Expected Returns with Cash Flow Trends and Cycles*

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

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

March 24, 2026

Abstract

We argue that cash flow growth contains a permanent trend component, rendering the price-dividend ratio a noisy proxy for expected returns. Jointly filtering trend growth, cash flow cycles, and expected returns via an Extended Kalman Filter, we show this cash flow noise generates severe attenuation bias in standard predictive regressions. Purging cash flow trends and cycles restores return predictability, delivering out-of-sample R^2 of 9% at one-year and 22% at five-year horizons, where traditional valuation ratios uniformly fall short. While discount rate variation dominates price-dividend movements at short and medium horizons, trend growth dominates at longer horizons.

JEL Classification: G4, G11, G12

Keywords: Return predictability, Trend-Cycles, Excess Volatility Puzzle, Variance Decompositions

*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, Andrea Tamoni, Riccardo Sabbatucci, Shrihari Santosh, Alexi Savov, Adi Sunderam, Erik Stafford, and Luis Viceira as well as seminar and conference participants at the Annual Valuation Workshop, the EFA, the Macro-Finance Junior Conference, University of Maryland Smith School of Business and Harvard Business School for helpful comments.

†shillenbrand@hbs.edu
‡otm210@nyu.edu

1 Introduction

Conventional wisdom in asset pricing holds that cash flow variation does not explain much of stock price movements (Shiller, 1981). Movements in valuation ratios are then largely attributable to time-varying discount rates (e.g., Cochrane, 2011). Several empirical findings, however, remain difficult to reconcile with this view.

A large literature documents that the predictive power of valuation ratios for stock-market returns is surprisingly low, both in-sample (e.g., Stambaugh, 1999; Ang and Bekaert, 2007) and out-of-sample (e.g., Goyal and Welch, 2008).[1] Corroborating this literature, we examine a ∼150-year sample (1871–2022) and find that standard valuation ratios offer little in-sample predictive power for stock returns and uniformly deliver negative out-of-sample R^2.

A related puzzle concerns the dynamics of valuation ratios themselves. Canonical asset pricing models (e.g., Campbell and Cochrane, 1999; Barro, 2006; Gabaix, 2012) generate discount rate variation at business-cycle frequencies, yet valuation ratios are far more persistent than these models predict, often struggling to pass standard unit-root tests (Campbell and Shiller, 1987). Moreover, over recent decades, valuation ratios have not merely been persistent—they have trended steadily higher. Lettau and Van Nieuwerburgh (2008) provide evidence that the price-dividend ratio has shifted multiple times to structurally higher levels, consistent with a unit-root component.

We propose a unified explanation for these findings. Cash flows, we argue, contain not only a stationary cycle component — consistent with short-run cash flow predictability (e.g., Lettau and Ludvigson, 2005; Koijen and Van Nieuwerburgh, 2011) — but also a permanent trend growth component. Since the price-dividend ratio reflects the present value of future expected returns and cash flow growth (Campbell and Shiller, 1988), it inherits both components. The trend makes valuation ratios persistent; the cycle adds further noise. Together, they contaminate the signal that valuation ratios carry about expected returns, explaining why these ratios are such weak return predictors.

To evaluate our hypothesis, we build a joint model of dividends, expected returns, and the price-dividend (pd) ratio. Building on prior state-space approaches (Pástor and Stambaugh, 2009; Koijen and Van Nieuwerburgh, 2011), we model expected returns as a stationary process. Unlike these studies, however, we also allow dividends to contain a permanent trend growth component alongside the

[1] See also, amongst others, Nelson and Kim (1993); Goetzmann and Jorion (1993); Amihud and Hurvich (2004); Boudoukh, Richardson, and Whitelaw (2008); Boudoukh, Israel, and Richardson (2022); Goyal, Welch, and Zafirov (2024). A number of studies show that modified predictors or alternative methodological choices can restore some evidence of predictability (see for example, Lettau and Ludvigson, 2001; Lewellen, 2004; Boudoukh, Michaely, Richardson, and Roberts, 2007; Campbell and Thompson, 2008; Lettau and Van Nieuwerburgh, 2008; Rapach, Strauss, and Zhou, 2010).

stationary cycle.[2] A unit-root growth component means the standard Campbell and Shiller (1988) log-linearization around a fixed mean is no longer adequate. Following Lettau and Van Nieuwerburgh (2008), we instead log-linearize around a time-varying steady-state price-dividend ratio that reflects movements in trend growth. We estimate the model via maximum likelihood using an Extended Kalman Filter,[3] yielding filtered estimates of expected returns, the cash-flow cycle, and trend growth.

Our model yields sharp testable predictions, which the data validate. The literature has worried that persistent valuation ratios can generate spurious evidence of predictability (Stambaugh, 1999). We argue the more fundamental problem runs in the opposite direction: because pd_t embeds cash-flow trend and cycle components, it is a noisy proxy for expected returns, attenuating predictive coefficients toward zero. Controlling for cash flow components should restore them. Specifically, at the one-year horizon, the model predicts an attenuated univariate slope of just -0.04; in the data, we find exactly -0.04 ($t = -1.33$). Controlling for the filtered cash-flow cycle and trend growth should remove the attenuation, raising the slope to -0.20; in the data, it rises to -0.24 ($t = -2.00$). Using the estimated expected-return component of *pd* directly—by construction purged of cash-flow noise—-as the sole predictor should yield this same unattenuated coefficient; we find -0.22 ($t = -2.20$). The unattenuated slopes are below one in magnitude because the *pd* ratio includes the full term structure of expected returns, of which one-year returns capture only a fraction. At the ten-year horizon, where a larger share of the term structure is captured, the unattenuated slope rises in magnitude accordingly—but the attenuation bias remains. The model predicts an attenuated slope of -0.18; in the data, we find -0.23 ($t = -1.35$). Controlling for cash-flow components, the slope should rise to -0.89; in the data, it rises to -1.17 ($t = -3.16$), and we cannot reject equality with the model prediction ($p = 0.45$). Where conventional regressions find no significant predictability over the $\sim$150 year sample, our approach—by purging cash flow noise—reveals a strongly significant expected-return signal at both short- and long-horizons.

We evaluate our model out of sample following the framework laid out by Goyal and Welch (2008). We estimate the state-space model in real time using an expanding window and use the model's expected-return estimate as the out-of-sample forecast. Our trend-cycle model persistently beats the historical mean benchmark, generating R^2_{OOS} of roughly 9% for one-year and 22% for five-year raw returns, both significant at the 5% level; excess return results are comparable at 9% and 17%. Traditional valuation ratios, by contrast, uniformly fail to beat the historical mean, producing negative R^2_{OOS}. Our model ex-

[2] A unit-root component in trend growth raises the familiar concern that valuations could become unbounded. We discuss this condition formally in Section 3 and show that it is comfortably satisfied in our data. Section 6 further relaxes the stationary discount rate assumption.

[3] Because the steady-state price-dividend ratio is a nonlinear function of trend growth g_t, the standard linear Kalman filter is inapplicable. We therefore employ an Extended Kalman Filter (EKF), which re-linearizes the measurement equation at each time step around the current estimate of the state.

plains why. As trend growth drifted permanently higher, valuation ratios rose in tandem, yet univariate predictive regressions—unable to distinguish higher growth from lower discount rates—systematically underpredicted returns. The trend-cycle model, by explicitly accounting for time variation in g, isolates a stationary expected-return signal that remains informative in real time, even as valuation ratios drift far from historical norms.

The fact that purging cash flow noise restores return predictability suggests that permanent trend growth is an important driver of valuation ratios. We are not the first to recognize this possibility. Kleidon (1986) and Marsh and Merton (1986) showed that Shiller's (1981) excess volatility analysis artificially suppressed low-frequency movements in cash flow growth by de-trending dividends ex-post with a "perfect-foresight" trend. Barsky and De Long (1993) argued that dividend growth may contain a small unit-root component that accounts for the bulk of long-run valuation ratio movements.

However, subsequent work has largely ignored the idea of permanent shifts in cash flow growth, likely due to a severe identification challenge. The cyclical component of cash flow growth is far more volatile than the trend over short-horizons, making it difficult to identify small but cumulating trend shocks in finite samples with cash flow data alone.[4] We make progress on this challenge by jointly modeling dividends and the price-dividend ratio, exploiting two asymmetries inherent in the state-space system. First, a shift in trend growth must simultaneously raise observed dividend growth and the pd ratio—while a shift in discount rates affects only the valuation ratio. Second, through present-value logic, small shifts in trend growth yield large movements in the price-dividend ratio, helping pin down trend growth.

The estimates from our state-space system suggest that trend growth drifted markedly, from roughly 2% per annum before 1900 to roughly 6% after 2000, consistent with the pronounced rise in valuation ratios. This secular increase reflects the cumulative effect of individually small shocks: estimated trend growth volatility is just 18bps per annum, but because trend growth follows a random walk, these small shocks accumulate over time.[5] The cyclical component of dividend growth, by contrast, is far more volatile (7.5% per annum) but largely transitory (AR(1) of -0.11). On the discount-rate side, expected returns exhibit pronounced annual volatility (6.1% per annum) but *crucially* with more moderate persistence (AR(1) of 0.83).

[4]The core problem is that cash flow data alone struggle to distinguish between a constant-growth process with i.i.d. shocks and one with a small but permanent component—with plausible parameters, standard time-series methods fail to reject constant trend growth in over a third of simulations (Barsky and De Long, 1993). Formally, building on Shephard and Harvey (1990), Barsky and De Long (1993) show that if dividend growth follows an IMA(1,1) process, $\Delta d_t = g_{t-1} + \epsilon_t$ where $g_t = g_{t-1} + (1-\theta)\epsilon_t$, then with a sample size of 120 and true $\theta = 0.97$, the likelihood function maximizes at the boundary ($\theta = 1$) in 36% of simulations—that is, standard methods fail to detect the permanent component over a third of the time.

[5]Over 150 years the standard deviation of cumulative trend growth is $0.18\% \cdot \sqrt{150} \approx 2.2\%$.

Our model implies that the permanent trend component in cash flows has important asset pricing implications. A key insight is that the decomposition of *pd* variation is horizon dependent. Over short and medium horizons, shifts in expected returns account for 72% of the variation in *pd*, while shifts in the trend-growth component contribute 16% and the stationary cash flow cycle contributes 12%. Over a 150-year horizon, however, the decomposition reverses—trend growth accounts for 74% of *pd* variation, expected returns for 22%, and the cash flow cycle just 4%. This reframes the Shiller (1981) excess volatility puzzle: discount rate variation is the dominant force at short and medium horizons, but at long-run horizons cash flows dominate.

This horizon dependence also helps reconcile the findings of Cochrane (2011), who concludes that returns account for nearly all *pd* variation using 15-year realized returns over 1947–2009. With an effective sample window of roughly 50 years, our model implies that discount rates account for roughly half of *pd* variation at this horizon—a substantial share, but not all. Moreover, as Boudoukh, Israel, and Richardson (2022) highlight, the exercise rests on only three to four independent observations, a setting in which overlapping-regression biases can inflate the return coefficient. As the sample lengthens, our model indicates the return share declines and the cash flow share rises. Extending the analysis to our full 150-year sample, 10-year returns and cash flow growth explain just 23% and 26% of *pd* variation, respectively (model-implied: 18% and 28%), with the remaining variation absorbed by *pd*–which the model attributes largely to the trend growth component. Consistent with this, 20-year cash flow growth explains 47% of *pd* variation (model-implied: 46%) as trend growth increasingly dominates.

Relatedly, Cochrane (2008) contends that the absence of cash flow predictability—the "dog that did not bark"—indirectly provides a more powerful case for return predictability. His argument relies on the present-value identity: if dividend growth is unforecastable, variation in the price-dividend ratio must reflect changing expected returns. However, as Cochrane (2008) himself acknowledges, this reasoning loses its force if the *pd* ratio contains a unit root, since a permanent component in expected cash flow growth can drive low-frequency variation in the *pd* ratio without generating detectable short-horizon predictability—precisely the conditions of our model.

Our analysis finds that the secular increase in valuation ratios observed over the past decades reflects a rise in trend growth. We confirm this with the following exercises. First, we re-estimate the model using an alternative cash flow measure—earnings—yielding nearly identical filtered trend growth and expected return paths. Second, instead of fixing steady-state expected returns, we re-estimate the model allowing for a random walk component in expected returns. We find that trend growth is largely unchanged; if anything, steady-state expected returns have increased from roughly 7% per annum in the pre-1900s to 8%-8.5% by the 1980s—the opposite of what one would expect from rising valuation ra-

tios. Importantly, our modeling approach is remarkably robust: both models deliver out-of-sample R^2 comparable to the baseline model.

Overall, our results offer a refinement of the Shiller (1981) excess volatility puzzle: stock prices are excessively volatile relative to fundamentals at business-cycle horizons, but over the last $\sim$150 years, trend growth can account for the bulk of *pd* movements. By jointly filtering cash flows and prices, we show that the apparent weakness of return predictability is an artifact of measurement error: once cash flow trends and cycles are purged, the stock market is, in fact, highly predictable.

The rest of the paper proceeds as follows. Section 2 documents the return predictability evidence, and presents motivating evidence on cash flow trends and cycles. Section 3 develops the state-space system and derives testable predictions. Section 4 tests the model predictions. Section 5 explores asset pricing implications, and Section 6 examines robustness. Section 7 concludes.

2 Motivating Evidence

We begin by documenting two central puzzles in empirical asset pricing. First, despite the theoretical consensus that discount rate movements largely drive asset price fluctuations (Campbell and Shiller, 1988; Cochrane, 2011), valuation ratios exhibit surprisingly low predictive power for stock-market returns, both in-sample (e.g., Stambaugh, 1999; Ang and Bekaert, 2007) and out-of-sample (e.g., Goyal and Welch, 2008; Goyal, Welch, and Zafirov, 2024). Second, valuation ratios have shifted to structurally higher levels over the past several decades (Lettau and Van Nieuwerburgh, 2008), a pattern difficult to reconcile with stationary models of discount rate variation.

We then present motivating evidence for our proposed explanation: that two distinct properties of cash flows—a permanent trend growth component and a stationary cycle—contaminate valuation ratios, simultaneously rendering them noisy predictors of returns and driving their secular rise.

2.1 Weak Return Predictability

Table 1 illustrates the weak return predictability tension. We conduct standard predictive regressions of one-year-ahead log raw and excess returns on 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*). Panel A reports in-sample estimates for the full history (1871–2022). To address the small-sample bias documented by Stambaugh (1999), we apply the bias-correction method of Amihud and Hurvich (2004). The predictive power of these valuation ratios is limited. The coefficients for the *pd* and *pe* ratios are statistically indistinguishable from zero. Only the *cape* ratio significantly predicts raw returns but fails to reliably predict excess returns.

Table 1: 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 p-values for the one-sided Clark and West (2007) test of the null hypothesis that predictive model ($\beta_1 \neq 0$) delivers a lower mean-squared forecast error than the historical-mean benchmark ($\beta_1 = 0$). Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Panel B reports results for an even stricter test: out-of-sample forecasting. As emphasized by Goyal and Welch (2008) and Goyal, Welch, and Zafirov (2024), even when a financial predictor forecasts returns in-sample, it often fails to translate into real-time forecasting power. To test this, we follow their methodology and estimate linear regressions recursively to predict one-year-ahead returns, $\hat{r}^{OOS}_{t+1} = \hat{\alpha}_t + \hat{\beta}_t x_t$, using only prior data. We evaluate performance using the out-of-sample R^2:

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

where $\bar{r}_t$ is the historical mean return up to time t. A negative R^2_{OOS} implies that the predictive model performs worse than the historical average benchmark. The predictive models fail to outperform the historical mean benchmark. Across every specification, the R^2_{oos} is negative, ranging from −0.57% to −7.11%. Furthermore, p-values from the one-sided Clark and West (2007) test (ranging from 0.114 to 0.444) indicate that we cannot reject the null hypothesis the historical mean is at least as good a predictor as the valuation model.

This establishes the paradox: while the consensus is that valuation ratios must reflect future dis-

count rates (Cochrane, 2011), in practice they are weak predictors of returns.

2.2 Trends in Valuation Ratios and Cash Flows

Why do valuation ratios fail to predict returns despite theoretically embedding discount rates? Our hypothesis is that cash flows contain both a permanent trend growth component and a stationary cycle, and that these two components contaminate valuation ratios, obscuring the true expected return signal. We now present motivating evidence for each in turn.

We begin with the low-frequency component. Panel A of Figure 1 plots the time series of the three valuation ratios. The visual evidence is difficult to reconcile with the traditional assumption of stationarity. While the ratios appear to fluctuate around a stable mean in the pre-war period, they exhibit a distinct upward trend in the post-war era.

To confirm this instability formally, we turn to the work of Lettau and Van Nieuwerburgh (2008). We replicate their analysis using the full 1871–2022 sample. Appendix Table A.1 reports tests for structural breaks using the Bai and Perron (1998) methodology across all three valuation ratios. The results confirm that the non-stationarity documented by Lettau and Van Nieuwerburgh (2008) persists through the full-sample. We detect statistically significant structural breaks for all measures. Consistent with Lettau and Van Nieuwerburgh (2008), we find that *pd* exhibits multiple structural breaks over the full sample—each shifting the ratio to a higher level: one break in 1954 and a second in 1992. This pattern of repeated upward shifts is consistent with a unit-root component in the *pd* ratio, but leaves open the economic source: do these shifts reflect a secular decline in expected returns, or a secular rise in trend growth?

Panel B provides direct evidence pointing toward the latter. Fifty-year rolling averages of one-year log growth in dividends and earnings have trended upward over the sample. This co-movement between long-run cash flow growth and valuation levels is consistent with our hypothesis that permanent shifts in trend growth are the primary driver of the secular rise in valuation ratios.

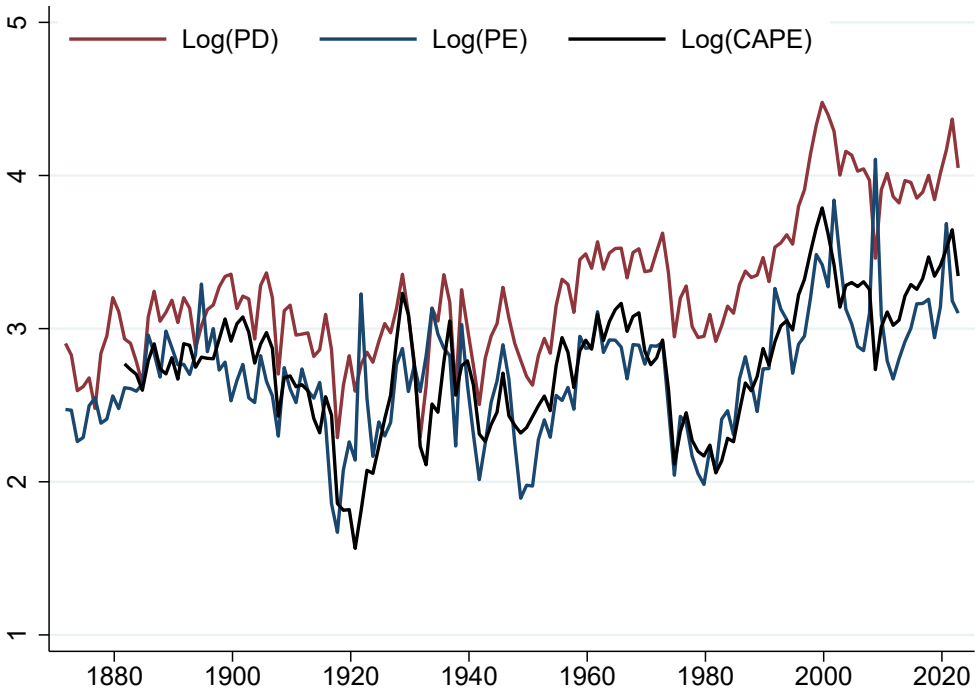
2.3 Cycles in Cash Flows

Trend growth is not the only source of cash flow contamination in valuation ratios. Cash flows also contain a stationary cycle component—consistent with the short-run cash flow predictability documented by Lettau and Ludvigson (2005) and Koijen and Van Nieuwerburgh (2011)—that introduces additional noise. To test for the presence of these cycles, we estimate a sequence of predictive regressions for horizons $h = 1, \ldots, 7$ years:

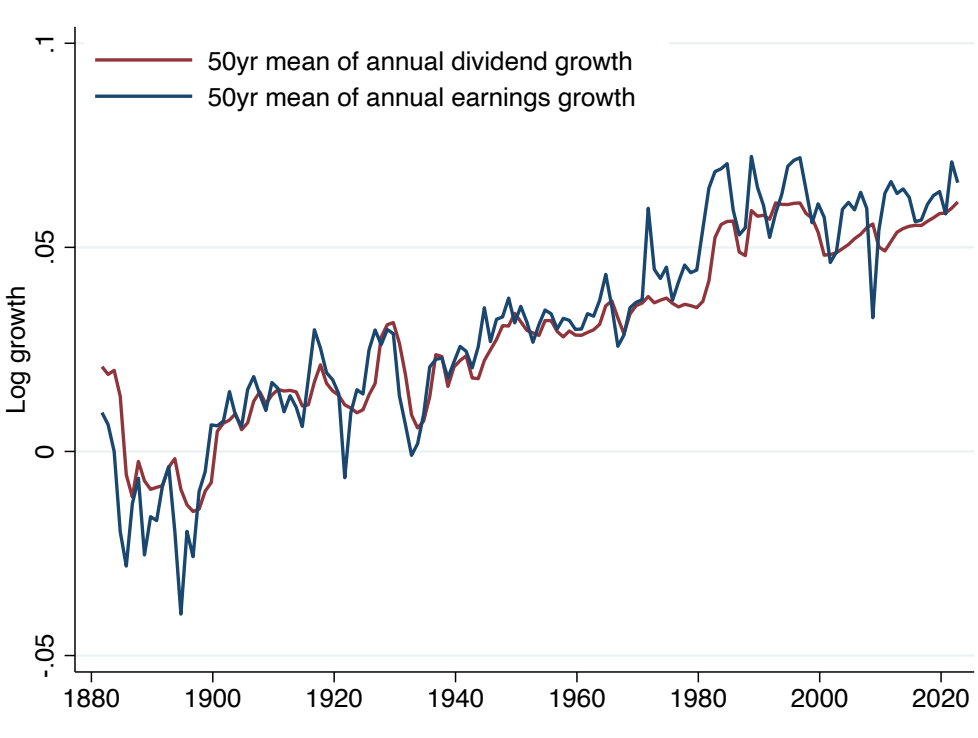
$$\Delta x_{t,t+h} = \alpha_h + \beta_h \Delta x_t + \nu_{t,t+h} \tag{2}$$

Figure 1: Trends in Valuation Ratios and Cash Flows

(A) Valuation Ratios



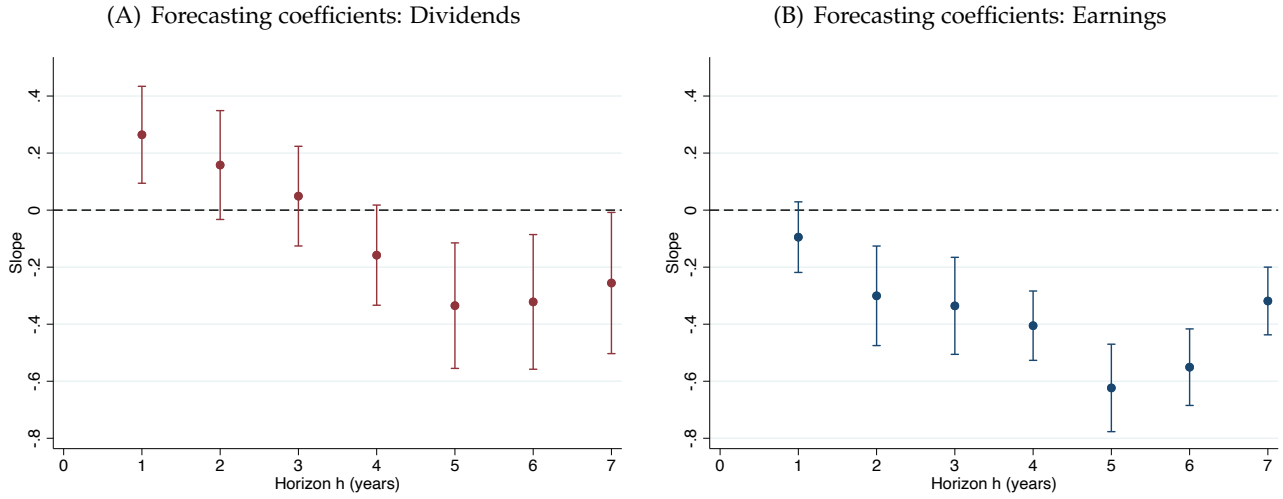
(B) Long-run Growth in Cash Flows



Note: Panel A plots the time series of the log price-dividend ratio, the log price-earnings ratio, and the log cyclically adjusted price-earnings ratio. Panel B plots 50-year rolling window averages of annual log growth in dividends and earnings. The sample is from 1871–2022.

where $\Delta x_{t,t+h}$ is cumulative cash flow growth over the next h years and Δx_t is the past one-year growth rate. Under i.i.d. growth, $\beta_h = 0$. A negative coefficient signals mean-reverting growth: if cash flows contain a transitory cyclical component, a burst of negative growth today will be followed by above-average growth as the variable reverts to trend. Assuming changes in the trend component are negligible at short horizons, these regressions primarily capture the dynamics of the cyclical component.

Figure 2: Cyclicality in Cash Flows



Note: This figure analyzes the cyclicality of annual log dividend and earnings growth, respectively. The figures display slope coefficients β_h from regressions of cumulative h-year growth on current 1-year growth ($\Delta x_{t,t+h} = \alpha_h + \beta_h \Delta x_t + \nu_{t,t+h}$) for log dividends (Panel A) and log earnings (Panel B) following eq. (2). Data are annual. The vertical bars represent 90% confidence intervals computed using Newey-West HAC standard errors.

Figure 2 plots the estimated slope coefficients. The results provide evidence of mean reversion in both dividends and earnings, with the pattern somewhat stronger for earnings—consistent with dividend smoothing by corporations (Lintner, 1956; Brav, Graham, Harvey, and Michaely, 2005). This cyclicality undermines the predictive power of the valuation ratio. As Van Binsbergen and Koijen (2010) demonstrate, when there is predictable variation in cash flows, standard predictive regressions cannot disentangle whether a movement in the valuation ratio is driven by discount rates or growth expectations. If the cash flow cycle is ignored, it obscures the return signal.

3 A Present-Value Model with Trends and Cycles

Motivated by these stylized facts, we develop a present-value model of expected returns and cash flows. To incorporate the stylized facts documented above, the model incorporates a non-stationary cash flow trend and a stationary cycle component.

Our model therefore integrates present-value models, common in finance (e.g., Van Binsbergen and

Koijen, 2010), with a trend-cycle model, widely used in macroeconomic applications (e.g., Beveridge and Nelson, 1981; Nelson and Plosser, 1982; Harvey, 1985).[6] Trend–cycle decompositions have long been standard in macroeconomics, providing a powerful way to separate persistent trends from transitory fluctuations in aggregate macroeconomic variables. Corporate finance has more recently adopted a similar perspective, highlighting the distinction between permanent and temporary cash-flow shocks in both theoretical (e.g., Gorbenko and Strebulaev, 2010) and empirical (e.g., Gryglewicz, Mancini, Morellec, Schroth, and Valta, 2022) work. Yet, asset pricing has made surprisingly little use of this idea, with the notable eception of Lettau and Ludvigson (2004), who decompose household wealth into trend and cycle components via the cointegrated relationship between consumption, wealth, and labor income.

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.

Cash Flows We decompose log dividends, d_t, into a non-stationary trend, τ_t, and a stationary 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 specification constitutes a "local linear" trend model since the trend τ_t is locally (from $t-1$ to t) linear in the growth rate g_t. The "slope" or growth rate g_t is itself a random walk, so the trend can permanently accelerate or decelerate over time; setting $\sigma_g^2 = 0$ recovers constant trend growth, as embedded in classical present-value models (Shiller, 1981). Since the trend level evolves deterministically given the growth rate—$\tau_t = \tau_{t-1} + g_{t-1}$ with no idiosyncratic shock—g_t is the sole source of trend non-stationarity, yielding a smooth (integrated random walk) trend. The cycle c_t follows a parsimonious and stationary ARMA(1,1) process ($|\rho_c| < 1$), capturing the mean-reverting cash flow dynamics documented in Figure 2. Relative to Van Binsbergen and Koijen (2010), who model expected dividend growth as a single stationary AR(1), our decomposition separates permanent shifts in trend growth from a transitory cash-flow cycle.

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

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 $\bar{\mu}$:

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

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.

Present-Value Identity Present-value identities specify stock prices as a function of future cash flows and returns. They typically start with the definition of log returns

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

and then replace the nonlinear term $\log(1 + \exp(pd_{t+1}))$ with a first-order linear Taylor approximation. Famously, Campbell and Shiller (1988) suggest linearizing around $\log(1 + \exp(\overline{pd}))$ where $\overline{pd}$ is the mean pd ratio. However, non-stationarity in cash flows and valuation ratios renders log-linearization around a fixed mean inappropriate in our setting.

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 $g_t < \bar{\mu}$,[7] it is defined by

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

Intuitively, higher trend growth g_t increases the trend pd ratio $\overline{pd}_t$, while higher average returns decrease it.

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:

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

where δ is a Jensen's-inequality constant close to zero and $\rho \approx 0.96$ (see the appendix for the full derivation). This looks similar to the familiar equation in Campbell and Shiller (1988) which specifies the price-

[7] This condition ensures that the price-dividend ratio is finite. It is always satisfied by our empirical estimates, since the estimated drift in g_t remains well below $\bar{\mu}$. In Section 6 we relax the stationary expected return assumption and allow a permanent component in expected returns, which provides additional flexibility. One could also impose this restriction directly on the cash-flow process (e.g., Barsky and De Long, 1993); we choose not to do so to keep the model parsimonious.

dividend ratio is a function of future expected cash flow growth and future returns. By contrast, equation 7 states that *deviations* of price-dividend ratio from its time-varying steady-state value, $pd_t - \overline{pd}_t$, are related to *deviations* in cash flows and returns from their steady-state value.

Substituting the cashflow and expected return dynamics into (7) allows us to express the valuation ratio purely in terms of the state variables:

$$pd_t = \delta + \overline{pd}_t - \frac{1}{1-\rho\rho_\mu}(\mu_t - \bar{\mu}) - \frac{1-\rho_c}{1-\rho\rho_c}\, c_t \;+\; \frac{1-\rho}{1-\rho\rho_c}\,\theta_c\,\varepsilon_{c,t}. \tag{8}$$

or more compactly:

$$pd_t = \delta + \overline{pd}_t + \mathcal{M}_t + \mathcal{C}_t \tag{9}$$

The three components have distinct roles. The non-stationary component, $\overline{pd}_t$, captures permanent 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. Thus, 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 has an ambiguous effect: a high dividend level relative to trend depresses the price-dividend ratio—as mean reversion implies lower future growth—while a positive dividend shock raises it.

3.2 Model Predictions

Our model makes direct predictions regarding why predictive regressions with valuation ratios have limited forecasting ability for returns: cash flow trends and cycles introduce measurement error into valuation ratios.

To understand this argument, suppose we want to predict the j-period return $r_{t,t+j} = \sum_{i=1}^{j} \rho^{i-1} r_{t+i}$. The conditional expectation $\mathbb{E}_t[r_{t,t+j}]$

$$r_{t,t+j} = \mathbb{E}_t[r_{t,t+j}] + \psi_{t,t+j}, \qquad \mathbb{E}_t(\psi_{t,t+j}) = 0. \tag{10}$$

is the best possible predictor of realized returns given time-t information and when regressing realized returns $r_{t,t+j}$ on this predictor we get a slope coefficient of 1. However, this predictor is generally not observed in practice. Instead, imagine we observe a noisy proxy

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

The predictive slope coefficient becomes

$$\frac{\text{Cov}(r_{t,t+j}, x_t)}{\text{Var}(x_t)} = \frac{\text{Var}(\mathbb{E}_t[r_{t,t+j}])}{\text{Var}(x_t)} = 1 - \frac{\text{Var}(\eta_t)}{\text{Var}(x_t)}. \tag{12}$$

Thus, the slope is no longer 1 but is attenuated toward zero by the presence of noise – the classical measurement error result.

The same logic applies in the present-value model. From equation (8), the price-dividend ratio can be written as

$$pd_t = \delta + \mathcal{M}_t + \eta_t, \tag{13}$$

so that the valuation ratio contains the long-run expected-return signal alongside a cash-flow noise component $\eta_t = \overline{pd}(g_t) + \mathcal{C}_t$. Forecasting $r_{t,t+j}$ with pd_t yields

$$b_j = \frac{\text{Cov}(r_{t,t+j}, pd_t)}{\text{Var}(pd_t)} = -[1 - (\rho\rho_\mu)^j]\frac{\text{Var}(\mathcal{M}_t)}{\text{Var}(pd_t)} = -[1 - (\rho\rho_\mu)^j]\left[1 - \frac{\text{Var}(\eta_t)}{\text{Var}(pd_t)}\right]. \tag{14}$$

The first term, $-[1 - (\rho\rho_\mu)^j]$, reflects the fact that the valuation ratio contains the expected-returns signal for the entire term structure of returns. As $j \to \infty$, this term converges to 1. The second term captures attenuation from cash-flow noise. If there were no noise in valuation ratios, this term would equal 1. In such case, the entire variance of pd_t would be driven by $\mathcal{M}_t$. As the variance of the noise component η_t increases, the predictive coefficient shrinks toward zero. In other words, the predictor suffers from classical measurement error.

Solving for $\text{Var}(\eta_t)$, we can write

$$b_j = -[1 - (\rho\rho_\mu)^j]\left[1 - \frac{\text{Var}(\overline{pd}(g_t)) + \text{Var}(\mathcal{C}_t)}{\text{Var}(pd_t)}\right]. \tag{15}$$

We provide the analytical formulas for $\text{Var}(\overline{pd}(g_t))$, $\text{Var}(\mathcal{C}_t$ and $\text{Var}(pd_t)$ in the appendix. The model leads to the following proposition regarding the attenuation bias in predictive regressions.

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

This proposition implies that univariate predictive regressions of realized returns on valuation ratios may recover little of the underlying expected-return signal whenever valuation ratios are substantially driven by cash flow trends and cycles. This offers a potential explanation for why the empirical return

predictability literature has found mixed evidence. The proposition also delivers precise quantitative predictions that we take to the data in the next section.

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

- *The univariate slope coefficient from regressing j-period realized returns $r_{t,t+j}$ on pd_t is given by eq.* (15) *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. (8). 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:

$$b_j \;=\; \frac{\mathrm{Cov}(r_{t,t+j},\, \mathcal{M}_t)}{\mathrm{Var}(\mathcal{M}_t)} \;=\; -[1-(\rho_\mu \rho)^j] \tag{16}$$

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.*

This provides a second testable prediction. Because $\mathcal{M}_t$ is purged of cash flow dynamics, its predictive slope is not attenuated, and augmenting the regression with trend and cycle controls has no effect.

4 Testing the Model on Return Predictability

We now estimate the model and evaluate its predictions. As a further validation exercise, we evaluate our model out of sample following the framework laid out by Goyal and Welch (2008).

4.1 Estimation

4.1.1 State-Space Representation

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

$$s_t = (\tau_t,\; g_t,\; c_t,\; \varepsilon_{c,t},\; \mu_t,\; 1)^\top \tag{17}$$

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

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & \rho_c & \theta_c & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \rho_\mu & (1-\rho_\mu)\bar{\mu} \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \qquad C = \begin{bmatrix} 0 & 0 & 0 \\ \sigma_g & 0 & 0 \\ 0 & \sigma_c & 0 \\ 0 & \sigma_c & 0 \\ 0 & 0 & \sigma_\mu \\ 0 & 0 & 0 \end{bmatrix}, \tag{18}$$

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:

$$x_t = (d_t, pd_t)^{\top} \tag{19}$$

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):

$$G_t = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & \lambda_t & -\dfrac{1-\rho_c}{1-\rho\rho_c} & \dfrac{1-\rho}{1-\rho\rho_c}\,\theta_c & -\dfrac{1}{1-\rho_\mu\rho} & \kappa_t \end{bmatrix}, \tag{20}$$

with

$$\lambda_t = 1 + \frac{\exp(\hat{g}_{t|t-1})}{\exp(\bar{\mu}) - \exp(\hat{g}_{t|t-1})}, \tag{21}$$

$$\kappa_t = \delta + \frac{\bar{\mu}}{1-\rho\rho_\mu} + \hat{g}_{t|t-1} - \log\big(\exp(\bar{\mu}) - \exp(\hat{g}_{t|t-1})\big) - \lambda_t\, \hat{g}_{t|t-1}. \tag{22}$$

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.

4.1.2 Estimation Results

We estimate the model at quarterly frequency using maximum likelihood.[8] For each candidate parameter vector, the EKF delivers filtered state estimates and an approximate log-likelihood; we select the parameters that maximize it. Appendix D.2 provides further details, including our choice of initial states. Table 2 reports the parameter estimates and bootstrapped standard errors.

Table 2: Trend-Cycle Model: Parameter Estimates

Parameter	ρ_c	θ_c	$\sigma_g(\cdot 100)$	σ_c	ρ_μ	σ_μ	$\bar{\mu}$
Estimate	0.9710	0.4927	0.0226	0.0273	0.9556	0.0048	0.0215
	(0.0080)	(0.0823)	(0.0166)	(0.0044)	(0.0357)	(0.0031)	(0.0030)

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 error are shown in parentheses.

Figure 3 plots the filtered states. Panel A reveals the key finding: trend growth drifts from roughly 2% per annum (0.5% per quarter) before 1900 to approximately 6% per annum (1.5% per quarter) in recent decades—consistent with the pronounced secular rise in valuation ratios documented in Section 2. Innovations to trend growth are, however, individually small with an annualized volatility of roughly 18bps—but because trend growth follows a random walk, they accumulate over time.[9] 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. The annualized AR(1) for cyclical dividend *levels* is $\rho_c^4 \approx 0.89$, implying the cyclical component of dividend *growth* has an AR(1) coefficient of approximately -0.11—close to i.i.d. but slowly mean-reverting—with an annualized volatility of 7.5%.[10]

To validate out-of-sample stability, we conduct a real-time filtering exercise over the post-war period. Beginning in 1946, we re-estimate all parameters by MLE at each subsequent quarter using only data available up to that date, retaining only the terminal filtered state from each vintage. The resulting real-time-filtered states (dashed blue in Figure 3) track the full-sample estimates (solid line) almost perfectly, confirming that the filter's dynamics are robust to recursive re-estimation.

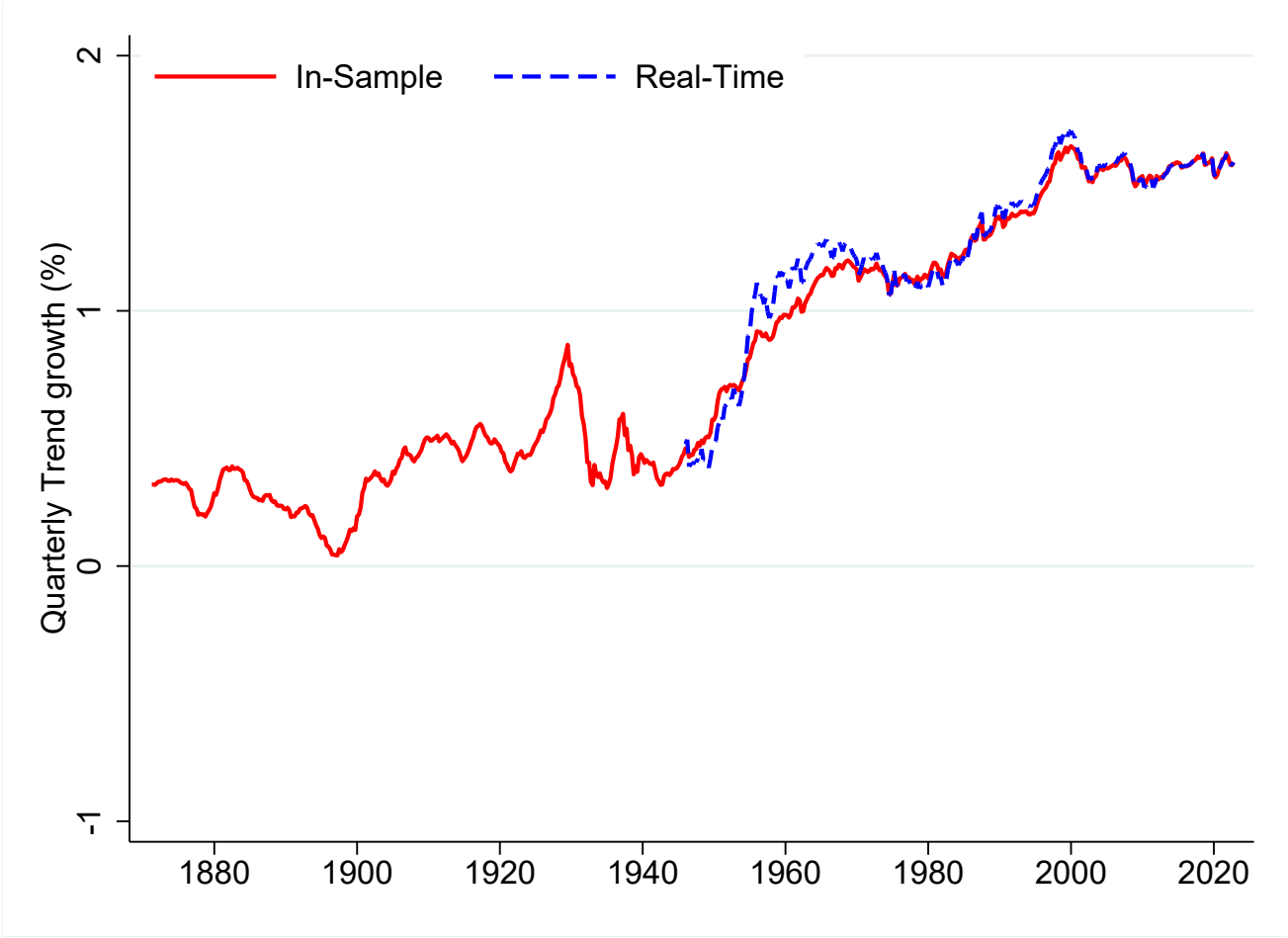
[8] We use annual dividends, take logs and subtract $\log(4)$ to convert them to a quarterly series. Following Cochrane (2011), the Campbell and Shiller (1988) identity requires dividends to be reinvested to the end of the measurement period rather than simply summed or reinvested at the risk-free rate; we approximate this by compounding quarterly dividends at a fixed annual rate of 8%.

[9] Annual trend growth is the sum of four quarterly trend-growth shocks, so its standard deviation is $4\,\sigma_g\,\sqrt{4}$.

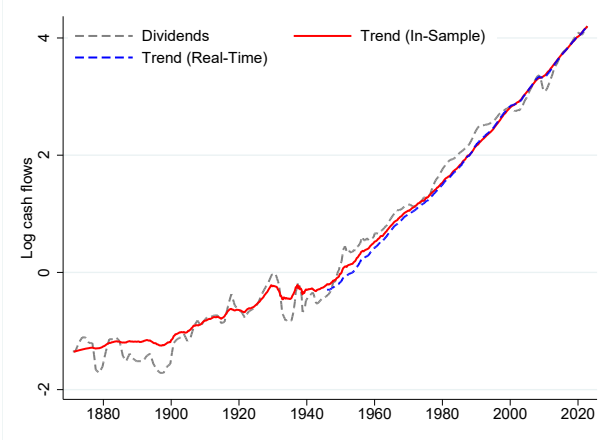
[10] Annual cycle growth $c_t - c_{t-4}$ has variance $2[\gamma(0) - \gamma(4)]$, where $\gamma(0) = (1 + \theta_c^2 + 2\rho_c\theta_c)\sigma_c^2/(1-\rho_c^2)$, $\gamma(1) = \rho_c\,\gamma(0) + \theta_c\,\sigma_c^2$, and $\gamma(h) = \rho_c\,\gamma(h-1)$ for $h \geq 2$. Substituting our estimates yields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

Figure 3: Trend-Cycle Model: Filtered States

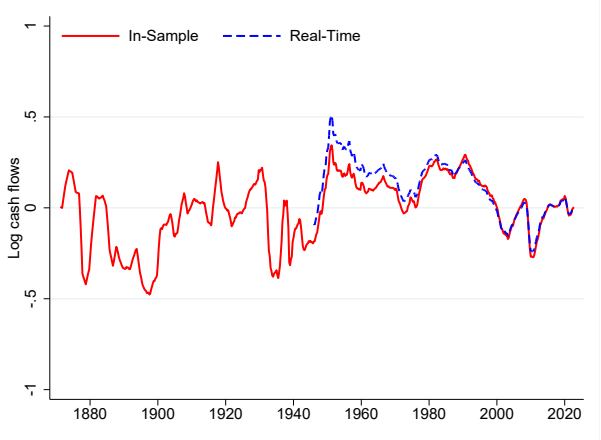
(A) Quarterly Trend growth



(B) Trend



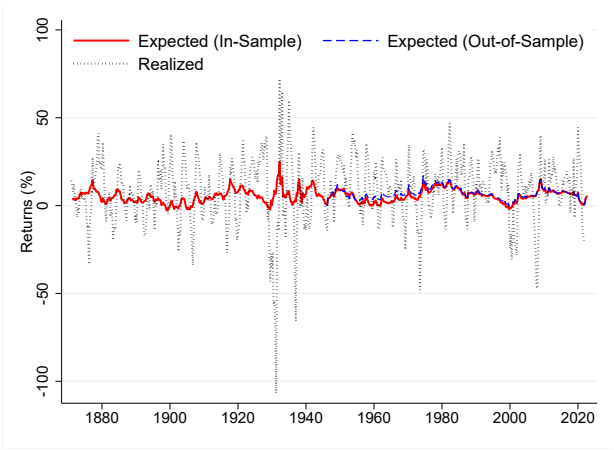
(C) Cycle



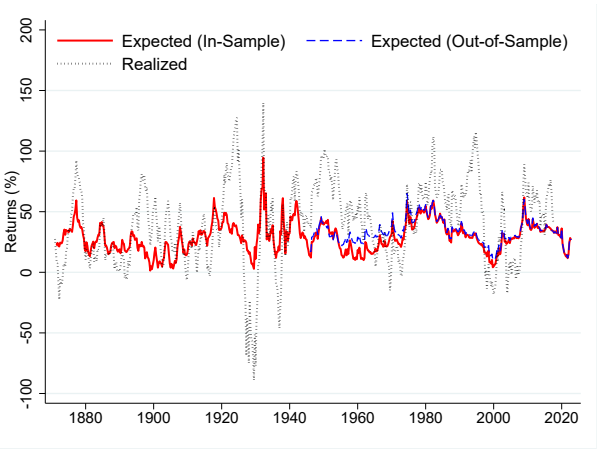
Note: This figure displays the filtered estimates for trend growth ($\hat{g}_t$), the trend ($\hat{\tau}_t$), and the cycle ($\hat{c}_t$) on a quarterly frequency. The solid red line plots the in-sample estimates derived from the full sample (1871–2022), while the dashed blue line plots the real-time estimates derived from an expanding training window starting in 1946.

Figure 4: Expected and Realized Returns

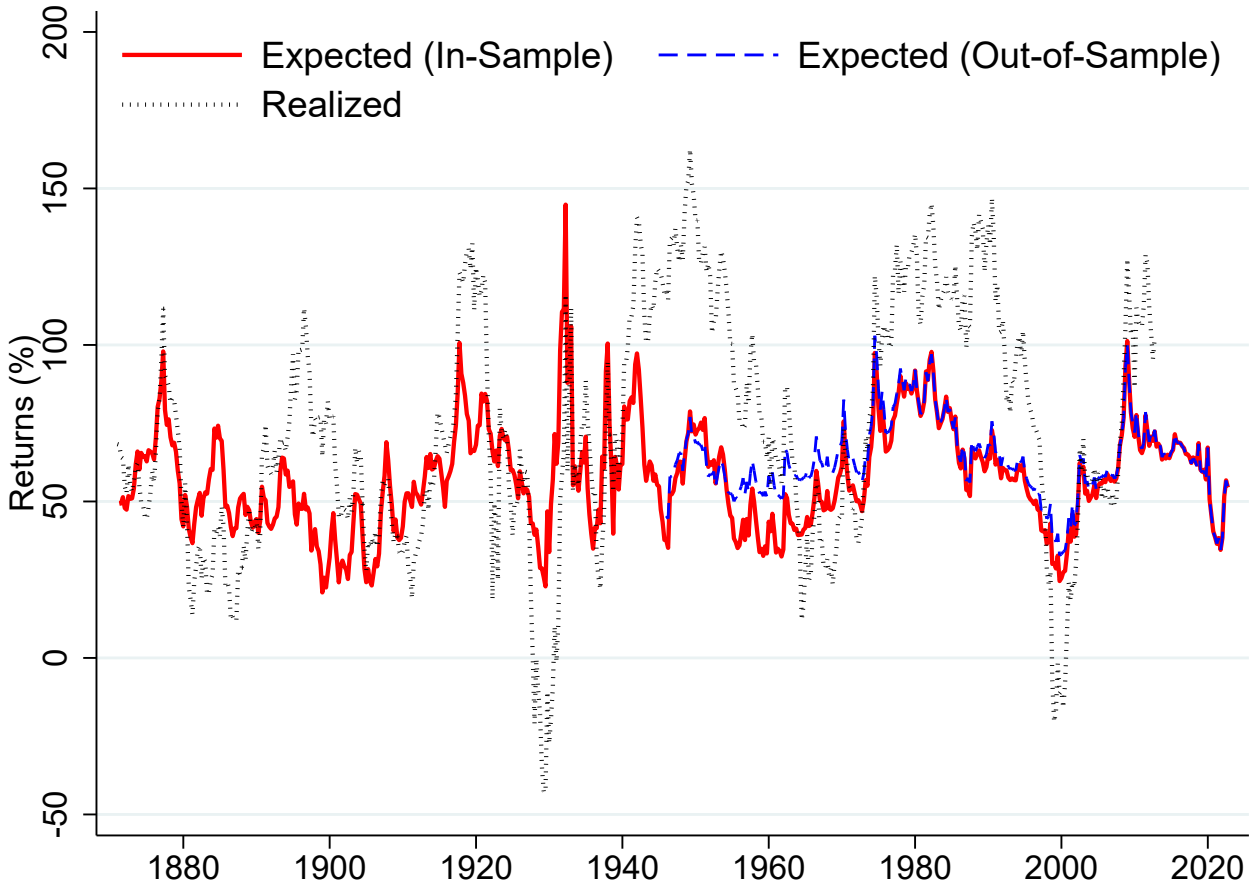
(A) 1-year



(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 (23) 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 1946. Realized returns are shown for comparison (dotted black line).

Given the parameter estimates and filtered states from the EKF, we use the theoretical identity in Equation (4) to obtain model-implied measures of expected returns over the next j years. Substituting estimated parameters and filtered states yields

$$\widehat{\mathbb{E}}_t[r_{t,t+j}] = \frac{1-\rho^j}{1-\rho}\hat{\bar{\mu}} + \frac{1-(\rho\,\hat{\rho}_\mu)^j}{1-\rho\,\hat{\rho}_\mu}(\hat{\mu}_t - \hat{\bar{\mu}}) \tag{23}$$

The first term is the present value of the long-run mean expected return; the second captures the cyclical deviation of expected returns from their long-run mean.

Figure 4 plots model-implied expected returns at the 1-, 5-, and 10-year horizons alongside realized returns. Three features stand out. First, unlike raw valuation ratios, the expected-return series is stationary (see Table A.2)—by construction, the nonstationary trend has been purged, leaving a moderately persistent (annual AR(1) of $\hat{\rho}_\mu \approx 0.83$; half-life of roughly four years) and volatile (annualized volatility of 6.1%) discount-rate signal. 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. Consistent with this persistence, the fit with realized returns improves markedly 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. Third, over the post-war period, the expanding-window estimates (dashed blue) track the full-sample estimates (solid red) very closely, confirming that the expected-return dynamics are robust to real-time re-estimation.

4.2 Testing the Model Predictions

We now test whether the model's predictions are borne out in the data. Table 3 reports predictive regressions of realized returns at the one-year and ten-year horizons. The top panel reports OLS coefficients with t-statistics based on both Newey-West (round brackets) and bootstrapped (square brackets) standard errors to deal with estimated regressors; inference throughout is based on the latter. The bottom panel reports the model-implied (analytical) coefficients and p-values for the null that the empirical and analytical coefficients are equal.

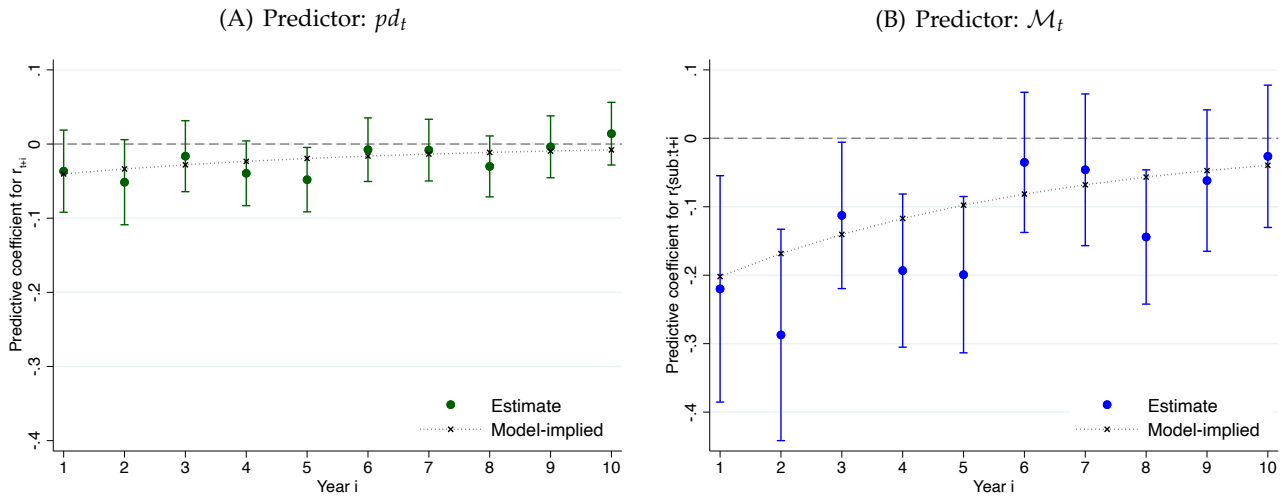
We begin with Prediction 1, which states that cash-flow noise in the pd ratio attenuates the slope in univariate predictive regressions. Column (1) shows the univariate regression of one-year returns on pd_t. The slope is -0.04 ($t = -1.09$)—statistically insignificant and exactly the value the model predicts. If this attenuation reflects cash flow noise, controlling for that noise should restore the slope. Column (2) adds the filtered cash flow components—$\widehat{\overline{pd}}_t$, $\hat{c}_t$, and $\hat{\varepsilon}_{c,t}$—to the regression. The slope on pd_t increases

six-fold to -0.24 ($t = -2.05$), statistically significant and close to the analytical prediction of -0.20. The remaining coefficients also align with the model predictions.

Prediction 2 offers an alternative route to restoring predictability: rather than controlling for cash flow noise, one can directly use the filtered expected-return component of pd, $\widehat{\mathcal{M}}_t$—by construction purged of cash-flow trends and cycles—as the sole predictor. Column (3) confirms this: $\widehat{\mathcal{M}}_t$ predicts one-year returns with a slope of -0.22 ($t = -2.19$), close to the predicted -0.20, and raises the R^2 from 1% to 6%. Column (4) shows that the slope on $\widehat{\mathcal{M}}_t$ is stable when cash flow controls are added, and these controls are themselves insignificant—confirming that $\widehat{\mathcal{M}}_t$ fully absorbs the expected-return signal of pd_t.[11]

Columns (5)–(8) repeat the analysis for ten-year returns.[12] The pattern mirrors the one-year results: the univariate slope on pd_t is small and insignificant (column 5); controlling for cash-flow noise produces a five-fold increase (column 6); $\widehat{\mathcal{M}}_t$ alone delivers comparable predictive power (column 7); and additional cash-flow controls leave the slope on $\widehat{\mathcal{M}}_t$ largely unchanged (column 8). Across both horizons, the analytical and empirical coefficients are closely aligned, and we cannot reject their equality at conventional levels.

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 error.

[11] Note that the empirical coefficients are subject to small-sample bias (Stambaugh, 1999). For the pd ratio, the downward bias is around 0.02 (comparing the OLS coefficient of -0.04 to the bias-adjusted coefficient in column (1) of Table 1). This bias is likely smaller for $\widehat{\mathcal{M}}_t$ given its lower persistence. Standard analytical bias corrections rely on stationary predictor dynamics and do not directly apply when regressors include nonstationary components.

[12] Despite using 150 years of data, these coefficients may be biased away from zero due to the overlapping nature of observations (Boudoukh, Israel, and Richardson, 2022). Consistent with this, the empirical coefficients exceed the model-implied coefficients in magnitude for the long-horizon regressions.

Table 3: Test of Model Predictions: Return Predictability

	$r_{t,t+j} = \sum_{j=1}^{T} \rho^{j-1} r_{t+j}$							
	1-year				10-year			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\widehat{\mathcal{M}}_t$			-0.22**	-0.24**			-1.16***	-1.17***
			(-3.80)	(-4.68)			(-5.82)	(-6.06)
			[-2.19]	[-2.07]			[-3.54]	[-3.18]
pd_t	-0.04	-0.24**			-0.23	-1.17***		
	(-1.11)	(-4.45)			(-1.91)	(-6.05)		
	[-1.09]	[-2.05]			[-1.38]	[-3.18]		
$\widehat{\overline{pd}}_t$		0.25*		0.01		1.20**		0.04
		(4.39)		(0.23)		(5.45)		(0.29)
		[1.94]		[0.15]		[2.50]		[0.14]
$\hat{c}_t$		-0.26		-0.08		-0.91**		-0.04
		(-2.64)		(-0.80)		(-2.53)		(-0.11)
		[-1.51]		[-0.59]		[-1.52]		[-0.07]
$\hat{\varepsilon}_{c,t}$		-0.21		-0.24		0.42		0.28
		(-0.61)		(-0.73)		(0.43)		(0.28)
		[-0.31]		[-0.36]		[0.45]		[0.30]
Observations	151	151	151	151	142	142	142	142
R^2	0.01	0.07	0.06	0.07	0.07	0.35	0.34	0.35
Adj. R^2	0.00	0.04	0.05	0.04	0.06	0.33	0.34	0.33
Model-implied coefficients:								
$\mathcal{M}_t$			-0.20	-0.20			-0.89	-0.89
			{0.83}	{0.72}			{0.41}	{0.45}
pd_t	-0.04	-0.20			-0.18	-0.89		
	{0.93}	{0.72}			{0.74}	{0.45}		
$\overline{pd}_t$		0.20		0.00		0.89		0.00
		{0.70}		{0.88}		{0.52}		{0.89}
c_t		-0.15		0.00		-0.67		0.00
		{0.52}		{0.56}		{0.69}		{0.94}
$\varepsilon_{c,t}$		-0.03		0.00		-0.11		0.00
		{0.78}		{0.72}		{0.57}		{0.77}

Note: This table tests Prediction 1 and 2 and reports predictive regressions of realized returns, $\sum_{j=1}^{T} \rho^{j-1} r_{t+j}$, over 1-year and 10-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, t-stats based on Newey-West standard errors in round brackets, and t-stats based on bootstrapped standard errors in square brackets. 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 $H_1: \beta \neq \beta^{\text{analytical}}$, i.e., the probability—under H_0—of a test statistic at least as extreme as the one observed. Significance levels based on bootstrapped standard errors: *(p<0.10), **(p<0.05), ***(p<0.01).

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 $\rho_\mu^{j-1}/(1-\rho\rho_\mu)$. Panel A uses pd_t as the predictor; Panel B uses $\widehat{\mathcal{M}}_t$. In both cases, the empirical slope coefficients decline with horizon, tracking the model-implied decay closely.

4.3 Out-of-Sample Return Predictability

We next test the out-of-sample performance of the trend-cycle model. We follow the methodology of Goyal and Welch (2008), estimating the model in real time using an expanding window and comparing forecasts against the expanding historical mean benchmark. We focus on their prominent post-1965 evaluation period—an era of materially higher trend growth where traditional ratios have performed most poorly.

Table 4 reports the results. Panels A and B report the in-sample R^2 and out-of-sample R^2_{OOS}, respectively. The trend-cycle model delivers R^2_{OOS} of 8.90% for one-year and 22.38% for five-year raw returns, both significant at the 5% level; excess return results are comparable at 9.12% and 16.78%. Traditional valuation ratios, by contrast, uniformly fail: *pd* and *cape* produce R^2_{OOS} of -1.82% and -2.50% for raw returns (Table 1), with neither approaching statistical significance.

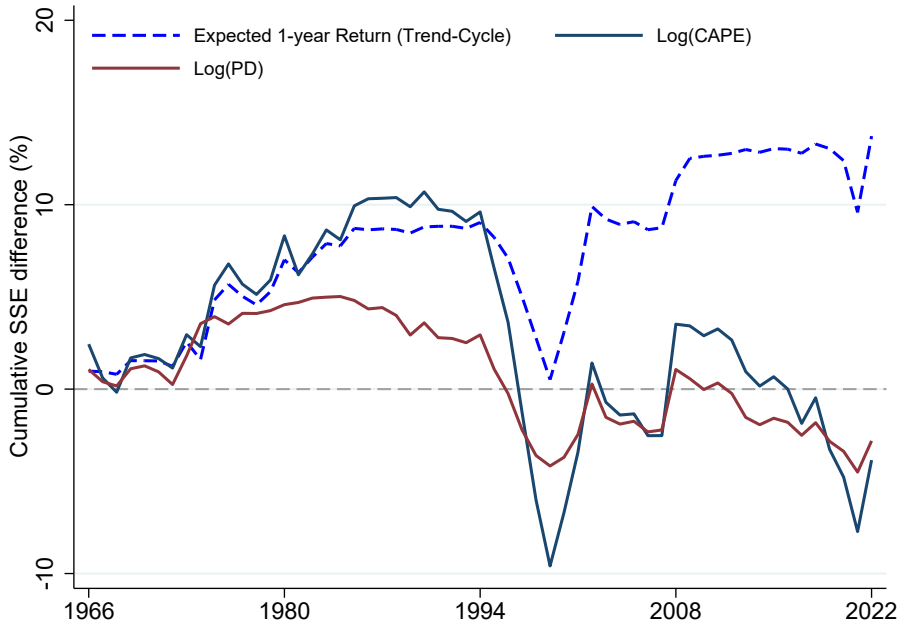
Notably, the out-of-sample R^2_{OOS} is remarkably close to—and in some cases exceeds—the post-1965 in-sample R^2: for example, 8.90% versus 7.65% for one-year raw returns and 22.38% versus 19.86% for five-year raw returns. This near-parity is striking, and reflects two offsetting forces: in-sample estimation benefits from forward-looking parameter estimates but faces a tougher full-sample mean benchmark, whereas the out-of-sample procedure uses only real-time information but is evaluated against the less demanding expanding historical mean.

Figure 6 illustrates the mechanism 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.

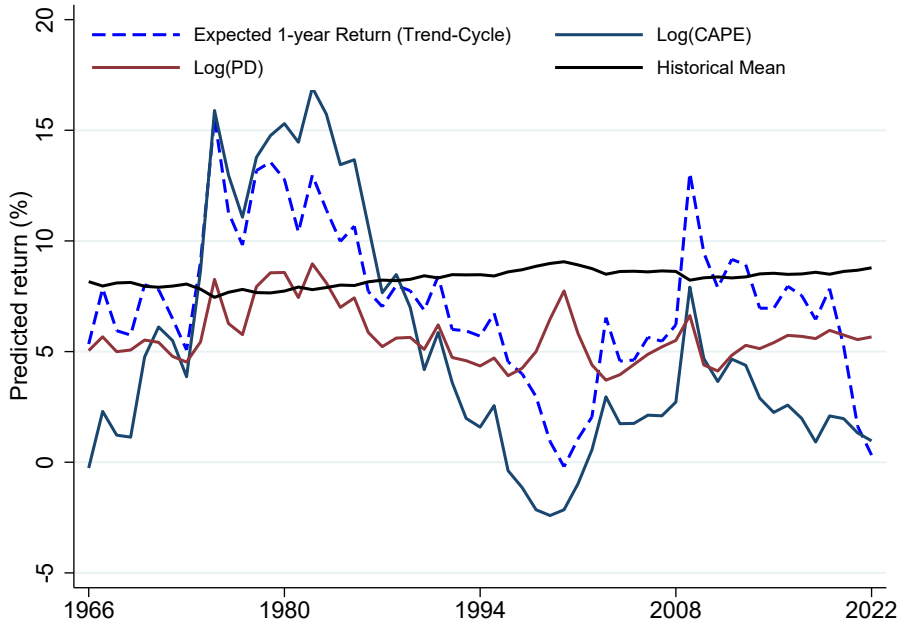
Panel (B) reveals the source of this divergence. The trend-cycle expected return is stationary by construction: it fluctuates around its long-run mean, spiking during market stress—the 1970s stagflation, the 2008 financial crisis—and compressing during booms, consistent with countercyclical risk premia. The trend-cycle model, by explicitly accounting for time variation in g_t, isolates this stationary signal even as valuation ratios drift far from historical norms. The *pd* prediction, by contrast, is nearly flat, its slope coefficients severely attenuated—a direct manifestation of Proposition 1. As trend growth

Figure 6: Out-of-Sample Return Predictability

(A) Cumulative SSE Differences



(B) 1-year Return Prediction



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, $\sum_t \left[(r_{t+1} - \bar{r}_t)^2 - (r_{t+1} - \hat{r}_{t+1}^{\text{OOS}})^2\right]$, so that positive values indicate the predictor outperforms the benchmark. Panel (B) plots the out-of-sample return predictions, $\hat{r}_{t+1}^{\text{OOS}}$, 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.

Table 4: Return Predictability: 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 (%)				
Full sample	2.37%	15.31%	3.87%	21.24%
Post-1965	7.65%	19.86%	8.13%	17.22%
Panel B: Out-of-Sample R^2_{OOS} (%)				
Post-1965	8.90%** [0.020]	22.38%** [0.021]	9.12%** [0.013]	16.78%** [0.033]

Note: Panel A reports in-sample R^2 and Panel B reports out-of-sample R^2_{OOS} for expected returns implied by the trend-cycle model at 1- and 5-year horizons. Multi-year returns are discounted at $\rho = 0.96$. The out-of-sample estimates are obtained by re-estimating the trend-cycle model using an expanding window starting in 1965. 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. $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

drifted permanently higher over the post-war period, valuation ratios rose in tandem, yet the univariate predictive regression—unable to distinguish higher growth from lower discount rates—systematically underpredicted returns.

In sum, where conventional regressions fail to beat the historical average out-of-sample, the trend-cycle model—by explicitly filtering out the nonstationary trend and cash-flow cycle—isolates a stationary expected-return signal that remains informative in real time, even as valuation ratios drift far from historical norms.

5 Asset Pricing Implications

Our model has several implications for expected returns and valuation ratios. We explore these next.

5.1 Model Implications

Time-variation in Expected Returns Our model produces an estimate for the amount of time-variation in expected returns. The volatility of expected returns over j periods is given by

$$\sigma(\mathbb{E}_t[r_{t,t+j}]) = \frac{1-(\rho\,\rho_\mu)^j}{1-\rho\,\rho_\mu}\,\frac{\sigma_\mu}{\sqrt{1-\rho_\mu^2}}. \tag{24}$$

The second term, $\sigma_\mu / \sqrt{1-\rho_\mu^2}$, is the unconditional volatility of the one-period expected return μ_t. The first term is a horizon multiplier, increasing in both the forecast horizon j and the persistence ρ_μ. Because expected returns are persistent, a shift in μ_t propagates into future periods, so that the cumulative j-period expected return inherits a magnified version of the same shock. As $j \to \infty$, the multiplier converges to $1/(1-\rho\,\rho_\mu)$.

Panel A of Table 5 reports the implied volatilities using our parameter estimates. The one-year expected return has a volatility of 6.1%. The horizon-multiplier effect raises the volatility of cumulative 5-year (10-year) expected returns to 20.4% (27.1%). These estimates imply substantially more time-variation in expected returns than standard predictive regressions suggest. Combining the univariate slope estimate of -0.04 (Table 3) with the standard deviation of the *pd* ratio of 0.46 implies a volatility of one-year expected returns of only 1.84%—more than three times smaller than our model-implied estimate.[13] The discrepancy is a direct consequence of the attenuation bias documented in Section 4: because the *pd* ratio embeds cash-flow trend and cycle components, the univariate slope is biased toward zero, compressing the implied variation in expected returns.

Variance Decompositions of Valuation Ratios. A central question in asset pricing is whether discount rate variation accounts for the bulk of stock price movements (e.g., Shiller, 1981; Cochrane, 2011). Our model reveals that the answer depends on the horizon.

Since $pd_t = \delta + \overline{pd}(g_t) + \mathcal{M}_t + \mathcal{C}_t$ in our model, we can decompose its variance as

$$\begin{aligned} &1 = b_{\overline{pd}} + b_{\mathcal{M}} + b_{\mathcal{C}}, \text{ where} \\ &b_{\overline{pd}} = \frac{\mathrm{Var}(\overline{pd}(g_t))}{\mathrm{Var}(pd_t)}, \quad b_{\mathcal{M}} = \frac{\mathrm{Var}(\mathcal{M}_t)}{\mathrm{Var}(pd_t)}, \quad \text{and} \quad b_{\mathcal{C}} = \frac{\mathrm{Var}(\mathcal{C}_t)}{\mathrm{Var}(pd_t)}. \end{aligned} \tag{25}$$

Variation in pd_t is thus explained by three components: the steady-state pd ratio $\overline{pd}_t$, the expected-return cycle $\mathcal{M}_t$, and the cash-flow cycle $\mathcal{C}_t$. Importantly, the processes for expected returns and the cash-flow cycle are stationary, so their variances are independent of the sample length T. By contrast, $\overline{pd}_t$ is a function of the random walk g_t, so its variance grows with the horizon T over which it is computed. We show in the appendix that

$$\mathrm{Var}\big(\overline{pd}(g_t)\big) = \frac{\sigma_g^2 \cdot T}{(1-\rho)^2} + \frac{1}{2}\,\frac{\rho^2}{(1-\rho)^4}\,\sigma_g^4 \cdot T^2, \tag{26}$$

which is increasing in T. Intuitively, because shocks to trend growth are permanent, they accumulate

[13] If $\mathbb{E}_t[r_{t+1}] = a + b\,pd_t$, then $\sigma(\mathbb{E}_t[r_{t+1}]) = |b| \cdot \sigma(pd_t)$.

Table 5: Time-variation in Expected Returns and Variance Decompositions

Panel A: Volatility of expected returns

	Expected return over horizon j			
	j=1-year	j=5-year	j=10-year	j=∞
$\sigma(\mathbb{E}_t[r_{t,t+j}])$	0.061	0.204	0.271	0.304

Panel B: Decomposition of pd_t

Horizon (years)	$\frac{Cov(\mathcal{M}_t, pd_t)}{\text{Var}(pd_t)}$	$\frac{Cov(\mathcal{C}_t, pd_t)}{\text{Var}(pd_t)}$	$\frac{Cov(\overline{pd}_t, pd_t)}{\text{Var}(pd_t)}$
T=1	0.84	0.14	0.02
T=10	0.72	0.12	0.16
T=50	0.44	0.08	0.48
T=150	0.22	0.04	0.74

Panel C: Decomposition of $pd_t - \overline{pd}_t$

Horizon (years)	$\frac{Cov(\mathcal{M}_t, pd_t - \overline{pd}_t)}{\text{Var}(pd_t - \overline{pd}_t)}$	$\frac{Cov(\mathcal{C}_t, pd_t - \overline{pd}_t)}{\text{Var}(pd_t - \overline{pd}_t)}$
$\forall T$	0.85	0.15

Panel D: Persistence

Horizon (years)	ϕ_{pd}	$\phi_{\mathcal{M}}$	$\phi_{\mathcal{C}}$	$\phi_{\overline{pd}}$
T=1	0.85	0.83	0.90	1.00
T=10	0.87	0.83	0.90	1.00
T=50	0.92	0.83	0.90	1.00
T=150	0.96	0.83	0.90	1.00

Note: This table reports model-implied estimates for the volatility of the expected returns (Panel A), variance decompositions of pd_t (Panel B), variance decompositions of $pd_t - \overline{pd}_t$ (Panel C), and annual AR(1) coefficients (Panel D).

rather than wash out.

Panel B of Table 5 implements these decompositions using our parameter estimates. The results reveal that the source of valuation-ratio variation depends critically on the horizon. Over short and medium horizons, expected returns dominate: at a 1-year horizon, shifts in expected returns account for 84% of the variation in pd_t, while the cash-flow cycle explains 14% and the trend pd ratio just 2%. As the horizon lengthens, however, the trend component becomes increasingly important. Over a 150-year horizon, 74% of the variation in pd_t is attributable to shifts in trend growth. The contributions of expected returns and the cash flow cycle shrink to 22% and 4%, respectively.

While the contribution of trend growth rises with the horizon, the relative shares of expected returns and the cash-flow cycle remain fixed. This observation connects to the influential Campbell and Shiller

(1988) variance decompositions of the *pd* ratio, which are typically performed around a fixed sample mean, implicitly assuming stationarity (e.g., Cochrane, 2011). Our results suggest these decompositions should instead be performed around a time-varying steady state. Decomposing the transitory deviation $pd_t - \overline{pd}_t$ yields

$$1 = \frac{\text{Var}(\mathcal{M}_t)}{\text{Var}(pd_t - \overline{pd}_t)} + \frac{\text{Var}(\mathcal{C}_t)}{\text{Var}(pd_t - \overline{pd}_t)}. \tag{27}$$

Panel C shows that, around this time-varying steady state, expected returns explain 85% of stationary valuation-ratio movements at all horizons, with the cash-flow cycle accounting for the remaining 15%.

Taken together, these results speak directly to the analysis of Shiller (1981) and the subsequent critiques by Kleidon (1986) and Marsh and Merton (1986). Shiller (1981) argued that the volatility of aggregate stock prices far exceeds that of ex-post rational prices P^*, constructed by discounting future realized cash flows at a constant rate. However, this conclusion rests on detrending dividends with a constant, full-sample growth rate. As Kleidon (1986) emphasizes, such ex-post detrending suppresses the variance of fundamentals by construction: imposing a deterministic trend on an inherently non-stationary process removes low-frequency variation in growth expectations, so that price movements reflecting stochastic shifts in g_t are misclassified as excess volatility.

Our results offer a quantitative resolution, bringing 150 years of stock-market data to bear within a present-value framework. Building on the minority view of Kleidon (1986) and Barsky and De Long (1993), we find that in the long run the bulk of valuation-ratio variation is driven by cash flow growth. At the same time, consistent with the prevailing interpretation following Shiller (1981), most short- to medium-run variation is "excessive" and reflects shifts in expected returns (or "discount rates" under rational expectations).

The Persistence of Valuation Ratios. Our model also yields a direct prediction for the persistence of the pd ratio. Define the AR(1) coefficient of a variable x as $\phi_x = \text{Cov}(x_t, x_{t-1})/\text{Var}(x_{t-1})$. In our model, the persistence of pd_t is a variance-weighted average of its components' individual persistences:

$$\phi_{pd} = \phi_{\overline{pd}} \cdot b_{\overline{pd}} + \phi_{\mathcal{M}} \cdot b_{\mathcal{M}} + \phi_{\mathcal{C}} \cdot b_{\mathcal{C}}. \tag{28}$$

The weights $b_{\overline{pd}}$, $b_{\mathcal{M}}$, and $b_{\mathcal{C}}$ are the variance shares from the decomposition above. The two stationary components, $\mathcal{M}_t$ and $\mathcal{C}_t$, have annual persistences of 0.83 and 0.90, respectively. Because $\overline{pd}_t$ is driven by the random walk g_t, its AR(1) coefficient is approximately 1.

An interesting result emerges. As the horizon T increases, the random-walk component $\overline{pd}_t$ ac-

counts for a growing share of the variance of pd_t, pulling the weighted-average persistence upward. At short horizons, most variation in pd_t is driven by expected returns, so the persistence of the pd ratio is close to that of $\mathcal{M}_t$—for example, 0.85 at a one-year horizon. Over a 150-year horizon, the near-unit-root trend component dominates, and the persistence of pd_t rises to approximately 0.96.

This addresses a long-standing puzzle in asset pricing: why are valuation ratios so extremely persistent when most theories predict discount-rate variation at business-cycle frequencies, implying substantially lower persistence? Our decomposition provides a resolution. The persistence of the expected-return signal $\mathcal{M}_t$ is moderate ($\phi_{\mathcal{M}} = 0.83$), consistent with models that link discount rates to cyclical risk factors. The high observed persistence of pd_t does not reflect slow-moving discount rates; rather, it is inherited from the permanent trend-growth component. Once this component is stripped out, the dynamics of expected returns align more closely with those implied by standard asset pricing theories.

5.2 Cash Flow Predictability

Our model implies that cash-flow growth is predictable at both short and long horizons. Short-run predictability is established (Van Binsbergen and Koijen, 2010; Koijen and Van Nieuwerburgh, 2011), and we confirm it in our setting (results not reported). We focus here on long-horizon cash-flow growth.

Table 6 shows that long-run cash flow growth is highly predictable. Column (1) regresses 10-year dividend growth, $\Delta d_{t,t+10} = \sum_{j=1}^{10} \rho^{j-1} \Delta d_{t+j}$, on the pd ratio and obtains a highly significant coefficient of 0.26 ($t = 2.73$), closely aligned with the model's prediction of 0.28.

The logic of our decomposition runs symmetrically to the return predictability results. Just as cash flow components of the *pd* ratio attenuate return forecasts, the expected-return component is noise when forecasting cash flows. Replacing the *pd* ratio with the filtered cash-flow states $\hat{g}_t$, $\hat{c}_t$, and $\hat{\varepsilon}_{c,t}$ should therefore strengthen predictive power. Column (2) confirms this: the R^2 doubles from 15% to 30%, with both trend growth and the cycle entering significantly. The coefficient on trend growth is 37, close to the model prediction; the large magnitude reflects the horizon multiplier inherent in projecting 40 quarters of cumulative growth onto quarterly trend growth. Column (3) adds the filtered expected-return signal $\widehat{\mathcal{M}}_t$. If the filter successfully separates cash-flow and discount-rate components, $\widehat{\mathcal{M}}_t$ should be uninformative for dividend growth. This is precisely what we find.

These results speak to a central debate in the literature. Cochrane (2008) argues that the absence of short-run cash flow predictability—the "dog that did not bark"—indirectly strengthens the case for return predictability. As Van Binsbergen and Koijen (2010) show, however, short-run cash flow predictability can be restored by leveraging the additional information embedded in a joint state-space system of dividends and the *pd* ratio—although, consistent with our results, it accounts for limited

Table 6: Test of Model Predictions: Cash Flow Predictability

	$\Delta d_{t,t+j} = \sum_{j=1}^{T} \rho^{j-1} \Delta d_{t+j}$						$\rho^j pd_{t+j}$	
	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***
	(2.87)			(3.26)			(7.08)	(3.16)
	[2.73]			[3.66]			[3.23]	[1.77]
$\hat{g}_t$		36.86***	37.51***		73.08***	75.42***		
		(3.34)	(3.39)		(3.65)	(4.20)		
		[2.08]	[1.91]		[2.89]	[2.76]		
$\hat{c}_t$		-0.70***	-0.76***		-1.29***	-1.38***		
		(-3.69)	(-4.00)		(-3.97)	(-3.95)		
		[-2.02]	[-2.15]		[-3.12]	[-3.23]		
$\hat{\varepsilon}_{c,t}$		-0.12	-0.06		0.71	0.78		
		(-0.25)	(-0.12)		(1.08)	(0.95)		
		[-0.16]	[-0.08]		[1.02]	[1.12]		
$\widehat{\mathcal{M}}_t$			-0.20			-0.18		
			(-1.34)			(-0.99)		
			[-0.80]			[-0.59]		
Observations	142	142	142	132	132	132	142	132
R^2	0.15	0.30	0.32	0.26	0.48	0.49	0.42	0.17
Adj. R^2	0.15	0.28	0.30	0.26	0.47	0.47	0.41	0.16
Model-implied coefficients:								
pd_t	0.28			0.46			0.54	0.34
	{0.85}			{0.90}			{0.59}	{0.33}
$\hat{g}_t$		33.10	33.10		55.25	55.25		
		{0.83}	{0.82}		{0.48}	{0.46}		
$\hat{c}_t$		-0.59	-0.59		-0.72	-0.72		
		{0.77}	{0.64}		{0.17}	{0.12}		
$\hat{\varepsilon}_{c,t}$		0.21	0.21		0.14	0.14		
		{0.67}	{0.73}		{0.42}	{0.36}		
$\widehat{\mathcal{M}}_t$			0.00			0.00		
			{0.43}			{0.56}		

Note: This table reports predictive regressions of realized dividend growth, $\sum_{j=1}^{T} \rho^{j-1} \Delta d_{t+j}$, over 10-year and 20-year horizons. 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 full-sample parameter estimates (1871–2022). The top panel reports OLS regression coefficients, Newey-West standard errors in round brackets, and bootstrapped standard errors in square brackets. 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 $H_1: \beta \neq \beta^{\text{analytical}}$, i.e., the probability—under H_0—of a test statistic at least as extreme as the one observed. Significance levels based on bootstrapped standard errors: *(p<0.10), **(p<0.05), ***(p<0.01).

variation at short horizons. We build on this approach but allow for a permanent component in trend growth, which reveals that cash flows are in fact highly predictable at longer horizons, as small trend shocks accumulate over time. Moreover, as Cochrane (2008) himself acknowledges, the "dog that did not bark" argument loses force if the *pd* ratio contains a unit root—precisely the structure our model identifies. Cash flow growth may explain little variation at a one-year horizon yet dominate longer-horizon variation through its permanent component.

Cochrane (2011) finds a return coefficient near one using 15-year returns over 1947–2009, leaving little room for cash flows. With an effective sample window of roughly 50 years, our model implies that discount rates account for roughly half of *pd* variation —a substantial share, but not all. However, as Boudoukh, Israel, and Richardson (2022) highlight, the short sample underlying Cochrane's result rests on only three to four independent observations, a setting in which overlapping-regression biases can inflate the return coefficient. Our model predicts that as the sample lengthens, the cash flow share should rise and the return share should decline, as trend shocks accumulate.

Consistent with this, in our ∼150-year sample, 10-year returns and cash flow growth explain just 23% and 26% of *pd* variation, respectively (model-implied: 18% and 28%). The remaining variation is absorbed by the future price-dividend ratio $\rho^T \, pd_{t+T}$, which accounts for 46% of *pd* variation at the 10-year horizon (column 7).[14] The model attributes this largely to the trend growth component. Consistent with this interpretation, at 20 years cash flow growth explains 47% of *pd* variation (model-implied: 46%), as trend growth increasingly dominates. Across all specifications, the data do not reject a single model-implied restriction.

6 Extensions

Our trend-cycle model makes two choices: it uses dividends as the cash flow measure and a constant steady-state expected return $\bar{\mu}$. This section examines sensitivity to each assumption and confirms that the core findings are robust.

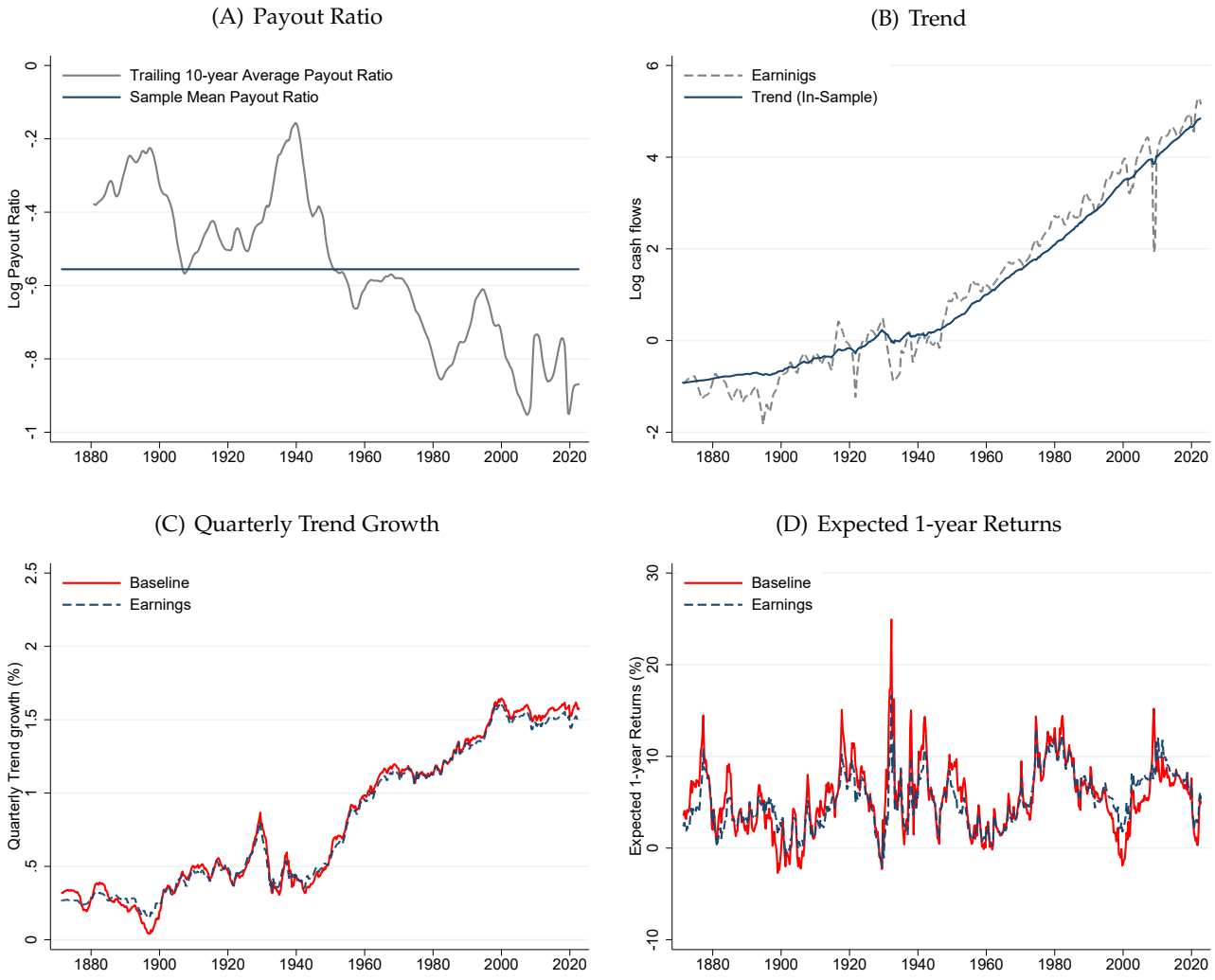
Section 6.1 re-estimates the model using earnings. Section 6.2 allows $\bar{\mu}$ to follow a random walk, addressing the possibility that the baseline misattributes declining expected returns to rising trend growth. In both cases, filtered trend growth is largely unchanged, confirming that the secular rise in valuation ratios is primarily a cash-flow phenomenon. When $\bar{\mu}$ is allowed to drift, it rises modestly from roughly 7% per annum in the pre-1900s to 8-8.5% by the 1980s—the opposite of what one would expect if declining discount rates were driving valuations higher. Section 6.3 shows that in- and out-of-sample return

[14] Note that if the price-dividend ratio contains a unit-root component, the Campbell and Shiller (1988) log-linearization becomes less accurate (Lettau and Van Nieuwerburgh, 2008). As a result, the predictive coefficients on cash-flow growth, returns, and the future price-dividend ratio need not sum exactly to one.

predictability is robust across specifications.

6.1 Trend-Cycle Model for Earnings

Figure 7: 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.

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. However, the secular shift toward share repurchases (Grullon and Michaely, 2002; Skinner, 2008) may raise the question of whether dividends fully capture the market's cash-flow generating capacity.

A shift towards repurchases has two offsetting effects in the dividend model. First, a shift away from

dividends towards repurchases depresses realized dividends growth, which the model may interpret as lower trend growth in dividends. Second, a declining payout ratio mechanically raises the price-dividend ratio, which the model may interpret as signaling higher trend growth in dividends. Because these effects work in opposite directions, the net impact on estimated trend growth is ambiguous in the dividend model.

To assess whether this biases our trend growth estimate, we re-estimate the model using earnings per share, which are unaffected by payout policy. The trade-off is that earnings are a noisy (and volatile) measure of the cash flow that is ultimately relevant to investors—future dividends. The key identity linking dividends and earnings is $d_t = e_t + de_t$, where $de_t \equiv d_t - e_t$ is the log payout ratio. If we assume a constant steady-state payout ratio $\overline{de}$,

$$de_{t+j} = \overline{de} + v_{t+j}, \qquad v_{t+j} \overset{i.i.d.}{\sim} (0, \sigma_v^2), \tag{29}$$

then 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 (see Appendix G for the full derivation):

$$pe_t - \overline{pe}_t \approx \delta + \sum_{j=1}^{\infty} \rho^{j-1} \, \mathbb{E}_t\left[\Delta e_{t+j} - g_t\right] - \sum_{j=1}^{\infty} \rho^{j-1} \, \mathbb{E}_t\left[r_{t+j-1,t+j} - \bar{\mu}\right], \tag{30}$$

where $\overline{pe}_t \equiv \overline{pd}_t + \overline{de}$ is the steady-state log price–earnings ratio. The payout ratio enters only through the constant $\overline{de}$, which shifts the steady-state valuation level. To isolate the effect on trend growth, we fix $\bar{\mu}$ at the baseline dividend-model estimate and $\overline{de}$ at the full-sample mean of $d_t - e_t$.

The constant-payout assumption is a deliberate modelling choice that makes the test conservative. Panel (A) of Figure 7 shows that the trailing 10-year average of de_t has trended downward, lying above $\overline{de}$ in the first half of the sample and below it in the second. Because the model uses a constant $\overline{de}$ throughout, the implied steady-state PE ratio $\overline{pe}_t$ is too high when the payout ratio has fallen below its mean. With $\bar{\mu}$ fixed, the filter can only reconcile this by lowering $\overline{pd}_t$—that is, by reducing estimated trend growth. The earnings specification therefore understates trend growth in the post-repurchase era, if anything.

Figure 7 confirms that our state estimates are robust. Filtered quarterly trend growth from the earnings-based specification closely tracks the dividend-based baseline (Panel C), though it lies very slightly below it in later periods—consistent with the conservative bias. Model-implied expected one-year returns are nearly identical across specifications (Panel D). In Section 6.3, we estimate the earnings-based model with $\bar{\mu}$ freely estimated by MLE and show that its in- and out-of-sample return predictabil-

ity is comparable to the dividend-based baseline.

6.2 Drifting Steady-State Expected Returns

A natural concern is that the secular rise in valuation ratios reflects declining expected returns rather than rising trend growth, and that by holding $\bar{\mu}$ fixed, the baseline misattributes the former to the latter. To address this, we replace the constant $\bar{\mu}$ with a random walk:

$$\begin{aligned} \mu_t &= \bar{\mu}_t + \rho_\mu \left(\bar{\mu}_{t-1} - \bar{\mu}_{t-1}\right) + \epsilon_{\mu,t}, \quad \varepsilon_{\mu,t} \sim \mathcal{N}(0, \sigma_\mu^2) \\ \bar{\mu}_t &= \bar{\mu}_{t-1} + \varepsilon_{\bar{\mu},t}, \quad \varepsilon_{\bar{\mu},t} \sim \mathcal{N}(0, \sigma_{\bar{\mu}}^2), \end{aligned} \tag{31}$$

so that expected returns μ_t mean-revert around a slowly evolving long-run level $\bar{\mu}_t$. The baseline is nested as $\sigma_{\bar{\mu}} = 0$. Full estimation details are in Appendix H.

Figure 8 compares the two specifications. Panel (A) shows that filtered trend growth is similar regardless of whether steady-state expected returns drift, confirming that trend growth is well-identified by the joint dynamics of dividends and prices.

Panel (B) plots steady-state expected returns. In the drifting specification, $\bar{\mu}_t$ rises from roughly 7% per annum at the start of the sample to 8–8.5% by the 1980s, closely tracking the expanding-window mean of realized returns. While rising expected returns and rising pd ratios may seem contradictory, rearranging the Gordon growth model to express returns as $r = g + D/P$ provides the intuition: the surge in trend growth is sufficiently large to allow both expected returns and valuations to rise simultaneously. This reinforces a central conclusion of the paper: the non-stationarity in valuation ratios is primarily a cash-flow phenomenon, consistent with Greenwald, Lettau, and Ludvigson (2025), who find cash flow fundamentals – particularly, rising profit shares – have been a primary driver of postwar equity growth.

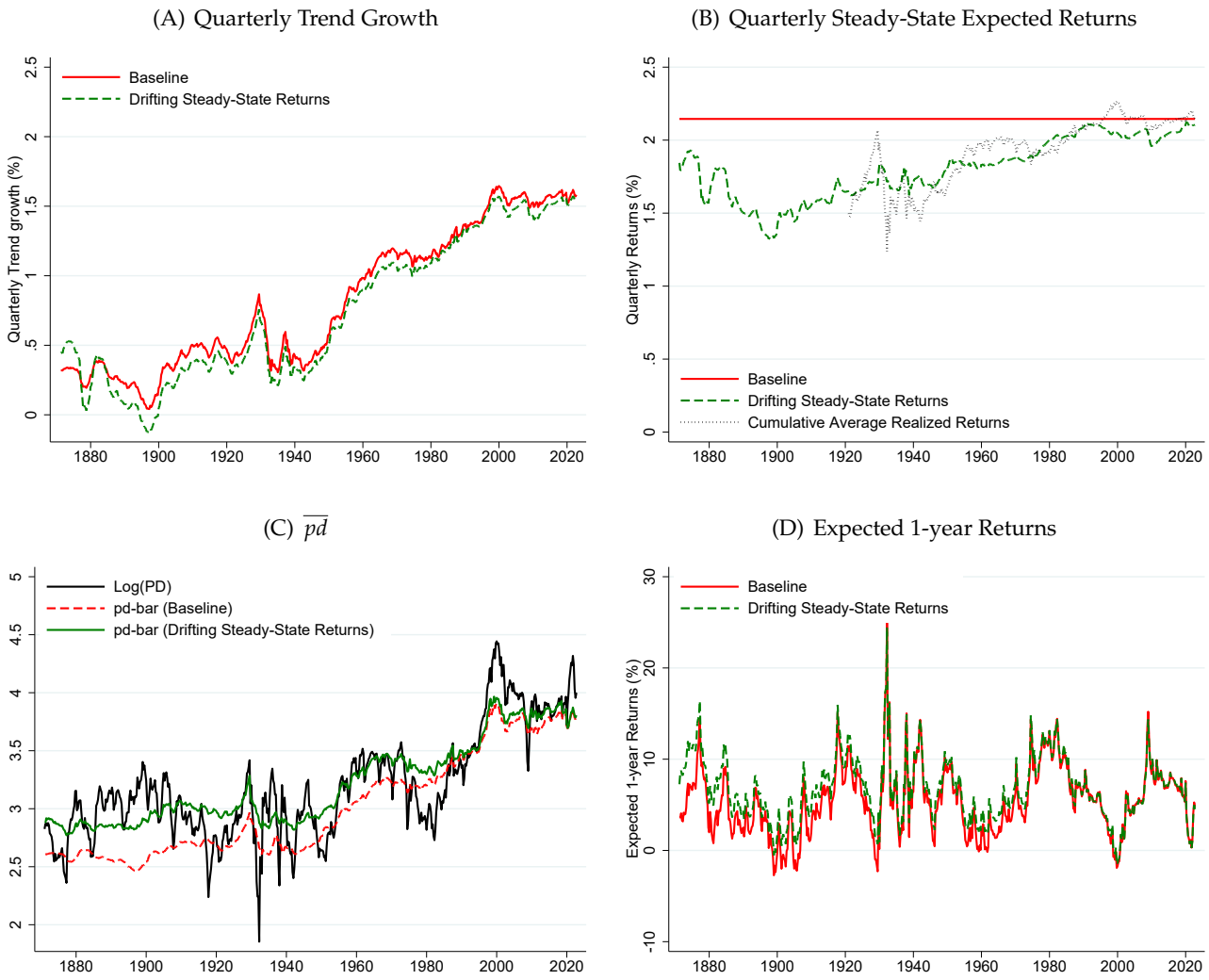
Although both models produce similar trend growth paths, the baseline assumes a higher $\bar{\mu}$ pre-1980. How can both satisfy the valuation identity? The answer lies in the expected return cycle. Rearranging the identity:[15]

$$\underbrace{pd_t - \overline{pd}_t}_{\text{Steady-state valuation gap}} \approx \underbrace{-\frac{1}{1-\rho\rho_\mu}\left(\mu_t - \bar{\mu}\right)}_{\text{Expected return cycle}} \underbrace{-\frac{1-\rho_c}{1-\rho\rho_c}\, c_t \;+\; \frac{1-\rho}{1-\rho\rho_c}\,\theta_c\,\varepsilon_{c,t}}_{\text{Cash flow cycle}}. \tag{32}$$

In the drifting specification, a lower $\bar{\mu}_t$ in the early sample produces a higher $\overline{pd}_t$, keeping the valuation gap near zero. In the baseline, the constant (and higher) $\bar{\mu}$ creates a positive valuation gap, which the model absorbs by driving the expected return cycle negative ($\mu_t - \bar{\mu} < 0$). Intuitively, the baseline

[15] For simplicity, we suppress the Jensen's inequality term δ, which is close to zero.

Figure 8: Trend-Cycle Model – Drifting Steady-State Expected Returns



Note: This figure compares the baseline model, in which the long-run mean of expected returns $\bar{\mu}$ is constant, to the extended model of Appendix H, in which $\bar{\mu}_t$ follows a random walk. 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) compares the observed log price–dividend ratio to the steady-state valuation ratio $\overline{pd}_t$ implied by each model. Panel (D) plots model-implied expected one-year returns under both specifications.

model offsets a steady-state expected return in the earlier periods that is too high by lowering the cyclical component on average.

This distinction matters most at short horizons, where the cycle carries more weight. Panel (D) shows that while one-year expected returns co-move strongly across models, their levels differ: pre-1980s one-year expected returns average around 5.2% per annum in the baseline versus 7% under the drifting specification. At longer horizons these differences diminish as the cycle attenuates. Section 6.3 evaluates return predictability across both specifications.

6.3 Model Extensions: Return Predictability

Table 7: 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 (%)				
Full sample				
Trend-cycle model (dividends)	2.37%	15.31%	3.87%	21.24%
Trend-cycle model (earnings)	2.58%	10.38%	3.77%	15.31%
Trend-cycle model (dividends; drifting $\bar{\mu}$)	3.04%	14.52%	4.70%	21.58%
Post-1965				
Trend-cycle model (dividends)	7.65%	19.86%	8.13%	17.22%
Trend-cycle model (earnings)	6.28%	17.70%	6.57%	16.19%
Trend-cycle model (dividends; drifting $\bar{\mu}$)	7.57%	18.98%	8.00%	15.99%
Panel B: Out-of-Sample R^2_{OOS} (%)				
Post-1965				
Trend-cycle model (dividends)	8.90%**	22.38%**	9.12%**	16.78%**
	[0.020]	[0.021]	[0.013]	[0.033]
Trend-cycle model (earnings)	7.90%**	15.62%**	7.92%***	11.32%*
	[0.018]	[0.023]	[0.006]	[0.056]
Trend-cycle model (dividends; drifting $\bar{\mu}$)	8.10%**	20.16%**	8.35%**	14.70%**
	[0.024]	[0.026]	[0.015]	[0.038]

Note: This table compares in-sample R^2 (Panel A) and out-of-sample R^2_{OOS} (Panel B) across three specifications of the trend-cycle model: the baseline using dividends, an earnings-based specification with $\bar{\mu}$ freely estimated by MLE, and a specification allowing $\bar{\mu}$ to follow a random walk. 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 starting in 1965. 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. $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Table 7 evaluates return predictability across three trend-cycle specifications: the baseline dividend model, an earnings-based variant, and a dividend model with drifting $\bar{\mu}$.[16] Forecasts are generated in real time using an expanding-window procedure beginning in 1965, and performance is measured by R^2_{OOS} relative to the expanding historical-mean benchmark.

The central finding is that predictability is robust across specifications. All three models deliver positive and statistically significant R^2_{OOS} at both horizons for both raw and excess returns. For raw returns,

[16] 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.

the baseline dividend model achieves 8.90% at one year and 22.38% at five years; the earnings-based model delivers 7.90% and 15.62%; and the drifting-$\bar{\mu}$ specification yields 8.10% and 20.16%. Excess return results are comparable, ranging from 7.92% to 9.12% at one year and 11.32% to 16.78% at five years. In all cases, we reject equal predictive accuracy at conventional levels using the Clark and West (2007) test. As with the baseline, out-of-sample R^2_{OOS} is remarkably close to—and in some cases exceeds—the post-1965 in-sample R^2 across all three specifications. Despite differences in the cash-flow measure and the assumed dynamics of steady-state expected returns, each specification isolates a real-time expected-return signal that materially improves upon the historical-mean forecast. The uniformity of these results suggests that the trend-cycle decomposition captures genuine features of the data-generating process, rather than artifacts of a particular specification.

7 Conclusion

Our results offer a refinement of the Shiller (1981) excess volatility puzzle. At short and medium horizons, discount rate variation accounts for roughly 75% of transitory movements in the price-dividend ratio—consistent with prices being excessively volatile relative to near-term fundamentals. At long-run horizons, however, the decomposition reverses: now shifts in permanent trend growth account for nearly three-quarters of price-dividend variation over the last $\sim$150 years. Our results suggest that both Shiller (1981) and his critics (Kleidon, 1986; Marsh and Merton, 1986) were right, but at different horizons.

Because valuation ratios conflate expected returns with cash flow trends and cycles, standard predictive regressions suffer from severe attenuation bias. By jointly filtering cash flows and prices within a present-value framework, we purge this noise and recover expected-return estimates that deliver substantial predictive power—both in- and out-of-sample—where standard valuation ratios uniformly fail. The apparent weakness of return predictability is an artifact of measurement error: once cash flow trends and cycles are filtered out, the stock market is, in fact, highly predictable.

Attributing the unit-root component of valuation ratios to trend growth also eases a long-standing challenge for discount-rate theories, which struggle to match the pronounced persistence of valuation ratios: it leaves a moderately persistent, stationary expected-return process. More broadly, models featuring permanent shifts in cash flow growth—whether through rational channels (Bansal and Yaron, 2004) or behavioral ones (Barsky and De Long, 1993; Bordalo, Gennaioli, Porta, and Shleifer, 2024)—are promising avenues for jointly explaining the long-run cash flow driven variation in stock price ratios and the excess volatility that remains at shorter horizons.

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.

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

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.

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.

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, Matthew Richardson, and Robert F Whitelaw, 2008. The myth of long-horizon predictability. The Review of Financial Studies 21, 1577–1605.

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.

Campbell, John Y, 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 Robert J. Shiller, 1987. Cointegration and tests of present value models. Journal of Political Economy 95, 1062–1088.

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.

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.

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

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

Goetzmann, William N, and Philippe Jorion, 1993. Testing the predictive power of dividend yields. The Journal of Finance 48, 663–679.

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

Gorbenko, Alexander S, and Ilya A Strebulaev, 2010. Temporary versus permanent shocks: Explaining corporate financial policies. The Review of Financial Studies 23, 2591–2647.

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.

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

Gryglewicz, Sebastian, Loriano Mancini, Erwan Morellec, Enrique Schroth, and Philip Valta, 2022. Understanding cash flow risk. The Review of Financial Studies 35, 3922–3972.

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.

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.

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, 2004. Understanding trend and cycle in asset values: Reevaluating the wealth effect on consumption. american economic review 94, 276–299.

——— , 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: 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.

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.

Nelson, Charles R, and Myung J Kim, 1993. Predictable stock returns: The role of small sample bias. The Journal of Finance 48, 641–661.

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.

Shephard, Neil G., and Andrew C. Harvey, 1990. On the probability of estimating a deterministic component in the local level model. Journal of Time Series Analysis 11, 339–347.

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

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.

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

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

APPENDIX

A Data Description

- **S&P 500 price index**: We obtain the S&P 500 index from Amit Goyal's website (replication data for Goyal and Welch (2008)). This is the S&P 500 as of year-end.
- **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.

B Structural Breaks in Valuation Ratios

Table A.1: Structural Breaks in Valuation Ratios

	Log(PD) (1)	Log(PE) (2)	Log(CAPE) (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).

C Derivations: Present-Value with Trends and Cycles

C.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:

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

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")

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

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 (but this over here.)

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

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

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

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

imposing a no-bubble condition $\mathbb{E}_t\left[\lim_{h\to\infty}(\prod_{i=1}^{h}\rho_{t+i})\left(pd_{t+h} - \overline{pd}_{t+h}\right)\right] = 0$ and separating the expectations

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

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

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 two approximating assumptions to keep the present-value formula tractable.

First, and following (Lettau and Van Nieuwerburgh, 2008), we make the assumption of a constant

$\rho = 0.99$, such that $\mathbb{E}_t\left[\prod_{i=1}^{j}\rho_{t+i}\right] = \rho^j$. This is likely a good approximation since ρ_{t+j} is generally around 0.99, but is expected to slow drift upward towards to 1 for large j. However, since future deviations in cash flows and returns from their trend values are 0 for large j, discounting them by a factor closer to one has a negligible overall effect.

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

$$\Delta g_{t+1+j} = g_{t+1+j} - g_{t+j}. \tag{A.18}$$

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

$$\overline{pd}(g_{t+j} + \Delta g) - \overline{pd}(g_{t+j}) \approx \overline{pd}'(g_{t+j})\,\Delta g + \frac{1}{2}\overline{pd}''(g_{t+j})(\Delta g)^2. \tag{A.19}$$

Taking conditional expectation at time $t + j$,

$$\mathbb{E}_{t+j}\left[\overline{pd}_{t+1+j} - \overline{pd}_{t+j}\right] \approx \overline{pd}'(g_{t+j})\mathbb{E}_{t+j}[\Delta g_{t+1+j}] + \frac{1}{2}\overline{pd}''(g_{t+j})\mathrm{Var}_{t+j}(\Delta g_{t+1+j}). \tag{A.20}$$

Since g_t follows a random walk with martingale increments,

$$\mathbb{E}_{t+j}[\Delta g_{t+1+j}] = 0, \tag{A.21}$$

so the first term drops out and

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

The second derivative is

$$\overline{pd}''(g) = \frac{\exp(\bar{\mu} + g)}{(\exp(\bar{\mu}) - \exp(g))^2}. \tag{A.23}$$

Since $\rho_j = \frac{\exp(\overline{pd}_j)}{1+\exp(\overline{pd}_j)}$ and therefore

$$\exp(\overline{pd}_j) = \frac{\exp(g_{t+j})}{\exp(\bar{\mu}) - \exp(g_{t+j})}. \tag{A.24}$$

it follows that

$$\rho_{t+j} = \frac{\dfrac{\exp(g_{t+j})}{\exp(\bar{\mu}) - \exp(g_{t+j})}}{1 + \dfrac{\exp(g_{t+j})}{\exp(\bar{\mu}) - \exp(g_{t+j})}} = \frac{\exp(g_{t+j})}{\exp(\bar{\mu})}. \tag{A.25}$$

Equivalently,

$$\exp(g_{t+j}) = \rho \exp(\bar{\mu}), \tag{A.26}$$

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

This means we can write

$$\overline{pd}''(g_{t+j}) = \frac{\exp(\bar{\mu} + g_{t+j})}{\left(\exp(\bar{\mu}) - \exp(g_{t+j})\right)^2} \tag{A.28}$$

$$= \frac{\exp(\bar{\mu})\exp(g_{t+j})}{\exp(\bar{\mu})^2(1-\rho)^2} \tag{A.29}$$

$$= \frac{\rho_{t+j}}{(1-\rho_{t+j})^2}. \tag{A.30}$$

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

$$\mathbb{E}_t\left[\overline{pd}_{t+1+j} - \overline{pd}_{t+j}\right] \approx \mathbb{E}_t\left[\frac{1}{2}\frac{\rho_{t+j}}{(1-\rho_{t+j})^2}\mathrm{Var}_{t+j}(\Delta g_{t+1+j})\right]. \tag{A.31}$$

Under the approximation that $\rho_{t+j} \approx \rho$

$$\mathbb{E}_t\left[\overline{pd}_{t+1+j} - \overline{pd}_{t+j}\right] \approx \frac{1}{2}\frac{\rho}{(1-\rho)^2}\sigma_g^2. \tag{A.32}$$

Thus, using these two assumptions we obtain

$$pd_t = \delta + \overline{pd}(g_t) + \mathcal{M}_t + \mathcal{C}_t \tag{A.33}$$

where

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

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

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

so that

$$\Delta d_{t+j} - g_t = \underbrace{\left(g_{t+j-1} - g_t\right)}_{\text{trend drift}} + \underbrace{\left(c_{t+j} - c_{t+j-1}\right)}_{\text{cycle change}}. \tag{A.34}$$

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:

$$\sum_{j=1}^{\infty} \rho^{j-1}\, \mathbb{E}_t\big[\Delta d_{t+j} - g_t\big] \;=\; \sum_{j=1}^{\infty} \rho^{j-1}\, \mathbb{E}_t\big[c_{t+j} - c_{t+j-1}\big]. \tag{A.35}$$

We split this into two sums:

$$\sum_{j=1}^{\infty} \rho^{j-1}\, \mathbb{E}_t\big[c_{t+j} - c_{t+j-1}\big] \;=\; \underbrace{\sum_{j=1}^{\infty} \rho^{j-1}\, \mathbb{E}_t[c_{t+j}]}_{\equiv\, S_1} \;-\; \underbrace{\sum_{j=1}^{\infty} \rho^{j-1}\, \mathbb{E}_t[c_{t+j-1}]}_{\equiv\, S_2}. \tag{A.36}$$

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

$$\mathbb{E}_t[c_{t+h}] = \rho_c^{h}\, c_t + \rho_c^{h-1}\, \theta_c\, \varepsilon_{c,t}. \tag{A.37}$$

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.37) and letting $\varphi \equiv \rho\rho_c$,

$$\begin{aligned} S_1 &= \sum_{j=1}^{\infty} \rho^{j-1} \Big[\rho_c^{j}\, c_t + \rho_c^{j-1}\, \theta_c\, \varepsilon_{c,t}\Big] \\ &= \rho_c\, c_t \sum_{j=1}^{\infty} \varphi^{j-1} \;+\; \theta_c\, \varepsilon_{c,t} \sum_{j=1}^{\infty} \varphi^{j-1} \\ &= \frac{\rho_c}{1-\varphi}\, c_t \;+\; \frac{\theta_c}{1-\varphi}\, \varepsilon_{c,t}. \end{aligned} \tag{A.38}$$

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

$$\begin{aligned} S_2 &= c_t + \sum_{j=2}^{\infty} \rho^{j-1}\, \mathbb{E}_t[c_{t+j-1}] \\ &= c_t + \sum_{j=2}^{\infty} \rho^{j-1} \Big[\rho_c^{j-1}\, c_t + \rho_c^{j-2}\, \theta_c\, \varepsilon_{c,t}\Big]. \end{aligned} \tag{A.39}$$

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.39):

$$\begin{aligned} S_2 &= c_t + \frac{\varphi}{1-\varphi}\, c_t + \frac{\rho}{1-\varphi}\, \theta_c\, \varepsilon_{c,t} \\ &= \frac{1}{1-\varphi}\, c_t + \frac{\rho\, \theta_c}{1-\varphi}\, \varepsilon_{c,t}. \end{aligned} \tag{A.40}$$

Combining. Subtracting (A.40) from (A.38):

$$\begin{aligned} S_1 - S_2 &= \frac{\rho_c - 1}{1-\varphi}\, c_t \;+\; \frac{(1-\rho)\, \theta_c}{1-\varphi}\, \varepsilon_{c,t} \\ &= -\frac{1-\rho_c}{1-\rho\rho_c}\, c_t \;+\; \frac{(1-\rho)\, \theta_c}{1-\rho\rho_c}\, \varepsilon_{c,t}. \end{aligned} \tag{A.41}$$

The cash-flow cycle is therefore

$$\mathcal{C}_t \;\equiv\; -\frac{1-\rho_c}{1-\rho\rho_c}\, c_t \;+\; \frac{(1-\rho)\, \theta_c}{1-\rho\rho_c}\, \varepsilon_{c,t}. \tag{A.42}$$

C.2 Variances

Prices are given by

$$pd_t = \delta + \overline{pd}(g_t) + \mathcal{M}_t + \mathcal{C}_t \tag{A.43}$$

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.

C.2.1 Variance of Cash-Flow Cycle Component

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

$$c_t = \rho_c\,c_{t-1} + \theta_c\,\varepsilon_{c,t-1} + \varepsilon_{c,t}, \quad , \quad \varepsilon_{c,t} \sim \mathcal{N}(0, \sigma_c^2) \tag{A.44}$$

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

$$\mathcal{C}_t = \underbrace{\frac{\rho_c - 1}{1-\rho\,\rho_c}}_{\beta_c}\,c_t \;+\; \underbrace{\frac{\theta_c\,(1-\rho)}{1-\rho\,\rho_c}}_{\beta_\varepsilon}\,\varepsilon_{c,t}. \tag{A.45}$$

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

$$\mathrm{Var}(\mathcal{C}_t) = \beta_c^2\,\mathrm{Var}(c_t) + \beta_\varepsilon^2\,\mathrm{Var}(\varepsilon_{c,t}) + 2\,\beta_c\,\beta_\varepsilon\,\mathrm{Cov}(c_t,\,\varepsilon_{c,t}). \tag{A.46}$$

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:

$$\mathrm{Var}(c_t) = \sigma_c^2\left(\frac{1 + \theta_c^2 + 2\,\rho_c\,\theta_c}{1-\rho_c^2}\right). \tag{A.47}$$

Substituting:

$$\mathrm{Var}(\mathcal{C}_t) = \sigma_c^2\left[\beta_c^2\left(\frac{1 + \theta_c^2 + 2\,\rho_c\,\theta_c}{1-\rho_c^2}\right) + \beta_\varepsilon^2 + 2\,\beta_c\,\beta_\varepsilon\right]. \tag{A.48}$$

C.2.2 Derivation of $Var(pd(\bar{g}_t))$

Process for the trend growth rate:

$$g_t = g_{t-1} + \varepsilon_{g,t}, \quad \varepsilon_{g,t} \sim \mathcal{N}(0, \sigma_g) \tag{A.49}$$

We want to know the variance of $\overline{pd}(g_t) = g_t - \log(\exp(\bar{\mu}) - \exp(g_t))$.

Step 1: We approximate the variance of $\overline{pd}(g_t)$ using a second-order Taylor expansion around $\bar{g}$.
Step 1a: Second-order Taylor expansion.

$$\overline{pd}(g_t) = \overline{pd}(\bar{g}) + \overline{pd}'(\bar{g})\,(g_t - \bar{g}) + \frac{1}{2}\overline{pd}''(\bar{g})\,(g_t - \bar{g})^2 + o\big((g_t - \bar{g})^2\big)\,. \tag{A.50}$$

Step 1b: Compute the mean.

Using $\mathbb{E}[g_t - \bar{g}] = 0$ and $\mathbb{E}[(g_t - \bar{g})^2] = \sigma_t^2$, we obtain

$$\mathbb{E}\big[\overline{pd}(g_t)\big] = \overline{pd}(\bar{g}) + \frac{1}{2}\overline{pd}''(\bar{g})\,\sigma_t^2 + O(\sigma_t^4). \tag{A.51}$$

Step 1c: Compute the second moment.

Squaring the Taylor expansion and keeping terms up to order σ_t^4, and using $\mathbb{E}[(g_t - \bar{g})^3] = 0$ and $\mathbb{E}[(g_t - \bar{g})^4] = 3\sigma_t^4$, gives

$$\mathbb{E}\big[\overline{pd}(g_t)^2\big] = \big(\overline{pd}(\bar{g})\big)^2 + \big(\overline{pd}'(\bar{g})\big)^2\sigma_t^2 + \overline{pd}(\bar{g})\,\overline{pd}''(\bar{g})\,\sigma_t^2 + \frac{3}{4}\big(\overline{pd}''(\bar{g})\big)^2\sigma_t^4 + o(\sigma_t^4). \tag{A.52}$$

Step 1d: Square of the mean.

From Step 1b,

$$\big(\mathbb{E}[\overline{pd}(g_t)]\big)^2 = \big(\overline{pd}(\bar{g})\big)^2 + \overline{pd}(\bar{g})\,\overline{pd}''(\bar{g})\,\sigma_t^2 + \frac{1}{4}\big(\overline{pd}''(\bar{g})\big)^2\sigma_t^4 + o(\sigma_t^4). \tag{A.53}$$

Step 1e: Variance.

Subtracting the squared mean from the second moment yields

$$\mathrm{Var}\big(\overline{pd}(g_t)\big) = \mathbb{E}\big[\overline{pd}(g_t)^2\big] - \big(\mathbb{E}[\overline{pd}(g_t)]\big)^2 \tag{A.54}$$

$$\approx \big(\overline{pd}'(\bar{g})\big)^2\sigma_t^2 + \frac{1}{2}\big(\overline{pd}''(\bar{g})\big)^2\sigma_t^4 + o(\sigma_t^4). \tag{A.55}$$

Thus, relative to the first-order delta approximation $\big(\overline{pd}'(\bar{g})\big)^2\sigma_t^2$, the second-order Taylor expansion adds the positive curvature term $\frac{1}{2}\big(\overline{pd}''(\bar{g})\big)^2\sigma_t^4$.

Step 2: Below we substitute the derivatives of

$$\overline{pd}(g) = g - \log\big(\exp(\bar{\mu}) - \exp(g)\big)$$

into the second-order variance approximation.

Step 2a: First derivative:

$$\overline{pd}'(g) = 1 - \frac{d}{dg}\log\big(\exp(\bar{\mu}) - \exp(g)\big) \tag{A.56}$$

$$= 1 - \frac{-\exp(g)}{\exp(\bar{\mu}) - \exp(g)} \tag{A.57}$$

$$= \frac{\exp(\bar{\mu})}{\exp(\bar{\mu}) - \exp(g)}. \tag{A.58}$$

Evaluating at $g = \bar{g}$:

$$\overline{pd}'(\bar{g}) = \frac{\exp(\bar{\mu})}{\exp(\bar{\mu}) - \exp(\bar{g})}. \tag{A.59}$$

Step 2b: Second derivative:

$$\overline{pd}''(g) = \frac{d}{dg}\left(\frac{\exp(\bar{\mu})}{\exp(\bar{\mu}) - \exp(g)}\right) \tag{A.60}$$

$$= \frac{\exp(\bar{\mu})\exp(g)}{\big(\exp(\bar{\mu}) - \exp(g)\big)^2}. \tag{A.61}$$

Evaluating at $g = \bar{g}$:

$$\overline{pd}''(\bar{g}) = \frac{\exp(\bar{\mu} + \bar{g})}{\big(\exp(\bar{\mu}) - \exp(\bar{g})\big)^2}. \tag{A.62}$$

Step 3: Rewrite everything in terms of $\rho = \frac{\exp(\overline{pd})}{1+\exp(\overline{pd})}$ where $\overline{pd} = \bar{g} - \log\big(\exp(\bar{\mu}) - \exp(\bar{g})\big)$.
Step 3a: Rewrite $\exp(\overline{pd})$.

$$\overline{pd} = \bar{g} - \log\big(\exp(\bar{\mu}) - \exp(\bar{g})\big) \tag{A.63}$$

$$= \log\left(\frac{\exp(\bar{g})}{\exp(\bar{\mu}) - \exp(\bar{g})}\right), \tag{A.64}$$

which implies

$$\exp(\overline{pd}) = \frac{\exp(\bar{g})}{\exp(\bar{\mu}) - \exp(\bar{g})}. \tag{A.65}$$

Step 3b: Express ρ in terms of $\bar{g}$ and $\bar{\mu}$.

$$\rho = \frac{\dfrac{\exp(\bar{g})}{\exp(\bar{\mu}) - \exp(\bar{g})}}{1 + \dfrac{\exp(\bar{g})}{\exp(\bar{\mu}) - \exp(\bar{g})}} \tag{A.66}$$

$$= \frac{\exp(\bar{g})}{\exp(\bar{\mu})}. \tag{A.67}$$

Equivalently,

$$\exp(\bar{g}) = \rho \, \exp(\bar{\mu}), \tag{A.68}$$

$$\exp(\bar{\mu}) - \exp(\bar{g}) = \exp(\bar{\mu})(1 - \rho). \tag{A.69}$$

Step 3c: First derivative in terms of ρ.

$$\overline{pd}'(\bar{g}) = \frac{\exp(\bar{\mu})}{\exp(\bar{\mu}) - \exp(\bar{g})} \tag{A.70}$$

$$= \frac{\exp(\bar{\mu})}{\exp(\bar{\mu})(1 - \rho)} \tag{A.71}$$

$$= \frac{1}{1 - \rho}. \tag{A.72}$$

Step 3d: Second derivative in terms of ρ.

$$\overline{pd}''(\bar{g}) = \frac{\exp(\bar{\mu} + \bar{g})}{\big(\exp(\bar{\mu}) - \exp(\bar{g})\big)^2} \tag{A.73}$$

$$= \frac{\exp(\bar{\mu})\exp(\bar{g})}{\exp(\bar{\mu})^2(1 - \rho)^2} \tag{A.74}$$

$$= \frac{\rho}{(1 - \rho)^2}. \tag{A.75}$$

Recall

$$\mathrm{Var}\big(\overline{pd}(g_t)\big) \approx \big(\overline{pd}'(\bar{g})\big)^2 \sigma_t^2 + \frac{1}{2}\big(\overline{pd}''(\bar{g})\big)^2 \sigma_t^4. \tag{A.76}$$

Substituting the expressions above gives

$$\mathrm{Var}\big(\overline{pd}(g_t)\big) \approx \frac{\sigma_{g,t}^2}{(1 - \rho)^2} + \frac{1}{2}\frac{\rho^2}{(1 - \rho)^4}\sigma_{g,t}^4 + o(\sigma_t^4). \tag{A.77}$$

For the random walk g_t, the variance is $\sigma^2_{g,t} = t \cdot \sigma^2_g$, and therefore grows linearly with sample size t.

C.2.3 Variance of Expected-Return Component

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

$$\mu_{t+1} = \bar{\mu} + \rho_\mu \left(\mu_t - \bar{\mu}\right) + \varepsilon_{\mu,t+1}, \quad , \quad \varepsilon_{\mu,t} \sim \mathcal{N}(0, \sigma^2_\mu) \tag{A.78}$$

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

$$\mathrm{Var}(\mathcal{M}_t) = \frac{1}{\left(1 - \rho\,\rho_\mu\right)^2} \, \frac{\sigma^2_\mu}{1 - \rho^2_\mu}. \tag{A.79}$$

D Model Estimation

D.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 $\bar{\mu}$:

$$\begin{aligned} d_t &= \tau_t + c_t, \\ \tau_t &= \tau_{t-1} + g_{t-1}, \\ g_t &= g_{t-1} + \varepsilon_{g,t},, \quad \varepsilon_{g,t} \sim \mathcal{N}(0, \sigma_g^2) \\ c_t &= \rho_c\, c_{t-1} + \theta_c\, \varepsilon_{c,t-1} + \varepsilon_{c,t},, \quad \varepsilon_{c,t} \sim \mathcal{N}(0, \sigma_c^2) \\ \mu_t &= \bar{\mu} + \rho_\mu\, (\mu_{t-1} - \bar{\mu}) + \varepsilon_{\mu,t},, \quad \varepsilon_{\mu,t} \sim \mathcal{N}(0, \sigma_\mu^2) \end{aligned} \tag{A.80}$$

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:

$$s_t = (\tau_t,\ g_t,\ c_t,\ \varepsilon_{c,t},\ \mu_t,\ 1)^\top. \tag{A.81}$$

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

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & \rho_c & \theta_c & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \rho_\mu & (1-\rho_\mu)\bar{\mu} \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \qquad C = \begin{bmatrix} 0 & 0 & 0 \\ \sigma_g & 0 & 0 \\ 0 & \sigma_c & 0 \\ 0 & \sigma_c & 0 \\ 0 & 0 & \sigma_\mu \\ 0 & 0 & 0 \end{bmatrix}, \tag{A.82}$$

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:

$$x_t = (d_t,\ pd_t)^\top. \tag{A.83}$$

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

$$pd_t = \overline{pd}_t + \delta - \frac{1-\rho_c}{1-\rho\rho_c}\, c_t \ + \ \frac{1-\rho}{1-\rho\rho_c}\, \theta_c\, \varepsilon_{c,t} - \frac{1}{1-\rho\rho_\mu}\, (\mu_t - \bar{\mu}). \tag{A.84}$$

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

equation 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, $\hat{g}_{t|t-1}$:

$$\overline{pd}_t \;\approx\; \widehat{\overline{pd}}_t + \lambda_t\,(g_t - \hat{g}_{t|t-1}), \tag{A.85}$$

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

$$\widehat{\overline{pd}}_t \equiv \hat{g}_{t|t-1} - \log\!\big(\exp(\bar{\mu}) - \exp(\hat{g}_{t|t-1})\big), \tag{A.86}$$

$$\lambda_t \equiv \left.\frac{\partial \overline{pd}_t}{\partial g_t}\right|_{g_t=\hat{g}_{t|t-1}} = 1 + \frac{\exp(\hat{g}_{t|t-1})}{\exp(\bar{\mu}) - \exp(\hat{g}_{t|t-1})}. \tag{A.87}$$

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.84) and rearranging to isolate the state vector s_t:

$$\begin{aligned} pd_t \;\approx\; & \delta + \lambda_t\, g_t - \frac{1-\rho_c}{1-\rho\rho_c}\, c_t + \frac{1-\rho}{1-\rho\rho_c}\,\theta_c\,\varepsilon_{c,t} - \frac{1}{1-\rho\rho_\mu}\,\mu_t \\ & + \underbrace{\frac{\bar{\mu}}{1-\rho\rho_\mu} + \widehat{\overline{pd}}_t - \lambda_t\,\hat{g}_{t|t-1}}_{\kappa_t}. \end{aligned} \tag{A.88}$$

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

$$G_t = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & \lambda_t & -\dfrac{1-\rho_c}{1-\rho\rho_c} & \dfrac{1-\rho}{1-\rho\rho_c}\,\theta_c & -\dfrac{1}{1-\rho_\mu\rho} & \kappa_t \end{bmatrix}, \tag{A.89}$$

with

$$\lambda_t = 1 + \frac{\exp(\hat{g}_{t|t-1})}{\exp(\bar{\mu}) - \exp(\hat{g}_{t|t-1})}, \tag{A.90}$$

$$\kappa_t = \delta + \frac{\bar{\mu}}{1-\rho\rho_\mu} + \hat{g}_{t|t-1} - \log\!\big(\exp(\bar{\mu}) - \exp(\hat{g}_{t|t-1})\big) - \lambda_t\,\hat{g}_{t|t-1}. \tag{A.91}$$

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.87) 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.

D.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. Broadly, we set the persistence of expected returns to match the AR(1) of the dividend-to-price ratio, the persistence of the dividend cycle level to match one plus the AR(1) of dividend growth, and the volatility of expected-return innovations to one-quarter of realized return volatility. The quarterly mean expected return $\bar{\mu}$ is initialized to the pre-sample average quarterly return. The volatility of trend growth is set to 10 basis points per quarter, reflecting a slowly evolving trend. 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%.

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.

D.3 Computation of Standard Errors

We compute standard errors for the parameter estimates in Table 3 using a parametric moving-block bootstrap that preserves the temporal dependence structure of the structural shocks. The procedure consists of three steps: recovering the structural innovations from the estimated model, resampling them in blocks to generate synthetic data, and re-estimating the model on each synthetic sample.

Step 1: Recover structural shocks. Using the maximum likelihood estimates $\hat{\Theta} = (\hat{\rho}_c, \hat{\theta}_c, \hat{\sigma}_g, \hat{\sigma}_c, \hat{\rho}_\mu, \hat{\sigma}_\mu, \hat{\bar{\mu}})$, we run the Rauch–Tung–Striebel (RTS) smoother to obtain the smoothed state sequence $\{s_{t|T}\}_{t=1}^{T}$, where $s_t = (\tau_t, g_t, c_t, \varepsilon_{c,t}, \mu_t, 1)^\top$. From the smoothed states we back out the standardized structural innovations $\hat{u}_t = (\hat{\epsilon}_{g,t},\ \hat{\epsilon}_{c,t},\ \hat{\epsilon}_{\mu,t})^\top$ for $t = 2, \ldots, T$ by inverting the state equations:

$$\hat{\epsilon}_{g,t} = \frac{g_{t|T} - g_{t-1|T}}{\hat{\sigma}_g}, \qquad \hat{\epsilon}_{c,t} = \frac{\varepsilon_{c,t|T}}{\hat{\sigma}_c}, \qquad \hat{\epsilon}_{\mu,t} = \frac{\mu_{t|T} - \hat{\rho}_\mu\, \mu_{t-1|T} - (1-\hat{\rho}_\mu)\hat{\bar{\mu}}}{\hat{\sigma}_\mu}.$$

Step 2: Moving-block bootstrap and simulation. For each bootstrap replication $b = 1, \ldots, B$, we resample the trivariate shock sequence $\{\hat{u}_t\}$ using a circular moving-block bootstrap with block length L. Specifically, we draw random starting indices and concatenate consecutive blocks of length L from the (circularly wrapped) shock series until the resampled sequence reaches the required length. This preserves the joint cross-sectional and short-range temporal dependence among the three shocks while allowing their long-run distribution to be explored.

Given the resampled shocks $\{u_t^{*(b)}\}$, we simulate a synthetic dataset $\{d_t^{(b)}, pd_t^{(b)}\}$ by iterating the state-space system forward using the estimated transition equation $s_{t+1} = A(\hat{\Theta})\, s_t + C(\hat{\Theta})\, u_{t+1}^{*(b)}$ and the (linearized) measurement equation, with the same initial conditions used in the original estimation.

Step 3: Re-estimation. On each synthetic sample $\{d_t^{(b)}, pd_t^{(b)}\}_{t=1}^{T}$, we re-estimate the full parameter vector $\hat{\Theta}^{(b)}$ by maximum likelihood, using the same Extended Kalman Filter procedure and optimizer as for the original data. The bootstrap standard error for each parameter θ_j is the sample standard deviation across the B successful replications:

$$\widehat{\mathrm{se}}(\hat{\theta}_j) = \left[\frac{1}{B-1}\sum_{b=1}^{B}\left(\hat{\theta}_j^{(b)} - \bar{\theta}_j\right)^2\right]^{1/2}, \qquad \bar{\theta}_j = \frac{1}{B}\sum_{b=1}^{B}\hat{\theta}_j^{(b)}.$$

In practice, we set $B = 1000$ and use a block length of $L = 20$ quarters (5 years), which is long enough to capture the persistence of the cyclical component ($\rho_c^4 \approx 0.89$ at the annual frequency) while still providing sufficient resampling variation. The resulting standard errors are reported in parentheses in Table 2.

D.4 Structural Break Test for Expected Return Estimate $\widehat{\mathcal{M}}_t$

Table A.2: Structural Break Test for Expected Return Estimate $\widehat{\mathcal{M}}_t$

	Expected Returns (1)
SupF(0,1)	3.497
SupF(0,2)	5.925
SupF(1,2)	2.141

Note: This table reports tests for structural breaks in the inifinite-horizon expected returns estimated from the Extended Kalman Filter (see Section D.1 for full details). 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).

E Model Implications: AR(1) Persistence

Price ratios are given by

$$pd_t = \delta + \overline{pd}(g_t) + \mathcal{M}_t + \mathcal{C}_t \tag{A.92}$$

where

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

Define the AR(1) coefficient of variable x as

$$\phi_x = \frac{\mathrm{Cov}(x_t, x_{t-1})}{\mathrm{Var}(x_{t-1})} = 1 + \frac{\mathrm{Cov}(\Delta x_t, x_{t-1})}{\mathrm{Var}(x_{t-1})} \tag{A.93}$$

Since everything is uncorrelated, the persistence of each component is a weighted average of the individual components:

$$\phi_{pd} = \phi_{\overline{pd}} \cdot \frac{\mathrm{Var}(\overline{pd}_{t-1})}{\mathrm{Var}(pd_{t-1})} + \phi_{\mathcal{M}} \cdot \frac{\mathrm{Var}(\mathcal{M}_{t-1})}{\mathrm{Var}(pd_{t-1})} + \phi_{\mathcal{C}} \cdot \frac{\mathrm{Var}(\mathcal{C}_{t-1})}{\mathrm{Var}(pd_{t-1})} \tag{A.94}$$

To obtain the persistence of the pd ratio, ϕ_{pd}, we therefore need to compute the persistence of the individual components, $\phi_{\overline{pd}}$, $\phi_{\mathcal{M}}$, and $\phi_{\mathcal{C}}$. We do this in the following.

E.1 Persistence of $\mathcal{M}_t$

Trivially $\phi_\mu = \phi_{\mathcal{M}} = \rho_\mu$.

Annual persistence. To go from the quarterly AR(1) to the annual AR(1) coefficient

$$\phi_x = \frac{\mathrm{Cov}(x_t, x_{t-4})}{\mathrm{Var}(x_{t-4})} = 1 + \frac{\mathrm{Cov}(\Delta_4 x_t, x_{t-4})}{\mathrm{Var}(x_{t-4})} \tag{A.95}$$

we use $\phi_\mu^{ANNUAL} = \phi_\mu^4 = \rho_\mu^4$.

E.2 Persistence of $\mathcal{C}_t$

Recall

$$\mathcal{C}_t = -\frac{1-\rho_c}{1-\rho\,\rho_c}\,c_t \;+\; \frac{\theta_c\,(1-\rho)}{1-\rho\,\rho_c}\,\varepsilon_{c,t}. \tag{A.96}$$

Define constants

$$a_c \equiv \frac{1-\rho_c}{1-\rho\,\rho_c}, \qquad b_c \equiv \frac{\theta_c\,(1-\rho)}{1-\rho\,\rho_c}, \tag{A.97}$$

so that

$$\mathcal{C}_t = -a_c\, c_t + b_c\, \varepsilon_{c,t}. \tag{A.98}$$

Quarterly persistence. The quarterly AR(1) coefficient of $\mathcal{C}_t$ is defined as

$$\phi_{\mathcal{C}} \equiv \frac{\mathrm{Cov}(\mathcal{C}_t, \mathcal{C}_{t-1})}{\mathrm{Var}(\mathcal{C}_{t-1})}. \tag{A.99}$$

The cross-moments are

$$\mathrm{Cov}(c_{t-1}, \varepsilon_{c,t-1}) = \sigma_c^2, \qquad \mathrm{Cov}(c_t, \varepsilon_{c,t-1}) = \rho_c\, \mathrm{Cov}(c_{t-1}, \varepsilon_{c,t-1}) + \theta_c\, \mathrm{Var}(\varepsilon_{c,t-1}) = (\rho_c + \theta_c)\sigma_c^2. \tag{A.100}$$

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

$$\begin{aligned}
\mathrm{Cov}(\mathcal{C}_t, \mathcal{C}_{t-1}) &= \mathrm{Cov}(-a_c\, c_t + b_c\, \varepsilon_{c,t},\ -a_c\, c_{t-1} + b_c\, \varepsilon_{c,t-1}) && \text{(A.101)}\\
&= a_c^2\, \mathrm{Cov}(c_t, c_{t-1}) -\ a_c b_c\, \mathrm{Cov}(c_t, \varepsilon_{c,t-1}) && \text{(A.102)}\\
&= a_c^2\, \gamma_1 -\ a_c b_c\, (\rho_c + \theta_c)\sigma_c^2, && \text{(A.103)}
\end{aligned}$$

where $\gamma_1 \equiv \mathrm{Cov}(c_t, c_{t-1})$. Similarly, the variance of $\mathcal{C}_{t-1}$ is

$$\begin{aligned}
\mathrm{Var}(\mathcal{C}_{t-1}) &= \mathrm{Var}(-a_c\, c_{t-1} + b_c\, \varepsilon_{c,t-1}) && \text{(A.104)}\\
&= a_c^2\, \mathrm{Var}(c_{t-1}) + b_c^2\, \mathrm{Var}(\varepsilon_{c,t-1}) - 2a_c b_c\, \mathrm{Cov}(c_{t-1}, \varepsilon_{c,t-1}) && \text{(A.105)}\\
&= a_c^2\, \gamma_0 + b_c^2\, \sigma_c^2 - 2a_c b_c\, \sigma_c^2, && \text{(A.106)}
\end{aligned}$$

where $\gamma_0 \equiv \mathrm{Var}(c_t) = \mathrm{Var}(c_{t-1})$ by stationarity. Hence the quarterly persistence of $\mathcal{C}_t$ is

$$\phi_{\mathcal{C}} = \frac{a_c^2\, \gamma_1 - a_c b_c (\rho_c + \theta_c)\sigma_c^2}{a_c^2\, \gamma_0 + b_c^2\, \sigma_c^2 - 2a_c b_c\, \sigma_c^2}. \tag{A.107}$$

ARMA(1,1) moments for c_t. Let $\gamma_k \equiv \mathrm{Cov}(c_t, c_{t-k})$. As shown above,

$$\gamma_1 = \rho_c\, \gamma_0 + \theta_c\, \sigma_c^2, \qquad \gamma_0 = \frac{1 + \theta_c^2 + 2\rho_c\theta_c}{1 - \rho_c^2}\, \sigma_c^2. \tag{A.108}$$

Therefore,

$$\phi_c \equiv \frac{\gamma_1}{\gamma_0} = \rho_c + \theta_c\, \frac{1 - \rho_c^2}{1 + \theta_c^2 + 2\rho_c\theta_c}. \tag{A.109}$$

Annual persistence. Define the annual (four-quarter) persistence of $\mathcal{C}_t$ as

$$\phi_{\mathcal{C}}^{ANNUAL} \equiv \frac{\mathrm{Cov}(\mathcal{C}_t, \mathcal{C}_{t-4})}{\mathrm{Var}(\mathcal{C}_{t-4})}. \tag{A.110}$$

Using $\mathcal{C}_{t-4} = -a_c c_{t-4} + b_c \varepsilon_{c,t-4}$ and orthogonality of $\varepsilon_{c,t}$ to $\mathcal{F}_{t-1}$, we obtain

$$\begin{aligned} \mathrm{Cov}(\mathcal{C}_t, \mathcal{C}_{t-4}) &= \mathrm{Cov}(-a_c\, c_t + b_c\, \varepsilon_{c,t},\ -a_c\, c_{t-4} + b_c\, \varepsilon_{c,t-4}) && \text{(A.111)} \\ &= a_c^2\, \mathrm{Cov}(c_t, c_{t-4}) - \ a_c b_c\, \mathrm{Cov}(c_t, \varepsilon_{c,t-4}). && \text{(A.112)} \end{aligned}$$

For an ARMA(1,1), $\gamma_4 = \mathrm{Cov}(c_t, c_{t-4}) = \rho_c^3 \gamma_1$ and

$$\mathrm{Cov}(c_t, \varepsilon_{c,t-4}) = \rho_c^3\, \mathrm{Cov}(c_{t-3}, \varepsilon_{c,t-4}) = \rho_c^3(\rho_c + \theta_c)\sigma_c^2, \tag{A.113}$$

so

$$\mathrm{Cov}(\mathcal{C}_t, \mathcal{C}_{t-4}) = \rho_c^3 \Big(a_c^2 \gamma_1 - a_c b_c (\rho_c + \theta_c)\sigma_c^2 \Big). \tag{A.114}$$

Moreover, by stationarity,

$$\mathrm{Var}(\mathcal{C}_{t-4}) = \mathrm{Var}(\mathcal{C}_t) = a_c^2\, \gamma_0 + b_c^2\, \sigma_c^2 - 2 a_c b_c\, \sigma_c^2. \tag{A.115}$$

Hence the annual persistence parameter is

$$\phi_{\mathcal{C}}^{ANNUAL} = \frac{\rho_c^3 \Big(a_c^2 \gamma_1 - a_c b_c (\rho_c + \theta_c)\sigma_c^2 \Big)}{a_c^2\, \gamma_0 + b_c^2\, \sigma_c^2 - 2 a_c b_c\, \sigma_c^2} = \rho_c^3\, \phi_{\mathcal{C}}. \tag{A.116}$$

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

E.3 Persistence of $\overline{pd}_t$

$$\overline{pd}_t = f(g_t), \qquad g_t = g_{t-1} + \varepsilon_{g,t}. \tag{A.117}$$

A first-order Taylor expansion of $f(g_t) = f(g_{t-1} + \varepsilon_{g,t})$ around g_{t-1} yields

$$\overline{pd}_t = f(g_{t-1} + \varepsilon_{g,t}) \approx f(g_{t-1}) + f'(g_{t-1})\, \varepsilon_{g,t}. \tag{A.118}$$

Hence

$$\Delta \overline{pd}_t = \overline{pd}_t - \overline{pd}_{t-1} \approx f'(g_{t-1})\, \varepsilon_{g,t}. \tag{A.119}$$

Using iterated expectations and $\mathbb{E}[\varepsilon_{g,t} \mid \mathcal{F}_{t-1}] = 0$,

$$\begin{aligned}
\mathrm{Cov}(\Delta\overline{pd}_t, \overline{pd}_{t-1}) &\approx \mathrm{Cov}\big(f'(g_{t-1})\,\varepsilon_{g,t},\ \overline{pd}_{t-1}\big) && \text{(A.120)}\\
&= \mathbb{E}\Big[f'(g_{t-1})\,\varepsilon_{g,t}\,\overline{pd}_{t-1}\Big] - \mathbb{E}\big[f'(g_{t-1})\,\varepsilon_{g,t}\big]\,\mathbb{E}[\overline{pd}_{t-1}] && \text{(A.121)}\\
&= \mathbb{E}\Big[f'(g_{t-1})\,\overline{pd}_{t-1}\,\mathbb{E}[\varepsilon_{g,t} \mid \mathcal{F}_{t-1}]\Big] - \mathbb{E}\big[f'(g_{t-1})\,\mathbb{E}[\varepsilon_{g,t} \mid \mathcal{F}_{t-1}]\big]\,\mathbb{E}[\overline{pd}_{t-1}] && \text{(A.122)}\\
&= 0. && \text{(A.123)}
\end{aligned}$$

Therefore

$$\phi_{\overline{pd}} = 1 + \frac{\mathrm{Cov}(\Delta\overline{pd}_t, \overline{pd}_{t-1})}{\mathrm{Var}(\overline{pd}_{t-1})} \approx 1. \tag{A.124}$$

Doing a second-order Taylor expansion instead implies that $\Delta\overline{pd}_t$ contains an additional curvature term proportional to $\overline{pd}''(g_{t-1})\,\varepsilon_{g,t}^2$. Since $\varepsilon_{g,t}^2$ has positive conditional mean σ_g^2, this term generates a positive conditional drift $\mathbb{E}[\Delta\overline{pd}_t \mid \mathcal{F}_{t-1}] \approx \frac{1}{2}\sigma_g^2\overline{pd}''(g_{t-1})$. Using $\rho_{t-1} = \exp(g_{t-1})/\exp(\bar{\mu})$, we have $\overline{pd}''(g_{t-1}) = \rho_{t-1}/(1-\rho_{t-1})^2$, which is increasing in g_{t-1} and therefore increasing in $\overline{pd}_{t-1}$. As a result, higher $\overline{pd}_{t-1}$ predicts larger expected $\Delta\overline{pd}_t$, implying $\mathrm{Cov}(\Delta\overline{pd}_t, \overline{pd}_{t-1}) > 0$. Plugging this into $\phi_{\overline{pd}} = 1 + \mathrm{Cov}(\Delta\overline{pd}_t, \overline{pd}_{t-1})/\mathrm{Var}(\overline{pd}_{t-1})$ yields $\phi_{\overline{pd}} > 1$ under the second-order approximation. However, empirically the second-order term tends to be small and we therefore neglect it in our analytical computations.

Annual persistence. Define the annual (four-quarter) persistence of $\overline{pd}_t$ as the population projection coefficient of $\overline{pd}_t$ onto $\overline{pd}_{t-4}$:

$$\phi_{\overline{pd}}^{ANNUAL} \equiv \frac{\mathrm{Cov}(\overline{pd}_t, \overline{pd}_{t-4})}{\mathrm{Var}(\overline{pd}_{t-4})} = 1 + \frac{\mathrm{Cov}(\overline{pd}_t - \overline{pd}_{t-4},\ \overline{pd}_{t-4})}{\mathrm{Var}(\overline{pd}_{t-4})}. \tag{A.125}$$

Iterating the first-order expansion forward four quarters yields

$$\overline{pd}_t - \overline{pd}_{t-4} \approx \sum_{j=0}^{3} f'(g_{t-1-j})\,\varepsilon_{g,t-j}. \tag{A.126}$$

Each innovation $\varepsilon_{g,t-j}$ is orthogonal to $\mathcal{F}_{t-4}$ (and hence orthogonal to $\overline{pd}_{t-4}$), implying

$$\mathrm{Cov}(\overline{pd}_t - \overline{pd}_{t-4},\ \overline{pd}_{t-4}) \approx 0. \tag{A.127}$$

Therefore,

$$\phi_{\overline{pd}}^{ANNUAL} \approx 1. \tag{A.128}$$

Thus, under the same first-order logic as at the quarterly horizon, annual persistence is also approximately one (rather than $(\phi_{\overline{pd}})^4$), reflecting that $\overline{pd}_t$ is locally random-walk-like under the first-order approximation.

F Cash Flow Predictability – Derivations

F.1 Regression I 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 derives the model-implied regression coefficients.

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

$$\Delta d_{t+k} = g_{t+k-1} + \big(c_{t+k} - c_{t+k-1}\big), \tag{A.129}$$

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

This implies

$$\Delta d_{t,t+j} = \underbrace{\sum_{k=1}^{j} \rho^{k-1}\, g_{t+k-1}}_{\text{trend component}} + \underbrace{\sum_{k=1}^{j} \rho^{k-1} \big(c_{t+k} - c_{t+k-1}\big)}_{\text{cycle component}}. \tag{A.130}$$

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:

$$\mathbb{E}_t\left[\sum_{k=1}^{j} \rho^{k-1}\, g_{t+k-1}\right] = g_t \sum_{k=1}^{j} \rho^{k-1} = g_t\, \frac{1-\rho^j}{1-\rho}. \tag{A.131}$$

Cycle component. We first simplify the weighted sum algebraically

$$\sum_{k=1}^{j} \rho^{k-1} \big(c_{t+k} - c_{t+k-1}\big) = \sum_{k=1}^{j} \rho^{k-1}\, c_{t+k} \;-\; \sum_{k=1}^{j} \rho^{k-1}\, c_{t+k-1}. \tag{A.132}$$

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:

$$\begin{aligned}
\sum_{k=1}^{j} \rho^{k-1} \big(c_{t+k} - c_{t+k-1}\big) &= \sum_{k=1}^{j-1} \rho^{k-1}\, c_{t+k} + \rho^{j-1}\, c_{t+j} - c_t - \sum_{k=1}^{j-1} \rho^{k}\, c_{t+k} \\
&= -\,c_t + \sum_{k=1}^{j-1} \big(\rho^{k-1} - \rho^{k}\big)\, c_{t+k} + \rho^{j-1}\, c_{t+j} \\
&= -\,c_t + (1-\rho) \sum_{k=1}^{j-1} \rho^{k-1}\, c_{t+k} + \rho^{j-1}\, c_{t+j}.
\end{aligned} \tag{A.133}$$

Conditional expectation of the cycle sum. Using the ARMA(1,1) forecast (A.37), 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.133).

Interior sum.

$$\begin{aligned}(1-\rho)\sum_{k=1}^{j-1}\rho^{k-1}\,\mathbb{E}_t[c_{t+k}] &= (1-\rho)\sum_{k=1}^{j-1}\rho^{k-1}\Big[\rho_c^k\, c_t + \rho_c^{k-1}\,\theta_c\,\varepsilon_{c,t}\Big] \\ &= (1-\rho)\,\rho_c\, c_t \sum_{k=1}^{j-1}\varphi^{k-1} \;+\; (1-\rho)\,\theta_c\,\varepsilon_{c,t}\sum_{k=1}^{j-1}\varphi^{k-1} \\ &= (1-\rho)\,\rho_c\,\frac{1-\varphi^{j-1}}{1-\varphi}\,c_t \;+\; (1-\rho)\,\frac{1-\varphi^{j-1}}{1-\varphi}\,\theta_c\,\varepsilon_{c,t}. \end{aligned} \tag{A.134}$$

Terminal term.

$$\rho^{j-1}\,\mathbb{E}_t[c_{t+j}] = \rho^{j-1}\,\rho_c^j\, c_t \;+\; \rho^{j-1}\,\rho_c^{j-1}\,\theta_c\,\varepsilon_{c,t}. \tag{A.135}$$

Assembling. Combining (A.134) and (A.135) with the $-c_t$ term from (A.133):

$$\begin{aligned}\mathbb{E}_t\left[\sum_{k=1}^{j}\rho^{k-1}\big(c_{t+k}-c_{t+k-1}\big)\right] &= \underbrace{\left[-1+(1-\rho)\,\rho_c\,\frac{1-\varphi^{j-1}}{1-\varphi}+\rho^{j-1}\rho_c^{j}\right]}_{\equiv\; A_j^{\rho}}\; c_t \\ &\quad+\; \underbrace{\left[(1-\rho)\,\frac{1-\varphi^{j-1}}{1-\varphi}+\rho^{j-1}\rho_c^{j-1}\right]}_{\equiv\; B_j^{\rho}}\;\theta_c\,\varepsilon_{c,t}. \end{aligned} \tag{A.136}$$

Adding the trend component (A.131) and the cycle component (A.136):

$$\mathbb{E}_t\big[\Delta d_{t,t+j}\big] \;=\; \frac{1-\rho^j}{1-\rho}\,g_t \;+\; A_j^{\rho}\,c_t \;+\; B_j^{\rho}\,\theta_c\,\varepsilon_{c,t}, \tag{A.137}$$

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

$$A_j^{\rho} = -1+(1-\rho)\,\rho_c\,\frac{1-\varphi^{j-1}}{1-\varphi}+\rho^{j-1}\rho_c^{j}, \qquad B_j^{\rho} = (1-\rho)\,\frac{1-\varphi^{j-1}}{1-\varphi}+\rho^{j-1}\rho_c^{j-1}. \tag{A.138}$$

F.2 Regression II with pd ratio

We run a predictive regression of $\Delta d_{t,t+j} \;\equiv\; \sum_{k=1}^{j}\rho^{k-1}\,\Delta d_{t+k}$ on pd_t. In the following, we derives the model-implied regression coefficient $\frac{\mathrm{Cov}(\Delta d_{t,t+j},\, pd_t)}{\mathrm{Var}(pd_t)}$.

Step 1: Decompose the numerator. From equation (9), the price–dividend ratio decomposes as

$$pd_t = \delta + \overline{pd}(g_t) + \mathcal{M}_t + \mathcal{C}_t,$$

where $\overline{pd}(g_t)$, $\mathcal{M}_t$, and $\mathcal{C}_t$ are mutually uncorrelated (since they are driven by independent shocks $\varepsilon_{g,t}$, $\varepsilon_{\mu,t}$, and $\varepsilon_{c,t}$, respectively). Therefore

$$\mathrm{Cov}(\Delta d_{t,t+j},\, pd_t) = \mathrm{Cov}\big(\Delta d_{t,t+j},\, \overline{pd}(g_t)\big) + \mathrm{Cov}\big(\Delta d_{t,t+j},\, \mathcal{M}_t\big) + \mathrm{Cov}\big(\Delta d_{t,t+j},\, \mathcal{C}_t\big). \tag{A.139}$$

Since dividend growth is driven entirely by cash-flow shocks ($\varepsilon_{g,t}$ and $\varepsilon_{c,t}$), which are independent of the expected-return shock $\varepsilon_{\mu,t}$, we have

$$\mathrm{Cov}\big(\Delta d_{t,t+j},\, \mathcal{M}_t\big) = 0. \tag{A.140}$$

The numerator thus reduces to two terms:

$$\mathrm{Cov}(\Delta d_{t,t+j},\, pd_t) = \mathrm{Cov}\big(\Delta d_{t,t+j},\, \overline{pd}(g_t)\big) + \mathrm{Cov}\big(\Delta d_{t,t+j},\, \mathcal{C}_t\big). \tag{A.141}$$

Step 2: Trend-growth contribution. The conditional expectation in (A.137) shows that the loading of $\mathbb{E}_t[\Delta d_{t,t+j}]$ on g_t is $\frac{1-\rho^j}{1-\rho}$. Since the only source of covariance between $\Delta d_{t,t+j}$ and $\overline{pd}(g_t)$ runs through g_t, we can write

$$\mathrm{Cov}\big(\Delta d_{t,t+j},\, \overline{pd}(g_t)\big) = \frac{1-\rho^j}{1-\rho}\,\mathrm{Cov}\big(g_t,\, \overline{pd}(g_t)\big). \tag{A.142}$$

Linearizing $\overline{pd}(g_t)$ around a reference value $\bar{g}$ as in equation (15) of the cash-flow predictability appendix,

$$\overline{pd}(g_t) \approx \text{const} + \lambda\, g_t, \qquad \lambda \equiv \overline{pd}'(\bar{g}) = \frac{1}{1-\rho},$$

gives $\mathrm{Cov}(g_t,\, \overline{pd}(g_t)) \approx \lambda\,\mathrm{Var}(g_t)$, and therefore

$$\mathrm{Cov}\big(\Delta d_{t,t+j},\, \overline{pd}(g_t)\big) \approx \frac{1-\rho^j}{1-\rho}\cdot\frac{1}{1-\rho}\,\mathrm{Var}(g_t) = (1-\rho^j)\,\mathrm{Var}\big(\overline{pd}(g_t)\big), \tag{A.143}$$

where the last equality uses $\mathrm{Var}(\overline{pd}(g_t)) \approx \lambda^2\,\mathrm{Var}(g_t) = \mathrm{Var}(g_t)/(1-\rho)^2$. Dividing by $\mathrm{Var}(\overline{pd}(g_t))$ confirms the component slope:

$$b_{g,j} \;\equiv\; \frac{\mathrm{Cov}\big(\Delta d_{t,t+j},\, \overline{pd}(g_t)\big)}{\mathrm{Var}\big(\overline{pd}(g_t)\big)} \;=\; 1-\rho^j. \tag{A.144}$$

Step 3: Cash-flow cycle contribution. From (A.137), the cycle-driven part of $\mathbb{E}_t[\Delta d_{t,t+j}]$ is $A_j^\rho\, c_t + B_j^\rho\,\theta_c\,\varepsilon_{c,t}$. Recall from equation (8) that the cash-flow cycle in prices is $\mathcal{C}_t = \beta_c\, c_t + \beta_\varepsilon\,\varepsilon_{c,t}$, where

$$\beta_c \equiv -\frac{1-\rho_c}{1-\rho\rho_c}, \qquad \beta_\varepsilon \equiv \frac{(1-\rho)\,\theta_c}{1-\rho\rho_c}.$$

Since the cycle states $(c_t, \varepsilon_{c,t})$ are orthogonal to g_t and μ_t, the covariance is

$$\begin{aligned}\operatorname{Cov}\big(\Delta d_{t,t+j},\, \mathcal{C}_t\big) &= \operatorname{Cov}\big(A_j^{\rho}\, c_t + B_j^{\rho}\, \theta_c\, \varepsilon_{c,t},\;\; \beta_c\, c_t + \beta_\varepsilon\, \varepsilon_{c,t}\big) \\ &= A_j^{\rho}\, \beta_c \operatorname{Var}(c_t) + B_j^{\rho}\, \theta_c\, \beta_\varepsilon \operatorname{Var}(\varepsilon_{c,t}) + \big(A_j^{\rho}\, \beta_\varepsilon + B_j^{\rho}\, \theta_c\, \beta_c\big) \operatorname{Cov}(c_t, \varepsilon_{c,t}). \end{aligned} \tag{A.145}$$

Using the ARMA(1,1) moments $\operatorname{Var}(c_t) = \sigma_c^2\, v_c$, $\operatorname{Var}(\varepsilon_{c,t}) = \sigma_c^2$, and $\operatorname{Cov}(c_t, \varepsilon_{c,t}) = \sigma_c^2$, where

$$v_c \equiv \frac{1 + \theta_c^2 + 2\rho_c\theta_c}{1 - \rho_c^2},$$

this simplifies to

$$\operatorname{Cov}\big(\Delta d_{t,t+j},\, \mathcal{C}_t\big) = \sigma_c^2 \Big[A_j^{\rho}\big(\beta_c\, v_c + \beta_\varepsilon\big) + B_j^{\rho}\, \theta_c \big(\beta_c + \beta_\varepsilon\big) \Big]. \tag{A.146}$$

The variance of $\mathcal{C}_t$ is (cf. Appendix C.2.1):

$$\operatorname{Var}(\mathcal{C}_t) = \sigma_c^2 \Big[\beta_c^2\, v_c + \beta_\varepsilon^2 + 2\, \beta_c\, \beta_\varepsilon \Big]. \tag{A.147}$$

Define the cycle-only predictive coefficient

$$a_{c,j} \;\equiv\; \frac{\operatorname{Cov}\big(\Delta d_{t,t+j},\, \mathcal{C}_t\big)}{\operatorname{Var}(\mathcal{C}_t)} \;=\; \frac{A_j^{\rho}\big(\beta_c\, v_c + \beta_\varepsilon\big) + B_j^{\rho}\, \theta_c \big(\beta_c + \beta_\varepsilon\big)}{\beta_c^2\, v_c + \beta_\varepsilon^2 + 2\, \beta_c\, \beta_\varepsilon}, \tag{A.148}$$

where σ_c^2 cancels from numerator and denominator.

Step 4: Variance decomposition. By the mutual orthogonality of $\overline{pd}(g_t)$, $\mathcal{M}_t$, and $\mathcal{C}_t$, the variance of pd_t decomposes as

$$\operatorname{Var}(pd_t) = \operatorname{Var}\big(\overline{pd}(g_t)\big) + \operatorname{Var}(\mathcal{M}_t) + \operatorname{Var}(\mathcal{C}_t). \tag{A.149}$$

Define the variance shares

$$b_{\overline{pd}} \equiv \frac{\operatorname{Var}\big(\overline{pd}(g_t)\big)}{\operatorname{Var}(pd_t)}, \qquad b_{\mathcal{M}} \equiv \frac{\operatorname{Var}(\mathcal{M}_t)}{\operatorname{Var}(pd_t)}, \qquad b_{\mathcal{C}} \equiv \frac{\operatorname{Var}(\mathcal{C}_t)}{\operatorname{Var}(pd_t)}, \tag{A.150}$$

so that $b_{\overline{pd}} + b_{\mathcal{M}} + b_{\mathcal{C}} = 1$.

Step 5: Univariate coefficient. Substituting (A.143) and (A.146) into (A.141) and dividing by $\operatorname{Var}(pd_t)$:

$$\begin{aligned}\frac{\operatorname{Cov}(\Delta d_{t,t+j},\, pd_t)}{\operatorname{Var}(pd_t)} &= \frac{\operatorname{Cov}\big(\Delta d_{t,t+j},\, \overline{pd}(g_t)\big)}{\operatorname{Var}(pd_t)} + \frac{\operatorname{Cov}\big(\Delta d_{t,t+j},\, \mathcal{C}_t\big)}{\operatorname{Var}(pd_t)} \\ &= \frac{\operatorname{Var}\big(\overline{pd}(g_t)\big)}{\operatorname{Var}(pd_t)} \cdot \frac{\operatorname{Cov}\big(\Delta d_{t,t+j},\, \overline{pd}(g_t)\big)}{\operatorname{Var}\big(\overline{pd}(g_t)\big)} + \frac{\operatorname{Var}(\mathcal{C}_t)}{\operatorname{Var}(pd_t)} \cdot \frac{\operatorname{Cov}\big(\Delta d_{t,t+j},\, \mathcal{C}_t\big)}{\operatorname{Var}(\mathcal{C}_t)} \\ &= b_{\overline{pd}} \cdot b_{g,j} \;+\; b_{\mathcal{C}} \cdot a_{c,j}. \end{aligned} \tag{A.151}$$

Using the component slopes from (A.144) and (A.148):

$$\frac{\mathrm{Cov}(\Delta d_{t,t+j},\ pd_t)}{\mathrm{Var}(pd_t)} \;=\; b_{\overline{pd}}\,(1-\rho^j) \;+\; b_{\mathcal{C}}\ a_{c,j}. \tag{A.152}$$

G 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 $\bar{\mu}$:

$$\begin{aligned} e_t &= \tau_t + c_t, \\ \tau_t &= \tau_{t-1} + g_{t-1}, \\ g_t &= g_{t-1} + \varepsilon_{g,t}, \quad \varepsilon_{g,t} \sim \mathcal{N}(0, \sigma_g^2) \\ c_t &= \rho_c \, c_{t-1} + \theta_c \, \varepsilon_{c,t-1} + \varepsilon_{c,t}, \quad \varepsilon_{c,t} \sim \mathcal{N}(0, \sigma_c^2) \\ \mu_t &= \bar{\mu} + \rho_\mu \, (\mu_{t-1} - \bar{\mu}) + \varepsilon_{\mu,t}, \quad \varepsilon_{\mu,t} \sim \mathcal{N}(0, \sigma_\mu^2) \end{aligned} \tag{A.153}$$

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.

G.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:

$$pd_t - \overline{pd}_t = \delta + \sum_{j=1}^{\infty} \rho^{j-1} \, \mathbb{E}_t\big[\Delta d_{t+j} - g_t\big] - \sum_{j=1}^{\infty} \rho^{j-1} \, \mathbb{E}_t\big[r_{t+j} - \bar{\mu}\big] \,, \tag{A.154}$$

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:

$$de_{t+j} = \overline{de} + v_{t+j}, \qquad v_{t+j} \overset{i.i.d.}{\sim} (0, \sigma_v^2), \quad \text{for all } j \geq 0, \tag{A.155}$$

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.154):

$$pe_t - de_t - \overline{pd}_t = \delta + \sum_{j=1}^{\infty} \rho^{j-1} \, \mathbb{E}_t\big[\Delta d_{t+j} - g_t\big] - \sum_{j=1}^{\infty} \rho^{j-1} \, \mathbb{E}_t\big[r_{t+j} - \bar{\mu}\big] \,. \tag{A.156}$$

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.155), $\mathbb{E}_t[\Delta de_{t+j}] = -v_t$ for $j = 1$ and $\mathbb{E}_t[\Delta de_{t+j}] = 0$ for $j \geq 2$. Substituting:

$$pe_t - de_t - \overline{pd}_t = \delta - v_t + \sum_{j=1}^{\infty} \rho^{j-1} \, \mathbb{E}_t\big[\Delta e_{t+j} - g_t\big] - \sum_{j=1}^{\infty} \rho^{j-1} \, \mathbb{E}_t\big[r_{t+j} - \bar{\mu}\big] \,. \tag{A.157}$$

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

$$pe_t - \overline{pe}_t = \delta + \sum_{j=1}^{\infty} \rho^{j-1}\, \mathbb{E}_t\left[\Delta e_{t+j} - g_t\right] - \sum_{j=1}^{\infty} \rho^{j-1}\, \mathbb{E}_t\left[r_{t+j} - \bar{\mu}\right], \tag{A.158}$$

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.153) yields:

$$pe_t = \overline{pe}_t + \delta - \frac{1-\rho_c}{1-\rho\rho_c}\, c_t \;+\; \frac{1-\rho}{1-\rho\rho_c}\, \theta_c\, \varepsilon_{c,t} - \frac{1}{1-\rho\rho_\mu}\, (\mu_t - \bar{\mu}). \tag{A.159}$$

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 D.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.89), with $\overline{de}$ absorbed into the intercept κ_t.

G.2 Estimation Details

We follow the same estimation procedure as the baseline case (Appendix D.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.

H 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 $\bar{\mu}_t$:

$$\begin{aligned}
d_t &= \tau_t + c_t, \\
\tau_t &= \tau_{t-1} + g_{t-1}, \\
g_t &= g_{t-1} + \varepsilon_{g,t}, \quad \varepsilon_{g,t} \sim \mathcal{N}(0, \sigma_g^2) \\
c_t &= \rho_c \, c_{t-1} + \theta_c \, \varepsilon_{c,t-1} + \varepsilon_{c,t}, \quad \varepsilon_{c,t} \sim \mathcal{N}(0, \sigma_c^2) \\
\mu_t &= \bar{\mu}_t + \rho_\mu \, (\mu_{t-1} - \bar{\mu}_{t-1}) + \varepsilon_{\mu,t}, \quad \varepsilon_{\mu,t} \sim \mathcal{N}(0, \sigma_\mu^2) \\
\bar{\mu}_t &= \bar{\mu}_{t-1} + \varepsilon_{\bar{\mu},t} \quad \varepsilon_{\bar{\mu},t} \sim \mathcal{N}(0, \sigma_{\bar{\mu}}^2)
\end{aligned} \tag{A.160}$$

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.

H.1 State-Space System

The state vector tracks μ_t directly rather than the cyclical deviation:

$$s_t = (\tau_t,\ g_t,\ c_t,\ \varepsilon_{c,t},\ \bar{\mu}_t,\ \mu_t,\ 1)^\top. \tag{A.161}$$

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

$$A = \begin{bmatrix}
1 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \rho_c & \theta_c & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 1-\rho_\mu & \rho_\mu & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix}, \qquad
C = \begin{bmatrix}
0 & 0 & 0 & 0 \\
\sigma_g & 0 & 0 & 0 \\
0 & \sigma_c & 0 & 0 \\
0 & \sigma_c & 0 & 0 \\
0 & 0 & \sigma_{\bar{\mu}} & 0 \\
0 & 0 & \sigma_{\bar{\mu}} & \sigma_\mu \\
0 & 0 & 0 & 0
\end{bmatrix}, \tag{A.162}$$

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.

H.2 Nonlinear Measurement Equation

The measurement vector remains $x_t = (d_t,\ pd_t)^\top$. The valuation ratio is:

$$pd_t = \delta + \overline{pd}_t - \frac{1-\rho_c}{1-\rho\rho_c}\, c_t \ + \ \frac{1-\rho}{1-\rho\rho_c}\, \theta_c\, \varepsilon_{c,t} - \frac{1}{1-\rho\rho_\mu}\,(\mu_t - \bar{\mu}_t), \tag{A.163}$$

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 $(\hat{g}_{t|t-1},\, \hat{\bar{\mu}}_{t|t-1})$:

$$\overline{pd}_t \;\approx\; \widehat{\overline{pd}}_t + \lambda_t^{g}\,(g_t - \hat{g}_{t|t-1}) + \lambda_t^{\bar{\mu}}\,(\bar{\mu}_t - \hat{\bar{\mu}}_{t|t-1}), \tag{A.164}$$

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

$$\widehat{\overline{pd}}_t \equiv \hat{g}_{t|t-1} - \log\big(\exp(\hat{\bar{\mu}}_{t|t-1}) - \exp(\hat{g}_{t|t-1})\big), \tag{A.165}$$

$$\lambda_t^{g} \equiv \left.\frac{\partial \overline{pd}_t}{\partial g_t}\right|_{\hat{g},\,\hat{\bar{\mu}}} = 1 + \frac{\exp(\hat{g}_{t|t-1})}{\exp(\hat{\bar{\mu}}_{t|t-1}) - \exp(\hat{g}_{t|t-1})}, \tag{A.166}$$

$$\lambda_t^{\bar{\mu}} \equiv \left.\frac{\partial \overline{pd}_t}{\partial \bar{\mu}_t}\right|_{\hat{g},\,\hat{\bar{\mu}}} = -\frac{\exp(\hat{\bar{\mu}}_{t|t-1})}{\exp(\hat{\bar{\mu}}_{t|t-1}) - \exp(\hat{g}_{t|t-1})}. \tag{A.167}$$

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.163) and collecting terms on the state vector:

$$pd_t \;\approx\; \lambda_t^{g}\, g_t - \frac{1-\rho_c}{1-\rho\rho_c}\, c_t + \frac{1-\rho}{1-\rho\rho_c}\,\theta_c\,\varepsilon_{c,t} + \left(\lambda_t^{\bar{\mu}} + \frac{1}{1-\rho\rho_\mu}\right)\bar{\mu}_t - \frac{1}{1-\rho\rho_\mu}\,\mu_t + \kappa_t, \tag{A.168}$$

where the intercept is:

$$\kappa_t = \delta + \widehat{\overline{pd}}_t - \lambda_t^{g}\,\hat{g}_{t|t-1} - \lambda_t^{\bar{\mu}}\,\hat{\bar{\mu}}_{t|t-1}. \tag{A.169}$$

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:

$$G_t = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & \lambda_t^{g} & -\dfrac{1-\rho_c}{1-\rho\rho_c} & \dfrac{1-\rho}{1-\rho\rho_c}\,\theta_c & \lambda_t^{\bar{\mu}} + \dfrac{1}{1-\rho_\mu\rho} & -\dfrac{1}{1-\rho_\mu\rho} & \kappa_t \end{bmatrix}. \tag{A.170}$$

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.

H.3 Estimation Details

We follow the same estimation procedure as the baseline case (Appendix D.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 joint reference point $(g_0,\ \mu_0)$. All other aspects of the optimisation, including the basinhopping algorithm remain unchanged.