portfolio allocation, market timing, and the hurdle rates of capital budgeting.

The approach extends in several directions. Equity markets are not unique: bond yields inherit slow-moving trends from expected inflation and the equilibrium real rate, and conditioning on those trends restores bond-return predictability and resolves the excess persistence of yields (Kozicki and Tinsley, 2001; Cieslak and Povala, 2015; Bauer and Rudebusch, 2020). Currencies are a natural next step: exchange rates approach a random walk precisely when fundamentals contain a near-permanent

[23] The earnings-based specification fixes the log payout ratio $\overline{de}$ at the expanding-window mean of $d_t - e_t$ and estimates $\bar{\mu}$ freely by MLE.

component (Engel and West, 2005). Whether trend growth was equally dominant in other countries or earlier eras is likewise open; the modern U.S. environment of high persistence and high equity duration is where it should matter most (Golez and Koudijs, 2018). Finally, at lower levels of aggregation the case is stronger still. Corporate cash flows inherit low-frequency variation not only from economic growth but from factor shares, market power, and other corporate-specific forces, and Vuolteenaho (2002) shows that cash flow news dominates firm-level returns. Firm and industry valuation ratios therefore likely confound persistent growth and expected returns more severely than the aggregate market, so filtering there could sharpen expected-return measurement where it is weakest.

References

Amihud, Yakov, and Clifford M Hurvich, 2004. Predictive regressions: A reduced-bias estimation method. Journal of Financial and Quantitative Analysis 39, 813–841.

Ang, Andrew, and Geert Bekaert, 2007. Stock return predictability: Is it there?. The Review of Financial Studies 20, 651–707.

Atkeson, Andrew, Jonathan Heathcote, and Fabrizio Perri, 2025. Reconciling macroeconomics and finance for the U.S. corporate sector: 1929 to present. NBER Working Paper 33459.

——— , 2026. A macroeconomic perspective on stock market valuation ratios. NBER Working Paper 34748.

Autor, David, David Dorn, Lawrence F Katz, Christina Patterson, and John Van Reenen, 2020. The fall of the labor share and the rise of superstar firms. The Quarterly journal of economics 135, 645–709.

Bai, Jushan, and Pierre Perron, 1998. Estimating and testing linear models with multiple structural changes. Econometrica pp. 47–78.

Bansal, Ravi, Dana Kiku, and Amir Yaron, 2012. An empirical evaluation of the long-run risks model for asset prices. Critical Finance Review 1, 183–221.

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

Barberis, Nicholas, Robin Greenwood, Lawrence Jin, and Andrei Shleifer, 2015. X-capm: An extrapolative capital asset pricing model. Journal of financial economics 115, 1–24.

Barkai, Simcha, 2020. Declining labor and capital shares. The Journal of Finance 75, 2421–2463.

Barro, Robert J, 2006. Rare disasters and asset markets in the twentieth century. The Quarterly Journal of Economics 121, 823–866.

Barsky, Robert B., and J. Bradford De Long, 1993. Why does the stock market fluctuate?. The Quarterly Journal of Economics 108, 291–311.

Bauer, Michael D, and Glenn D Rudebusch, 2020. Interest rates under falling stars. American Economic Review 110, 1316–1354.

Bernstein, William J, and Robert D Arnott, 2003. Earnings growth: The two percent dilution. Financial Analysts Journal 59, 47–55.

Beveridge, Stephen, and Charles R Nelson, 1981. A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the 'business cycle'. Journal of Monetary economics 7, 151–174.

Bordalo, Pedro, Nicola Gennaioli, Rafael La Porta, and Andrei Shleifer, 2024. Belief overreaction and stock market puzzles. Journal of Political Economy 132, 1450–1484.

Boudoukh, Jacob, Ronen Israel, and Matthew Richardson, 2022. Biases in long-horizon predictive regressions. Journal of financial economics 145, 937–969.

Boudoukh, Jacob, Roni Michaely, Matthew Richardson, and Michael R Roberts, 2007. On the importance of measuring payout yield: Implications for empirical asset pricing. The Journal of Finance 62, 877–915.

Boudoukh, Jacob, and Matthew Richardson, 1993. Stock returns and inflation: A long-horizon perspective. The American economic review 83, 1346–1355.

Brav, Alon, John R Graham, Campbell R Harvey, and Roni Michaely, 2005. Payout policy in the 21st century. Journal of financial economics 77, 483–527.

Breeden, Douglas T, 1979. An intertemporal asset pricing model with stochastic consumption and investment opportunities. Journal of financial Economics 7, 265–296.

Campbell, John Y, 2008. Estimating the equity premium. Canadian Journal of Economics 41, 1–21.

——— , 2017. Financial decisions and markets: a course in asset pricing Princeton University Press.

——— , and John H Cochrane, 1999. By force of habit: A consumption-based explanation of aggregate stock market behavior. Journal of political Economy 107, 205–251.

Campbell, John Y, and Pierre Perron, 1991. Pitfalls and opportunities: what macroeconomists should know about unit roots. NBER macroeconomics annual 6, 141–201.

Campbell, John Y, Carolin Pflueger, and Luis M Viceira, 2020. Macroeconomic drivers of bond and equity risks. Journal of Political Economy 128, 3148–3185.

Campbell, John Y, and Robert J Shiller, 1988. The dividend-price ratio and expectations of future dividends and discount factors. The 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?. The Review of Financial Studies 21, 1509–1531.

Cieslak, Anna, and Carolin Pflueger, 2023. Inflation and asset returns. Annual Review of Financial Economics 15, 433–448.

Cieslak, Anna, and Pavol Povala, 2015. Expected returns in treasury bonds. The Review of Financial Studies 28, 2859–2901.

Clark, Peter K., 1987. The cyclical component of U.S. economic activity. Quarterly Journal of Economics 102, 797–814.

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. The Review of Financial Studies 21, 1533–1575.

——— , 2011. Presidential address: Discount rates. The Journal of finance 66, 1047–1108.

Croce, Mariano M, Martin Lettau, and Sydney C Ludvigson, 2015. Investor information, long-run risk, and the term structure of equity. The Review of Financial Studies 28, 706–742.

Crouzet, Nicolas, and Janice Eberly, 2023. Rents and intangible capital: A q+ framework. The Journal of Finance 78, 1873–1916.

De Loecker, Jan, Jan Eeckhout, and Gabriel Unger, 2020. The rise of market power and the macroeconomic implications. The Quarterly journal of economics 135, 561–644.

Durbin, James, and Siem Jan Koopman, 2001. Time Series Analysis by State Space Methods Oxford University Press: Oxford.

Eggertsson, Gauti B, Jacob A Robbins, and Ella Getz Wold, 2021. Kaldor and piketty's facts: The rise of monopoly power in the united states. Journal of Monetary Economics 124, S19–S38.

Engel, Charles, and Kenneth D West, 2005. Exchange rates and fundamentals. Journal of political Economy 113, 485–517.

Fama, Eugene F, and Kenneth R French, 2002. The equity premium. The journal of finance 57, 637–659.

Fama, Eugene F, and G William Schwert, 1977. Asset returns and inflation. Journal of financial economics 5, 115–146.

Farhi, Emmanuel, and François Gourio, 2018. Accounting for macro-finance trends: Market power, intangibles, and risk premia. Brookings Papers on Economic Activity.

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

Golez, Benjamin, and Peter Koudijs, 2018. Four centuries of return predictability. Journal of Financial Economics 127, 248–263.

Gormsen, Niels Joachim, and Eben Lazarus, 2026. Interest rates and equity valuations. .

Goyal, Amit, and Ivo Welch, 2008. A comprehensive look at the empirical performance of equity premium prediction. The Review of Financial Studies 21, 1455–1508.

——— , and Athanasse Zafirov, 2024. A comprehensive 2022 look at the empirical performance of equity premium prediction. The Review of Financial Studies 37, 3490–3557.

Greenwald, Daniel L, Martin Lettau, and Sydney C Ludvigson, 2025. How the wealth was won: Factor shares as market fundamentals. Journal of Political Economy 133, 1083–1132.

Greenwood, Robin, and Andrei Shleifer, 2014. Expectations of returns and expected returns. The Review of Financial Studies 27, 714–746.

Grullon, Gustavo, and Roni Michaely, 2002. Dividends, share repurchases, and the substitution hypothesis. the Journal of Finance 57, 1649–1684.

Gutiérrez, Germán, and Thomas Philippon, 2017. Investmentless growth: An empirical investigation. Brookings Papers on Economic Activity 48, 89–169.

Hansen, Lars Peter, John C Heaton, and Nan Li, 2008. Consumption strikes back? measuring long-run risk. Journal of Political economy 116, 260–302.

Hansen, Lars Peter, and Kenneth J Singleton, 1982. Generalized instrumental variables estimation of nonlinear rational expectations models. Econometrica: Journal of the Econometric Society pp. 1269–1286.

Harvey, Andrew C., 1985. Trends and cycles in macroeconomic time series. Journal of Business & Economic Statistics 3, 216–227.

——— , 1989. Forecasting, Structural Time Series Models and the Kalman Filter Cambridge University Press: Cambridge.

He, Zhiguo, and Arvind Krishnamurthy, 2013. Intermediary asset pricing. American Economic Review 103, 732–770.

Hillenbrand, Sebastian, and Odhrain McCarthy, 2024. Street earnings: Implications for asset pricing. .

Johnston, Louis D., and Samuel H. Williamson, 2023. What was the u.s. gdp then?. MeasuringWorth.

Karabarbounis, Loukas, and Brent Neiman, 2014. The global decline of the labor share. The Quarterly journal of economics 129, 61–103.

Katz, Michael, Hanno Lustig, and Lars Nielsen, 2017. Are stocks real assets? sticky discount rates in stock markets. The Review of Financial Studies 30, 539–587.

Kendall, Maurice G, 1954. Note on bias in the estimation of autocorrelation. Biometrika 41, 403–404.

King, Robert G., Charles I. Plosser, James H. Stock, and Mark W. Watson, 1991. Stochastic trends and economic fluctuations. American Economic Review 81, 819–840.

Kleidon, Allan W, 1986. Variance bounds tests and stock price valuation models. Journal of Political Economy 94, 953–1001.

Koijen, Ralph SJ, and Stijn Van Nieuwerburgh, 2011. Predictability of returns and cash flows. Annu. Rev. Financ. Econ. 3, 467–491.

Kozicki, Sharon, and Peter A Tinsley, 2001. Shifting endpoints in the term structure of interest rates. Journal of monetary Economics 47, 613–652.

Kuvshinov, Dmitry, and Kaspar Zimmermann, 2022. The big bang: Stock market capitalization in the long run. Journal of Financial Economics 145, 527–552.

Kwiatkowski, Denis, Peter CB Phillips, Peter Schmidt, and Yongcheol Shin, 1992. Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root?. Journal of econometrics 54, 159–178.

Laubach, Thomas, and John C Williams, 2003. Measuring the natural rate of interest. Review of Economics and Statistics 85, 1063–1070.

Lettau, Martin, and Sydney Ludvigson, 2001. Consumption, aggregate wealth, and expected stock returns. the Journal of Finance 56, 815–849.

Lettau, Martin, and Sydney C Ludvigson, 2005. Expected returns and expected dividend growth. Journal of Financial Economics 76, 583–626.

Lettau, Martin, and Stijn Van Nieuwerburgh, 2008. Reconciling the return predictability evidence. The Review of Financial Studies 21, 1607–1652.

Lewellen, Jonathan, 2004. Predicting returns with financial ratios. Journal of Financial Economics 74, 209–235.

Lintner, John, 1956. Distribution of incomes of corporations among dividends, retained earnings, and taxes. The American economic review 46, 97–113.

Liu, Yukun, and Ben Matthies, 2022. Long-run risk: Is it there?. Journal of Finance 77, 1587–1633.

Ljung, G. M., and G. E. P. Box, 1978. On a measure of lack of fit in time series models. Biometrika 65, 297–303.

Lucas Jr, Robert E, 1978. Asset prices in an exchange economy. Econometrica: journal of the Econometric Society pp. 1429–1445.

Marsh, Terry A, and Robert C Merton, 1986. Dividend variability and variance bounds tests for the rationality of stock market prices. The American Economic Review 76, 483–498.

Mehra, Rajnish, and Edward C Prescott, 1985. The equity premium: A puzzle. Journal of monetary Economics 15, 145–161.

Modigliani, Franco, and Richard A Cohn, 1979. Inflation, rational valuation and the market. Financial Analysts Journal 35, 24–44.

Nelson, Charles R., and Charles I. Plosser, 1982. Trends and random walks in macroeconomic time series: Some evidence and implications. Journal of Monetary Economics 10, 139–162.

Pástor, L'uboš, and Robert F Stambaugh, 2009. Predictive systems: Living with imperfect predictors. The Journal of Finance 64, 1583–1628.

Rapach, David E, Jack K Strauss, and Guofu Zhou, 2010. Out-of-sample equity premium prediction: Combination forecasts and links to the real economy. The Review of Financial Studies 23, 821–862.

Schorfheide, Frank, Dongho Song, and Amir Yaron, 2018. Identifying long-run risks: A bayesian mixed-frequency approach. Econometrica 86, 617–654.

Shiller, Robert J, 1981. Do stock prices move too much to be justified by subsequent changes in dividends?. The American Economic Review.

Skinner, Douglas J, 2008. The evolving relation between earnings, dividends, and stock repurchases. Journal of financial economics 87, 582–609.

Stambaugh, Robert F, 1999. Predictive regressions. Journal of financial economics 54, 375–421.

Stock, James H, and Mark W Watson, 2007. Why has us inflation become harder to forecast?. Journal of Money, Credit and banking 39, 3–33.

Torous, Walter, Rossen Valkanov, and Shu Yan, 2004. On predicting stock returns with nearly integrated explanatory variables. Journal of Business 77, 937–966.

Van Binsbergen, Jules H, and Ralph SJ Koijen, 2010. Predictive regressions: A present-value approach. The Journal of Finance 65, 1439–1471.

Vuolteenaho, Tuomo, 2002. What drives firm-level stock returns?. The Journal of Finance 57, 233–264.

Watson, Mark W, 1986. Univariate detrending methods with stochastic trends. Journal of monetary economics 18, 49–75.

Weil, Philippe, 1989. The equity premium puzzle and the risk-free rate puzzle. Journal of monetary economics 24, 401–421.

ONLINE APPENDIX

A Data

A.1 Main Data

- **S&P 500 price index**: We obtain the S&P 500 index from Amit Goyal's website (replication data for Goyal and Welch (2008)).
- **S&P 500 dividends**: We obtain the S&P 500 annual dividends from Robert Shiller's website.
- **S&P 500 earnings**: We obtain the S&P 500 annual earnings from Robert Shiller's website.
- **S&P 500 returns**: Following Goyal and Welch (2008) and a vast literature, we compute one-year returns as $R_{t+1} = \frac{P_{t+1}+D_{t+1}}{P_t}$ where D_{t+1} denotes dividends per S&P share over the past year (paid out in year $t+1$) and P_t denotes the S&P 500 at the end of year t.
- **Risk-free returns**: Risk-free returns are from Amit Goyal's website.

A.2 Dividend-Per-Share Decomposition

The decomposition in Section 6.2 requires the four components. Nominal GDP and the price level come from Johnston and Williamson (2023), whose annual nominal-GDP series extends back to 1871. We use its GDP deflator index as our measure of prices, so that inflation and real GDP growth are measured on a single consistent source over the full sample.

The dividend share of GDP and the net dilution yield require more steps, because both require an aggregate dividend series and a share count consistent with the S&P index over a century. We construct both from CRSP. We build aggregate dollar dividends from the CRSP daily file as the difference between cum- and ex-dividend returns, multiplied by lagged market value, summed across the firms in the index and then over trailing twelve-month windows. The one complication is the index's own history. In March 1957, the S&P index expanded from 90 to 500 constituents, mechanically jumping the level of aggregate dividends. We locate the transition and compute the implied splice factor from the ratio of the index divisor immediately before and after, and rescale the pre-1957 series so the level is continuous across the break. The result is a dollar dividend series on a consistent S&P basis from 1926 onward.

We recover net dilution from the same data. Net issuance is the change in shares outstanding, adjusted for splits and other distributions through the CRSP cumulative adjustment factor, valued at lagged prices and cumulated into a share-count series $\ln S$. Dividing aggregate dividends by this share count yields our constructed dividend-per-share index, which combines, by construction, inflation, real GDP growth, the dividend share of GDP, and net dilution.

Our constructed index does not match Shiller's published dividend-per-share series exactly because we do not always observe the full set of index constituents, and we do not separately track index additions and deletions or distinguish true deletions from mergers and delistings. We collect this remaining wedge—the log difference between Shiller's dividend per share and our constructed index—into the residual ω, in Table 9.

B Derivations: Present-Value with Trends and Cycles

B.1 Extended Campbell-Shiller identity

In this section, we detail the derivation of the log-linear approximation allowing for a random walk component of growth rates and returns, following Lettau and Van Nieuwerburgh (2008).

We start with the definition of the gross return of an asset:

$$R_{t+1} = \frac{P_{t+1} + D_{t+1}}{P_t} = \frac{D_{t+1}}{D_t} \frac{1 + PD_{t+1}}{PD_t} \tag{A.1}$$

where $PD_{t+1} = P_{t+1}/D_{t+1}$ or in log terms

$$r_{t+1} = \Delta d_{t+1} + \log(1 + \exp(pd_{t+1})) - pd_t \tag{A.2}$$

Campbell-Shiller perform a log-linearization around a constant pd ratio. Crucially, this hinges on stationary stochastic processes and is therefore not applicable to our model. In this section, we build on the extension of Lettau and Van Nieuwerburgh (2008) allowing for unit root processes.

Following Lettau and Van Nieuwerburgh (2008), we define the steady-state price–dividend ratio (fixing returns and growth rates at their trend values) as

$$\overline{R}_t = \overline{\Delta D}_t \frac{1 + \overline{PD}_t}{\overline{PD}_t}. \tag{A.3}$$

Re-arranging yields

$$\overline{PD}_t = \frac{\overline{\Delta D}_t}{\overline{R}_t - \overline{\Delta D}_{t+1}} \tag{A.4}$$

Plugging in $\overline{\Delta D}_t = \exp(g_t)$, $\overline{R}_t = \exp(\bar{\mu})$ (and assuming $g_t < \bar{\mu}$), we get that the steady-state log price-to-dividend ratio is given by:

$$\overline{pd}_t = \Delta \bar{d}_{t+1} - \log(\exp(\bar{r}_{t+1}) - \exp(\Delta \bar{d}_{t+1})) \tag{A.5}$$

Same as saying

$$\overline{pd}_t = g_t - \log(\exp(\bar{\mu}) - \exp(g_t)) \tag{A.6}$$

Therefore it must also hold that

$$\bar{r}_{t+1} = \Delta \bar{d}_{t+1} + \log(1 + \exp(\overline{pd}_t)) - \overline{pd}_t \tag{A.7}$$

Taking differences between the identities of pd_t and $\overline{pd}_t$ we get

$$r_{t+1} - \bar{r}_{t+1} = (\Delta d_{t+1} - \Delta \bar{d}_{t+1}) + \left[\log(1 + \exp(pd_{t+1})) - \log(1 + \exp(\overline{pd}_t))\right] - (pd_t - \overline{pd}_t) \tag{A.8}$$

A first-order approximation of $\log(1 + \exp(pd_{t+1}))$ around $\log(1 + \exp(\overline{pd}_t))$ yields (this is the usual

"Campbell-Shiller trick")

$$\log(1+\exp(pd_{t+1})) \;\approx\; \log\Big(1+\exp(\overline{pd}_t)\Big) + \rho_t\left(pd_{t+1}-\overline{pd}_t\right). \tag{A.9}$$

where the first derivative $\rho_t = \frac{\exp(\overline{pd}_t)}{1+\exp(\overline{pd}_t)}$. Re-arranging, we get

$$pd_t - \overline{pd}_t = (\Delta d_{t+1} - \Delta\bar{d}_{t+1}) - (r_{t+1}-\bar{r}_{t+1}) + \rho_t\left(pd_{t+1}-\overline{pd}_t\right) \tag{A.10}$$

Since $pd_{t+1}-\overline{pd}_t = \left(pd_{t+1}-\overline{pd}_{t+1}\right)+\left(\overline{pd}_{t+1}-\overline{pd}_t\right)$, we can express this as

$$pd_t - \overline{pd}_t = (\Delta d_{t+1} - \Delta\bar{d}_{t+1}) - (r_{t+1}-\bar{r}_{t+1}) + \rho_t\left(pd_{t+1}-\overline{pd}_{t+1}\right) + \rho_t\left(\overline{pd}_{t+1}-\overline{pd}_t\right) \tag{A.11}$$

Iterating forward on $pd_{t+1}-\overline{pd}_{t+1}$ gives

$$pd_t - \overline{pd}_t = \sum_{j=0}^{\infty}(\prod_{i=1}^{j}\rho_{t+i})(\Delta d_{t+1+j} - \Delta\bar{d}_{t+1+j}) - \sum_{j=0}^{\infty}(\prod_{i=1}^{j}\rho_{t+i})(r_{t+1+j}-\bar{r}_{t+1+j}) \tag{A.12}$$

$$+\sum_{j=0}^{\infty}(\prod_{i=1}^{j}\rho_{t+i})\left(\overline{pd}_{t+1+j}-\overline{pd}_{t+j}\right) + \lim_{h\to\infty}(\prod_{i=1}^{h}\rho_{t+i})\left(pd_{t+h}-\overline{pd}_{t+h}\right) \tag{A.13}$$

Taking expectations at time t yields

$$pd_t - \overline{pd}_t = \mathbb{E}_t\left[\sum_{j=0}^{\infty}(\prod_{i=1}^{j}\rho_{t+i})(\Delta d_{t+1+j} - \Delta\bar{d}_{t+1+j})\right] - \mathbb{E}_t\left[\sum_{j=0}^{\infty}(\prod_{i=1}^{j}\rho_{t+i})(r_{t+1+j}-\bar{r}_{t+1+j})\right] \tag{A.14}$$

$$+\mathbb{E}_t\left[\sum_{j=0}^{\infty}(\prod_{i=0}^{j}\rho_{t+i})\left(\overline{pd}_{t+1+j}-\overline{pd}_{t+j}\right)\right] + \mathbb{E}_t\left[\lim_{h\to\infty}(\prod_{i=0}^{h}\rho_{t+i})\left(pd_{t+h}-\overline{pd}_{t+h}\right)\right] \tag{A.15}$$

imposing a no-bubble condition $\mathbb{E}_t\left[\lim_{h\to\infty}(\prod_{i=0}^{h}\rho_{t+i})\left(pd_{t+h}-\overline{pd}_{t+h}\right)\right]=0$ and separating the expectations

$$pd_t - \overline{pd}_t = \sum_{j=0}^{\infty}\mathbb{E}_t\left[\prod_{i=1}^{j}\rho_{t+i}\right]\cdot\mathbb{E}_t\left[\Delta d_{t+1+j} - \Delta\bar{d}_{t+1+j}\right] - \sum_{j=0}^{\infty}\mathbb{E}_t\left[\prod_{i=1}^{j}\rho_{t+i}\right]\mathbb{E}_t\left[r_{t+1+j}-\bar{r}_{t+1+j}\right] \tag{A.16}$$

$$+\sum_{j=0}^{\infty}\mathbb{E}_t\left[\prod_{i=0}^{j}\rho_{t+i}\right]\mathbb{E}_t\left[\overline{pd}_{t+1+j}-\overline{pd}_{t+j}\right] \tag{A.17}$$

The first summation term is determined by the expectation of future deviation of cash flow growth from trend growth, i.e., the sum of future cash flow cycles. The second summation term is the expected deviation of returns from their mean values. The third term considers the drift in the steady-state of pd_t. This third term is positive since the function $\overline{pd}(g) = g - \log\big(\exp(\bar{\mu})-\exp(g)\big)$ is convex in the random walk g_t.

We make three approximating assumptions to arrive at the measurement equation.

First, following Lettau and Van Nieuwerburgh (2008), we anchor future linearization weights at today's coefficient, $\mathbb{E}_t[\rho_{t+j}] = \rho_t$, so that $\mathbb{E}_t\left[\prod_{i=1}^{j}\rho_{t+i}\right] = \rho_t^j$. Three substitutions then simplify the identity. In the first summation, $\mathbb{E}_t[\Delta\bar{d}_{t+1+j}] = \mathbb{E}_t[g_{t+j}] = g_t$, since g_t is a martingale. In the second, $\bar{r}_{t+1+j} = \bar{\mu}$ by construction of the steady state. And the third summation, a discounted sum of expected anchor drifts, collapses to a scalar,

$$\delta_t \;\equiv\; \sum_{j=0}^{\infty}\rho_t^{j+1}\,\mathbb{E}_t\left[\overline{pd}_{t+1+j} - \overline{pd}_{t+j}\right], \tag{A.18}$$

which is strictly positive by the convexity argument above and which the second assumption evaluates in closed form. Thus, after re-indexing, we obtain

$$pd_t - \overline{pd}_t = \delta_t + \sum_{j=1}^{\infty}\rho_t^{j-1}\,\mathbb{E}_t\left[\Delta d_{t+j} - g_t\right] - \sum_{j=1}^{\infty}\rho_t^{j-1}\,\mathbb{E}_t\left[r_{t+j-1,t+j} - \bar{\mu}\right], \tag{A.19}$$

Second, we assume that $\mathbb{E}_t\left[\overline{pd}_{t+1+j} - \overline{pd}_{t+j}\right]$ is well-approximated by a second-order Taylor expansion. This avoids assuming that $\overline{pd}_{t+j}$ follows a Martingale which is incorrect (Lettau and Van Nieuwerburgh, 2008). To derive the second-order Taylor expansion, let

$$\Delta g_{t+1+j} = g_{t+1+j} - g_{t+j}. \tag{A.20}$$

A second–order Taylor expansion of $\overline{pd}(g_{t+j} + \Delta g_{t+1+j})$ around g_{t+j} gives

$$\overline{pd}(g_{t+j} + \Delta g) - \overline{pd}(g_{t+j}) \approx \overline{pd}'(g_{t+j})\,\Delta g + \frac{1}{2}\overline{pd}''(g_{t+j})(\Delta g)^2. \tag{A.21}$$

Taking conditional expectation at time $t+j$,

$$\mathbb{E}_{t+j}\left[\overline{pd}_{t+1+j} - \overline{pd}_{t+j}\right] \approx \overline{pd}'(g_{t+j})\mathbb{E}_{t+j}[\Delta g_{t+1+j}] + \frac{1}{2}\overline{pd}''(g_{t+j})\mathrm{Var}_{t+j}(\Delta g_{t+1+j}). \tag{A.22}$$

Since g_t follows a random walk with martingale increments,

$$\mathbb{E}_{t+j}[\Delta g_{t+1+j}] = 0, \tag{A.23}$$

so the first term drops out and

$$\mathbb{E}_{t+j}\left[\overline{pd}_{t+1+j} - \overline{pd}_{t+j}\right] \approx \frac{1}{2}\overline{pd}''(g_{t+j})\mathrm{Var}_{t+j}(\Delta g_{t+1+j}). \tag{A.24}$$

The second derivative is

$$\overline{pd}''(g) = \frac{\exp(\bar{\mu} + g)}{\left(\exp(\bar{\mu}) - \exp(g)\right)^2}. \tag{A.25}$$

Since $\rho_{t+j} = \frac{\exp(\overline{pd}_{t+j})}{1+\exp(\overline{pd}_{t+j})}$ and therefore

$$\exp(\overline{pd}_{t+j}) = \frac{\exp(g_{t+j})}{\exp(\bar{\mu}) - \exp(g_{t+j})}. \tag{A.26}$$

it follows that

$$\rho_{t+j} = \frac{\dfrac{\exp(g_{t+j})}{\exp(\bar{\mu}) - \exp(g_{t+j})}}{1 + \dfrac{\exp(g_{t+j})}{\exp(\bar{\mu}) - \exp(g_{t+j})}} = \frac{\exp(g_{t+j})}{\exp(\bar{\mu})}. \tag{A.27}$$

Equivalently,

$$\exp(g_{t+j}) = \rho_{t+j} \exp(\bar{\mu}), \tag{A.28}$$

$$\exp(\bar{\mu}) - \exp(g_{t+j}) = \exp(\bar{\mu})(1 - \rho_{t+j}). \tag{A.29}$$

This means we can write

$$\overline{pd}''(g_{t+j}) = \frac{\exp(\bar{\mu} + g_{t+j})}{\big(\exp(\bar{\mu}) - \exp(g_{t+j})\big)^2} \tag{A.30}$$

$$= \frac{\exp(\bar{\mu})\exp(g_{t+j})}{\exp(\bar{\mu})^2(1-\rho_{t+j})^2} \tag{A.31}$$

$$= \frac{\rho_{t+j}}{(1-\rho_{t+j})^2}. \tag{A.32}$$

Plugging this in for the second derivative and taking expectations at time t, we get

$$\mathbb{E}_t\left[\overline{pd}_{t+1+j} - \overline{pd}_{t+j}\right] \approx \mathbb{E}_t\left[\frac{1}{2}\frac{\rho_{t+j}}{(1-\rho_{t+j})^2}\mathrm{Var}_{t+j}(\Delta g_{t+1+j})\right]. \tag{A.33}$$

Under the approximation, $\rho_{t+j} \approx \rho_t$, we obtain

$$\mathbb{E}_t\left[\overline{pd}_{t+1+j} - \overline{pd}_{t+j}\right] \approx \frac{1}{2}\frac{\rho_t}{(1-\rho_t)^2}\sigma_g^2, \tag{A.34}$$

and summing over the discounted horizon gives the drift constant

$$\delta_t = \frac{1}{2}\frac{\rho_t^2}{(1-\rho_t)^3}\,\sigma_g^2. \tag{A.35}$$

Third, we assume $\rho_t = \rho$ is constant and equal to its sample average. Thus, substituting the cash flow and expected-return dynamics, we obtain

$$pd_t = \delta + \overline{pd}(g_t) + \mathcal{M}_t + \mathcal{C}_t \tag{A.36}$$

where

- $\delta = \frac{1}{2}\frac{\rho^2}{(1-\rho)^3}\sigma_g^2$
- $\overline{pd}(g_t) = g_t - \log(\exp(\bar{\mu}) - \exp(g_t))$
- $\mathcal{M}_t = -\frac{1}{1-\rho\rho_\mu}(\mu_t - \bar{\mu})$
- $\mathcal{C}_t = -\frac{1-\rho_c}{1-\rho\,\rho_c}\,c_t \;+\; \frac{\theta_c\,(1-\rho)}{1-\rho\,\rho_c}\,\varepsilon_{c,t}$.

Derivation of $\mathcal{C}_t$. From the cash-flow specification, one-period log dividend growth satisfies

$$\Delta d_{t+j} = g_{t+j-1} + \big(c_{t+j} - c_{t+j-1}\big),$$

so that

$$\Delta d_{t+j} - g_t = \underbrace{\big(g_{t+j-1} - g_t\big)}_{\text{trend drift}} + \underbrace{\big(c_{t+j} - c_{t+j-1}\big)}_{\text{cycle change}}. \tag{A.37}$$

Since g_t is a random walk, $\mathbb{E}_t[g_{t+j-1}] = g_t$ for all $j \geq 1$, so the trend-drift term vanishes in expectation:

$$\mathbb{E}_t\big[g_{t+j-1} - g_t\big] = 0 \qquad \forall\, j \geq 1.$$

The cash-flow sum therefore reduces to the discounted sum of expected cycle changes:

$$\sum_{j=1}^{\infty} \rho^{j-1}\,\mathbb{E}_t\big[\Delta d_{t+j} - g_t\big] \;=\; \sum_{j=1}^{\infty} \rho^{j-1}\,\mathbb{E}_t\big[c_{t+j} - c_{t+j-1}\big]. \tag{A.38}$$

We split this into two sums:

$$\sum_{j=1}^{\infty} \rho^{j-1}\,\mathbb{E}_t\big[c_{t+j} - c_{t+j-1}\big] \;=\; \underbrace{\sum_{j=1}^{\infty} \rho^{j-1}\,\mathbb{E}_t[c_{t+j}]}_{\equiv\, S_1} \;-\; \underbrace{\sum_{j=1}^{\infty} \rho^{j-1}\,\mathbb{E}_t[c_{t+j-1}]}_{\equiv\, S_2}. \tag{A.39}$$

Forecasting the cycle. Since the cycle c_t follows an ARMA(1,1),

$$c_{t+1} = \rho_c\, c_t + \theta_c\, \varepsilon_{c,t} + \varepsilon_{c,t+1},$$

the h-step-ahead conditional expectation for any $h \geq 1$ is

$$\mathbb{E}_t[c_{t+h}] = \rho_c^h\, c_t + \rho_c^{h-1}\,\theta_c\,\varepsilon_{c,t}. \tag{A.40}$$

This follows by iterating the ARMA recursion forward and noting that all future innovations $\varepsilon_{c,t+s}$ ($s \geq 1$) have zero conditional mean.

Evaluating S_1. Substituting (A.40) and letting $\varphi \equiv \rho\rho_c$,

$$\begin{aligned} S_1 &= \sum_{j=1}^{\infty} \rho^{j-1} \Big[\rho_c^j \, c_t + \rho_c^{j-1} \, \theta_c \, \varepsilon_{c,t} \Big] \\ &= \rho_c \, c_t \sum_{j=1}^{\infty} \varphi^{j-1} \; + \; \theta_c \, \varepsilon_{c,t} \sum_{j=1}^{\infty} \varphi^{j-1} \\ &= \frac{\rho_c}{1-\varphi} \, c_t \; + \; \frac{\theta_c}{1-\varphi} \, \varepsilon_{c,t}. \end{aligned} \tag{A.41}$$

Evaluating S_2. Separate the $j = 1$ term (for which $c_{t+j-1} = c_t$ is known) from the remaining terms ($j \geq 2$):

$$\begin{aligned} S_2 &= c_t + \sum_{j=2}^{\infty} \rho^{j-1} \, \mathbb{E}_t[c_{t+j-1}] \\ &= c_t + \sum_{j=2}^{\infty} \rho^{j-1} \Big[\rho_c^{j-1} \, c_t + \rho_c^{j-2} \, \theta_c \, \varepsilon_{c,t} \Big]. \end{aligned} \tag{A.42}$$

The two geometric series are

$$\sum_{j=2}^{\infty} \rho^{j-1} \rho_c^{j-1} = \sum_{j=2}^{\infty} \varphi^{j-1} = \frac{\varphi}{1-\varphi},$$

and

$$\sum_{j=2}^{\infty} \rho^{j-1} \rho_c^{j-2} = \frac{1}{\rho_c} \sum_{j=2}^{\infty} \varphi^{j-1} = \frac{1}{\rho_c} \cdot \frac{\varphi}{1-\varphi} = \frac{\rho}{1-\varphi}.$$

Substituting into (A.42):

$$\begin{aligned} S_2 &= c_t + \frac{\varphi}{1-\varphi} \, c_t + \frac{\rho}{1-\varphi} \, \theta_c \, \varepsilon_{c,t} \\ &= \frac{1}{1-\varphi} \, c_t + \frac{\rho \, \theta_c}{1-\varphi} \, \varepsilon_{c,t}. \end{aligned} \tag{A.43}$$

Combining. Subtracting (A.43) from (A.41):

$$\begin{aligned} S_1 - S_2 &= \frac{\rho_c - 1}{1-\varphi} \, c_t \; + \; \frac{(1-\rho) \, \theta_c}{1-\varphi} \, \varepsilon_{c,t} \\ &= - \frac{1-\rho_c}{1-\rho\rho_c} \, c_t \; + \; \frac{(1-\rho) \, \theta_c}{1-\rho\rho_c} \, \varepsilon_{c,t}. \end{aligned} \tag{A.44}$$

The cash-flow cycle is therefore

$$\mathcal{C}_t \; \equiv \; - \frac{1-\rho_c}{1-\rho\rho_c} \, c_t \; + \; \frac{(1-\rho) \, \theta_c}{1-\rho\rho_c} \, \varepsilon_{c,t}. \tag{A.45}$$

C Model Estimation

C.1 Extended Kalman Filter

Model Specification We collect the full structural model here for reference. Log dividends, d_t, are decomposed into a non-stationary trend, τ_t, and a stationary cycle, c_t. Expected returns, μ_t, follow an AR(1) process around a constant steady-state level $\bar{\mu}$:

$$\begin{aligned} d_t &= \tau_t + c_t, \\ \tau_t &= \tau_{t-1} + g_{t-1}, \\ g_t &= g_{t-1} + \varepsilon_{g,t},, \quad \varepsilon_{g,t} \sim \mathcal{N}(0, \sigma_g^2) \\ c_t &= \rho_c\, c_{t-1} + \theta_c\, \varepsilon_{c,t-1} + \varepsilon_{c,t},, \quad \varepsilon_{c,t} \sim \mathcal{N}(0, \sigma_c^2) \\ \mu_t &= \bar{\mu} + \rho_\mu\, (\mu_{t-1} - \bar{\mu}) + \varepsilon_{\mu,t},, \quad \varepsilon_{\mu,t} \sim \mathcal{N}(0, \sigma_\mu^2) \end{aligned} \tag{A.46}$$

All shocks are mutually uncorrelated. The trend level τ_t evolves deterministically given the growth rate g_t, which is itself a random walk; g_t is therefore the sole source of trend non-stationarity, yielding a smooth (integrated random walk) trend. The cycle c_t follows a parsimonious ARMA(1,1) process.

State-Space System The state vector collects all unobserved quantities:

$$s_t = (\tau_t,\ g_t,\ c_t,\ \varepsilon_{c,t},\ \mu_t,\ 1)^\top. \tag{A.47}$$

The transition equation $s_{t+1} = A\, s_t + C\, \epsilon_{t+1}$ is governed by:

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & \rho_c & \theta_c & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \rho_\mu & (1-\rho_\mu)\bar{\mu} \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \qquad C = \begin{bmatrix} 0 & 0 & 0 \\ \sigma_g & 0 & 0 \\ 0 & \sigma_c & 0 \\ 0 & \sigma_c & 0 \\ 0 & 0 & \sigma_\mu \\ 0 & 0 & 0 \end{bmatrix}, \tag{A.48}$$

where C is a 6×3 shock-loading matrix and the state innovations $\epsilon_{t+1}^\top = (\epsilon_{g,t+1},\ \epsilon_{c,t+1},\ \epsilon_{\mu,t+1}) \sim \mathcal{N}(0, I_3)$ are mutually independent.

Nonlinear Measurement Equation The measurement vector stacks dividends and the valuation ratio:

$$x_t = (d_t,\ pd_t)^\top. \tag{A.49}$$

The valuation ratio is pinned down by the present-value identity:

$$pd_t = \overline{pd}_t + \delta - \frac{1-\rho_c}{1-\rho\rho_c}\, c_t \ + \ \frac{1-\rho}{1-\rho\rho_c}\, \theta_c\, \varepsilon_{c,t} - \frac{1}{1-\rho\rho_\mu}\, (\mu_t - \bar{\mu}). \tag{A.50}$$

where $\overline{pd}_t = g_t - \log\big(\exp(\bar{\mu}) - \exp(g_t)\big)$ denote the steady-state log price–dividend ratio and $\delta = \frac{\rho^2}{2(1-\rho)^3\sigma_g^2}$ is a Jensen's inequality term. Because it is a nonlinear function of g_t, the valuation-ratio equa-

tion is nonlinear in the state vector s_t. A standard Kalman filter—which requires a linear mapping from states to observables—cannot be applied directly. We therefore estimate the model using an Extended Kalman Filter (EKF), which replaces the nonlinear measurement function with a first-order Taylor approximation, re-evaluated at each time step.

To implement the EKF within a linear state-space framework ($x_t = G_t\, s_t$), we take the first-order Taylor expansion of $\overline{pd}_t$ around the filter's prior estimate of trend growth, $\hat{g}_{t|t-1}$:

$$\overline{pd}_t \;\approx\; \widehat{\overline{pd}}_t + \lambda_t\,(g_t - \hat{g}_{t|t-1}), \tag{A.51}$$

where the evaluated steady-state valuation ratio and its Jacobian are:

$$\widehat{\overline{pd}}_t \equiv \hat{g}_{t|t-1} - \log\big(\exp(\bar{\mu}) - \exp(\hat{g}_{t|t-1})\big), \tag{A.52}$$

$$\lambda_t \equiv \left.\frac{\partial \overline{pd}_t}{\partial g_t}\right|_{g_t=\hat{g}_{t|t-1}} = 1 + \frac{\exp(\hat{g}_{t|t-1})}{\exp(\bar{\mu}) - \exp(\hat{g}_{t|t-1})}. \tag{A.53}$$

The derivative λ_t captures the local sensitivity of the steady-state valuation ratio to trend growth. It equals unity when g_t is far below $\bar{\mu}$ and diverges as $g_t \to \bar{\mu}$, reflecting the convexity of the Gordon-growth present-value function.

Substituting the linearized $\overline{pd}_t$ into (A.50) and rearranging to isolate the state vector s_t:

$$\begin{aligned} pd_t \;\approx\; & \delta + \lambda_t\, g_t - \frac{1-\rho_c}{1-\rho\rho_c}\, c_t + \frac{1-\rho}{1-\rho\rho_c}\,\theta_c\,\varepsilon_{c,t} - \frac{1}{1-\rho\rho_\mu}\,\mu_t \\ & + \underbrace{\frac{\bar{\mu}}{1-\rho\rho_\mu} + \widehat{\overline{pd}}_t - \lambda_t\,\hat{g}_{t|t-1}}_{\kappa_t}. \end{aligned} \tag{A.54}$$

Thus, we have $x_t = G_t\, s_t$, where

$$G_t = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & \lambda_t & -\dfrac{1-\rho_c}{1-\rho\rho_c} & \dfrac{1-\rho}{1-\rho\rho_c}\,\theta_c & -\dfrac{1}{1-\rho_\mu\rho} & \kappa_t \end{bmatrix}, \tag{A.55}$$

with

$$\lambda_t = 1 + \frac{\exp(\hat{g}_{t|t-1})}{\exp(\bar{\mu}) - \exp(\hat{g}_{t|t-1})}, \tag{A.56}$$

$$\kappa_t = \delta + \frac{\bar{\mu}}{1-\rho\rho_\mu} + \hat{g}_{t|t-1} - \log\big(\exp(\bar{\mu}) - \exp(\hat{g}_{t|t-1})\big) - \lambda_t\,\hat{g}_{t|t-1}. \tag{A.57}$$

At each step of the filter, G_t is re-evaluated at the current prior estimate $\hat{g}_{t|t-1}$, so that the linearization tracks the evolving state of the economy. The Jacobian term λ_t in (A.53) is crucial for correctly identifying the model. If it were instead fixed at some constant value, the filter would misattribute valuation surprises. In periods when g_t is close to $\bar{\mu}$, the true sensitivity of valuations to growth news is much

larger than average, and a fixed linearization would systematically underweight growth-rate revisions. The EKF addresses this by re-linearizing at each time step, ensuring that the filter's updating weights correctly reflect the local curvature of the present-value function. In practice, because the estimated volatility of trend growth σ_g is small, the nonlinearity is mild and the first-order Taylor approximation provides an accurate representation of the measurement equation.

C.2 Estimation Details

To estimate the parameters $\Theta = (\rho_c, \theta_c, \sigma_g, \sigma_c, \rho_\mu, \sigma_\mu, \bar{\mu})$, we employ Maximum Likelihood Estimation (MLE) using the Extended Kalman Filter. Because the likelihood surface is generally non-convex, we utilize a basinhopping algorithm combined with L-BFGS-B as the local optimizer to ensure robust convergence.

We calibrate starting values for the optimizer using pre-sample moments from Golez and Koudijs (2018), pooling their Netherlands and UK (1700–1812) and UK (1813–1870) subsamples. We set the persistence of expected returns to the quarterly equivalent of the annual AR(1) of the dividend-to-price ratio ($\rho_\mu = 0.93$), the persistence of the dividend cycle to the quarterly equivalent of the annual cycle-level persistence, computed as one plus the AR(1) of annual dividend growth ($\rho_c = 0.95$), the volatility of expected-return innovations to one-quarter of quarterly realized return volatility ($\sigma_\mu = 0.005$), and the volatility of cycle innovations (σ_c) to its value implied by quarterly dividend-growth volatility under the short-horizon approximation of a constant trend ($\sigma_c = 0.012$). The quarterly mean expected return $\bar{\mu}$ is initialized to the pre-sample average quarterly return. The volatility of trend growth is set to 5 basis points per quarter. Estimates are invariant to initialization across the range of starting values we examined. The Campbell–Shiller present-value identity requires dividends to be reinvested to the end of the measurement period rather than simply summed, or reinvested at the risk-free rate (Cochrane, 2011). We approximate this by compounding quarterly dividends at a fixed annual rate of 8%. The quarterly pd is then the log of the S&P500 price index less the log of one quarter of the reinvestment-adjusted annual dividend.

We initialize the state vector $\hat{s}_{0|0} = (\tau_0,\ g_0,\ c_0,\ \varepsilon_{c,0},\ \mu_0,\ 1)^\top$ by anchoring the trend τ_0 to the first observed log-dividend value and the initial trend growth rate g_0 to the pre-sample average quarterly dividend growth rate from Golez and Koudijs (2018). The initial expected return μ_0 is set to the pre-sample mean $\bar{\mu}$, and the remaining cyclical components are set to zero.

To determine the initial state covariance matrix $P_{0|0}$, we assume the system is in its steady state at the beginning of the sample. We therefore set $P_{0|0}$ equal to the steady-state variance of the state vector implied by the model parameters. Calculating this variance requires fixing the measurement matrix G_t. However, G_t is time-varying as it depends on the steady-state valuation ratio $\overline{pd}_t$, which is a nonlinear function of the non-stationary trend growth g_t. To address this dependency, we evaluate the measurement matrix at a constant reference point, calculating a constant matrix G_{ref} where $g = g_0$ and $\mu_0 = \bar{\mu}$. Using G_{ref}, we solve for the steady-state covariance by iterating the discrete-time Riccati equation, given the transition matrix A and state noise covariance $Q = CC^\top$. The resulting $P_{0|0}$ is computed once from the initial parameter values and held fixed throughout the MLE optimization.

C.3 Bootstrap Inference

The bootstrap delivers the inference for Tables 4 and 6 and the standard errors for the parameter estimates in Table 3. The baseline model cannot generate the replications. Given enough simulated histories, its unit-root trend g_t eventually reaches the long-run expected return $\bar{\mu}$, and the price-dividend ratio diverges. We therefore simulate the stationary-trend variant (constant-ρ) of Appendix G, which replaces the random-walk trend with an AR(1) of persistence $\rho_g = 0.999$ and is defined on every sample path, requiring only $\bar{g} < \bar{\mu}$. On the actual data the two models deliver very similar filtered states (Appendix 7.3), and very similar predictability coefficients (Tables A.8 and A.9). Re-estimation moves the parameters little. Table A.1 reports the stationary estimates. Each is within a third of a bootstrap standard error of its baseline value (Table 3), and the largest change is the trend innovation, $\sigma_g \times 100 = 0.0266$ against the baseline's 0.0225—the direction mean reversion implies, since part of each innovation now decays and matching the same trend path requires slightly larger shocks. Because the initial states are unobserved, the likelihood must assume a distribution for them, $N(s_0, S_0)$, and each replication draws its initial states from exactly the distribution the estimation itself uses. The mean s_0 sets initial trend growth to the pre-sample estimate g_{init} and initial expected returns to the estimated long-run mean $\hat{\bar{\mu}}$. The covariance S_0 is the steady-state filter covariance, computed at the estimation's initial parameter values and held fixed, exactly as in estimation, and the states are drawn jointly. Each replication then evolves under independent Gaussian shocks at the estimated parameters over the 1871Q1–2022Q4 grid, builds dividends, prices, and returns from the simulated states, and is re-estimated and re-filtered by the procedure applied to the actual data. We use $B = 1{,}000$ replications.

Table A.1: Parameter Estimates: Stationary Trend Growth

Parameter	ρ_c	θ_c	$\sigma_g(\cdot 100)$	σ_c	ρ_μ	σ_μ	$\bar{\mu}$
Estimate	0.9710	0.4924	0.0266	0.0273	0.9509	0.0052	0.0210

Note: This table reports maximum-likelihood parameter estimates for the stationary trend-growth (constant-ρ) variant of Appendix G, estimated on the 1871Q1–2022Q4 sample.

Inference for the predictability tables We test each coefficient against zero and against its model-implied value. Replication b reruns the regression on its own simulated history for $\hat{\beta}_b^{\text{ols}}$, and recomputes the model-implied coefficient β_b^{model} at its own re-estimated parameters and filtered states, including the historical component contributions (A.120).

The zero test compares t-statistics. Its benchmark is the projection coefficient π_b^*, the slope from regressing replication b's true conditional mean of the outcome on its filtered regressors—the coefficient the regression would recover with no sampling noise, given the filter. Each replication's error is scaled by its own Newey–West standard error,

$$t_b^* = \frac{\hat{\beta}_b^{\text{ols}} - \pi_b^*}{\widehat{\text{se}}_b}, \tag{A.58}$$

computed by the same estimator and bandwidth rule as in the actual data. The p-value counts where

the observed $t = \hat{\beta}^{\text{ols}}/\widehat{\text{se}}$ falls among them,

$$p = \min\left\{1,\ 2\min\left(\frac{1+\#\{t_b^* \le t\}}{B+1},\ \frac{1+\#\{t_b^* \ge t\}}{B+1}\right)\right\}. \tag{A.59}$$

The square brackets report this p-value.

The equality test compares coefficients directly. The benchmark β_b^{model} is itself estimated, so the discrepancies

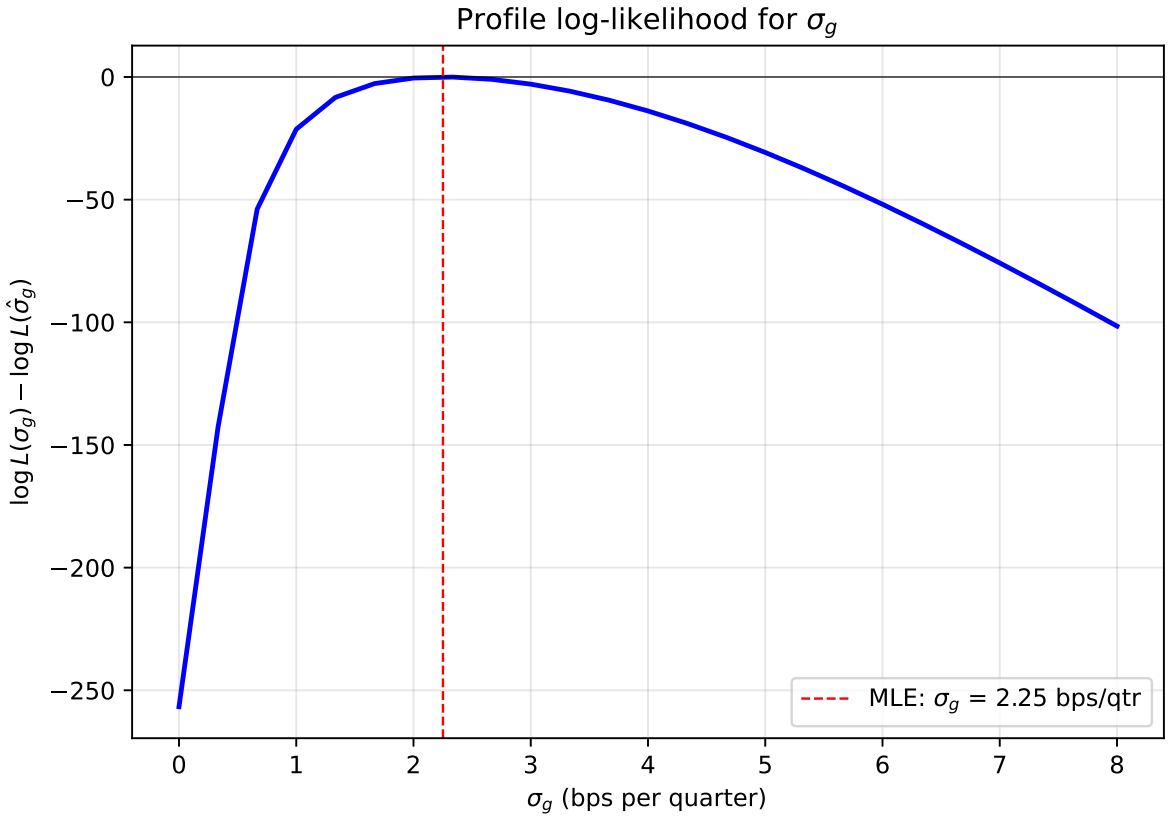
$$D_{\text{obs}} = \hat{\beta}^{\text{ols}} - \beta^{\text{model}}, \qquad D_b^* = \hat{\beta}_b^{\text{ols}} - \beta_b^{\text{model}} \tag{A.60}$$

carry the benchmark's sampling variation as well as the regression's. The curly brackets report the count (A.59) applied to D_{obs} among the D_b^*.

Parameter standard errors The same replications deliver the sampling distributions of the parameter estimates. Table 3 reports, for each parameter, the standard deviation of the re-estimated values across replications.

C.4 Profile Likelihood for σ_g

Figure A.1: Profile Log-Likelihood for σ_g



Note: This figure plots the profile log-likelihood for σ_g, normalized by the unconstrained maximum. At each grid value of σ_g, all other parameters are held at their unconstrained MLE values. The vertical dashed line marks the MLE, $\hat{\sigma}_g \times 100 = 0.0225$. The horizontal axis is in basis points per quarter.

C.5 Persistence of Filtered Expected Returns

Table A.2: Tests for Persistence and Stationarity

	ADF	KPSS	Sample
Filtered Expected Returns ($-\widehat{\mathcal{M}}_t$)	-3.84***	0.22	1871–2022

Note: This table reports unit-root and stationarity tests for the filtered expected-return component, $-\widehat{\mathcal{M}}_t$, obtained from the state-space decomposition. Data are annual. The augmented Dickey-Fuller (ADF) test has a unit root as the null hypothesis, so rejection is evidence of stationarity. The Kwiatkowski, Phillips, Schmidt, and Shin (1992) (KPSS) test has stationarity as the null hypothesis, so a failure to reject is evidence of stationarity. The two tests thus point in the same direction from opposite nulls. The ADF augmentation lag is selected by BIC and increased, if necessary, until the residuals satisfy a Ljung and Box (1978) test at the 10% level; the KPSS truncation lag is set to the same value. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

D Model Implications

D.1 Variances

Prices are given by

$$pd_t = \delta + \overline{pd}(g_t) + \mathcal{M}_t + \mathcal{C}_t \tag{A.61}$$

where

- $\overline{pd}(g_t) = g_t - \log(\exp(\bar{\mu}) - \exp(g_t))$
- $\mathcal{M}_t = -\frac{1}{1-\rho\rho_\mu}(\mu_t - \bar{\mu})$
- $\mathcal{C}_t = -\frac{1-\rho_c}{1-\rho\,\rho_c}\,c_t \;+\; \frac{\theta_c\,(1-\rho)}{1-\rho\,\rho_c}\,\varepsilon_{c,t}$

In the following, we derive the variance of all three terms.

D.1.1 Unconditional Variance of Cash Flow Cycle

The cash flow cycle follows an ARMA(1,1)

$$c_t = \rho_c\,c_{t-1} + \theta_c\,\varepsilon_{c,t-1} + \varepsilon_{c,t}, \quad , \quad \varepsilon_{c,t} \sim \mathcal{N}(0, \sigma_c^2) \tag{A.62}$$

The cash-flow component of prices is defined as $\mathcal{C}_t$

$$\mathcal{C}_t = \underbrace{\frac{\rho_c - 1}{1-\rho\,\rho_c}}_{\beta_c}\,c_t \;+\; \underbrace{\frac{\theta_c\,(1-\rho)}{1-\rho\,\rho_c}}_{\beta_\varepsilon}\,\varepsilon_{c,t}. \tag{A.63}$$

To compute $\mathrm{Var}(\mathcal{C}_t)$, we expand the variance of this linear combination:

$$\mathrm{Var}(\mathcal{C}_t) = \beta_c^2\,\mathrm{Var}(c_t) + \beta_\varepsilon^2\,\mathrm{Var}(\varepsilon_{c,t}) + 2\,\beta_c\,\beta_\varepsilon\,\mathrm{Cov}(c_t,\,\varepsilon_{c,t}). \tag{A.64}$$

Given the ARMA(1,1) structure of c_t, the relevant moments are:

- Variance of the shock: $\mathrm{Var}(\varepsilon_{c,t}) = \sigma_c^2$.
- Covariance: $\mathrm{Cov}(c_t,\,\varepsilon_{c,t}) = \sigma_c^2$.
- Unconditional variance of the cycle:

$$\mathrm{Var}(c_t) = \sigma_c^2\left(\frac{1 + \theta_c^2 + 2\,\rho_c\,\theta_c}{1-\rho_c^2}\right). \tag{A.65}$$

Substituting:

$$\mathrm{Var}(\mathcal{C}_t) = \sigma_c^2\left[\beta_c^2\left(\frac{1 + \theta_c^2 + 2\,\rho_c\,\theta_c}{1-\rho_c^2}\right) + \beta_\varepsilon^2 + 2\,\beta_c\,\beta_\varepsilon\right]. \tag{A.66}$$

D.1.2 Unconditional Variance of Expected Returns

Expected returns follow an AR(1) around a constant steady state:

$$\mu_{t+1} = \bar{\mu} + \rho_\mu \left(\mu_t - \bar{\mu}\right) + \varepsilon_{\mu,t+1}, \quad , \quad \varepsilon_{\mu,t} \sim \mathcal{N}(0, \sigma_\mu^2) \tag{A.67}$$

We use the definition of $\mathcal{M}_t = -\frac{1}{1-\rho\rho_\mu}(\mu_t - \bar{\mu})$ which reflects the expected-return signal contained in stock prices. It is the expectation for the entire stream of future returns. The variance of the expected-return component is

$$\mathrm{Var}(\mathcal{M}_t) = \frac{1}{\left(1 - \rho\,\rho_\mu\right)^2} \, \frac{\sigma_\mu^2}{1 - \rho_\mu^2}. \tag{A.68}$$

D.1.3 Derivation of $\mathrm{Var}_T(\overline{pd}(g_t))$

We compute $\mathrm{Var}_T(\overline{pd}(g_t)) = \mathbb{E}\big[T^{-1}\sum_{t=1}^{T}\left(\overline{pd}(g_t) - m_T\right)^2\big]$, the expected sample variance of the trend component over a window of length T, where m_T is the window mean. Starting from $g_0 = \bar{g} \equiv \bar{\mu} + \log\rho$ – the trend state at which the linearization coefficient $\rho(g) = e^{g-\bar{\mu}}$ equals the constant discounting coefficient ρ. The random walk gives $g_t - \bar{g} = \sum_{s=1}^{t}\varepsilon_{g,s} \equiv w_t$. To first order, $\overline{pd}(g_t) = \overline{pd}(\bar{g}) + \lambda w_t$ with $\lambda = \overline{pd}'(\bar{g}) = 1/(1-\rho)$, and the constant drops out under demeaning.

The sample variance is an average over pairs of dates,

$$\frac{1}{T}\sum_{t=1}^{T}\left(x_t - \bar{x}_T\right)^2 = \frac{1}{2T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\left(x_t - x_s\right)^2. \tag{A.69}$$

Under the random walk, $\mathbb{E}(w_t - w_s)^2 = \sigma_g^2|t-s|$, and $\sum_t\sum_s|t-s| = T(T^2-1)/3$, so

$$\mathrm{Var}_T\big(\overline{pd}(g_t)\big) \approx \lambda^2\sigma_g^2\,\frac{T^2-1}{6T}. \tag{A.70}$$

Curvature adds further variation, well-defined under the globally defined stationary anchor of Appendix G. The shares reported in Table 7 are computed by simulation from that anchor and include it.

D.1.4 Derivation of $\mathrm{Var}_T(M_t)$ and $\mathrm{Var}_T(C_t)$

For the stationary components, write $\gamma_X(h) \equiv \mathrm{Cov}(X_t, X_{t-h})$ for the autocovariance at lag h, so $\gamma_X(0) = \mathrm{Var}(X)$. Expanding the demeaned sample variance of a mean-zero process and taking expectations,

$$\mathrm{Var}_T(X) = \mathbb{E}\Big[\frac{1}{T}\sum_{t=1}^{T}X_t^2\Big] - \mathbb{E}\big[\bar{X}_T^2\big] = \gamma_X(0) - \mathrm{Var}\big(\bar{X}_T\big) \tag{A.71}$$

That variance sums the autocovariances over all T^2 pairs of dates; lag 0 occurs T times and each lag $h \geq 1$ occurs $2(T-h)$ times, so

$$\mathrm{Var}\big(\bar{X}_T\big) = \frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\gamma_X(|t-s|) = \frac{1}{T^2}\Big[T\gamma_X(0) + 2\sum_{h=1}^{T-1}(T-h)\,\gamma_X(h)\Big]. \tag{A.72}$$

For a geometric autocovariance structure, $\gamma_X(h) = \gamma_X(0)\, a^h$, the sum in (A.72) evaluates to $a\big[T(1-a) - (1-a^T)\big]/(1-a)^2$, giving $\mathrm{Var}_T(X) = \gamma_X(0)\, K_T(a)$ with

$$K_T(a) \equiv 1 - \frac{1}{T} - \frac{2a\big(T(1-a) - (1-a^T)\big)}{T^2(1-a)^2}. \tag{A.73}$$

Since $\gamma_{\mathcal{M}}(h) = \mathrm{Var}(\mathcal{M})\rho_\mu^h$, we obtain $\mathrm{Var}_T(\mathcal{M}) = \mathrm{Var}(\mathcal{M})\, K_T(\rho_\mu)$.

The cycle component is not an AR(1), but its autocovariances are geometric from lag one, with $\gamma_c(1) = \rho_c\gamma_c(0) + \theta_c\sigma_c^2$,

$$\gamma_{\mathcal{C}}(1) = \beta_c^2\,\gamma_c(1) + \beta_c\beta_\varepsilon\,(\rho_c + \theta_c)\,\sigma_c^2, \qquad \gamma_{\mathcal{C}}(h) = \rho_c\,\gamma_{\mathcal{C}}(h-1) \quad \text{for } h \geq 2, \tag{A.74}$$

so substituting into (A.72),

$$\mathrm{Var}_T(\mathcal{C}) = \gamma_{\mathcal{C}}(0)\Big(1 - \frac{1}{T}\Big) - \frac{2\,\gamma_{\mathcal{C}}(1)}{T^2}\,\frac{T(1-\rho_c) - (1-\rho_c^T)}{(1-\rho_c)^2}, \tag{A.75}$$

with $\gamma_{\mathcal{C}}(0) = \mathrm{Var}(\mathcal{C}_t)$ as derived above. Both converge to their unconditional variances as T grows.

D.1.5 Derivation of $\mathrm{Var}_T(pd_t)$

Because the component paths are mutually independent, the expected cross moments vanish, so

$$\mathrm{Var}_T(pd_t) = \mathrm{Var}_T\big(\overline{pd}(g_t)\big) + \mathrm{Var}_T(\mathcal{M}_t) + \mathrm{Var}_T(\mathcal{C}_t). \tag{A.76}$$

D.2 Persistence

Price ratios are given by

$$pd_t = \delta + \overline{pd}(g_t) + \mathcal{M}_t + \mathcal{C}_t \tag{A.77}$$

where

- $\overline{pd}(g_t) = g_t - \log(\exp(\bar{\mu}) - \exp(g_t))$
- $\mathcal{M}_t = -\frac{1}{1-\rho\rho_\mu}(\mu_t - \bar{\mu})$
- $\mathcal{C}_t = -\frac{1-\rho_c}{1-\rho\,\rho_c}\,c_t \;+\; \frac{\theta_c\,(1-\rho)}{1-\rho\,\rho_c}\,\varepsilon_{c,t}$

D.2.1 Finite-Window Lag-h persistence

For a window $x_1, \ldots, x_T$ and $1 \le h < T$, define the sample autocovariance

$$\widehat{\gamma}_{x,T}(h) \equiv \frac{1}{T}\sum_{t=h+1}^{T}\big(x_t - \bar{x}_T\big)\big(x_{t-h} - \bar{x}_T\big), \qquad \bar{x}_T \equiv \frac{1}{T}\sum_{t=1}^{T} x_t, \tag{A.78}$$

and let $\mathrm{Cov}_T(x_t, x_{t-h}) \equiv \mathbb{E}\big[\widehat{\gamma}_{x,T}(h)\big]$ and $\mathrm{Var}_T(x_t) \equiv \mathbb{E}\big[\widehat{\gamma}_{x,T}(0)\big]$ denote its model expectations. The finite-window lag-h persistence of x is

$$\phi^{(h)}_{x,T} = \frac{\mathrm{Cov}_T(x_t, x_{t-h})}{\mathrm{Var}_T(x_t)}. \tag{A.79}$$

Because the components are driven by mutually independent shock sequences and the stationary components have constant means, the expected cross-lag moments vanish at every lag, and the persistence of pd_t is a weighted average of the component persistences:

$$\phi^{(h)}_{pd,T} = \phi^{(h)}_{\overline{pd},T}\cdot\frac{\mathrm{Var}_T(\overline{pd}(g_t))}{\mathrm{Var}_T(pd_t)} + \phi^{(h)}_{\mathcal{M},T}\cdot\frac{\mathrm{Var}_T(\mathcal{M}_t)}{\mathrm{Var}_T(pd_t)} + \phi^{(h)}_{\mathcal{C},T}\cdot\frac{\mathrm{Var}_T(\mathcal{C}_t)}{\mathrm{Var}_T(pd_t)}. \tag{A.80}$$

with the variance shares as weights. Both the weights and the component persistences depend on the window length.

Conventions For a stationary X_t with autocovariances $\gamma_X(k) \equiv \mathrm{Cov}(X_t, X_{t-k})$, the variance of the window mean sums the autocovariances over all T^2 pairs of dates,

$$\mathrm{Var}\big(\bar{X}_T\big) = \frac{1}{T^2}\Big[T\,\gamma_X(0) + 2\sum_{k=1}^{T-1}(T-k)\,\gamma_X(k)\Big], \tag{A.81}$$

By definition $\mathrm{Var}_T(X_t) = \gamma_X(0) - \mathrm{Var}(\bar{X}_T)$. The expected demeaned lag-1 covariance follows from the same sums:

$$\mathrm{Cov}_T(X_t, X_{t-1}) = \frac{1}{T}\Big[(T-1)\,\gamma_X(1) - (T+1)\,\mathrm{Var}\big(\bar{X}_T\big)\Big] + \frac{2}{T^2}\sum_{k=0}^{T-1}\gamma_X(k). \tag{A.82}$$

The first bracket is the population covariance less the variance of the window mean; the last term is a boundary correction from the dates that appear in only one lag pair. Both corrections vanish as T grows, so $\phi^{(1)}_{X,T} \to \gamma_X(1)/\gamma_X(0)$. At lag h, for $1 \le h < T$,

$$\mathrm{Cov}_T(X_t, X_{t-h}) = \frac{1}{T}\Big[(T-h)\,\gamma_X(h) - (T+h)\,\mathrm{Var}\big(\bar{X}_T\big)\Big] + \frac{2}{T^2}\sum_{j=1}^{h}\Big[\sum_{k=0}^{j-1}\gamma_X(k) + \sum_{k=1}^{T-j}\gamma_X(k)\Big], \tag{A.83}$$

with the boundary sum now running over the h first and h last dates; (A.82) is the case $h = 1$.

D.2.2 Persistence of $\mathcal{M}_t$

Population $\phi^{(1)}_{\mathcal{M}} = \rho_\mu$ and $\phi^{(4)}_{\mathcal{M}} = \rho_\mu^4$.

Window With $\gamma_{\mathcal{M}}(k) = \mathrm{Var}(\mathcal{M})\,\rho_\mu^k$, the sums in (A.81)–(A.82) are geometric:

$$\frac{\mathrm{Var}\big(\bar{\mathcal{M}}_T\big)}{\mathrm{Var}(\mathcal{M})} = \frac{1}{T} + \frac{2\rho_\mu\big[T(1-\rho_\mu) - (1-\rho_\mu^T)\big]}{T^2(1-\rho_\mu)^2} \equiv 1 - K_T(\rho_\mu), \qquad \sum_{k=0}^{T-1}\rho_\mu^k = \frac{1-\rho_\mu^T}{1-\rho_\mu}, \tag{A.84}$$

and the quarterly window persistence is

$$\phi^{(1)}_{\mathcal{M},T} = \frac{(T-1)\,\rho_\mu - (T+1)\big[1-K_T(\rho_\mu)\big] + \dfrac{2}{T}\,\dfrac{1-\rho_\mu^T}{1-\rho_\mu}}{T\,K_T(\rho_\mu)}. \tag{A.85}$$

The annual version substitutes lag 4:

$$\phi^{(4)}_{\mathcal{M},T} = \frac{(T-4)\,\rho_\mu^4 - (T+4)\big[1-K_T(\rho_\mu)\big] + \dfrac{2\,(1+\rho_\mu)}{T\,(1-\rho_\mu)}\Big[4 - \rho_\mu\,(1+\rho_\mu^2)\,(1+\rho_\mu^{T-4})\Big]}{T\,K_T(\rho_\mu)}. \tag{A.86}$$

D.3 Persistence of $\mathcal{C}_t$

Population, quarterly Recall

$$\mathcal{C}_t = -a_c\,c_t + b_c\,\varepsilon_{c,t}, \qquad a_c \equiv \frac{1-\rho_c}{1-\rho\,\rho_c}, \qquad b_c \equiv \frac{\theta_c\,(1-\rho)}{1-\rho\,\rho_c}. \tag{A.87}$$

The cross-moments are

$$\mathrm{Cov}(c_{t-1}, \varepsilon_{c,t-1}) = \sigma_c^2, \qquad \mathrm{Cov}(c_t, \varepsilon_{c,t-1}) = (\rho_c + \theta_c)\sigma_c^2. \tag{A.88}$$

Using these moment restrictions (and $\mathrm{Cov}(\varepsilon_{c,t}, c_{t-1}) = 0$, $\mathrm{Cov}(\varepsilon_{c,t}, \varepsilon_{c,t-1}) = 0$),

$$\gamma_{\mathcal{C}}(1) = a_c^2\,\gamma_1 - \ a_c b_c\,(\rho_c + \theta_c)\sigma_c^2, \qquad \gamma_{\mathcal{C}}(0) = a_c^2\,\gamma_0 + b_c^2\,\sigma_c^2 - 2a_c b_c\,\sigma_c^2, \tag{A.89}$$

where $\gamma_k \equiv \mathrm{Cov}(c_t, c_{t-k})$, with

$$\gamma_1 = \rho_c\,\gamma_0 + \theta_c\,\sigma_c^2, \qquad \gamma_0 = \frac{1+\theta_c^2+2\rho_c\theta_c}{1-\rho_c^2}\,\sigma_c^2. \tag{A.90}$$

Hence $\phi_{\mathcal{C}}^{(1)} = \gamma_{\mathcal{C}}(1)/\gamma_{\mathcal{C}}(0)$.

Population, annual For an ARMA(1,1), $\gamma_4 = \rho_c^3\gamma_1$ and $\mathrm{Cov}(c_t, \varepsilon_{c,t-4}) = \rho_c^3(\rho_c+\theta_c)\sigma_c^2$, so

$$\phi_{\mathcal{C}}^{(4)} = \rho_c^3\,\phi_{\mathcal{C}}^{(1)}. \tag{A.91}$$

In general, $\phi_{\mathcal{C}}^{(4)} \neq (\phi_{\mathcal{C}}^{(1)})^4$ because $\mathcal{C}_t$ inherits ARMA(1,1) dynamics from c_t, rather than following a pure AR(1).

Window The autocovariances of $\mathcal{C}_t$ are geometric from lag one, $\gamma_{\mathcal{C}}(k) = \rho_c^{k-1}\gamma_{\mathcal{C}}(1)$ for $k \geq 1$, so substituting into (A.81)–(A.82),

$$\mathrm{Var}\big(\bar{\mathcal{C}}_T\big) = \frac{1}{T^2}\left[T\,\gamma_{\mathcal{C}}(0) + \frac{2\,\gamma_{\mathcal{C}}(1)}{1-\rho_c}\Big(T - \frac{1-\rho_c^T}{1-\rho_c}\Big)\right], \tag{A.92}$$

$$\phi_{\mathcal{C},T}^{(1)} = \frac{\frac{1}{T}\Big[(T-1)\,\gamma_{\mathcal{C}}(1) - (T+1)\,\mathrm{Var}\big(\bar{\mathcal{C}}_T\big)\Big] + \frac{2}{T^2}\Big[\gamma_{\mathcal{C}}(0) + \gamma_{\mathcal{C}}(1)\,\frac{1-\rho_c^{T-1}}{1-\rho_c}\Big]}{\gamma_{\mathcal{C}}(0) - \mathrm{Var}\big(\bar{\mathcal{C}}_T\big)}, \tag{A.93}$$

The annual version substitutes lag 4, using $\gamma_{\mathcal{C}}(4) = \rho_c^3\,\gamma_{\mathcal{C}}(1)$:

$$\phi_{\mathcal{C},T}^{(4)} = \frac{\frac{1}{T}\Big[(T-4)\,\rho_c^3\,\gamma_{\mathcal{C}}(1) - (T+4)\,\mathrm{Var}\big(\bar{\mathcal{C}}_T\big)\Big] + \frac{2}{T^2}\left[4\,\gamma_{\mathcal{C}}(0) + \frac{\gamma_{\mathcal{C}}(1)}{1-\rho_c}\Big(8 - (1+\rho_c)(1+\rho_c^2)\big(1+\rho_c^{T-4}\big)\Big)\right]}{\gamma_{\mathcal{C}}(0) - \mathrm{Var}\big(\bar{\mathcal{C}}_T\big)}. \tag{A.94}$$

D.3.1 Persistence of $\overline{pd}_t$

From D.1, we use the fact that, to a first order, $\overline{pd}(g_t) = \overline{pd}(\bar{g}) + \lambda w_t$ with $\lambda = \overline{pd}'(\bar{g}) = 1/(1-\rho)$ and $w_t = \sum_{s\leq t}\varepsilon_{g,s}$. Under the unit root the trend component has no stationary population autocorrelation, so only the window statistic is defined. We compute it for the linearized component; the loading λ cancels from the ratio.

Window Writing the lag-1 products as

$$(w_t - \bar{w}_T)(w_{t-1} - \bar{w}_T) = \tfrac{1}{2}(w_t - \bar{w}_T)^2 + \tfrac{1}{2}(w_{t-1} - \bar{w}_T)^2 - \tfrac{1}{2}(w_t - w_{t-1})^2 \tag{A.95}$$

and summing over $t = 2, \ldots, T$,

$$T\,\widehat{\mathrm{Cov}}_T(w_t, w_{t-1}) = T\,\widehat{\mathrm{Var}}_T(w_t) - \tfrac{1}{2}\big[(w_1 - \bar{w}_T)^2 + (w_T - \bar{w}_T)^2\big] - \tfrac{1}{2}\sum_{t=2}^{T}\varepsilon_{g,t}^2. \tag{A.96}$$

Taking expectations, using $\mathrm{Var}_T(w_t) = \sigma_g^2(T^2-1)/6T$, $\mathbb{E}\big[(w_1-\bar{w}_T)^2\big] = \mathbb{E}\big[(w_T-\bar{w}_T)^2\big] = \sigma_g^2(2T-1)(T-1)/6T$, and $\mathbb{E}\big[\sum_{t=2}^{T}\varepsilon_{g,t}^2\big] = \sigma_g^2(T-1)$,

$$\phi_{\overline{pd},T}^{(1)} = \frac{T^2-4T+1}{T\,(T+1)}. \tag{A.97}$$

The same calculation at lag h gives

$$\phi_{\overline{pd},T}^{(h)} = 1 - \frac{h\left(5T^2-6hT+2h^2-1\right)}{T\left(T^2-1\right)}, \tag{A.98}$$

So at lag 4,

$$\phi_{\overline{pd},T}^{(4)} = 1 - \frac{4\left(5T^2-24T+31\right)}{T\left(T^2-1\right)}. \tag{A.99}$$

E Return Predictability – Derivations

We run predictive regressions of $r_{t,t+j} \equiv \sum_{k=1}^{j} \rho^{k-1} r_{t+k}$ on the filtered states. The derivations below use conditional expected returns and the true states; the regressions use realized returns and filtered states, and the bootstrap accounts for the difference. From equation (4),

$$\mathbb{E}_t[r_{t,t+j}] \;=\; \frac{1-\rho^j}{1-\rho}\,\bar{\mu} \;-\; q_j\,\mathcal{M}_t, \qquad q_j \equiv 1-(\rho\rho_\mu)^j. \tag{A.100}$$

E.1 Coefficient of a Regression of $\mathbb{E}_t[r_{t,t+j}]$ on pd_t

With the window moments Cov_T and Var_T of Appendix D, define

$$b^r_{j,T} \;\equiv\; \frac{\mathrm{Cov}_T\big(\mathbb{E}_t[r_{t,t+j}],\, pd_t\big)}{\mathrm{Var}_T(pd_t)}. \tag{A.101}$$

The model period is a quarter, so T measures the regression window in quarters; the regressions themselves use annual observations. The forecast (A.100) loads only on $\mathcal{M}_t$, and the components of pd_t in (10) are driven by independent shocks, so

$$b^r_{j,T} \;=\; -\,q_j\,b_{\mathcal{M},T}, \qquad b_{\mathcal{M},T} \equiv \frac{\mathrm{Cov}_T(\mathcal{M}_t,\, pd_t)}{\mathrm{Var}_T(pd_t)}. \tag{A.102}$$

Cash flows do not forecast returns, so they are absent from the numerator. They attenuate the coefficient through the denominator: $b_{\mathcal{M},T} = 1 - b_{\overline{pd},T} - b_{\mathcal{C},T}$. The univariate slope is the expected-return slope $-q_j$ scaled by the expected-return share of pd_t variation.

E.2 Coefficient of a Regression of $\mathbb{E}_t[r_{t,t+j}]$ on pd_t, $\overline{pd}_t$, c_t, and $\varepsilon_{c,t}$

No moments are needed here. Solving the decomposition (10) for $\mathcal{M}_t$ and substituting into (A.100):

$$\mathbb{E}_t[r_{t,t+j}] \;=\; \text{const} \;-\; q_j\,pd_t \;+\; q_j\,\overline{pd}(g_t) \;+\; q_j\,\beta_c\,c_t \;+\; q_j\,\beta_\varepsilon\,\varepsilon_{c,t}. \tag{A.103}$$

Once the cash-flow components are controlled, pd_t identifies $\mathcal{M}_t$ algebraically, and its coefficient is the full slope $-q_j$.

E.3 Coefficient of a Regression of $\mathbb{E}_t[r_{t,t+j}]$ on $\mathcal{M}_t$

In the regression on $\mathcal{M}_t$, the slope is $-q_j$. Adding the cash-flow states changes nothing – expected returns do not load on them – so their model-implied coefficients are exactly zero. These loadings involve only the model's dynamic parameters and do not depend on the sample.

Implementation in the regression tables The tables use realized returns and the historical filtered states. In every specification, the reported model-implied coefficient is the window-T coefficient above, evaluated at the estimated parameters and, where weights enter, at the historical variance shares (A.120). For the univariate regression,

$$\widehat{b}^r_{j,T} \;=\; -\,\widehat{q}_j\,\widehat{b}_{\mathcal{M},T}. \tag{A.104}$$

The controlled and $\mathcal{M}_t$ regressions involve no weights; their model-implied coefficients are the slopes of (A.103) and (A.100) at the estimated parameters. The bootstrap provides the finite-sample distribution of the empirical coefficients around these benchmarks.

F Cash Flow Predictability – Derivations

F.1 Regression with trend growth and cash-flow cycle

We run a predictive regression of $\Delta d_{t,t+j} \equiv \sum_{k=1}^{j} \rho^{k-1} \Delta d_{t+k}$ on g_t, c_t and $\varepsilon_{c,t}$. In the following, we derive the model-implied regression coefficients.

From the cash-flow specification, one-period log dividend growth satisfies

$$\Delta d_{t+k} = g_{t+k-1} + \big(c_{t+k} - c_{t+k-1}\big), \tag{A.105}$$

where g_t is a random walk ($g_{t+1} = g_t + \varepsilon_{g,t+1}$, $\varepsilon_{g,t} \sim \mathcal{N}(0, \sigma_g^2)$) and c_t follows an ARMA(1,1) ($c_{t+1} = \rho_c\, c_t + \theta_c\, \varepsilon_{c,t} + \varepsilon_{c,t+1}$, $\varepsilon_{c,t} \sim \mathcal{N}(0, \sigma_c^2)$).

This implies

$$\Delta d_{t,t+j} = \underbrace{\sum_{k=1}^{j} \rho^{k-1}\, g_{t+k-1}}_{\text{trend component}} + \underbrace{\sum_{k=1}^{j} \rho^{k-1} \big(c_{t+k} - c_{t+k-1}\big)}_{\text{cycle component}}. \tag{A.106}$$

We evaluate the conditional expectation of each component in turn.

Trend component. Since g_t is a random walk, $\mathbb{E}_t[g_{t+k-1}] = g_t$ for all $k \geq 1$. Therefore:

$$\mathbb{E}_t\left[\sum_{k=1}^{j} \rho^{k-1}\, g_{t+k-1}\right] = g_t \sum_{k=1}^{j} \rho^{k-1} = g_t\, \frac{1-\rho^j}{1-\rho}. \tag{A.107}$$

Cycle component. We first simplify the weighted sum algebraically

$$\sum_{k=1}^{j} \rho^{k-1} \big(c_{t+k} - c_{t+k-1}\big) = \sum_{k=1}^{j} \rho^{k-1}\, c_{t+k} \;-\; \sum_{k=1}^{j} \rho^{k-1}\, c_{t+k-1}. \tag{A.108}$$

Re-index the second sum by setting $\ell = k - 1$:

$$\sum_{k=1}^{j} \rho^{k-1}\, c_{t+k-1} = \sum_{\ell=0}^{j-1} \rho^{\ell}\, c_{t+\ell} = c_t + \sum_{\ell=1}^{j-1} \rho^{\ell}\, c_{t+\ell}.$$

Subtracting:

$$\begin{aligned}
\sum_{k=1}^{j} \rho^{k-1} \big(c_{t+k} - c_{t+k-1}\big) &= \sum_{k=1}^{j-1} \rho^{k-1}\, c_{t+k} + \rho^{j-1}\, c_{t+j} - c_t - \sum_{k=1}^{j-1} \rho^{k}\, c_{t+k} \\
&= -\,c_t + \sum_{k=1}^{j-1} \big(\rho^{k-1} - \rho^{k}\big)\, c_{t+k} + \rho^{j-1}\, c_{t+j} \\
&= -\,c_t + (1-\rho) \sum_{k=1}^{j-1} \rho^{k-1}\, c_{t+k} + \rho^{j-1}\, c_{t+j}.
\end{aligned} \tag{A.109}$$

Conditional expectation of the cycle sum. Using the ARMA(1,1) forecast (A.40), namely $\mathbb{E}_t[c_{t+h}] = \rho_c^h\, c_t + \rho_c^{h-1}\,\theta_c\,\varepsilon_{c,t}$ for $h \geq 1$, and letting $\varphi \equiv \rho\rho_c$, we evaluate the three terms in (A.109).

Interior sum.

$$\begin{aligned}(1-\rho)\sum_{k=1}^{j-1}\rho^{k-1}\,\mathbb{E}_t[c_{t+k}] &= (1-\rho)\sum_{k=1}^{j-1}\rho^{k-1}\Big[\rho_c^k\, c_t + \rho_c^{k-1}\,\theta_c\,\varepsilon_{c,t}\Big] \\ &= (1-\rho)\,\rho_c\, c_t \sum_{k=1}^{j-1}\varphi^{k-1} \;+\; (1-\rho)\,\theta_c\,\varepsilon_{c,t}\sum_{k=1}^{j-1}\varphi^{k-1} \\ &= (1-\rho)\,\rho_c\,\frac{1-\varphi^{j-1}}{1-\varphi}\,c_t \;+\; (1-\rho)\,\frac{1-\varphi^{j-1}}{1-\varphi}\,\theta_c\,\varepsilon_{c,t}. \end{aligned} \tag{A.110}$$

Terminal term.

$$\rho^{j-1}\,\mathbb{E}_t[c_{t+j}] = \rho^{j-1}\,\rho_c^j\, c_t \;+\; \rho^{j-1}\,\rho_c^{j-1}\,\theta_c\,\varepsilon_{c,t}. \tag{A.111}$$

Assembling. Combining (A.110) and (A.111) with the $-c_t$ term from (A.109):

$$\begin{aligned}\mathbb{E}_t\left[\sum_{k=1}^{j}\rho^{k-1}\big(c_{t+k}-c_{t+k-1}\big)\right] &= \underbrace{\left[-1+(1-\rho)\,\rho_c\,\frac{1-\varphi^{j-1}}{1-\varphi}+\rho^{j-1}\rho_c^j\right]}_{\equiv\; A_j^\rho}\; c_t \\ &\quad+\; \underbrace{\left[(1-\rho)\,\frac{1-\varphi^{j-1}}{1-\varphi}+\rho^{j-1}\rho_c^{j-1}\right]}_{\equiv\; B_j^\rho}\;\theta_c\,\varepsilon_{c,t}. \end{aligned} \tag{A.112}$$

Adding the trend component (A.107) and the cycle component (A.112):

$$\mathbb{E}_t\big[\Delta d_{t,t+j}\big] \;=\; \frac{1-\rho^j}{1-\rho}\,g_t \;+\; A_j^\rho\, c_t \;+\; B_j^\rho\,\theta_c\,\varepsilon_{c,t}, \tag{A.113}$$

where $\varphi \equiv \rho\rho_c$ and

$$A_j^\rho = -1+(1-\rho)\,\rho_c\,\frac{1-\varphi^{j-1}}{1-\varphi}+\rho^{j-1}\rho_c^j < 0, \qquad B_j^\rho = (1-\rho)\,\frac{1-\varphi^{j-1}}{1-\varphi}+\rho^{j-1}\rho_c^{j-1} > 0. \tag{A.114}$$

The sign of B_j^ρ is immediate. For A_j^ρ, note $\rho^{j-1}\rho_c^j = \varphi^{j-1}\rho_c$, so $A_j^\rho = \rho_c\, B_j^\rho - 1$; since $1-\rho < 1-\varphi$ implies $B_j^\rho \leq 1$, we have $A_j^\rho \leq \rho_c - 1 < 0$. The loadings in (A.113) involve only the model's dynamic parameters and do not depend on the sample.

F.2 Coefficient of a Regression of $\mathbb{E}_t[\Delta d_{t,t+j}]$ on pd_t

With the window moments Cov_T and Var_T of Appendix D, define

$$b_{j,T}^d \;\equiv\; \frac{\mathrm{Cov}_T\big(\mathbb{E}_t[\Delta d_{t,t+j}],\, pd_t\big)}{\mathrm{Var}_T(pd_t)}. \tag{A.115}$$

The model period is a quarter, so T here measures the regression window in quarters. The components of pd_t in (10) are driven by mutually independent shocks, so cross moments vanish and expected returns drop from the numerator:

$$\mathrm{Cov}_T\big(\mathbb{E}_t[\Delta d_{t,t+j}],\, pd_t\big) = \mathrm{Cov}_T\big(\mathbb{E}_t[\Delta d_{t,t+j}],\, \overline{pd}(g_t)\big) + \mathrm{Cov}_T\big(\mathbb{E}_t[\Delta d_{t,t+j}],\, \mathcal{C}_t\big). \tag{A.116}$$

The coefficient is therefore a variance-weighted average of two component slopes, one for the trend and one for the cycle. We derive each, then assemble.

Trend slope Linearize $\overline{pd}(g_t)$ around $g_0 \equiv \bar{\mu} + \log\rho$ as in the derivation of (A.70), $\overline{pd}(g_t) \approx \text{const} + \lambda\,(g_t - g_0)$ with $\lambda = 1/(1-\rho)$. The forecast (A.113) loads on g_t with $\frac{1-\rho^j}{1-\rho}$, so

$$\mathrm{Cov}_T\big(\mathbb{E}_t[\Delta d_{t,t+j}],\, \overline{pd}(g_t)\big) \;\approx\; \frac{1-\rho^j}{1-\rho}\,\lambda\,\mathrm{Var}_T(g_t) \;=\; (1-\rho^j)\,\mathrm{Var}_T\big(\overline{pd}(g_t)\big),$$

using $\mathrm{Var}_T(\overline{pd}(g_t)) \approx \lambda^2\mathrm{Var}_T(g_t)$, i.e. (A.70). The λ's cancel in the slope: the trend slope is $1-\rho^j$, whatever the window.

Cycle slope The cycle states are orthogonal to g_t and μ_t, so only the cycle part of the forecast, $A_j^\rho\, c_t + B_j^\rho\,\theta_c\,\varepsilon_{c,t}$, covaries with $\mathcal{C}_t = \beta_c\, c_t + \beta_\varepsilon\,\varepsilon_{c,t}$ of Appendix D.1. Using the ARMA(1,1) moments $\mathrm{Var}(c_t) = \sigma_c^2\, v_c$ with $v_c \equiv (1+\theta_c^2+2\rho_c\theta_c)/(1-\rho_c^2)$, $\mathrm{Var}(\varepsilon_{c,t}) = \sigma_c^2$, and $\mathrm{Cov}(c_t, \varepsilon_{c,t}) = \sigma_c^2$, the contemporaneous covariance is

$$\mathrm{Cov}\big(A_j^\rho\, c_t + B_j^\rho\,\theta_c\,\varepsilon_{c,t},\; \mathcal{C}_t\big) \;=\; \sigma_c^2\Big[A_j^\rho\big(\beta_c\, v_c + \beta_\varepsilon\big) + B_j^\rho\,\theta_c\big(\beta_c + \beta_\varepsilon\big)\Big]. \tag{A.117}$$

The window moment adds lead–lag terms, yielding the finite sample estimator $a_{c,j,T}$. At the estimated parameters and sample lengths the two differ by less than 10^{-3}, and so we evaluate $a_{c,j,T}$ at the population slope,

$$a_{c,j} \;=\; \frac{A_j^\rho\big(\beta_c\, v_c + \beta_\varepsilon\big) + B_j^\rho\,\theta_c\big(\beta_c + \beta_\varepsilon\big)}{\beta_c^2\, v_c + \beta_\varepsilon^2 + 2\,\beta_c\,\beta_\varepsilon};$$

Assembling. Divide (A.116) by $\mathrm{Var}_T(pd_t) = \mathrm{Var}_T\big(\overline{pd}(g_t)\big) + \mathrm{Var}_T(\mathcal{M}_t) + \mathrm{Var}_T(\mathcal{C}_t)$ from Appendix D.1:

$$b_{j,T}^d \;\approx\; b_{\overline{pd},T}\,\big(1-\rho^j\big) \;+\; b_{\mathcal{C},T}\; a_{c,j,T}, \tag{A.118}$$

where $b_{X,T} = \mathrm{Cov}_T(X_t, pd_t)/\mathrm{Var}_T(pd_t)$. Under orthogonality, we have $b_{X,T} = \mathrm{Var}_T(X_t)/\mathrm{Var}_T(pd_t)$ and they correspond to the window variance shares. Expected returns do not forecast dividend growth, so $\mathcal{M}_t$ is absent from the numerator. It attenuates the coefficient through the denominator: $b_{\overline{pd},T} + b_{\mathcal{C},T} = 1 - b_{\mathcal{M},T}$.

Terminal pd The terminal-pd coefficient follows from an identity. Iterating the loglinear identity over j periods,

$$pd_t \;\approx\; \kappa_j + \Delta d_{t,t+j} - r_{t,t+j} + \rho^j\, pd_{t+j}, \qquad r_{t,t+j} \equiv \sum_{k=1}^{j} \rho^{\,k-1}\, r_{t+k},$$

and projecting each side on pd_t over the same sample gives $1 = b_j^d - b_j^r + b_j^{pd}$, so

$$b_j^{pd} \;=\; 1 - b_j^d + b_j^r. \tag{A.119}$$

Implementation in the regression tables. The tables use realized dividend growth and the historical filtered states. The reported model-implied coefficient conditions on the historical composition of pd_t. Define the contribution of component $X_t \in \{\overline{pd}(g_t),\, \mathcal{C}_t,\, \mathcal{M}_t\}$ over the regression sample as

$$\widehat{b}_{X,T} \;\equiv\; \frac{\widehat{\mathrm{Cov}}_T\big(\widehat{X}_t,\, pd_t\big)}{\widehat{\mathrm{Var}}_T(pd_t)}, \tag{A.120}$$

This is the realized window-anologue of $b_{X,T}$. Under the model's component-orthogonality restrictions, the corresponding ratio of expected window moments reduces to the variance share $\mathrm{Var}_T(X_t)/\mathrm{Var}_T(pd_t)$. Because pd_t sums to its components, the contributions sum to one exactly. The model-implied coefficient applies the component slopes to these weights,

$$\widehat{b}_{j,T}^{\,d} \approx \widehat{b}_{\overline{pd},T}\,\big(1-\rho^j\big) \;+\; \widehat{b}_{\mathcal{C},T}\;\hat{a}_{c,j,T}, \tag{A.121}$$

evaluated at the estimated parameters. The expected-return component has slope zero, so it enters only through the denominator: $\widehat{b}_{\overline{pd},T} + \widehat{b}_{\mathcal{C},T} = 1 - \widehat{b}_{\mathcal{M},T}$, exactly. The analytical coefficient combines the model-implied trend and cycle slopes with the variance contributions of those components to pd_t. The empirical regression instead uses realized dividend growth. Its coefficient may differ from the analytical benchmark because of realized long-horizon cash-flow shocks, overlapping outcomes, finite-sample cross-component covariances, and parameter and filtering uncertainty. The bootstrap reproduces the complete estimation and regression procedure and provides the finite-sample distribution of the empirical coefficient relative to the analytical benchmark.

G A Stationary Present-Value Equation from Dividend Strips

This appendix derives a present-value measurement equation for pd_t under a stationary AR(1) trend. The equation is valid for every value of the current trend-growth state, including $g_t \geq \bar{\mu}$, provided long-run growth lies below the steady-state expected return, $\bar{g} < \bar{\mu}$. It solves the model both for the constant campbell–shiller weight, ρ, and for state-dependent strip weights.

G.1 Setup

Log dividends are $d_t = \tau_t + c_t$, with trend $\tau_t = \tau_{t-1} + g_{t-1}$. Trend growth, the cash-flow cycle, and expected returns follow

$$g_t = \bar{g} + \rho_g(g_{t-1} - \bar{g}) + \varepsilon_{g,t}, \qquad c_t = \rho_c c_{t-1} + \theta_c \varepsilon_{c,t-1} + \varepsilon_{c,t}, \qquad \mu_t = \bar{\mu} + \rho_\mu(\mu_{t-1} - \bar{\mu}) + \varepsilon_{\mu,t}, \tag{A.122}$$

with mutually uncorrelated Gaussian shocks and $\rho_g, \rho_c, \rho_\mu \in (0,1)$. We maintain $\bar{g} < \bar{\mu}$.

G.2 The price–dividend ratio as a strip sum

Under the usual no-bubble condition, the stock is a portfolio of claims to its individual future dividends. Let $P_t^{(j)}$ be the time-t price of a claim to the single dividend D_{t+j}—the j-period dividend strip. Then

$$P_t = \sum_{j=1}^{\infty} P_t^{(j)}, \qquad PD_t \equiv \frac{P_t}{D_t} = \sum_{j=1}^{\infty} \frac{P_t^{(j)}}{D_t}. \tag{A.123}$$

Let $R_{t,t+j}^{(j)}$ be the cumulative log return from buying strip j at t and holding it to maturity. Its terminal payoff is the dividend itself, so the return is the payoff over the price,

$$R_{t,t+j}^{(j)} = d_{t+j} - p_t^{(j)}, \qquad p_t^{(j)} \equiv \log P_t^{(j)}. \tag{A.124}$$

Rearranging, subtracting d_t, and using the fact that cumulative log dividend growth telescopes,

$$\log\left(\frac{P_t^{(j)}}{D_t}\right) = \sum_{i=1}^{j} \Delta d_{t+i} - R_{t,t+j}^{(j)}. \tag{A.125}$$

This holds path by path. Taking time-t expectations:

$$\log\left(\frac{P_t^{(j)}}{D_t}\right) = \mathbb{E}_t\left[\sum_{i=1}^{j} \Delta d_{t+i}\right] - \mathbb{E}_t\left[R_{t,t+j}^{(j)}\right]. \tag{A.126}$$

Thus, the log strip price relative to the current dividend equals expected cumulative log dividend growth less the strip's expected cumulative log return.

The cumulative strip return is the sum of its one-period returns:

$$R_{t,t+j}^{(j)} = \sum_{i=1}^{j} r_{t+i-1,t+i}^{(j-i+1)}. \tag{A.127}$$

Assume every strip earns the same expected return over the next period, whatever its remaining maturity,

$$\mathbb{E}_s\left[r^{(m)}_{s,s+1}\right] = \mathbb{E}_s[r_{s,s+1}], \qquad m \geq 1. \tag{A.128}$$

This is an expectations-hypothesis-style restriction for dividend strips. By the law of iterated expectations,

$$\mathbb{E}_t\left[R^{(j)}_{t,t+j}\right] = \mathbb{E}_t\left[\sum_{i=1}^{j} r_{t+i-1,t+i}\right]. \tag{A.129}$$

Substituting into equation (A.126), exponentiating, and summing across strips gives

$$PD_t = \sum_{j=1}^{\infty} \exp\left[\mathbb{E}_t\left(\textstyle\sum_{i=1}^{j} \Delta d_{t+i}\right) - \mathbb{E}_t\left(\textstyle\sum_{i=1}^{j} r_{t+i-1,t+i}\right)\right], \qquad pd_t = \log PD_t. \tag{A.130}$$

We split each strip's exponent into a trend part and a cyclical deviation. One-period log dividend growth is

$$\Delta d_{t+i} = g_{t+i-1} + (c_{t+i} - c_{t+i-1}),$$

so cumulated growth is

$$\sum_{i=1}^{j} \Delta d_{t+i} = \sum_{i=1}^{j} g_{t+i-1} + (c_{t+j} - c_t). \tag{A.131}$$

The cyclical changes telescope to the endpoint term. Taking expectations,

$$\mathbb{E}_t\left[\sum_{i=1}^{j} g_{t+i-1}\right] = j\bar{g} + \Lambda_j(g_t - \bar{g}), \qquad \Lambda_j \equiv \sum_{i=0}^{j-1} \rho_g^i = \frac{1-\rho_g^j}{1-\rho_g}.$$

The return term is

$$\mathbb{E}_t\left[\sum_{i=1}^{j} r_{t+i-1,t+i}\right] = j\bar{\mu} + \sum_{i=1}^{j} \mathbb{E}_t[r_{t+i-1,t+i} - \bar{\mu}].$$

The strip sum can therefore be written as

$$PD_t = \sum_{j=1}^{\infty} \exp(a_j + x_j), \qquad a_j = j(\bar{g} - \bar{\mu}) + \Lambda_j(g_t - \bar{g}), \qquad \Lambda_j = \frac{1-\rho_g^j}{1-\rho_g}, \tag{A.132}$$

where a_j is the value of strip j when the cycles are off, $\mu_t = \bar{\mu}$ and $c_t = \varepsilon_{c,t} = 0$. The cyclical deviation is

$$x_j = \underbrace{\left(\mathbb{E}_t[c_{t+j}] - c_t\right)}_{\text{cash-flow cycle}} - \underbrace{\sum_{i=1}^{j} \mathbb{E}_t[r_{t+i-1,t+i} - \bar{\mu}]}_{\text{expected-return cycle}}. \tag{A.133}$$

G.3 The measurement equation

Write $A_j \equiv e^{a_j}$ for the trend strip, so that

$$pd_t = \log \sum_j A_j e^{x_j}.$$

The deviation x_j is zero when the cycles are off. We expand the cycle factor around $x_j = 0$,

$$e^{x_j} = 1 + x_j + O(x_j^2), \tag{A.134}$$

and keep the linear term. Substituting (A.134),

$$pd_t = \log \sum_j A_j e^{x_j} \approx \log \sum_j A_j (1 + x_j) = \log \Big[\Big(\textstyle\sum_j A_j \Big) \Big(1 + \textstyle\sum_j \omega_j x_j \Big) \Big], \qquad \omega_j \equiv \tfrac{A_j}{\sum_k A_k}. \tag{A.135}$$

The final equality is exact algebra after the first-order expansion. The weights ω_j are the shares of the no-cycle price-dividend ratio accounted for by each strip. They are non-negative and sum to one.

Splitting the logarithm and using $\log(1 + u) \approx u$ for $u = \sum_j \omega_j x_j$ gives

$$pd_t \approx \log \sum_j A_j + \sum_j \omega_j x_j. \tag{A.136}$$

The first term defines the stationary trend anchor:

$$F_{\rho_g}(g_t) \equiv \log \sum_{j=1}^{\infty} \exp\big[j(\bar{g} - \bar{\mu}) + \Lambda_j (g_t - \bar{g}) \big], \qquad \Lambda_j = \frac{1 - \rho_g^j}{1 - \rho_g}. \tag{A.137}$$

This anchor is finite for every current value of g_t as long as $\bar{g} < \bar{\mu}$. A high current g_t eventually mean-reverts toward $\bar{g}$, so its effect on far-dated dividends is bounded.

The second term in (A.136) is a strip-share-weighted average of the expected-return and cash-flow-cycle deviations. The shares depend on the current trend state through A_j, so we write them as $\omega_{j,t} \equiv \omega_j(g_t)$.

Expected-return block. Since $\mathbb{E}_t[r_{t+i-1,t+i}] = \mathbb{E}_t[\mu_{t+i-1}]$ and μ_t follows an AR(1),

$$\mathbb{E}_t[r_{t+i-1,t+i} - \bar{\mu}] = \rho_\mu^{i-1} (\mu_t - \bar{\mu}).$$

Thus

$$\sum_{i=1}^{j} \mathbb{E}_t[r_{t+i-1,t+i} - \bar{\mu}] = \Lambda_j^\mu (\mu_t - \bar{\mu}), \qquad \Lambda_j^\mu \equiv \sum_{i=0}^{j-1} \rho_\mu^i = \frac{1 - \rho_\mu^j}{1 - \rho_\mu}. \tag{A.138}$$

Weighting by the strip shares, the expected-return block is

$$\mathcal{M}_t^F = -A_\mu^F(g_t)(\mu_t - \bar{\mu}), \qquad A_\mu^F(g_t) \equiv \sum_{j=1}^{\infty} \omega_{j,t} \Lambda_j^\mu. \tag{A.139}$$

Cash-flow-cycle block. The ARMA(1,1) forecast is

$$\mathbb{E}_t[c_{t+j}] = \rho_c^j c_t + \rho_c^{j-1}\theta_c \varepsilon_{c,t}.$$

The endpoint deviation in (A.133) is therefore

$$\mathbb{E}_t[c_{t+j}] - c_t = -(1-\rho_c)\Lambda_j^c c_t + \rho_c^{j-1}\theta_c\varepsilon_{c,t}, \qquad \Lambda_j^c \equiv \frac{1-\rho_c^j}{1-\rho_c}. \tag{A.140}$$

Weighting by the strip shares, the cash-flow-cycle block is

$$\mathcal{C}_t^F = -A_c^F(g_t)c_t + A_\varepsilon^F(g_t)\varepsilon_{c,t}, \qquad A_c^F(g_t) \equiv (1-\rho_c)\sum_{j=1}^{\infty}\omega_{j,t}\Lambda_j^c, \qquad A_\varepsilon^F(g_t) \equiv \theta_c\sum_{j=1}^{\infty}\omega_{j,t}\rho_c^{j-1}. \tag{A.141}$$

The Jensen term. The baseline measurement equation carries a scalar Jensen term. The stationary analogue applies the same second-order correction to the trend anchor F_{ρ_g} around its conditional mean, discounted at the steady-state strip ratio $\bar\rho$ and evaluated at $g_t = \bar g$:

$$\delta_{\rho_g} = \frac{1}{2}\frac{\bar\rho}{1-\bar\rho}F''_{\rho_g}(\bar g)\sigma_g^2, \qquad F''_{\rho_g}(\bar g) = \mathrm{Var}_\omega(\Lambda_j), \tag{A.142}$$

where the variance is taken under the steady-state strip weights $\omega_j(\bar g) = (1-\bar\rho)\bar\rho^{j-1}$.[1] The correction nests the baseline. As $\rho_g \to 1$, $\Lambda_j \to j$, so

$$F''_{\rho_g}(\bar g) \to \mathrm{Var}_\omega(j) = \frac{\bar\rho}{(1-\bar\rho)^2},$$

and

$$\delta_{\rho_g} \to \frac{1}{2}\frac{\bar\rho^2}{(1-\bar\rho)^3}\sigma_g^2,$$

the baseline Jensen term evaluated at $\rho = \bar\rho$.

The stationary-strip measurement equation is therefore

$$pd_t \approx \delta_{\rho_g} + F_{\rho_g}(g_t) + \mathcal{M}_t^F + \mathcal{C}_t^F, \tag{A.143}$$

with $\mathcal{M}_t^F$ and $\mathcal{C}_t^F$ given by (A.139) and (A.141). The expected-return and cash-flow-cycle loadings are state-dependent strip-share-weighted averages. The only fixed discount object is the steady-state strip ratio $\bar\rho$ inside the scalar Jensen correction.

G.4 Two special cases

The loadings nest the familiar constant-coefficient blocks. At the steady state $g_t = \bar g$, the strip ratio A_{j+1}/A_j collapses to $\bar\rho$, the shares become geometric, $\omega_j(\bar g) = (1-\bar\rho)\bar\rho^{\,j-1}$, and the loadings reduce to

[1] In closed form,

$$F''_{\rho_g}(\bar g) = (1-\rho_g)^{-2}\left[\frac{(1-\bar\rho)\rho_g^2}{1-\bar\rho\rho_g^2} - \left(\frac{(1-\bar\rho)\rho_g}{1-\bar\rho\rho_g}\right)^2\right].$$

the constant-ρ Campbell–Shiller blocks evaluated at $\rho = \bar{\rho}$:

$$A_\mu^F(\bar{g}) = \frac{1}{1-\bar{\rho}\rho_\mu}, \qquad A_c^F(\bar{g}) = \frac{1-\rho_c}{1-\bar{\rho}\rho_c}, \qquad A_\varepsilon^F(\bar{g}) = \frac{(1-\bar{\rho})\theta_c}{1-\bar{\rho}\rho_c}. \tag{A.144}$$

The constant-ρ variant The constant-ρ stationary model reported in Section 7.3 keeps the moving anchor $F_{\rho_g}(g_t)$ but freezes the cycle loadings and the Jensen term at the baseline constant $\rho = 0.99$ rather than at the steady-state ratio $\bar{\rho}$ of (A.144). The cyclical blocks $\mathcal{M}_t$ and $\mathcal{C}_t$ therefore take the same form as in the baseline model, and the specification differs in one place only—the trend law. The parameters themselves are re-estimated. The approximation of $\bar{\rho}$ by 0.99 is small—$\bar{\rho} = 0.988$ at the estimated parameters.

The unit-root limit. As $\rho_g \to 1$, $\Lambda_j \to j$ and the shares become geometric at the state-dependent ratio $\rho_t = \exp(g_t - \bar{\mu})$—the dynamic coefficient of Section 3. Away from these two limits, the exact objects are the strip shares $\omega_{j,t}$.

G.5 The valuation slope and the loading sensitivities

Differentiating (A.137),

$$F'_{\rho_g}(g_t) = \sum_{j=1}^{\infty} \omega_{j,t}\Lambda_j, \qquad \omega_{j,t} = \frac{\exp[\, j(\bar{g}-\bar{\mu}) + \Lambda_j(g_t - \bar{g})\,]}{\sum_{k\geq 1}\exp[\, k(\bar{g}-\bar{\mu}) + \Lambda_k(g_t-\bar{g})\,]}. \tag{A.145}$$

The slope of the trend anchor is the strip-share-weighted average of the trend multipliers Λ_j.

The cycle loadings are strip-share-weighted averages of fixed horizon profiles. Their sensitivities to the trend state share one form. Since $\partial a_j / \partial g_t = \Lambda_j$,

$$\frac{\partial \omega_{j,t}}{\partial g_t} = \omega_{j,t}\big(\Lambda_j - F'_{\rho_g}(g_t)\big). \tag{A.146}$$

Therefore, for any loading

$$A_L(g_t) = \sum_j \omega_{j,t} L_j,$$

where the horizon profile L_j does not depend on g_t,

$$A'_L(g_t) = \mathrm{Cov}_{\omega_t}(\Lambda_j, L_j). \tag{A.147}$$

The slope of the trend anchor is the strip-weighted mean of the multipliers. Its curvature is their variance. The sensitivity of each cycle loading is the covariance of the trend multiplier with that loading's horizon profile. In particular,

$$\big(A_\mu^F\big)'(g_t) = \mathrm{Cov}_{\omega_t}(\Lambda_j, \Lambda_j^\mu) > 0, \qquad \big(A_c^F\big)'(g_t) = (1-\rho_c)\mathrm{Cov}_{\omega_t}(\Lambda_j, \Lambda_j^c) > 0, \qquad \big(A_\varepsilon^F\big)'(g_t) = \theta_c \mathrm{Cov}_{\omega_t}(\Lambda_j, \rho_c^{j-1}). \tag{A.148}$$

The first two signs follow because Λ_j, Λ_j^μ, and Λ_j^c all rise with maturity. The last covariance is negative because ρ_c^{j-1} falls with maturity; for positive θ_c, the derivative $(A_\varepsilon^F)'$ is therefore negative as well. Higher

trend growth raises the duration of the trend anchor, the expected-return block, and the cycle-level block. It need not raise the loading on the impact cash-flow shock.

G.6 Estimation

The state vector is

$$s_t = (\tau_t, g_t, c_t, \varepsilon_{c,t}, \mu_t, 1)^\top,$$

with transition equation implied by (A.122). The measurement vector is

$$x_t = (d_t, pd_t)^\top.$$

The dividend equation $d_t = \tau_t + c_t$ is linear. The valuation equation (A.143) is nonlinear in g_t through $F_{\rho_g}(g_t)$ and through the state-dependent loadings. Conditional on g_t, it remains linear in $(c_t, \varepsilon_{c,t}, \mu_t)$.

To evaluate the likelihood with an Extended Kalman Filter, we replace the nonlinear measurement function each period by its first-order Taylor expansion at the prior state estimate $\hat{s}_{t|t-1}$. Let $h(s_t)$ denote the right-hand side of (A.143). The partial derivatives evaluated at the prior are

$$\frac{\partial h}{\partial c_t} = -A_c^F(\hat{g}_{t|t-1}), \qquad \frac{\partial h}{\partial \varepsilon_{c,t}} = A_\varepsilon^F(\hat{g}_{t|t-1}), \qquad \frac{\partial h}{\partial \mu_t} = -A_\mu^F(\hat{g}_{t|t-1}). \tag{A.149}$$

For the trend state, the derivative is the anchor slope plus the chain-rule terms from the moving loadings:

$$D_t \equiv \frac{\partial h}{\partial g_t} = F'_{\rho_g}(\hat{g}_{t|t-1}) - \left(A_c^F\right)'(\hat{g}_{t|t-1})\hat{c}_{t|t-1} + \left(A_\varepsilon^F\right)'(\hat{g}_{t|t-1})\hat{\varepsilon}_{c,t|t-1} - \left(A_\mu^F\right)'(\hat{g}_{t|t-1})\left(\hat{\mu}_{t|t-1} - \bar{\mu}\right), \tag{A.150}$$

with the derivatives given by (A.148). This Taylor approximation is used only to run the filter. It is distinct from the cycle expansion in Section G.3, which derives the measurement equation itself. Both the level and every slope are finite, so the linearization is executable at any state the filter visits, including $\hat{g}_{t|t-1} \geq \bar{\mu}$.

The linearized measurement system is $x_t = G_t s_t$, with

$$G_t = \begin{pmatrix} 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & D_t & -A_c^F(\hat{g}_{t|t-1}) & A_\varepsilon^F(\hat{g}_{t|t-1}) & -A_\mu^F(\hat{g}_{t|t-1}) & \kappa_t \end{pmatrix}, \tag{A.151}$$

and

$$\kappa_t = \delta_{\rho_g} + F_{\rho_g}(\hat{g}_{t|t-1}) + A_\mu^F(\hat{g}_{t|t-1})\bar{\mu} - D_t\hat{g}_{t|t-1}. \tag{A.152}$$

The matrix G_t is re-evaluated at the prior state estimate each period. Relative to the baseline filter, the second row now varies not only in the trend-anchor slope but also in the stationary-block loadings; through the chain-rule terms in D_t, it also depends on the prior cycle and expected-return states. The state vector is unchanged. The parameter count is unchanged.

We estimate the remaining parameters by maximum likelihood,

$$\Theta = (\rho_c, \theta_c, \sigma_g, \sigma_c, \rho_\mu, \sigma_\mu, \bar{\mu}).$$

We do not estimate ρ_g and $\bar{g}$ jointly with the rest. At near-unit-root values, ρ_g and $\bar{g}$ are weakly separated by the data, yet both enter valuation through F_{ρ_g}. Estimating them freely alongside σ_g and $\bar{\mu}$ over roughly 600 quarters produces flat likelihood ridges and unstable estimates. We therefore fix $\rho_g = 0.999$ and pin $\bar{g}$ at the sample mean. In the likelihood maximization, we impose $\bar{\mu} > \bar{g}$ so that the strip sum converges.

The infinite sums are evaluated numerically. Since $\Lambda_j \leq (1-\rho_g)^{-1}$ and $\bar{g} < \bar{\mu}$, the terms decay geometrically at long maturities. We truncate at $J = 10{,}000$. All sums are computed using log-sum-exp normalization:

$$\omega_{j,t} = \frac{\exp(a_{j,t} - m_t)}{\sum_{k=1}^{J} \exp(a_{k,t} - m_t)}, \qquad m_t = \max_{1 \leq k \leq J} a_{k,t}.$$

The loadings (A.139)–(A.141) and the sensitivities (A.148) are means and covariances under the same truncated strip shares. The Jensen correction δ_{ρ_g} is recomputed whenever $\bar{\mu}$ changes during likelihood maximization, since $\bar{\rho} = \exp(\bar{g} - \bar{\mu})$. These calculations add little computational burden beyond the sums already required for F_{ρ_g} and F'_{ρ_g}. At $\rho_g = 0.999$, and across the range of g_t visited by the filter, strips beyond $J = 10{,}000$ are immaterial.

H A Variance Decomposition of Returns

This appendix derives the variance decomposition of one- and multi-period returns, in the baseline and under stationary trend growth. Returns are stationary in both—differencing removes the unit root—so every object below has population moments, the baseline included.

H.1 Setup and conventions

One-period ahead dividend growth is given by

$$\Delta d_{t+1} = g_t + (c_{t+1} - c_t), \tag{A.153}$$

so the only surprise in Δd_{t+1} is $\varepsilon_{c,t+1}$. Second, we write the linearized valuation solution (10) as

$$pd_t = \text{const} + \Lambda\,(g_t - \bar{g}) + \mathcal{M}_t + \mathcal{C}_t, \qquad \mathcal{M}_t = -\frac{\mu_t - \bar{\mu}}{1 - \rho\rho_\mu}, \qquad \mathcal{C}_t = b_c\, c_t + b_e\, \varepsilon_{c,t}, \tag{A.154}$$

with $b_c = -(1-\rho_c)/(1-\rho\rho_c)$ and $b_e = (1-\rho)\theta_c/(1-\rho\rho_c)$ as in eq. (A.45) of Appendix B. The trend loading is $\Lambda = \lambda = 1/(1-\rho)$ in the baseline—the first-order expansion of $\overline{pd}(g_t)$ used throughout Appendix D.1—and $\Lambda = 1/(1-\rho\rho_g)$ under stationary trend growth, the slope of F_{ρ_g} in eq. (A.145).

H.2 One-period return news

Take the log-linear return identity, $r_{t,t+1} = \kappa + \rho\, pd_{t+1} - pd_t + \Delta d_{t+1}$, and subtract its time-$t$ expectation. The constant and pd_t drop out, leaving

$$r_{t,t+1} - \mu_t = \rho\,\big(pd_{t+1} - \mathbb{E}_t\, pd_{t+1}\big) + \big(\Delta d_{t+1} - \mathbb{E}_t\, \Delta d_{t+1}\big). \tag{A.155}$$

The return surprise comes from two places: the surprise in tomorrow's valuation and the surprise in tomorrow's dividend. This yields,

$$r_{t,t+1} - \mu_t = \Psi_g\, \varepsilon_{g,t+1} + \Psi_c\, \varepsilon_{c,t+1} - \Psi_\mu\, \varepsilon_{\mu,t+1}, \tag{A.156}$$

with

$$\Psi_g = \rho\,\Lambda, \qquad \Psi_\mu = \frac{\rho}{1-\rho\rho_\mu}, \qquad \Psi_c = 1 + \rho\,(b_c + b_e) = \frac{(1-\rho)(1+\rho\theta_c)}{1-\rho\rho_c}. \tag{A.157}$$

The last equality uses $(1-\rho\rho_c) - \rho(1-\rho_c) = 1-\rho$. In the baseline $\Psi_g = \rho\lambda = \rho/(1-\rho)$; under stationary trend growth $\Psi_g = \rho/(1-\rho\rho_g)$.

The same loadings come from news definitions. Discount-rate news is

$$N^{DR}_{t+1} = \sum_{j\geq 1} \rho^j\,(\mathbb{E}_{t+1} - \mathbb{E}_t)\, r_{t+j,t+j+1} = \varepsilon_{\mu,t+1} \sum_{j\geq 1} \rho^j \rho_\mu^{j-1} = \Psi_\mu\, \varepsilon_{\mu,t+1}. \tag{A.158}$$

Trend cash-flow news, using $(\mathbb{E}_{t+1} - \mathbb{E}_t)\,\Delta d_{t+1+j} = \rho_g^{j-1}\varepsilon_{g,t+1}$ for $j \geq 1$ and zero for $j = 0$ (by (A.153)), is

$$N^{CF,g}_{t+1} = \varepsilon_{g,t+1} \sum_{j\geq 1} \rho^j \rho_g^{j-1} = \frac{\rho}{1-\rho\rho_g}\, \varepsilon_{g,t+1} = \Psi_g\, \varepsilon_{g,t+1}, \tag{A.159}$$

and cycle news is $N_{t+1}^{CF,c} = \Psi_c\, \varepsilon_{c,t+1}$ by the telescoping argument of eqs. (A.38)–(A.44). So (A.156) is equivalently stated by $r_{t+1} - \mathbb{E}_t r_{t+1} = N_{t+1}^{CF,g} + N_{t+1}^{CF,c} - N_{t+1}^{DR}$, driven by the model's three orthogonal shocks. Assuming the shocks are independent, the one-period variance is given by:

$$\mathrm{Var}\big(r_{t,t+1} - \mu_t\big) = \Psi_g^2\sigma_g^2 + \Psi_c^2\sigma_c^2 + \Psi_\mu^2\sigma_\mu^2. \tag{A.160}$$

H.3 Multi-period returns

Each period realized return can be stated as,

$$r_{t+k-1,t+k} = \mathbb{E}_{t+k-1}\big[r_{t+k-1,t+k}\big] + v_{t+k}, \tag{A.161}$$

where v_{t+k} is the innovation (A.156). Sum over the window and subtract the date-t expectation; since $\mathbb{E}_t\, r_{t+k-1,t+k} = \mathbb{E}_t\big[\mathbb{E}_{t+k-1}[r_{t+k-1,t+k}]\big]$ by iterated expectations,

$$\sum_{k=1}^{h} r_{t+k-1,t+k} - \mathbb{E}_t \sum_{k=1}^{h} r_{t+k-1,t+k} = \sum_{k=1}^{h} v_{t+k} + \sum_{k=1}^{h} \Big(\mathbb{E}_{t+k-1}\big[r_{t+k-1,t+k}\big] - \mathbb{E}_t\big[r_{t+k-1,t+k}\big]\Big). \tag{A.162}$$

The forecast of a one-period return is the state μ, and only ε_μ moves it. Iterating (4),

$$\mathbb{E}_{t+k-1}\big[r_{t+k-1,t+k}\big] - \mathbb{E}_t\big[r_{t+k-1,t+k}\big] = \sum_{i=1}^{k-1} \rho_\mu^{k-1-i}\, \varepsilon_{\mu,t+i}. \tag{A.163}$$

The shocks are independent across type and date, so

$$\mathrm{Var}_t\Big(\sum_{k=1}^{h} r_{t+k-1,t+k}\Big) = h\,\Psi_g^2\sigma_g^2 + h\,\Psi_c^2\sigma_c^2 + \sigma_\mu^2 \sum_{m=0}^{h-1} A_\mu(m)^2, \tag{A.164}$$

where,

$$A_\mu(m) = \frac{1-\rho_\mu^m}{1-\rho_\mu} - \Psi_\mu, \qquad A_\mu(0) = -\Psi_\mu, \qquad A_\mu(\infty) = -\frac{1-\rho}{(1-\rho_\mu)(1-\rho\rho_\mu)}. \tag{A.165}$$

At $h = 1$ this is (A.160). As $h \to \infty$ the per-period contributions converge to $\Psi_g^2\sigma_g^2$, $\Psi_c^2\sigma_c^2$, and $\sigma_\mu^2 A_\mu(\infty)^2$.

H.4 Decomposition results

Table A.3: Variance decomposition of returns

	Baseline			Stationary trend growth		
Window	DR	CF_c	CF_g	DR	CF_c	CF_g
T=1	0.92	0.02	0.07	0.93	0.02	0.06
T=10	0.73	0.05	0.22	0.76	0.05	0.19
T=50	0.50	0.09	0.41	0.54	0.10	0.36
T=150	0.44	0.10	0.45	0.48	0.11	0.41

Note: This table decomposes the variance of unexpected returns cumulated over windows of length T years into the shares of discount-rate news (DR), cash-flow cycle news (CF_c), and trend-growth news (CF_g), following Appendix H. Both models use the estimated parameters (Table 3).

I Present-Value Model for Earnings

We collect the full structural model here for reference. Log earnings, e_t, are decomposed into a non-stationary trend, τ_t, and a stationary cycle, c_t. Expected returns, μ_t, follow an AR(1) process around a constant steady-state level $\bar{\mu}$:

$$\begin{aligned} e_t &= \tau_t + c_t, \\ \tau_t &= \tau_{t-1} + g_{t-1}, \\ g_t &= g_{t-1} + \varepsilon_{g,t}, \quad \varepsilon_{g,t} \sim \mathcal{N}(0, \sigma_g^2) \\ c_t &= \rho_c \, c_{t-1} + \theta_c \, \varepsilon_{c,t-1} + \varepsilon_{c,t}, \quad \varepsilon_{c,t} \sim \mathcal{N}(0, \sigma_c^2) \\ \mu_t &= \bar{\mu} + \rho_\mu \, (\mu_{t-1} - \bar{\mu}) + \varepsilon_{\mu,t}, \quad \varepsilon_{\mu,t} \sim \mathcal{N}(0, \sigma_\mu^2) \end{aligned} \tag{A.166}$$

The structural model is identical to the dividend-based baseline, with log earnings e_t replacing log dividends d_t. The trend growth rate g_t now captures trend earnings-per-share growth rather than dividend-per-share growth.

I.1 Present-Value Identity for the PE Ratio

We derive the present-value identity for the price–earnings ratio by starting from the log price–dividend ratio. In the baseline model, the present-value decomposition of the log price–dividend ratio is:

$$pd_t - \overline{pd}_t = \delta + \sum_{j=1}^{\infty} \rho^{j-1} \, \mathbb{E}_t\big[\Delta d_{t+j} - g_t\big] - \sum_{j=1}^{\infty} \rho^{j-1} \, \mathbb{E}_t\big[r_{t+j} - \bar{\mu}\big] \,, \tag{A.167}$$

where $\overline{pd}_t = g_t - \log\big(\exp(\bar{\mu}) - \exp(g_t)\big)$ is the steady-state log price–dividend ratio and δ is a Jensen's inequality term.

Define the log payout ratio as $de_t \equiv d_t - e_t$, so that $d_t = e_t + de_t$. We assume:

$$de_{t+j} = \overline{de} + v_{t+j}, \qquad v_{t+j} \overset{i.i.d.}{\sim} (0, \sigma_v^2), \quad \text{for all } j \geq 0, \tag{A.168}$$

where $\overline{de}$ is the steady-state log payout ratio. This assumption implies that payout policy is stationary around a constant level, so that earnings and dividends share the same non-stationary trend.

Since $p_t - e_t = (p_t - d_t) + (d_t - e_t) = pd_t + de_t$, the log price–earnings ratio is $pe_t = pd_t + de_t$. Substituting into (A.167):

$$pe_t - de_t - \overline{pd}_t = \delta + \sum_{j=1}^{\infty} \rho^{j-1} \, \mathbb{E}_t\big[\Delta d_{t+j} - g_t\big] - \sum_{j=1}^{\infty} \rho^{j-1} \, \mathbb{E}_t\big[r_{t+j} - \bar{\mu}\big] \,. \tag{A.169}$$

Next, we re-express the cash-flow sum in terms of earnings growth. Since $d_{t+j} = e_{t+j} + de_{t+j}$, we have $\Delta d_{t+j} = \Delta e_{t+j} + \Delta de_{t+j}$. Under the i.i.d. assumption in (A.168), $\mathbb{E}_t[\Delta de_{t+j}] = -v_t$ for $j = 1$ and $\mathbb{E}_t[\Delta de_{t+j}] = 0$ for $j \geq 2$. Substituting:

$$pe_t - de_t - \overline{pd}_t = \delta - v_t + \sum_{j=1}^{\infty} \rho^{j-1} \, \mathbb{E}_t\big[\Delta e_{t+j} - g_t\big] - \sum_{j=1}^{\infty} \rho^{j-1} \, \mathbb{E}_t\big[r_{t+j} - \bar{\mu}\big] \,. \tag{A.170}$$

Rearranging the left-hand side using $de_t = \overline{de} + v_t$:

$$pe_t - \overline{pe}_t = \delta + \sum_{j=1}^{\infty} \rho^{j-1}\, \mathbb{E}_t\left[\Delta e_{t+j} - g_t\right] - \sum_{j=1}^{\infty} \rho^{j-1}\, \mathbb{E}_t\left[r_{t+j} - \bar{\mu}\right], \tag{A.171}$$

where $\overline{pe}_t \equiv \overline{pd}_t + \overline{de} = g_t - \log\!\big(\exp(\bar{\mu}) - \exp(g_t)\big) + \overline{de}$ is the steady-state log price–earnings ratio.

Evaluating the conditional expectations under the model in (A.166) yields:

$$pe_t = \overline{pe}_t + \delta - \frac{1-\rho_c}{1-\rho\rho_c}\, c_t \;+\; \frac{1-\rho}{1-\rho\rho_c}\, \theta_c\, \varepsilon_{c,t} - \frac{1}{1-\rho\rho_\mu}\, (\mu_t - \bar{\mu}). \tag{A.172}$$

The nonlinearity in g_t enters through $\overline{pe}_t$ and is handled identically to the dividend-based model via the Extended Kalman Filter linearisation described in Appendix C.1. The measurement vector is $x_t = (e_t,\ pe_t)^\top$, and the linearised measurement matrix G_t takes the same form as in equation (A.55), with $\overline{de}$ absorbed into the intercept κ_t.

I.2 Estimation Details

We follow the same estimation procedure as the baseline case (Appendix C.2), with the parameter vector $\Theta = (\rho_c, \theta_c, \sigma_g, \sigma_c, \rho_\mu, \sigma_\mu, \bar{\mu})$. The steady-state log payout ratio $\overline{de}$ is set to the full-sample mean of $d_t - e_t$ and held fixed throughout the estimation, entering only through the intercept κ_t of the measurement equation. The initial state vector, Riccati-based initialisation of $P_{0|0}$, and all other aspects of the optimisation remain unchanged.

J Steady-State Drift in Expected Returns

We generalise the baseline model by allowing the long-run mean of expected returns to have a random walk component. Concretely, expected returns mean-revert around a drifting long-run level $\bar{\mu}_t$:

$$\begin{aligned} d_t &= \tau_t + c_t, \\ \tau_t &= \tau_{t-1} + g_{t-1}, \\ g_t &= g_{t-1} + \varepsilon_{g,t}, \quad \varepsilon_{g,t} \sim \mathcal{N}(0, \sigma_g^2) \\ c_t &= \rho_c \, c_{t-1} + \theta_c \, \varepsilon_{c,t-1} + \varepsilon_{c,t}, \quad \varepsilon_{c,t} \sim \mathcal{N}(0, \sigma_c^2) \\ \mu_t &= \bar{\mu}_t + \rho_\mu \, (\mu_{t-1} - \bar{\mu}_{t-1}) + \varepsilon_{\mu,t}, \quad \varepsilon_{\mu,t} \sim \mathcal{N}(0, \sigma_\mu^2) \\ \bar{\mu}_t &= \bar{\mu}_{t-1} + \varepsilon_{\bar{\mu},t} \quad \varepsilon_{\bar{\mu},t} \sim \mathcal{N}(0, \sigma_{\bar{\mu}}^2) \end{aligned} \tag{A.173}$$

All four shocks $(\varepsilon_{g,t},\ \varepsilon_{c,t},\ \varepsilon_{\bar{\mu},t},\ \varepsilon_{\mu,t})$ are mutually independent. Expected returns thus combine a persistent cyclical component—governed by ρ_μ as in the baseline—with a slowly drifting attractor $\bar{\mu}_t$. The baseline model is nested as the special case $\sigma_{\bar{\mu}} = 0$, in which $\bar{\mu}_t = \bar{\mu}$ for all t and μ_t reduces to the baseline's AR(1) process around a constant mean.

J.1 State-Space System

The state vector tracks μ_t directly rather than the cyclical deviation:

$$s_t = (\tau_t,\ g_t,\ c_t,\ \varepsilon_{c,t},\ \bar{\mu}_t,\ \mu_t,\ 1)^\top. \tag{A.174}$$

The transition equation $s_{t+1} = A\, s_t + C\, \epsilon_{t+1}$ is governed by:

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \rho_c & \theta_c & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1-\rho_\mu & \rho_\mu & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \qquad C = \begin{bmatrix} 0 & 0 & 0 & 0 \\ \sigma_g & 0 & 0 & 0 \\ 0 & \sigma_c & 0 & 0 \\ 0 & \sigma_c & 0 & 0 \\ 0 & 0 & \sigma_{\bar{\mu}} & 0 \\ 0 & 0 & \sigma_{\bar{\mu}} & \sigma_\mu \\ 0 & 0 & 0 & 0 \end{bmatrix}, \tag{A.175}$$

where C is a 7×4 shock-loading matrix and the state innovations $\epsilon_{t+1}^\top = (\varepsilon_{g,t+1},\ \varepsilon_{c,t+1},\ \varepsilon_{\bar{\mu},t+1},\ \varepsilon_{\mu,t+1}) \sim \mathcal{N}(0, I_4)$ are mutually independent. The sixth row of A encodes the expected-return dynamics: $\mu_{t+1} = (1-\rho_\mu)\,\bar{\mu}_t + \rho_\mu\,\mu_t + \text{shocks}$. The μ_t row of C loads on both the trend shock $\varepsilon_{\bar{\mu}}$ and the cyclical shock ε_μ, reflecting the fact that a shift in $\bar{\mu}_t$ contemporaneously moves μ_t by the same amount.

J.2 Nonlinear Measurement Equation

The measurement vector remains $x_t = (d_t,\ pd_t)^\top$. The valuation ratio is:

$$pd_t = \delta + \overline{pd}_t - \frac{1-\rho_c}{1-\rho\rho_c}\, c_t \ + \ \frac{1-\rho}{1-\rho\rho_c}\, \theta_c\, \varepsilon_{c,t} - \frac{1}{1-\rho\rho_\mu}\,(\mu_t - \bar{\mu}_t), \tag{A.176}$$

where $\overline{pd}_t = g_t - \log\big(\exp(\bar{\mu}_t) - \exp(g_t)\big)$ is the steady-state log price–dividend ratio, a nonlinear function of both the dividend-growth trend g_t and the long-run expected return $\bar{\mu}_t$. We take a first-order Taylor expansion of $\overline{pd}_t$ around the filter's joint prior $(\hat{g}_{t|t-1},\, \hat{\bar{\mu}}_{t|t-1})$:

$$\overline{pd}_t \;\approx\; \widehat{\overline{pd}}_t + \lambda_t^{g}\,(g_t - \hat{g}_{t|t-1}) + \lambda_t^{\bar{\mu}}\,(\bar{\mu}_t - \hat{\bar{\mu}}_{t|t-1}), \tag{A.177}$$

where the evaluated steady-state valuation ratio and its partial derivatives are:

$$\widehat{\overline{pd}}_t \equiv \hat{g}_{t|t-1} - \log\big(\exp(\hat{\bar{\mu}}_{t|t-1}) - \exp(\hat{g}_{t|t-1})\big), \tag{A.178}$$

$$\lambda_t^{g} \equiv \left.\frac{\partial \overline{pd}_t}{\partial g_t}\right|_{\hat{g},\,\hat{\bar{\mu}}} = 1 + \frac{\exp(\hat{g}_{t|t-1})}{\exp(\hat{\bar{\mu}}_{t|t-1}) - \exp(\hat{g}_{t|t-1})}, \tag{A.179}$$

$$\lambda_t^{\bar{\mu}} \equiv \left.\frac{\partial \overline{pd}_t}{\partial \bar{\mu}_t}\right|_{\hat{g},\,\hat{\bar{\mu}}} = -\frac{\exp(\hat{\bar{\mu}}_{t|t-1})}{\exp(\hat{\bar{\mu}}_{t|t-1}) - \exp(\hat{g}_{t|t-1})}. \tag{A.180}$$

The derivative λ_t^{g} captures the sensitivity of steady-state valuations to trend growth, as in the baseline. The new derivative $\lambda_t^{\bar{\mu}}$ captures the sensitivity to shifts in the long-run discount rate: a higher $\bar{\mu}_t$ raises the rate at which future cash flows are discounted, lowering valuations. Note that $\lambda_t^{\bar{\mu}} < 0$ always, with $\lambda_t^{\bar{\mu}} \to -1$ when g_t is far below $\bar{\mu}_t$ and $|\lambda_t^{\bar{\mu}}| \to \infty$ as $g_t \to \bar{\mu}_t$.

Substituting the linearised $\overline{pd}_t$ into (A.176) and collecting terms on the state vector:

$$pd_t \;\approx\; \lambda_t^{g}\, g_t - \frac{1-\rho_c}{1-\rho\rho_c}\, c_t + \frac{1-\rho}{1-\rho\rho_c}\,\theta_c\,\varepsilon_{c,t} + \left(\lambda_t^{\bar{\mu}} + \frac{1}{1-\rho\rho_\mu}\right)\bar{\mu}_t - \frac{1}{1-\rho\rho_\mu}\,\mu_t + \kappa_t, \tag{A.181}$$

where the intercept is:

$$\kappa_t = \delta + \widehat{\overline{pd}}_t - \lambda_t^{g}\,\hat{g}_{t|t-1} - \lambda_t^{\bar{\mu}}\,\hat{\bar{\mu}}_{t|t-1}. \tag{A.182}$$

The loading on $\bar{\mu}_t$ combines two effects: the Taylor-expansion sensitivity $\lambda_t^{\bar{\mu}}$ from the steady-state valuation, and $+1/(1-\rho\rho_\mu)$ from the present value of the deviation $-(\mu_t - \bar{\mu}_t)/(1-\rho\rho_\mu)$.

The linearised measurement equation $x_t = G_t\, s_t$ is therefore:

$$G_t = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & \lambda_t^{g} & -\dfrac{1-\rho_c}{1-\rho\rho_c} & \dfrac{1-\rho}{1-\rho\rho_c}\,\theta_c & \lambda_t^{\bar{\mu}} + \dfrac{1}{1-\rho_\mu\rho} & -\dfrac{1}{1-\rho_\mu\rho} & \kappa_t \end{bmatrix}. \tag{A.183}$$

As in the baseline, G_t is re-evaluated at each step of the filter at the current prior estimates $(\hat{g}_{t|t-1},\, \hat{\bar{\mu}}_{t|t-1})$. The key difference is that the linearisation now operates over a two-dimensional surface rather than a single curve, allowing the filter to jointly update its beliefs about trend growth and the long-run discount rate in response to valuation surprises.

J.3 Estimation Details

We follow the same estimation procedure as the baseline case (Appendix C.2), with the parameter vector $\Theta = (\rho_c, \theta_c, \sigma_g, \sigma_c, \rho_\mu, \sigma_{\bar{\mu}}, \sigma_\mu,)$. The initial state vector is $\hat{s}_{0|0} = (\tau_0,\ g_0,\ 0,\ 0, \bar{\mu_0}, \mu_0,\ 1)^\top$, where $\bar{\mu_0}$ and μ_0 are both set to the pre-sample average quarterly return from Golez and Koudijs (2018). The Riccati-based initialisation of $P_{0|0}$ evaluates G_{ref} at the reference point $(g_0,\ \mu_0)$. All other aspects of the optimisation, including the basinhopping algorithm, remain unchanged.

K Steady-State Drift in Expected Returns and Correlated Shocks

We generalise the steady-state drifting model in Appendix J to allow its innovation to be correlated with the trend-growth shock. Formally, the model is:

$$\begin{aligned} d_t &= \tau_t + c_t, \\ \tau_t &= \tau_{t-1} + g_{t-1}, \\ g_t &= g_{t-1} + \varepsilon_{g,t}, \\ c_t &= \rho_c\, c_{t-1} + \theta_c\, \varepsilon_{c,t-1} + \varepsilon_{c,t}, \\ \bar{\mu}_t &= \bar{\mu}_{t-1} + \varepsilon_{\bar{\mu},t}, \\ \mu_t &= \bar{\mu}_t + \rho_\mu\,(\mu_{t-1} - \bar{\mu}_{t-1}) + \varepsilon_{\mu,t}. \end{aligned} \tag{A.184}$$

The cash-flow cycle shock $\varepsilon_{c,t}$ and the expected-return cycle shock $\varepsilon_{\mu,t}$ are mutually independent and independent of the two permanent shocks. The trend-growth and long-run discount-rate innovations are jointly Gaussian and correlated,

$$\begin{pmatrix} \varepsilon_{g,t} \\ \varepsilon_{\bar{\mu},t} \end{pmatrix} \sim \mathcal{N}\left(0,\, \begin{bmatrix} \sigma_g^2 & \chi\,\sigma_g\sigma_{\bar{\mu}} \\ \chi\,\sigma_g\sigma_{\bar{\mu}} & \sigma_{\bar{\mu}}^2 \end{bmatrix}\right). \tag{A.185}$$

We represent them through a Cholesky factor in mutually independent standardized innovations $\epsilon_{g,t},\, \epsilon_{\nu,t} \sim \mathcal{N}(0,1)$:

$$\varepsilon_{g,t} = \sigma_g\, \epsilon_{g,t}, \qquad \varepsilon_{\bar{\mu},t} = \sigma_{\bar{\mu}}\big(\chi\, \epsilon_{g,t} + \sqrt{1-\chi^2}\, \epsilon_{\nu,t}\big), \tag{A.186}$$

and likewise $\varepsilon_{c,t} = \sigma_c\, \epsilon_{c,t}$ and $\varepsilon_{\mu,t} = \sigma_\mu\, \epsilon_{\mu,t}$, so that $\epsilon_t = (\epsilon_{g,t},\, \epsilon_{c,t},\, \epsilon_{\nu,t},\, \epsilon_{\mu,t})^\top \sim \mathcal{N}(0, I_4)$. The standardized vector ϵ_t is what the shock-loading matrix C multiplies, so the standard deviations and the correlation enter the state-space system only through C. The correlation χ governs the comovement of trend-growth and long-run discount-rate shocks. At $\chi = 0$ recovers independent permanent shocks (the uncorrelated drifting-$\bar{\mu}$ model per Appendix J).

K.1 State-Space System

The state vector is:

$$s_t = (\tau_t,\, g_t,\, c_t,\, \varepsilon_{c,t},\, \bar{\mu}_t,\, \mu_t,\, 1)^\top. \tag{A.187}$$

The transition equation $s_{t+1} = A\, s_t + C\, \epsilon_{t+1}$ is governed by:

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \rho_c & \theta_c & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1-\rho_\mu & \rho_\mu & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \qquad C = \begin{bmatrix} 0 & 0 & 0 & 0 \\ \sigma_g & 0 & 0 & 0 \\ 0 & \sigma_c & 0 & 0 \\ 0 & \sigma_c & 0 & 0 \\ \chi\sigma_{\bar\mu} & 0 & \sqrt{1-\chi^2}\,\sigma_{\bar\mu} & 0 \\ \chi\sigma_{\bar\mu} & 0 & \sqrt{1-\chi^2}\,\sigma_{\bar\mu} & \sigma_\mu \\ 0 & 0 & 0 & 0 \end{bmatrix}, \tag{A.188}$$

where C is a 7×4 shock-loading matrix and the standardized innovations $\epsilon_{t+1}^\top = (\epsilon_{g,t+1},\ \epsilon_{c,t+1},\ \epsilon_{v,t+1},\ \epsilon_{\mu,t+1}) \sim \mathcal{N}(0, I_4)$ are mutually independent.

K.2 Nonlinear Measurement Equation

The measurement vector remains $x_t = (d_t,\ pd_t)^\top$. The valuation ratio is:

$$pd_t = \delta(\chi) + \overline{pd}_t - \frac{1-\rho_c}{1-\rho\rho_c}\, c_t + \frac{1-\rho}{1-\rho\rho_c}\, \theta_c\, \varepsilon_{c,t} - \frac{1}{1-\rho\rho_\mu}\, (\mu_t - \bar\mu_t), \tag{A.189}$$

with $\overline{pd}_t$ the steady-state ratio. Because the steady-state ratio depends on $(g_t, \bar\mu_t)$ only through the spread $sp_t \equiv \bar\mu_t - g_t$, the Jensen correction is evaluated using the innovation variance of the *spread* rather than that of growth alone. From (A.186),

$$\Delta sp_t = (\chi\sigma_{\bar\mu} - \sigma_g)\, \epsilon_{g,t} + \sqrt{1-\chi^2}\, \sigma_{\bar\mu}\, \epsilon_{v,t}, \qquad \sigma_{sp}^2(\chi) \equiv \mathrm{Var}(\Delta s_t) = \sigma_g^2 + \sigma_{\bar\mu}^2 - 2\chi\, \sigma_g \sigma_{\bar\mu}, \tag{A.190}$$

so that

$$\delta(\chi) \approx \frac{1}{2}\, \frac{\rho^2}{(1-\rho)^3}\, \sigma_s^2(\chi) = \frac{1}{2}\, \frac{\rho^2}{(1-\rho)^3}\, \big[\sigma_g^2 + \sigma_{\bar\mu}^2 - 2\chi\, \sigma_g \sigma_{\bar\mu}\big]. \tag{A.191}$$

This nests the relevant cases. At $\chi = 0$ it reduces to the uncorrelated correction $\frac{1}{2}\frac{\rho^2}{(1-\rho)^3}(\sigma_g^2 + \sigma_{\bar\mu}^2)$, while at $\chi = 1$ with $\sigma_{\bar\mu} = \sigma_g$ the spread is constant and the correction vanishes.

We take a first-order Taylor expansion of $\overline{pd}_t$ around the filter's joint prior $(\hat g_{t|t-1},\ \hat{\bar\mu}_{t|t-1})$:

$$\overline{pd}_t \;\approx\; \widehat{\overline{pd}}_t + \lambda_t^g\, (g_t - \hat g_{t|t-1}) + \lambda_t^{\bar\mu}\, (\bar\mu_t - \hat{\bar\mu}_{t|t-1}), \tag{A.192}$$

where the evaluated steady-state valuation ratio and its partial derivatives are:

$$\widehat{\overline{pd}}_t \equiv \hat{g}_{t|t-1} - \log\big(\exp(\hat{\bar{\mu}}_{t|t-1}) - \exp(\hat{g}_{t|t-1})\big), \tag{A.193}$$

$$\lambda_t^g \equiv \left.\frac{\partial \overline{pd}_t}{\partial g_t}\right|_{\hat{g},\,\hat{\bar{\mu}}} = \frac{\exp(\hat{\bar{\mu}}_{t|t-1})}{\exp(\hat{\bar{\mu}}_{t|t-1}) - \exp(\hat{g}_{t|t-1})}, \tag{A.194}$$

$$\lambda_t^{\bar{\mu}} \equiv \left.\frac{\partial \overline{pd}_t}{\partial \bar{\mu}_t}\right|_{\hat{g},\,\hat{\bar{\mu}}} = -\frac{\exp(\hat{\bar{\mu}}_{t|t-1})}{\exp(\hat{\bar{\mu}}_{t|t-1}) - \exp(\hat{g}_{t|t-1})}. \tag{A.195}$$

The two derivatives satisfy the identity

$$\lambda_t^g = -\lambda_t^{\bar{\mu}}, \tag{A.196}$$

which makes explicit that the steady-state valuation component loads on g_t and $\bar{\mu}_t$ only through their spread. Substituting the linearised $\overline{pd}_t$ into (A.176) and collecting terms on the state vector:

$$pd_t \;\approx\; \lambda_t^g\, g_t - \frac{1-\rho_c}{1-\rho\rho_c}\, c_t + \frac{1-\rho}{1-\rho\rho_c}\,\theta_c\,\varepsilon_{c,t} + \left(\lambda_t^{\bar{\mu}} + \frac{1}{1-\rho\rho_\mu}\right)\bar{\mu}_t - \frac{1}{1-\rho\rho_\mu}\,\mu_t + \kappa_t, \tag{A.197}$$

where the intercept is:

$$\kappa_t = \delta(\chi) + \widehat{\overline{pd}}_t - \lambda_t^g\,\hat{g}_{t|t-1} - \lambda_t^{\bar{\mu}}\,\hat{\bar{\mu}}_{t|t-1}. \tag{A.198}$$

The linearised measurement equation $x_t = G_t\, s_t$ is therefore:

$$G_t = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & \lambda_t^g & -\dfrac{1-\rho_c}{1-\rho\rho_c} & \dfrac{1-\rho}{1-\rho\rho_c}\,\theta_c & \lambda_t^{\bar{\mu}} + \dfrac{1}{1-\rho_\mu\rho} & -\dfrac{1}{1-\rho_\mu\rho} & \kappa_t \end{bmatrix}. \tag{A.199}$$

As in the baseline, G_t is re-evaluated at each step of the filter at the current prior estimates $(\hat{g}_{t|t-1},\, \hat{\bar{\mu}}_{t|t-1})$. The key difference is that the linearisation now operates over the two-dimensional surface $(g_t, \bar{\mu}_t)$ rather than a single curve, allowing the filter to jointly update its beliefs about trend growth and the long-run discount rate in response to valuation surprises.

K.3 Estimation Details

The correlation χ between the two permanent innovations is weakly identified from a single historical sample. The price–dividend ratio identifies the permanent shocks only through the spread-shock variance $\sigma_s^2(\chi) = \sigma_g^2 + \sigma_{\bar{\mu}}^2 - 2\chi\,\sigma_g\sigma_{\bar{\mu}}$, so σ_g, $\sigma_{\bar{\mu}}$, and χ trade off along a ridge. We therefore do not estimate χ freely. Instead we fix χ on a grid $\chi \in \{0.25, 0.50, 0.75\}$ and, at each node, estimate the remaining parameters

$$\Theta(\chi) = (\rho_c,\, \theta_c,\, \sigma_g,\, \sigma_c,\, \rho_\mu,\, \sigma_{\bar{\mu}},\, \sigma_\mu)$$

by maximum likelihood with the extended Kalman filter, following the same procedure as the baseline case (Appendix C.2). The initial state vector is $\hat{s}_{0|0} = (\tau_0,\ g_0,\ 0,\ 0,\ \bar{\mu}_0,\ \mu_0,\ 1)^\top$, where $\bar{\mu}_0$ and μ_0 are both set to the pre-sample average quarterly return from Golez and Koudijs (2018). The Riccati-based initialisation of $P_{0|0}$ evaluates G_{ref} at the reference point $(g_0,\ \bar{\mu}_0)$. All other aspects of the optimisation, including the basinhopping algorithm, remain unchanged.

Table A.4 reports the maximized log-likelihood and the estimated permanent volatilities at each node, alongside the independent shock case ($\chi = 0$).

Table A.4: Correlated Permanent-Shock Profile

χ	$\ell(\chi)$	σ_g	$\sigma_{\bar{\mu}}$	$\sigma_s(\chi)$
0.00	3002.44	0.000215	0.000093	0.000235
0.25	3004.05	0.000232	0.000133	0.000237
0.50	3005.53	0.000258	0.000158	0.000225
0.75	3006.81	0.000312	0.000198	0.000210

Note: Each row fixes χ and re-estimates the remaining parameters by maximum likelihood with the extended Kalman filter. $\ell(\chi)$ is the maximized log-likelihood at each node. $\sigma_s(\chi) = (\sigma_g^2 + \sigma_{\bar{\mu}}^2 - 2\chi\sigma_g\sigma_{\bar{\mu}})^{1/2}$ is the implied volatility of the innovation to the long-run spread $\bar{\mu}_t - g_t$.

L Drifting Expected Returns without a Cash Flow Trend

This appendix considers a counterfactual restriction of the drifting-mean model of Appendix J. We shut down the permanent component of cash-flow growth—setting $\sigma_g = 0$—so that the only non-stationary force in the system is the drifting long-run expected return $\bar{\mu}_t$.

L.1 State-Space System

With $\sigma_g = 0$, the law of motion $g_t = g_{t-1} + \varepsilon_{g,t}$ collapses to a constant, $g_t = \bar{g}$ for all t. The cash-flow trend is then a deterministic linear trend, $\tau_t = \tau_{t-1} + \bar{g}$, rather than an integrated random walk. All non-stationarity in the price–dividend ratio must now be absorbed by $\bar{\mu}_t$. The DGP becomes:

$$\begin{aligned} d_t &= \tau_t + c_t, \\ \tau_t &= \tau_{t-1} + \bar{g}, \\ c_t &= \rho_c\, c_{t-1} + \theta_c\, \varepsilon_{c,t-1} + \varepsilon_{c,t}, \quad \varepsilon_{c,t} \sim \mathcal{N}(0, \sigma_c^2) \\ \mu_t &= \bar{\mu}_t + \rho_\mu\, (\mu_{t-1} - \bar{\mu}_{t-1}) + \varepsilon_{\mu,t}, \quad \varepsilon_{\mu,t} \sim \mathcal{N}(0, \sigma_\mu^2) \\ \bar{\mu}_t &= \bar{\mu}_{t-1} + \varepsilon_{\bar{\mu},t}, \quad \varepsilon_{\bar{\mu},t} \sim \mathcal{N}(0, \sigma_{\bar{\mu}}^2) \end{aligned} \tag{A.200}$$

The growth rate $\bar{g}$ is no longer a fixed initialisation but is estimated as a free parameter, so the cash-flow trend is identified by the data rather than imposed.

Because g_t is now a constant, it need not be carried as a state. The state vector drops to six elements:

$$s_t = (\tau_t,\ c_t,\ \varepsilon_{c,t},\ \bar{\mu}_t,\ \mu_t,\ 1)^\top, \tag{A.201}$$

with $\bar{g}$ entering the transition through the constant state. The transition equation $s_{t+1} = A\, s_t + C\, \epsilon_{t+1}$ is governed by:

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & \bar{g} \\ 0 & \rho_c & \theta_c & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1-\rho_\mu & \rho_\mu & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \qquad C = \begin{bmatrix} 0 & 0 & 0 \\ \sigma_c & 0 & 0 \\ \sigma_c & 0 & 0 \\ 0 & \sigma_{\bar{\mu}} & 0 \\ 0 & \sigma_{\bar{\mu}} & \sigma_\mu \\ 0 & 0 & 0 \end{bmatrix}, \tag{A.202}$$

where C is a 6×3 shock-loading matrix and the state innovations $\epsilon_{t+1}^\top = (\epsilon_{c,t+1},\ \epsilon_{\bar{\mu},t+1},\ \epsilon_{\mu,t+1}) \sim \mathcal{N}(0, I_3)$ are mutually independent. The trend shock ε_g is absent; the deterministic trend slope $\bar{g}$ now sits in the $(\tau,\ 1)$ position of A. As in Appendix J, the μ_t row of C loads on both $\varepsilon_{\bar{\mu}}$ and ε_μ.

L.2 Nonlinear Measurement Equation

The measurement vector remains $x_t = (d_t,\ pd_t)^\top$. The valuation ratio is:

$$pd_t = \delta + \overline{pd}_t - \frac{1-\rho_c}{1-\rho\rho_c}\, c_t \ + \ \frac{1-\rho}{1-\rho\rho_c}\, \theta_c\, \varepsilon_{c,t} - \frac{1}{1-\rho\rho_\mu}\, (\mu_t - \bar{\mu}_t), \tag{A.203}$$

where $\overline{pd}_t = \bar{g} - \log\big(\exp(\bar{\mu}_t) - \exp(\bar{g})\big)$.

The measurement equation is still nonlinear, but now only through $\bar{\mu}_t$: with $\bar{g}$ constant, $\overline{pd}_t$ traces a one-dimensional curve in $\bar{\mu}_t$ rather than a surface. This mirrors the baseline of Appendix C.2, which linearises over g_t with $\bar{\mu}$ fixed; here we linearise over $\bar{\mu}_t$ with g fixed. We take a first-order Taylor expansion of $\overline{pd}_t$ around the filter's prior $\hat{\bar{\mu}}_{t|t-1}$:

$$\overline{pd}_t \;\approx\; \widehat{\overline{pd}}_t + \lambda_t^{\bar{\mu}}\,(\bar{\mu}_t - \hat{\bar{\mu}}_{t|t-1}), \tag{A.204}$$

where

$$\widehat{\overline{pd}}_t \equiv \bar{g} - \log\big(\exp(\hat{\bar{\mu}}_{t|t-1}) - \exp(\bar{g})\big), \tag{A.205}$$

$$\lambda_t^{\bar{\mu}} \equiv \left.\frac{\partial \overline{pd}_t}{\partial \bar{\mu}_t}\right|_{\hat{\bar{\mu}}} = -\frac{\exp(\hat{\bar{\mu}}_{t|t-1})}{\exp(\hat{\bar{\mu}}_{t|t-1}) - \exp(\bar{g})}. \tag{A.206}$$

As in Appendix J, $\lambda_t^{\bar{\mu}} < 0$ always: a higher long-run discount rate lowers steady-state valuations. Substituting into (A.203) and collecting terms on the state vector gives the linearised measurement equation $x_t = G_t\, s_t$:

$$G_t = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & -\dfrac{1-\rho_c}{1-\rho\rho_c} & \dfrac{1-\rho}{1-\rho\rho_c}\,\theta_c & \lambda_t^{\bar{\mu}} + \dfrac{1}{1-\rho_\mu\rho} & -\dfrac{1}{1-\rho_\mu\rho} & \kappa_t \end{bmatrix}, \tag{A.207}$$

where the intercept is

$$\kappa_t = \delta + \widehat{\overline{pd}}_t - \lambda_t^{\bar{\mu}}\,\hat{\bar{\mu}}_{t|t-1}. \tag{A.208}$$

The constant sensitivity of $\overline{pd}_t$ to the fixed growth rate $\bar{g}$ is absorbed entirely into κ_t. As in the baseline, G_t is re-evaluated at each step at the current prior estimate $\hat{\bar{\mu}}_{t|t-1}$.

L.3 Estimation Details

We follow the same estimation procedure as the baseline case (Appendix C.2), with the parameter vector $\Theta = (\bar{g}, \rho_c, \theta_c, \sigma_c, \rho_\mu, \sigma_{\bar{\mu}}, \sigma_\mu)$. Relative to Appendix J, the trend-growth volatility σ_g is dropped and the constant growth rate $\bar{g}$ is added; the parameter count is unchanged. The initial state vector is $\hat{s}_{0|0} = (\tau_0,\ 0,\ 0,\ \bar{\mu}_0,\ \mu_0,\ 1)^\top$, where $\bar{\mu}_0$ and μ_0 are both set to the pre-sample average quarterly return from Golez and Koudijs (2018). The growth rate $\bar{g}$ is initialised at the pre-sample average quarterly dividend growth rate from Golez and Koudijs (2018). The Riccati-based initialisation of $P_{0|0}$ evaluates G_{ref} at the reference point $\bar{\mu}_0$. All other aspects of the optimisation, including the basinhopping algorithm, remain unchanged.

M Tables

Table A.5: Structural Breaks in Valuation Ratios

	Log(PD)	Log(PE)	Log(CAPE)
	(1)	(2)	(3)
Break Date(s)	1954, 1992	1990	1991
SupF(0,1)	56.539***	12.541**	14.644***
SupF(0,2)	54.980***	6.652	8.166**
SupF(1,2)	11.382**	0.714	3.855

Note: This table reports tests for structural breaks in the mean of three valuation ratios following Lettau and Van Nieuwerburgh (2008): the log price-dividend ratio, the log price-earnings ratio, and the log cyclically adjusted price-earnings ratio. The break dates and test statistics are estimated using the Bai and Perron (1998) methodology. The first row reports the estimated break dates selected by the sequential procedure. The rows labeled SupF(0,1) and SupF(0,2) report the statistics for testing the null hypothesis of no structural breaks against the alternative of one and two breaks, respectively. The row labeled SupF(1,2) reports the sequential test statistic for the null hypothesis of one break against the alternative of two breaks. The tests allow for serial correlation and heteroskedasticity in the residuals (using quadratic spectral kernel pre-whitening), and the trimming parameter is set to 10% of the sample. Significance levels: *(p<0.10), **(p<0.05), ***(p<0.01).

Table A.6: Post-CRSP Return Predictability

$r_{t,t+j} = \sum_{i=1}^{j} \rho^{i-1} r_{t+i}$

	1-year				10-year				20-year
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$\widehat{\mathcal{M}}_t$			-0.24	-0.24			-1.28	-1.22	
			(0.00)	(0.00)			(0.00)	(0.00)	
pd_t	-0.06	-0.23			-0.41	-1.22			-0.62
	(0.07)	(0.00)			(0.01)	(0.00)			(0.00)
$\widehat{\widetilde{pd}}_t$		0.22		-0.01		1.09		-0.13	
		(0.00)		(0.70)		(0.00)		(0.54)	
$\hat{c}_t$		-0.16		0.02		-0.67		0.25	
		(0.19)		(0.90)		(0.03)		(0.45)	
$\hat{\varepsilon}_{c,t}$		-0.23		-0.26		0.29		0.14	
		(0.48)		(0.43)		(0.61)		(0.81)	
Observations	96	96	96	96	87	87	87	87	77
R^2	0.02	0.06	0.06	0.06	0.21	0.39	0.37	0.39	0.34
Adj. R^2	0.01	0.02	0.05	0.02	0.20	0.36	0.36	0.36	0.33

Note: This table reports predictive regressions of realized returns, $r_{t,t+j} = \sum_{i=1}^{j} \rho^{i-1} r_{t+i}$, over 1-year and 10-year horizons for the post-CRSP (i.e., post-1926) subsample. Independent variables include the filtered expected-return signal ($\widehat{\mathcal{M}}_t$), log price-dividend ratio (pd_t), and filtered cash-flow noise components ($\widehat{\widetilde{pd}}_t, \hat{c}_t, \hat{\varepsilon}_{c,t}$), derived from the baseline model. It reports OLS regression coefficients with p-values based on Newey-West t-stats in round brackets.

Table A.7: Post-CRSP Cash Flow Predictability

	$\Delta d_{t,t+j} = \sum_{i=1}^{j} \rho^{i-1} \Delta d_{t+i}$						$\rho^j pd_{t+j}$	
	10-year			20-year			10-year	20-year
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
pd_t	0.12			0.16			0.42	0.15
	(0.24)			(0.06)			(0.00)	(0.05)
$\hat{g}_t$		18.73	20.75		31.50	36.00		
		(0.24)	(0.20)		(0.12)	(0.00)		
$\hat{c}_t$		-0.66	-0.69		-1.07	-1.16		
		(0.00)	(0.00)		(0.00)	(0.00)		
$\hat{\varepsilon}_{c,t}$		-0.30	-0.34		0.04	0.01		
		(0.64)	(0.39)		(0.90)	(0.99)		
$\widehat{\mathcal{M}}_t$			-0.25			-0.23		
			(0.15)			(0.27)		
Observations	87	87	87	77	77	77	87	77
R^2	0.05	0.21	0.25	0.07	0.34	0.37	0.37	0.11
Adj. R^2	0.03	0.19	0.21	0.05	0.32	0.34	0.36	0.09

Note: This table reports predictive regressions of realized dividend growth, $\Delta d_{t,t+j} = \sum_{i=1}^{j} \rho^{i-1} \Delta d_{t+i}$, over 10-year and 20-year horizons for the post-CRSP (i.e., post-1926) subsample. Independent variables include the log price-dividend ratio (pd_t), the filtered cash-flow states ($\hat{g}_t, \hat{c}_t, \hat{\varepsilon}_{c,t}$) and the filtered expected-return signal ($\widehat{\mathcal{M}}_t$), derived from the baseline model. It reports OLS regression coefficients with p-values based on Newey-West t-stats in round brackets.

Table A.8: Stationary trend growth: In-Sample Return Predictability

	$r_{t,t+j} = \sum_{i=1}^{j} \rho^{i-1} r_{t+i}$								
	1-year				10-year				20-year
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$\widehat{\mathcal{M}}_t$			-0.21	-0.24			-1.12	-1.15	
pd_t	-0.04	-0.24			-0.23	-1.15			-0.25
$\widehat{pd}_t$		0.26		0.02		1.27		0.12	
$\hat{c}_t$		-0.25		-0.07		-0.88		-0.01	
$\hat{\varepsilon}_{c,t}$		-0.22		-0.25		0.39		0.24	
Observations	151	151	151	151	142	142	142	142	132
R^2	0.01	0.06	0.05	0.06	0.07	0.31	0.30	0.31	0.05
Adj. R^2	0.00	0.03	0.04	0.03	0.06	0.29	0.29	0.29	0.04

Note: The table reports OLS coefficients under the stationary trend growth (constant-ρ) model of Appendix G.

Table A.9: Stationary trend growth: Long-Horizon Cash Flow Predictability

	$\Delta d_{t,t+j} = \sum_{i=1}^{j} \rho^{i-1} \Delta d_{t+i}$						$\rho^j pd_{t+j}$	
	10-year			20-year			10-year	20-year
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
pd_t	0.26			0.47			0.46	0.22
$\hat{g}_t$		35.21	37.52		70.22	74.31		
$\hat{c}_t$		-0.62	-0.69		-1.15	-1.25		
$\hat{\varepsilon}_{c,t}$		-0.16	-0.10		0.64	0.72		
$\widehat{\mathcal{M}}_t$			-0.20			-0.19		
Observations	142	142	142	132	132	132	142	132
R^2	0.15	0.26	0.28	0.26	0.42	0.43	0.42	0.17
Adj. R^2	0.15	0.25	0.26	0.26	0.41	0.41	0.41	0.16

Note: The table reports OLS coefficients under the stationary trend growth (constant-ρ) model of Appendix G.

Table A.10: Sample-Length Dependence: Persistence

	Baseline			Stationary trend growth			Implied ϕ_{pd}, by shares	
Window	$\phi_{\mathcal{M}}$	$\phi_{\mathcal{C}}$	$\phi_{\overline{pd}}$	$\phi_{\mathcal{M}}$	$\phi_{\mathcal{C}}$	$\phi_{\overline{pd}}$	Model	Historical
T=10	0.48	0.54	0.55	0.48	0.54	0.55	0.49	0.50
T=50	0.78	0.84	0.90	0.78	0.84	0.90	0.81	0.83

Note: This table repeats Panel B of Table 7 at shorter windows. $T = 1$ is undefined. $T = 150$ and $T = \infty$ appear in the main table.

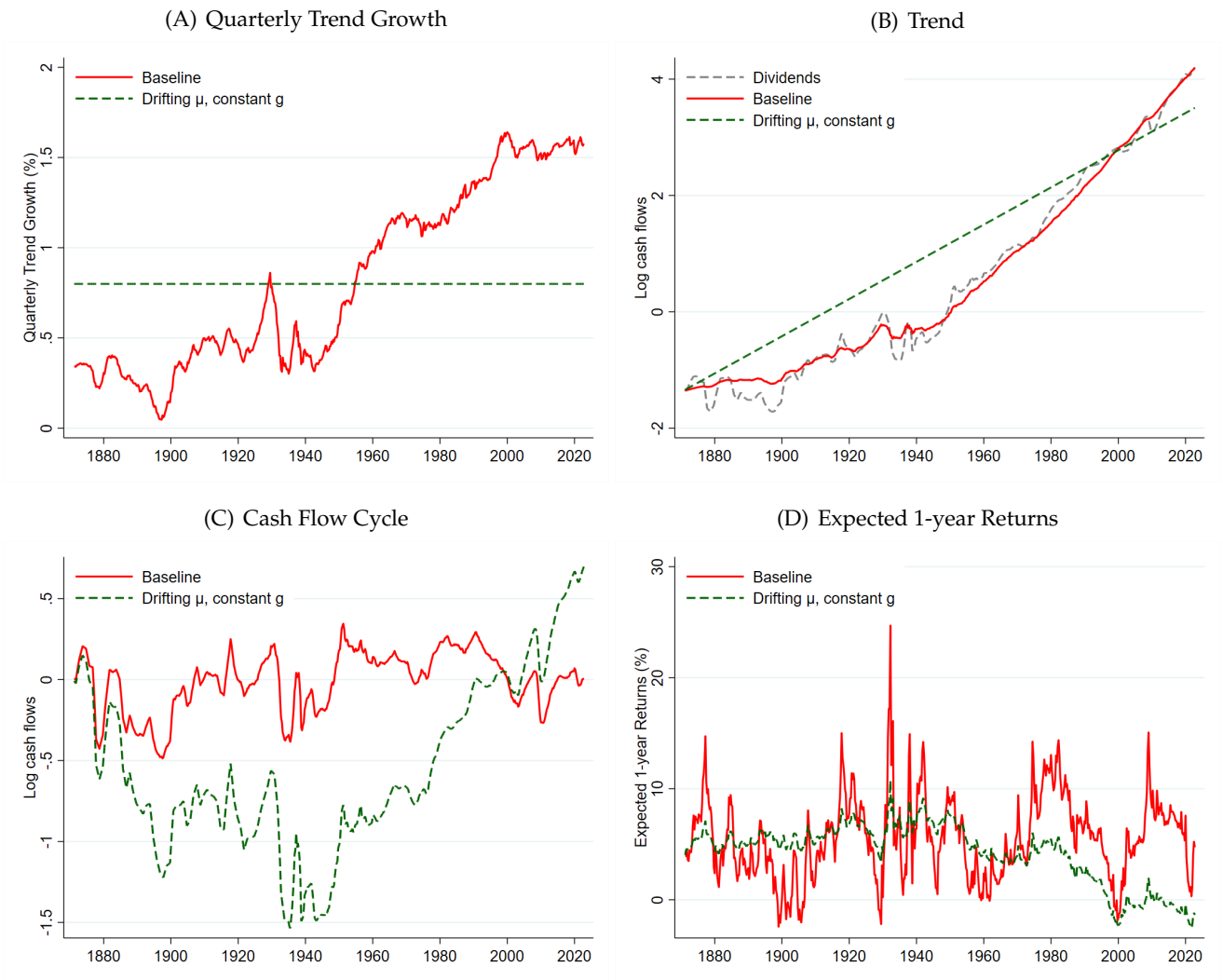
Table A.11: Model Extensions: In and Out-of-Sample R^2

	Raw Returns		Excess Returns	
	1-yr	5-yr	1-yr	5-yr
Panel A: In-Sample R^2 (%)				
1965–2022				
Trend-cycle model (dividends)	7.67%	19.88%	8.14%	17.26%
Trend-cycle model (dividends; drifting $\bar{\mu}$)	7.73%	19.92%	8.15%	17.03%
Trend-cycle model (dividends; drifting $\bar{\mu}$, constant g)	−19.38%	−120.18%	−18.57%	−125.11%
Trend-cycle model (earnings)	6.25%	17.70%	6.54%	16.24%
Trend-cycle model (dividends; stationary trend growth)	6.81%	17.65%	7.32%	14.50%
Panel B: Out-of-Sample R^2_{OOS} (%)				
1965–2022				
Trend-cycle model (dividends)	8.93%**	22.20%**	9.15%**	16.57%**
	[0.020]	[0.022]	[0.013]	[0.033]
Trend-cycle model (dividends; drifting $\bar{\mu}$)	8.10%**	20.16%**	8.35%**	14.70%**
	[0.024]	[0.026]	[0.015]	[0.038]
Trend-cycle model (earnings)	7.93%**	15.29%**	7.94%***	10.94%*
	[0.017]	[0.023]	[0.006]	[0.057]
Trend-cycle model (dividends; stationary trend growth)	7.69%**	19.11%**	7.94%**	13.21%**
	[0.030]	[0.036]	[0.023]	[0.042]

Note: This table compares in-sample R^2 (Panel A) and out-of-sample R^2_{OOS} (Panel B) across five specifications of the trend-cycle model: the baseline (dividends); drifting $\bar{\mu}$; constant g; earnings (with $\bar{\mu}$ freely estimated by MLE); and stationary trend growth. The constant g counterfactual is reported in Panel A only. The drifting-$\bar{\mu}$ specification assumes that discount-rate and trend-growth shocks are independent. All specifications report expected returns at 1- and 5-year horizons, with multi-year returns discounted at $\rho = 0.96$. The out-of-sample estimates are obtained by re-estimating the trend-cycle model using an expanding window. Realized excess returns subtract cumulated future three-month Treasury bill rates over the forecast horizon. Model-implied excess returns subtract the current annualized three-month bill rate projected over the horizon assuming a flat term structure, discounted at ρ. Stars denote significance of the Clark and West (2007) MSFE-adjusted statistic with Newey–West standard errors ($h-1$ lags); p-values in brackets. $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

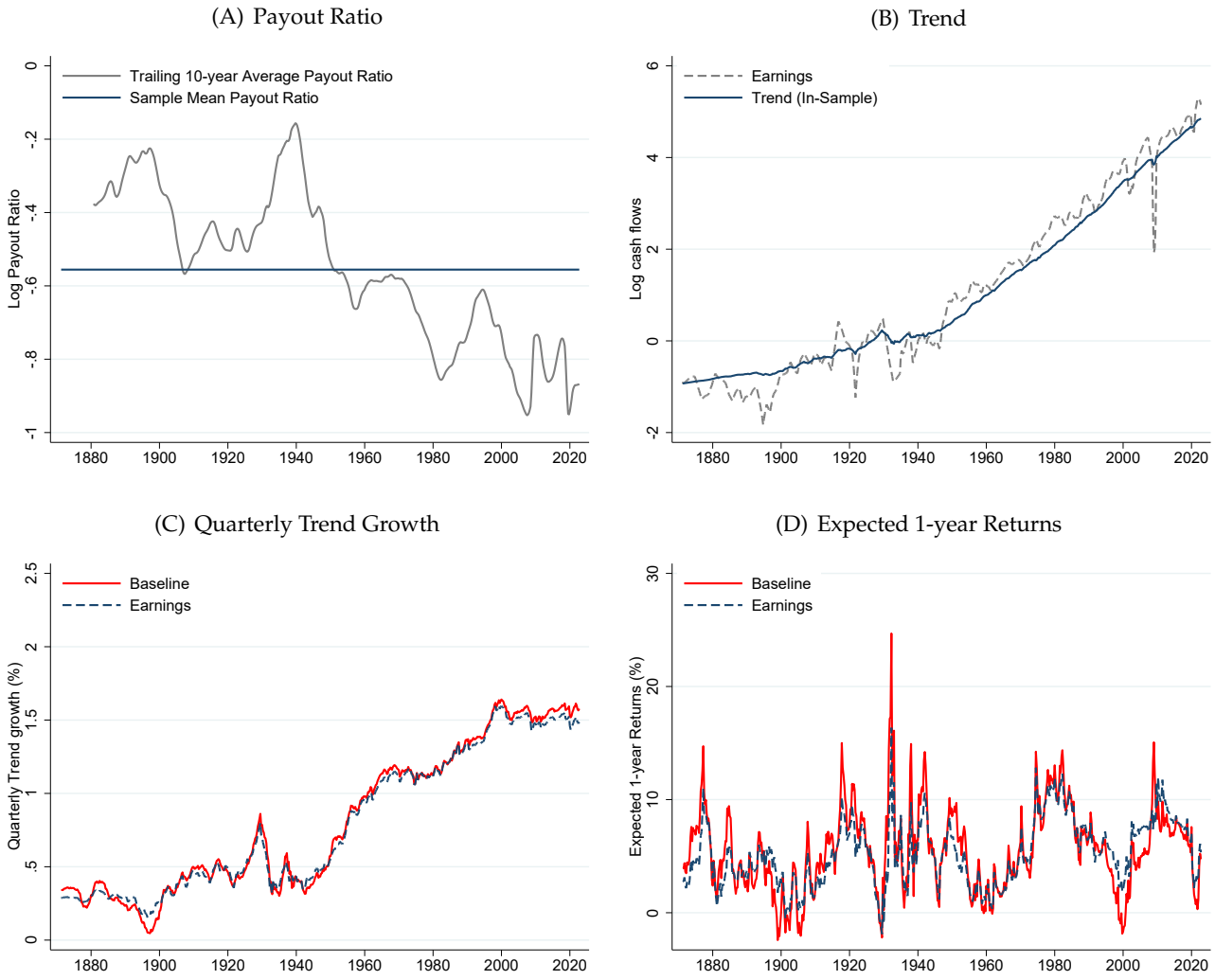
N Figures

Figure A.2: Drifting Expected Returns without a Cash-Flow Trend



Note: This figure compares the baseline trend-cycle model to the counterfactual specification with $\sigma_g = 0$, in which $\bar{g}$ is estimated as a free constant and the steady-state expected return $\bar{\mu}_t$ has a random walk component as in Section 7.1. Panel (A) plots quarterly trend growth under each specification. Panel (B) plots log dividends against the fitted trend. Panel (C) plots the filtered cash-flow cycle. Panel (D) plots model-implied expected one-year returns. Estimates are based on the full sample 1871–2022.

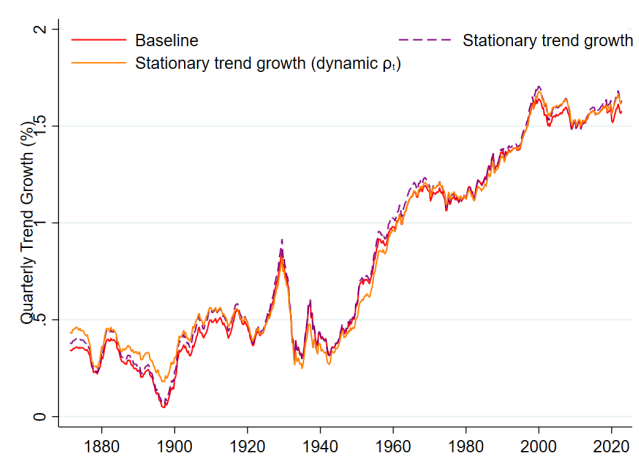
Figure A.3: Trend-Cycle Model – Earnings



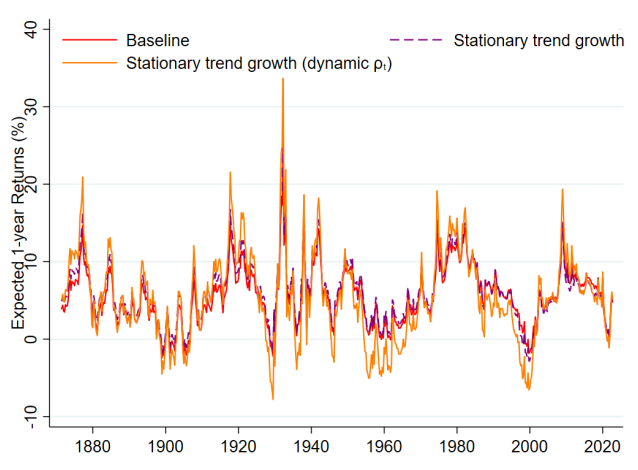
Note: This figure compares the trend-cycle model based on dividends (Baseline) to an earnings-based specification that assumes a constant log payout ratio $\overline{de}$, set to the full-sample mean of $d_t - e_t$, with steady-state expected returns $\bar{\mu}$ fixed at the baseline estimate. Panel (A) plots the trailing 10-year average of the log payout ratio against the sample-mean payout ratio $\overline{de}$. Panel (B) plots log earnings against the filtered earnings trend. Panel (C) plots filtered quarterly trend growth under each specification. Panel (D) plots model-implied expected one-year returns under both specifications.

Figure A.4: Baseline versus Stationary trend growth

(A) Trend growth



(B) Expected 1-year Returns



Note: The figure compares the unit-root baseline with two stationary specifications in which trend growth follows a highly persistent AR(1), $g_t = \bar{g} + \rho_g(g_{t-1} - \bar{g}) + \varepsilon_{g,t}$ with $\rho_g = 0.999$. The first keeps ρ constant; the second also lets the cycle loadings move with trend growth ("dynamic ρ_t") — see Appendix G. Panel (A) plots filtered trend growth and Panel (B) the implied one-year expected return under each specification.