Crimes Against Campbell-Shiller[*]

Itzhak Ben-David[†] and Alex Chinco[‡]

April 17, 2026
[Click here for the latest version]

Abstract

The Campbell and Shiller (1988) log-linear approximation is widely viewed as a model-free accounting identity that always holds: in sample, in expectation, and under arbitrary subjective beliefs. None of these claims is true. The formula is far from automatic even in realized data. Many companies do not pay dividends, making the calculation ill-defined. For dividend-paying stocks, the results are not always what they seem. The formula registers buybacks and new issuance as phantom cash-flow shocks. Taking expectations comes with its own complications. Researchers describe Campbell-Shiller as a dynamic Gordon model, but this interpretation requires investors to consistently think in present-value terms and to know the cap rate with basis-point precision. The data do not support either prerequisite. Finally, the formula generically fails under arbitrary subjective beliefs. Exceptions represent knife-edge cases where forecast errors obey a precise adding-up condition. By insisting that Campbell-Shiller always holds, researchers make it harder to learn about the true pricing model.

Keywords: Campbell-Shiller, Log-Linearization, Present-Value Relationships, Variance Decomposition, Subjective Beliefs, Forecast Errors

[*]This paper has benefited from careful feedback by John Campbell on earlier drafts.
[†]The Ohio State University and NBER. ben-david.1@osu.edu
[‡]Michigan State University. alexchinco@gmail.com

Introduction

Campbell and Shiller (1988) is one of the most influential articles in empirical asset pricing. The paper approximates a stock's current price-to-dividend (PD) ratio by log-linearizing the definition of realized returns

$$\log \mathrm{PD}_t \;=\; \text{constant} - \sum_{h=1}^{\infty} \rho^{h-1} \cdot \left\{ \mathbb{E}_t[\mathrm{R}_{t+h}] - \mathbb{E}_t[\Delta \log \mathrm{Div}_{t+h}] \right\} \tag{1}$$

$\mathbb{E}_t[\mathrm{R}_{t+h}]$ is the stock's expected return h years from now, and $\mathbb{E}_t[\Delta \log \mathrm{Div}_{t+h}]$ is its expected dividend growth. $\rho = 1/(1 + e^{-\overline{\log \mathrm{PD}}})$ stems from the paper's approximation procedure, which is centered around the average log PD ratio.

The original analysis is sharp, self-aware, and careful about assumptions. This paper is not a critique of Campbell-Shiller. It is an inventory of the many qualifications needed to apply the formula. The derivation of Campbell-Shiller starts with the definition of realized returns. It is always possible to calculate a return. So researchers have come to believe the formula somehow inherits the same sort of universality. It does not. We focus on three widely held views.

First, researchers believe it is straightforward to apply Campbell-Shiller in realized data. Bernanke and Kuttner (2005) describes the formula as "nothing more than a dynamic accounting identity" that "contains no real economic content." Greenwood and Shleifer (2014) emphasizes the same talking points, calling Campbell-Shiller "an accounting identity from the viewpoint of the econometrician." Gourinchas and Rey (2007) writes that, if you remove the two expectation operators, Equation (1) "must hold along every sample path."

This is not true. To start with, 20% of S&P 500 companies do not pay dividends, and Equation (1) is undefined for these companies. Researchers could apply Campbell-Shiller at the firm level using total corporate payouts. Instead, they typically look at per-share dividends. This seemingly innocuous change introduces two sources of confounding variation: composition (investors receive payouts via buybacks rather than dividends) and dilution (changes in the share count from repurchases and issuance). Neither has anything to do with fundamentals. Both register as phantom shocks in Campbell-Shiller.

To illustrate, consider a company that increases buybacks while maintaining the same total payout. Campbell-Shiller sees this change as a negative cash-flow shock because the firm's dividend has dropped. New issuance mechanically dilutes per-share dividends. Suppose a company issues equity to fund a positive-NPV project. If the board immediately raises the firm's total dividend to reflect the project's earnings boost, then the increase offsets the dilution, so Equation (1) sees a positive cash-flow shock. But, if the board delays for any reason, the same positive-NPV project will show up as a negative cash-flow shock. What matters is the timing of board approval, not the present value.

Second, researchers believe that Campbell-Shiller must hold in expectation. Cochrane (2011) writes that "movements in the price-dividend ratio must reflect changes in expected future dividends or changes in discount rates." Lettau and Ludvigson (2001) argues that when the dividend yield is high, "agents must be expecting either high returns on assets in the future or low dividend growth rates." Cohen, Polk, and Vuolteenaho (2003) insists that all differences in the current multiple "must necessarily be accounted for by cross-sectional variation in expected long-horizon returns and/or growth."

This is not true. From 1963 through 1982, Exxon Mobil paid a regular dividend, did not repurchase shares, and maintained a stable share count. It would be reasonable to think the in-sample version of Campbell-Shiller might hold in this data sample. But this does not imply that Exxon Mobil's \$29.75/sh price on December 31st 1982 was determined by the forward-looking version of Campbell-Shiller. These are two very different claims.

The expectation operators in Equation (1) are doing real work. They completely change the economic interpretation of the formula. When looking at realized data, Campbell-Shiller summarizes observed prices. After taking expectations, the same equation transforms into a statement about how investors determine prices. The "ex post versus ex ante" language makes this sound like a technicality. It is not. The distinction is between describing a price path and claiming to understand the mechanism that produced it.

It is also hard to see how real-world investors could know the forecasts on the right-hand side of Equation (1) with enough precision. To make things easy on investors, let's assume expected returns and dividend growth do not vary. The present value of a stock's future dividend stream simplifies to

$$\text{Price}_t \;=\; \mathbb{E}_t[\text{Div}_{t+1}] \times \left(\frac{1}{\text{R} - \text{G}}\right) \tag{2}$$

This is the Gordon model—i.e. the thing Campbell-Shiller is approximating.

The Gordon model prices stocks by multiplying next year's expected dividend times one over a tiny number, $(\frac{1}{\text{R}-\text{G}})$. We show that, to use this formula, an investor would need to know the denominator, (R–G), to within $\pm\text{DY}^2$ or less. To put this in perspective, the S&P 500 has an average dividend yield of $\overline{\text{DY}} \approx 2\%$, implying a minimum precision of $\overline{\text{DY}}^2 \approx 0.04\%$pt! Anything larger, and the uncertainty about (R−G) would have a bigger price impact than forgetting to include the present value of next year's expected dividend.

You cannot call something an "accounting identity" if you cannot keep track of whether you have "counted" the single largest present-value term in the price. Standard estimates of the equity risk premium range from 4% to 8% per year. This spread is 100× larger than the required level of precision about the S&P 500's cap rate, $\overline{\text{DY}}^2 \approx 0.04\%$pt. The "must" in the quotes above is completely unwarranted. If investors were using something similar to the Gordon model, then prices would move for lots of random reasons.

Third, researchers believe that Campbell-Shiller holds under arbitrary subjective beliefs. Conventional wisdom says that the two expectation operators in Equation (1) do not have to be rational. De la O and Myers (2024) states that, "because the relationship comes from an accounting identity, it holds even if subjective expectations are not rational." Bordalo, Gennaioli, La Porta, and Shleifer (2024) writes that the formula holds "for any deviation from RE, e.g., learning, rational inattention, bounded rationality, and behavioral biases."

This is not true. Here is a counterexample. For the sake of argument, imagine that Campbell-Shiller holds under correct expectations. Now suppose that investors unilaterally increase their S&P 500 return forecast for next year by +1%pt. The left-hand side of Equation (1) will not move. The right-hand side will be −1%pt lower. Equality has been violated.

If it is acceptable for investors to hold objectively correct beliefs, then only being +1%pt wrong about next year's return forecast is also a well-posed belief specification. You might think this is a silly mistake to make, and we can have that debate. But it will be a debate about what researchers think is reasonable, not what is permitted by the math. When researchers use Campbell-Shiller to study subjective beliefs, they are focusing on a subset of belief specifications.

We characterize this subset. Suppose that Campbell-Shiller holds under the correct expectation operator, $\mathbb{E}_t[\cdot]$. For Equation (1) to remain valid under some other subjective forecasting rule, $\widetilde{\mathbb{F}}_t[\cdot]$, the resulting errors must satisfy

$$\underbrace{\sum_{h=1}^{\infty} \frac{(\widetilde{\mathbb{F}}_t - \mathbb{E}_t)[\Delta \log \mathrm{Div}_{t+h}]}{(1+\overline{\mathrm{DY}})^{h-1}}}_{\substack{\text{Present value of errors about future} \\ \text{dividend growth discounted at } \overline{\mathrm{DY}}}} = \underbrace{\sum_{h=1}^{\infty} \frac{(\widetilde{\mathbb{F}}_t - \mathbb{E}_t)[\mathrm{R}_{t+h}]}{(1+\overline{\mathrm{DY}})^{h-1}}}_{\substack{\text{Present value of return} \\ \text{errors discounted at } \overline{\mathrm{DY}}}} \tag{3}$$

For Campbell-Shiller to hold, the present value of mistakes about future dividend growth and mistakes about future returns must cancel out when discounted at $\overline{\mathrm{DY}}$. The time-cost of money is R, but mistakes need to get discounted at $\overline{\mathrm{DY}}$.

The adding-up condition clearly rules out unilateral changes. If investors increase their next-twelve-month return forecast for the S&P 500 by +1%pt, then Campbell-Shiller requires some offsetting mistake. Investors might overestimate dividend growth. Or they could underestimate returns at some longer horizon. Either way, the magnitude of this mistake needs to be carefully calibrated. A +1%pt error about expected dividend growth next year would net out one for one. But to accomplish the same thing with a mistake about the S&P 500's dividend growth three years out, the error would have to be $(1{+}0.02)^2 \cdot 1\%\text{pt} \approx 1.04\%\text{pt}$.

Consistency with Campbell-Shiller is not a property of beliefs. It is a property of beliefs together with the way that they are modeled. Notice that the unilateral +1%pt error in next year's S&P 500 return forecast can be brought back in line with Campbell-Shiller by adding an additional assumption: investors also overestimate the S&P 500's dividend growth next year by +1%pt.

We work through numerous examples from the literature. For example, return extrapolation assumes that investors overweight recent price changes when forecasting returns, $(\widetilde{\mathbb{F}} - \mathbb{E})[\mathrm{R}] \neq 0$. On its own, this univariate bias is inconsistent with the Campbell-Shiller formula. But Barberis, Greenwood, Jin, and Shleifer (2015) show how to embed this belief specification in a more elaborate model with two investor types and a market-clearing condition. The resulting equilibrium does satisfy the adding-up condition in Equation (3).

The modeling overhead needed to bring subjective beliefs in line with Campbell-Shiller can be considerable. For instance, Jin and Sui (2022) is titled "Asset Pricing with Return Extrapolation," but the description of return extrapolation takes up less than a page. The authors spend the next 22 pages introducing additional structure to ensure internal consistency: a Lucas-tree economy, Epstein-Zin preferences, two coupled differential equations for equilibrium prices, a numerical solver, 11 calibrated parameters. It is this machinery, not any particular property of extrapolative beliefs, that determines whether the Campbell-Shiller formula holds.

This observation is deeply problematic for the belief-based asset-pricing literature. Researchers in this area have seen Campbell-Shiller as a model-free way to learn about the asset-pricing implications of subjective beliefs. Instead, they have been getting a signal that was roughly $\frac{22}{23} \approx 96\%$ unstated modeling assumptions and only $\frac{1}{23} \approx 4\%$ subjective beliefs. These are page counts from Jin and Sui (2022), but the fact generalizes. Whenever the Campbell-Shiller formula is imposed rather than tested, the modeling apparatus needed to satisfy Equation (3) contaminates any inference about beliefs.

Our point is not that the Campbell-Shiller formula never holds. Our point is that it does not have to hold. It is not an "identity." The equation "contains real economic content." When a researcher uses Equation (1), they are making a strong claim. Multiples can move for reasons unrelated to "changes in expected future dividends or changes in discount rates." Campbell-Shiller is not automatically satisfied "even if subjective expectations are not rational."

All these issues stem from the same root cause. Equation (1) generalizes the Gordon model to allow for time-varying parameters. This is a useful extension, but only in situations where prices come from a time-varying Gordon model. If investors are not applying present-value logic, then allowing R and G to change over time does not help. If investors do not know these parameters with enough precision, then this dynamic extension makes things worse. It does not make sense to approximate a special case of the dividend discount model in a world where the majority of corporate payouts are buybacks.

Why haven't researchers noticed these violations of Campbell-Shiller? One reason has to do with how the formula gets used. Researchers do not plug numbers into Equation (1) and compare the predicted PD ratio to the observed data. They assume the formula holds and use this restriction to estimate something else. Often, the calculation involves a variance decomposition.

The S&P 500's log PD ratio changes over time. Suppose you regressed the index's log PD ratio at time t on the weighted-sum of its expected future returns, $-\sum_{h=1}^{\infty} \rho^{h-1} \cdot \mathbb{E}_t[\mathrm{R}_{t+h}]$. Further suppose you ran the same regression with $\sum_{h=1}^{\infty} \rho^{h-1} \cdot \mathbb{E}_t[\Delta \log \mathrm{Div}_{t+h}]$ on the right-hand side. Researchers interpret these two slope coefficients, $\beta_{\mathrm{R}}(\infty)$ and $\beta_{\mathrm{Div}}(\infty)$, as the share of valuation changes explained by discount-rate shocks and cash-flow news, respectively. If Campbell-Shiller holds, then all variation in the S&P 500's log PD ratio should come from one of these two sources, $\beta_{\mathrm{R}}(\infty) + \beta_{\mathrm{Div}}(\infty) = 100\%$.

It is not possible to directly estimate $\beta_{\mathrm{Div}}(\infty)$. Running the above regressions would require summing an infinite series at each point in time t. To operationalize the restriction imposed by Campbell-Shiller, researchers typically regress a stock's current log PD ratio on next year's dividend growth rather than the full infinite sum. Then, they scale up the slope coefficient, $\beta_{\mathrm{Div}}(1)$, to arrive at the full dividend-growth variance share, $\beta_{\mathrm{Div}}(\infty)$.

For example, the headline result in Cochrane (2008) is that most of the variation in the S&P 500's log PD ratio is due to fluctuations in future discount rates, $\beta_{\mathrm{R}}(\infty) \approx 92\%$. He arrived at this number by first regressing the S&P 500's current log PD ratio on next year's dividend growth to get $\beta_{\mathrm{Div}}(1) \approx 0.8\%$. Then, he scaled up this one-period-ahead estimate by a factor of ~10× to arrive at

$$\beta_{\mathrm{Div}}(\infty) \;=\; \overset{0.8\%}{\beta_{\mathrm{Div}}(1)} \times \overbrace{\left(\frac{1}{1 - \underset{0.98}{\rho} \cdot \underset{0.92}{\phi}}\right)}^{\sim 10\times} \;\approx\; 8\% \tag{4}$$

The S&P 500 has $\rho = \left(\frac{1}{1+2\%}\right) \approx 0.98$ and Cochrane estimated a persistence parameter of $\phi \approx 0.92$. If there are only two sources of variation and one explains 8%, the remaining 92% of variation must come from the other source.

The logic is pulled straight from Sherlock Holmes: "When you have eliminated the impossible, whatever remains, however improbable, must be the truth." (Doyle, 1890) By treating Campbell-Shiller as an identity, Cochrane restricts his attention to only two sources of variation. Once he ruled out dividend growth, discount rates were the only remaining option. But the restriction is doing all the work. If Campbell-Shiller does not hold in sample as an identity in per-share dividends, then other sources of variation exist. Cochrane just never considered them as suspects. That is bad detective work.

Calling the Campbell-Shiller formula an "identity" is not an innocuous mistake. Terminology matters. It can create real problems. This example shows why. Cochrane's variance-share calculation (Equation 4) is doing the exact same thing as Gordon's pricing rule (Equation 2). Both formulas use a theory-implied multiple to scale up a short-term estimate.

Researchers recognize that the Gordon model is a model. When a paper uses Equation (2) to value a stock, researchers know to inspect the underlying assumptions. They ask things like: "Do R and G really remain constant like the model assumes?" That line of questioning is precisely what led to the Campbell-Shiller formula in the first place. By contrast, the calculation in Equation (4) comes pre-credentialed as a fact. It does not receive the same scrutiny. That is wrong. Given how the formula gets used, Campbell-Shiller is a model in all but name. It should be treated as such. Doing otherwise is assumption laundering.

Paper Outline. The remainder of this paper is organized as follows. Each section builds on the last. Section 1 shows that Campbell-Shiller does not automatically hold as an accounting identity in-sample. There are some situations where it does apply. Section 2 shows that, even in these cases, the formula does not automatically hold under correct expectations. Section 3 covers subjective beliefs. For the sake of argument, we assume that Campbell-Shiller holds in expectation under correct beliefs. We show that the formula will generically be violated by alternative specifications absent further assumptions.

Related Work. Campbell and Shiller (1988) derives the log-linear approximation and applies it to aggregate US stock-market data. Campbell (1991) uses the framework to decompose unexpected returns into cash-flow news and discount-rate news, finding that most aggregate return variation comes from discount-rate revisions. Vuolteenaho (2002) finds the opposite at the firm level. Larrain and Yogo (2008) shows that Campbell-Shiller holds at the firm level when total payouts are used, a point we build on. Boudoukh, Michaely, Richardson, and Roberts (2007) proposes the payout yield as a more comprehensive cash flow measure.

A large empirical literature uses Campbell-Shiller to decompose the sources of variation in price-dividend ratios. Cochrane (2011) argues that virtually all variation reflects changing discount rates. Lettau, Ludvigson, and Manoel (2018) revisits the finding with improved econometric methods. We do not take a position in this debate. We argue that the framework within which it takes place rests on assumptions that are routinely violated.

Nagel and Xu (2022) surveys the broader literature on expectations in asset pricing. See Bordalo, Gennaioli, and Shleifer (2018), Greenwood and Shleifer (2014), Barberis et al. (2015), Fuster, Laibson, and Mendel (2010), and Adam, Marcet, and Nicolini (2016); Adam, Marcet, and Beutel (2017). Our paper does not argue against any of these models. We show that using them with Campbell-Shiller requires more infrastructure than is typically acknowledged. Much of the belief-based asset-pricing literature assumes that Campbell-Shiller holds under arbitrary subjective expectations. We derive the conditions under which Campbell-Shiller holds for biased beliefs. Meeting these conditions requires substantial modeling overhead.

Other papers have investigated whether Campbell-Shiller's approximation is accurate. Engsted, Pedersen, and Tanggaard (2012) puts an upper bound on the average error. Cho, Kremens, Lee, and Polk (2024) proposes an alternative present-value formula with smaller approximation errors. Chen and Zhao (2009) argues that VAR-based return decompositions are sensitive to the choice of state variables. Our question is logically prior: even a perfect approximation is useless if the object being approximated does not apply.

1 Realized Data

It is widely believed that Campbell-Shiller holds as an accounting identity in realized data. Remove the expectation operators from Equation (1), and there is no way the formula can fail. It has to be correct in-sample. That is the conventional wisdom. This section shows why it is wrong.

Even if you remove the expectation operators, the left-hand side of Equation (1) must be well-defined. Subsection 1.1 notes that roughly 20% of S&P 500 constituents pay no cash dividend, making the formula undefined for these firms. Researchers typically apply Campbell-Shiller to per-share data rather than firm-level quantities, and Subsection 1.2 characterizes the errors introduced by this substitution. Subsections 1.3 and 1.4 provide concrete examples showing how these errors show up in response to repurchases and new issuance.

The bottom line is simple: Campbell-Shiller is not an accounting identity in realized data. There is nothing easy or automatic about applying Equation (1). It requires numerous assumptions even before taking expectations.

1.1 Zero Dividends

Campbell-Shiller cannot be applied to stocks with $\text{Div}_t = \$0/\text{sh}$. This number is the denominator in the price-to-dividend (PD) ratio, $\text{PD}_t = \left(\frac{\text{Price}_t}{\text{Div}_t}\right)$. Log dividend growth is also undefined when there is no current or future dividend, $\Delta \log \text{Div}_{t+1} = \log \text{Div}_{t+1} - \log \text{Div}_t$. These problems are not edge cases. Roughly 20% of S&P 500 constituents pay no cash dividend at all. Alphabet, Amazon, Meta, and Berkshire Hathaway all paid no dividend for years or decades. Microsoft went public in 1986 and paid its first dividend in January 2003. In the intervening years, its equity valuation exceeded \$500B while the company returned little cash to shareholders in any form.

On some level, researchers know all this. They use the original dividend-based Campbell-Shiller formula when studying the S&P 500 at the index level. But it is common to replace $\mathbb{E}[\text{Div}]$ with $\mathbb{E}[\text{EPS}]$ when studying individual stocks. Vuolteenaho (2002) showed how to make the jump from dividends to earnings per share. The paper has been cited more than 1,500 times. The paper must be adding something to the original formula that warrants all these citations.

We are not claiming that Campbell-Shiller cannot be modified to handle non-dividend payers. It clearly can. Our point is that the necessity of such modifications tells us something important: there is nothing automatic about applying Campbell-Shiller in-sample. It is not an accounting identity that always holds. Barring further assumptions, it is literally undefined for 20% of stocks in the S&P 500 even before taking expectations.

1.2 Per-Share Data

"Asset-pricing theory all stems from one simple concept: price equals expected discounted payoff." (Cochrane, 2001, page 1) It should not matter how the payoff is subdivided. Hence, the right realized return to look at is the firm-level calculation

$$1 + \mathrm{R}_{t+1} \;=\; \frac{\mathrm{MCap}_{t+1} + \mathrm{TotPmt}_{t+1}}{\mathrm{MCap}_t} \tag{5}$$

MCap_t is the ex-payout firm value and TotPmt_{t+1} is the total cash paid to shareholders between t and $(t+1)$, which includes both dividends and repurchases.

When applied to Equation (5), Campbell-Shiller would relate the total-payout multiple $\left(\frac{\mathrm{MCap}_t}{\mathrm{TotPmt}_t}\right)$ to the present-value-weighted sums of total-payout growth $\Delta \log \mathrm{TotPmt}_{t+h}$ and returns R_{t+h}. The correct growth-rate variable is $\Delta \log \mathrm{TotPmt}$. The left-hand side should have $\log\left(\frac{\mathrm{MCap}}{\mathrm{TotPmt}}\right)$ as the valuation ratio.

But researchers do not typically use these inputs. Instead, the standard empirical convention replaces this firm-level calculation with the per-share version

$$1 + \mathrm{R}_{t+1} \;=\; \frac{\mathrm{Price}_{t+1} + \mathrm{Div}_{t+1}}{\mathrm{Price}_t} \tag{6}$$

Price_t is the price per share, and Div_{t+1} is the per-share dividend.

What are the consequences of this substitution? We give the conventional approach the benefit of the doubt. The remainder of this section assumes that R_{t+1} is still correct. Modigliani-Miller guarantees that per-share returns are unaffected by payout policy when total payout and investment are the same.

Even so, researchers who plug per-share dividend growth into Campbell-Shiller are using a distorted version of total-payout growth.

Assume that a stock paid a dividend this year and last year. Let TotPmt_{t+1} denote the firm's total payout next year, and let TotDiv_{t+1} denote the total dividend payout next year. By contrast, Div_{t+1} denotes the firm's dividend on a per-share basis. Before repurchasing shares, the company has $\#\mathrm{Shares}_{t+1}$ outstanding at the end of next year.

Proposition 1.2. *Suppose that Equations (5) and (6) yield the same* R_{t+1}*. The gap between per-share dividend growth and total-payout growth decomposes as*

$$\Delta \log \mathrm{Div}_{t+1} - \Delta \log \mathrm{TotPmt}_{t+1} = \overbrace{\big(\Delta \log \mathrm{TotDiv}_{t+1} - \Delta \log \mathrm{TotPmt}_{t+1}\big)}^{\text{Payout composition}} - \underbrace{\Delta \log \#\mathrm{Shares}_{t+1}}_{\text{Share dilution}} \tag{7}$$

The first term is non-zero whenever the dividend share of total payout changes. The second term is non-zero whenever the share count changes.

The first term is payout composition. It captures the gap between dividend growth and total-payout growth. When a firm shifts cash from dividends to repurchases, $\Delta \log \mathrm{TotDiv}$ falls even though $\Delta \log \mathrm{TotPmt}$ is unchanged. The second term is about changing share counts. Repurchases and new issuance change the denominator of $\mathrm{Div}_{t+1} = \frac{\mathrm{TotDiv}_{t+1}}{\#\mathrm{Shares}_{t+1}}$, leading to mechanical dilution and accretion effects even if the numerator remains constant.

We are giving the conventional wisdom about Campbell-Shiller being an in-sample identity the benefit of the doubt. We assume the switch from total firm-level payouts to per-share dividends does not impact returns, R_{t+1}. Nevertheless, problems remain. The confounds enter through $\Delta \log \mathrm{Div}_{t+h}$ and through PD_t, which absorbs whatever the growth term does not. The "cash-flow news" and "discount-rate news" that researchers typically extract from the Campbell-Shiller formula are not pure-tone signals. Even under ideal conditions, both signals get blended with payout-policy changes in non-trivial ways.

1.3 Repurchases

Researchers typically reason that all variation in the price-dividend ratio must reflect either future dividend growth or future returns. Proposition 1.2 shows that the switch from firm-level total payouts to per-share dividends introduces two additional suspects: payout composition and share dilution. To see how payout composition works in practice, we look at three otherwise identical firms that deliver the same total payout in different ways.

Each firm has $\#\text{Shares}_{t-1} = 100\text{M}$ shares outstanding and a share price of $\text{Price}_{t-1} = \$3.125/\text{sh}$. Each period, every firm earns \$25M and distributes all of it to shareholders. All three firms have the same 8% discount rate, and since there is no investment, the ex-dividend firm value is a flat perpetuity, $\$25\text{M} \times \left(\frac{1}{8\%}\right) = \$312.5\text{M}$. The three firms differ only in how they split the \$25M payout between dividends and repurchases:

1. Baseline Inc. This company always pays \$25M in dividends and repurchases \$0 (100/0 split). With 100M shares at time t, the firm has $\text{Div}_t = \$0.25/\text{sh}$ and $\text{Price}_t = \$3.125/\text{sh}$. At time $(t{+}1)$, the numbers are the same: $\text{Div}_{t+1} = \$0.25/\text{sh}$, $\text{Price}_{t+1} = \$3.125/\text{sh}$, $\text{R}_{t+1} = 8\%$, and $\Delta \log \text{Div}_{t+1} = 0\%$.

2. Always Co. This company maintains a steady 80/20 split, paying out \$20M in dividends and repurchasing \$5M every year. At time t, the firm has $\text{Div}_t = \frac{\$20\text{M}}{100\text{M}} = \$0.20/\text{sh}$. The \$5M repurchase happens at the ex-dividend price. After dividends, the firm holds \$5M in cash earmarked for repurchase, so the ex-dividend equity value is \$312.5M + \$5M = \$317.5M across 100M shares, giving $\text{Price}_t = \$3.175/\text{sh}$. The firm buys back $\frac{\$5\text{M}}{\$3.175/\text{sh}} \approx 1.575\text{M}$ shares, leaving $\#\text{Shares}_t \approx 98.43\text{M}$.

   At time $(t{+}1)$, Always Co maintains the same 80/20 split on a smaller base, leading to $\text{Div}_{t+1} \approx \$0.2032/\text{sh}$, $\text{Price}_{t+1} \approx \$3.226/\text{sh}$, $\text{R}_{t+1} = 8\%$, and $\Delta \log \text{Div}_{t+1} = \log\left(\frac{\$0.2032/\text{sh}}{\$0.20/\text{sh}}\right) \approx +1.6\%$. The dividend per share grows by 1.6% not because the firm's fundamentals improved, but because the shrinking share count mechanically boosts the firm's dividend payout on a per-share basis. The PD ratio is 15.9× versus Baseline Inc's 12.5×, a 27% relative difference caused entirely by the lower per-share dividend.

	Baseline Inc	Always Co	Switch Ltd
Total earnings	\$25.0M	\$25.0M	\$25.0M
Total payout	\$25.0M	\$25.0M	\$25.0M
Total dividend$_t$	\$25.0M	\$20.0M	\$25.0M
Buyback amount$_t$	\$0.0M	\$5.0M	\$0.0M
Total dividend$_{t+1}$	\$25.0M	\$20.0M	\$20.0M
Buyback amount$_{t+1}$	\$0.0M	\$5.0M	\$5.0M
Payout split$_t$	100/0	80/20	100/0
Payout split$_{t+1}$	100/0	80/20	80/20
#Shares$_t$	100.0M	98.4M	100.0M
#Shares$_{t+1}$	100.0M	96.9M	98.4M
PD$_t$	12.5×	15.9×	12.5×
R$_{t+1}$	8.0%	8.0%	8.0%
$\Delta \log \text{Div}_{t+1}$	0.0%	+1.6%	−22.3%

Table 1. Three firms with the same earnings, total payout, and return. Baseline Inc and Switch Ltd are indistinguishable at time t. Both firms have the same PD$_t$, Div$_t$, Price$_t$, and R$_t$. Yet one delivers $\Delta \log \text{Div}_{t+1} = 0\%$ while the other has −22.3%. Always Co maintains a stable 80/20 payout split. It shows a mechanical +1.6% growth in dividends per-share due to its shrinking share count.

3. Switch Ltd. This company switches payout policies. At time t, Switch Ltd only pays \$25M in dividends (100/0 split). At time $(t+1)$, the firm switches to an 80/20 split (\$20M dividends, \$5M repurchases). The firm's numbers at time t match Baseline Inc: $\text{Div}_t = \$0.25/\text{sh}$, and $\text{Price}_t = \$3.125/\text{sh}$. Its new numbers at time $(t+1)$ are: $\text{Div}_{t+1} = \$0.20/\text{sh}$, $\text{Price}_{t+1} = \$3.175/\text{sh}$, $\text{R}_{t+1} = 8\%$, and $\Delta \log \text{Div}_{t+1} = \log\left(\frac{\$0.20/\text{sh}}{\$0.25/\text{sh}}\right) = -22.3\%$.

Table 1 summarizes the results. Notice that Baseline Inc and Switch Ltd are indistinguishable at time t. They have the same $\text{PD}_t = 12.5\times$, the same $\text{Div}_t = \$0.25/\text{sh}$, the same $\text{Price}_t = \$3.125/\text{sh}$, and the same return $\text{R}_t = 8\%$. A forecasting model estimated on per-share data must produce the same forecast for both firms. Yet one delivers $\Delta \log \text{Div}_{t+1} = 0\%$ and the other has −22.3%.

This −22.3% shock in year $(t+1)$ is not a cash-flow shock. Switch Ltd has the same earnings, the same total payout, and the same return as Baseline

Inc. Nothing fundamental has changed. The firm simply redirected \$5M from dividends to buybacks. But Campbell-Shiller has no notion of share count or payout policy. The formula does not distinguish between a -22.3% decline in $\Delta \log \text{Div}$ caused by deteriorating earnings and a -22.3% decline caused by a shift toward buybacks. This is a distinction that makes a difference to investors.

Even the steady-state case is affected. Always Co maintains the same 80/20 payout policy throughout. This firm sees $\Delta \log \text{Div} \approx +1.6\%$ purely from the shrinking share count, and a PD ratio of $15.9\times$ versus Baseline Inc's $12.5\times$. Campbell-Shiller is internally consistent for Always Co. But the numbers encode payout-policy decisions, not firm fundamentals. To a researcher, Always Co looks like a firm with $\text{G} = +1.6\%$. To know otherwise requires data on buybacks, which the Campbell-Shiller framework does not use.

Researchers currently believe that all variation in PD_t must reflect either future returns or future dividend growth. In this example, none of it does. All three firms earn the same amount, pay out the same amount, and deliver the same 8% return. Total-payout growth is zero every period. The 27% gap in PD ratios between Always Co and Baseline Inc, and the -22.3% per-share dividend shock at Switch Ltd, come entirely from the payout-composition term in Proposition 1.2. A researcher using the Campbell-Shiller formula with per-share dividends would attribute these differences to fundamentals. They are not. Getting from the textbook formula to an empirical implementation requires choices about how to measure payouts, and those choices have first-order consequences. Calling the result an "identity" obscures this fact.

1.4 New Issuance

The previous subsection isolated payout composition. This subsection isolates share dilution, the second channel from Proposition 1.2. All payouts are dividends, so the composition term drops out. We now consider two new firms that issue equity to fund the same positive-NPV project. The new earnings are real. So you would think the Campbell-Shiller formula would register a positive cash-flow shock next year. That is not always what happens.

The two firms in this example look identical to Baseline Inc to begin with. Each starts out earning \$25M per period, distributes all of it as dividends, has an 8% discount rate, and trades at Price_t = \$3.125/sh on $\#\text{Shares}_t$ = 100M shares. However, between t and $(t{+}1)$, both firms invest in the same positive-NPV project that earns 12% and generates \$9.375M per year in additional earnings. Both firms cover the up-front cost by issuing 25M new shares at \$3.125/sh. The two firms differ only in how quickly they raise their dividend.

1. Now Corp. This company immediately raises its total dividend to \$25M + \$9.375M = \$34.375M. The ex-dividend firm value is $\$34.375\text{M} \times \left(\frac{1}{8\%}\right) =$ \$429.6875M. With $\#\text{Shares}_t$ = 125M, the company has Div_{t+1} = \$0.275/sh and $\text{Price}_{t+1} = \frac{\$429.6875\text{M}}{125\text{M}} = \$3.4375/\text{sh}$. The firm's other Campbell-Shiller inputs are R_{t+1} = 18.8% and $\Delta \log \text{Div}_{t+1} = \log\left(\frac{\$0.275/\text{sh}}{\$0.25/\text{sh}}\right) = +9.5\%$.
2. Delay PLC. This firm issues the same number of new shares and invests in the same project as Now Corp. However, Delay PLC keeps its total dividend at \$25M at time $(t{+}1)$ and retains \$9.375M in earnings. The ex-dividend value consists of operating assets (\$429.6875M) plus retained cash (\$9.375M). With $\#\text{Shares}_t$ = 125M, Delay PLC has $\text{Div}_{t+1} = \frac{\$25\text{M}}{125\text{M}} = \$0.20/\text{sh}$ and $\text{Price}_{t+1} = \frac{\$429.6875\text{M}+\$9.375\text{M}}{125\text{M}} = \$3.5125/\text{sh}$. The cum-dividend value per share is \$3.7125/sh. The other Campbell-Shiller inputs are R_{t+1} = 18.8% and $\Delta \log \text{Div}_{t+1} = \log\left(\frac{\$0.20/\text{sh}}{\$0.25/\text{sh}}\right) = -22.3\%$.

Table 2 summarizes the results. Now Corp and Delay PLC did the same thing: they issued the same number of shares, invested in the same project, earned the same total earnings, and delivered the same return to existing shareholders. The cum-dividend value per share is \$3.7125/sh in both cases. The only difference is a board-level decision about when to raise the dividend. Yet the Campbell-Shiller formula registers Now Corp as a +9.5% positive cash-flow shock and Delay PLC as a −22.3% negative cash-flow shock. That is a 32%pt swing from dividend timing alone. This is not what researchers have in mind when they say "cash-flow shock."

Firms typically adjust their dividend slowly (Lintner, 1956). Most firms behave like Delay PLC. This does not matter for valuation purposes. The market

	Baseline Inc	Now Corp	Delay PLC
New shares issued	0.0M	25.0M	25.0M
Project earnings		\$9.4M	\$9.4M
Total earnings	\$25.0M	\$34.4M	\$34.4M
Total payout	\$25.0M	\$34.4M	\$25.0M
Retained cash	\$0.0M	\$0.0M	\$9.4M
Div_{t+1}	\$0.25/sh	\$0.27/sh	\$0.20/sh
Price_{t+1}	\$3.13/sh	\$3.44/sh	\$3.51/sh
$\text{Div}_{t+1} + \text{Price}_{t+1}$	\$3.38/sh	\$3.71/sh	\$3.71/sh
PD_t	12.5×	12.5×	12.5×
R_{t+1}	8.0%	18.8%	18.8%
$\Delta \log \text{Div}_{t+1}$	0.0%	+9.5%	−22.3%

Table 2. Now Corp and Delay PLC issue the same shares, invest in the same project, earn the same total earnings, and deliver the same return to shareholders. The only difference is a board-level decision about when to raise the dividend. Yet the Campbell-Shiller formula registers one as a +9.5% positive cash-flow shock and the other as a −22.3% negative cash-flow shock. That is a 32%pt swing from dividend timing alone.

prices the same project the same way regardless of when the board raises the dividend. Now Corp and Delay PLC both have cum-dividend prices of \$3.7125/sh. But it does matter for Campbell-Shiller. One company invests in a positive-NPV project and has a positive cash-flow shock. The other company makes the same investment and records a negative shock. Instead of capturing present-value logic, the formula is picking up an economically irrelevant timing choice.

Researchers currently believe that all variation in PD_t must reflect either changes in future returns or changes in future dividend growth. This is simply not true when looking at per-share data. Now Corp and Delay PLC both invest in the same project, have the same return, and receive the same cum-dividend valuation. The 32%pt swing in per-share dividend growth comes from a board decision about dividend timing. Getting from the textbook formula to an empirical implementation requires choices about how to measure payouts, when to recognize them, and how to handle share-count changes. Each choice has first-order consequences. Calling the result an "identity" obscures this fact.

2 Correct Expectations

The previous section gave a number of reasons why Campbell-Shiller can fail to hold in-sample. The formula is clearly not an identity that always has to hold. There are, however, empirical settings in which the conditions needed to apply the formula in-sample are satisfied. Suppose we restrict attention to one of these cases. Researchers currently believe that, whenever Campbell-Shiller holds in sample, it must also hold in expectation. This is not true.

Taking expectations changes the economics. The in-sample version of Equation (1) describes observed prices. The forward-looking version says why those prices were observed. Moreover, it is not just any claim. Campbell-Shiller extends the Gordon model to allow for time-varying parameters. This interpretation requires the other assumptions behind the Gordon model to remain valid. Subsection 2.1 catalogs these assumptions. The first two require investors to consistently apply present-value logic. Subsection 2.2 shows that many do not. Subsection 2.3 shows that, even if investors did think this way, calling Campbell-Shiller an "accounting identity" requires an implausible level of precision.

2.1 Gordon Model

Campbell-Shiller extends the Gordon model to allow for time-varying coefficients. Equation (1) is a log-linear approximation to Equation (2). Researchers understand this. Campbell and Shiller (1988) describes their own model as "a dynamic version of the Gordon (1962) model." Gao and Martin (2021) writes that "Campbell and Shiller (1988) generalizes the Gordon growth model to the empirically relevant case in which these quantities are time-varying."

What's more, the Gordon model (Equation 2) is a special case of the Dividend Discount Model (DDM). It offers a simpler implementation in a world where the expected dividend-growth rate is constant. Campbell-Shiller relaxes this constant-parameter assumption. But to interpret the formula as a statement about how prices are set, the remaining steps in the derivation of the DDM must still work. We now review what those steps are.

Step 1. Assume current price obeys one-period-ahead present-value logic.

Campbell-Shiller may start with the definition of a realized return, but the model it approximates starts with the one-period-ahead present-value formula

$$\text{Price}_t = \frac{\mathbb{E}_t[\text{Div}_{t+1}]}{1 + \text{R}} + \frac{\mathbb{E}_t[\text{Price}_{t+1}]}{1 + \text{R}} \tag{8}$$

The DDM assumes that a stock's current price will reflect the discounted value of its total payout next year. Part of that payout will come in the form of a cash dividend. Part will come from the resale price.

Step 2. Assume investors think about future prices in present-value terms.

Suppose you buy a 1-year \$100 bond with a 2% coupon payment. Next year you will receive a 2% × \$100 = \$2 coupon and a \$100 face-value payment. You do not need to have a theory of how large the face-value payment will be. The \$100 amount is stated on the bond at the time of purchase. Things are different for stocks. The anticipated resale price is analogous to the bond's \$100 face value, but you do not know Price_{t+1} ahead of time.

Hence, a theory of current stock prices requires a theory of future stock prices. The DDM assumes that they are the same theory. The model says that investors reason about future prices at every horizon $h = 1, 2, 3, \ldots$ in the same way as the current price level

$$\mathbb{E}_t[\text{Price}_{t+h}] = \frac{\mathbb{E}_t[\text{Div}_{(t+h)+1}]}{1 + \text{R}} + \frac{\mathbb{E}_t[\text{Price}_{(t+h)+1}]}{1 + \text{R}} \tag{9}$$

This second present-value assumption is much stronger than the first. Price_t is an observed outcome that is subject to market forces. One could argue that the invisible hand somehow enforces Equation (8) even if no individual investor uses it. The same argument does not work for Equation (9). $\mathbb{E}_t[\text{Price}_{t+h}]$ is a number that exists in an investor's head, not on a Bloomberg Terminal. Prognostications and guesses do not have to satisfy budget constraints or clear markets.

Step 3. Iterate forward, pushing unknown resale price far into the future.

The DDM assumes that investors think about the current price level and future resale prices using the same present-value logic. This self-similarity makes it possible to iterate forward, recursively replacing the future resale price with its one-step-ahead present-value equivalent

$$\text{Price}_t = \sum_{h=1}^{H} \frac{\mathbb{E}_t[\text{Div}_{t+h}]}{(1+\text{R})^h} + \frac{\mathbb{E}_t[\text{Price}_{t+H}]}{(1+\text{R})^H} \tag{10}$$

The first term is the present value of the stock's dividends over the next H years. The second is the present value of its resale price in year H.

Step 4. Use transversality condition to eliminate resale price completely.

If H is large enough, discounting takes over and the present value of the future resale price can be ignored. Researchers call this a "transversality condition"

$$\lim_{H \to \infty} \frac{\mathbb{E}_t[\text{Price}_{t+H}]}{(1+\text{R})^H} = \$0/\text{sh} \tag{11}$$

Romans did not care about prices today when trading stocks in 1AD.

If we can ignore the second term in Equation (10) when taking the limit $H \to \infty$, then we arrive at the Dividend Discount Model (DDM)

$$\text{Price}_t = \sum_{h=1}^{\infty} \frac{\mathbb{E}_t[\text{Div}_{t+h}]}{(1+\text{R})^h} \tag{12}$$

This classic model says that a stock's current price should reflect the discounted value of its expected future dividend stream.

Recall that pricing a stock required a theory of the current price level and a theory of the future resale price. The DDM assumes that both theories are one and the same. As a result, it is possible to iterate forward and remove the future resale price from Equation (12). The DDM effectively prices stocks by pretending they are fixed-income products. This is why the discount rate in the DDM behaves like a bond yield, moving in the opposite direction as the current price.

Step 5. Assume expected annual dividend-growth rate is constant.

The Gordon model is a special case of the DDM. To operationalize Equation (12), you must make some sort of assumption about a stock's future dividend growth. It is not physically possible to sum an infinite series of arbitrary forecasts. Gordon (1962) points out a particularly convenient special case. Suppose that, in addition to R being constant, expected dividend growth is also constant, $\mathbb{E}_t[\mathrm{Div}_{t+h}] = (1{+}\mathrm{G})^h \cdot \mathrm{Div}_t$. It is then possible to price a stock by multiplying next year's dividend forecast times

$$\mathrm{PD} \;=\; \left(\frac{1}{\mathrm{R}-\mathrm{G}}\right) \;=\; \left(\sum_{h=1}^{\infty} \frac{\mathbb{E}_t[\mathrm{Div}_{t+h}]}{(1+\mathrm{R})^h}\right) \Big/ \; \mathbb{E}_t[\mathrm{Div}_{t+1}] \tag{13}$$

This is the pricing formula for a growing perpetuity. Gordon says to price stocks by pretending they are really simple bonds.

There is a lot more to the Gordon model than assuming constant parameters. Step 5 is just the icing on the cake. Before you get there, investors must set today's price using one-step-ahead present-value logic. They must believe every future price also obeys the same present-value relationship. You must iterate this logic forward indefinitely. The present value of the resale price must vanish in the limit. Only after all of this does the constant-parameter assumption enter, converting the general DDM into a tractable special case.

Campbell-Shiller is meant to relax the constant-parameter assumption. The log of Equation (13) is directly comparable to the Campbell-Shiller formula

$$\log(\mathrm{R}-\mathrm{G}) \;\sim\; \sum_{h=1}^{\infty} \rho^{h-1} \cdot \left\{ \mathbb{E}_t[\mathrm{R}_{t+h}] - \mathbb{E}_t[\Delta \log \mathrm{Div}_{t+h}] \right\} \tag{14}$$

The formula allows R and G to vary over time, replacing the constant (R–G) with a weighted sum of future expected values. This is a useful extension in situations where prices are governed by an approximate Gordon model. However, serious problems arise when researchers try to interpret Equation (1) as a generalized Gordon multiple outside of the original model's effective range.

2.2 Present-Value Logic

The forward-looking version of Campbell-Shiller is meant to be a generalized Gordon model. For that interpretation to hold, Steps 1 through 4 of the previous subsection still have to hold. Researchers sometimes flag Step 4 as the substantive assumption. It is not. Transversality is a mopping-up condition. The heavy lifting is done by the forward iteration in Step 3, and this step relies on the prior two. Step 1 says investors price today's share using one-period-ahead present-value logic. Step 2 says they forecast every future resale price the same way. For Equation (1) to always hold under correct expectations, investors must reliably think in present-value terms. They do not.

Sell-side equity analysts are professional forecasters whose price targets move billions of dollars of trading volume. Ben-David and Chinco (2025) documents that these market participants typically value companies using a backward-looking multiple rather than a forward-looking one like the Gordon model says to do. Over 3/4 of analyst reports explicitly state that they used some version of the pricing rule below

$$\text{PriceTarget} \;=\; \mathbb{E}[\text{EPS}] \times \text{TrailingPE} \tag{15}$$

where TrailingPE = Price/EPS is the trailing twelve-month PE ratio. Analysts take their short-term EPS forecast, $\mathbb{E}[\text{EPS}]$, and scale it by the stock's trailing PE. They do not pick a PE ratio that reflects the present value of the company's expected earnings stream in years $t = 2, 3, 4, \ldots$ The result is a price target anchored to the stock's recent valuation history, not to a discounted cash flow model. There is no discount rate. There is no terminal value. There is no infinite sum over future cash flows. The entire pricing exercise is a two-variable calculation: expected earnings times a backward-looking multiple.

It is not just sell-side analysts, either. In February 2013, David Einhorn of Greenlight Capital Management gave a public presentation to fellow Apple shareholders in an effort to convince the company to distribute its cash hoard.[1] Figure 1 shows a key slide from this presentation. Einhorn's argument was that

[1] David Einhorn. "iPrefs: Unlocking Value." Conference Call. February 21, 2013. [link]

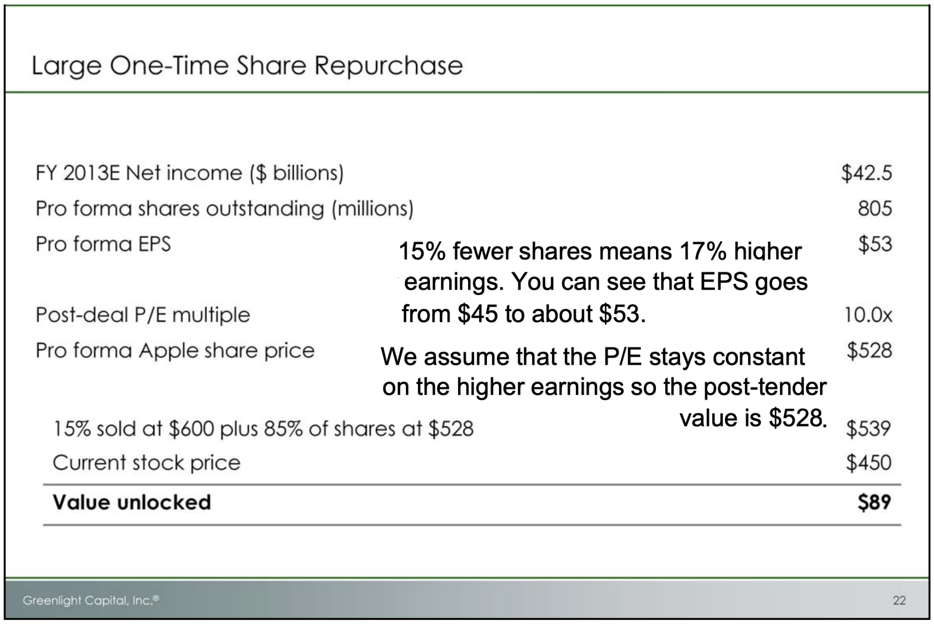

Figure 1. Slide from a February 2013 presentation by David Einhorn of Greenlight Capital Management to Apple (AAPL) shareholders. Einhorn calculates that an \$84B cash-financed repurchase program would generate a +7.80/sh EPS pop and a share price increase of $\$528/\text{sh} - \$450/\text{sh} = \$7.80/\text{sh} \times 10 \approx +\$78/\text{sh}$.

an \$84B cash-financed repurchase program would mechanically boost Apple's per-share earnings from \$45/sh to \$52.80/sh, and that this \$7.80/sh EPS gain should be valued at Apple's existing 10× PE ratio. His pitch was that repurchases would increase share prices by $\$528/\text{sh} - \$450/\text{sh} = \$7.80/\text{sh} \times 10 \approx +\$78/\text{sh}$.

Einhorn is a sophisticated investor, managing billions of dollars. He was making a public case to the shareholders of one of the world's most valuable companies. The valuation framework he reached for was $\mathbb{E}[\text{EPS}] \times \text{TrailingPE}$, not a discounted cash flow model. Einhorn did not discount the future earnings stream that the repurchase would generate. He did not estimate how the buyback would change Apple's cost of capital or long-run growth rate. He took his EPS forecast and multiplied it by a backward-looking multiple, exactly the same pricing rule that sell-side analysts rely on.

Hartzmark and Sussman (2025) provides perhaps the most direct test of whether investors apply present-value logic. If people reliably set prices by discounting expected future payoffs, then you should be able to hand someone information about a company's fundamentals and have them give you back the correct price. This is exactly what the authors do. They show online participants, MBA students, and asset-management professionals information about fundamentals. Then the authors ask them to identify the market price. Across all three groups, subjects cannot detect even extreme deviations from actual prices. When asked directly, both professionals and large language models say that it is not possible to reliably estimate the level of the market to within 10% using fundamental information alone, without first seeing the prevailing price.

One reaction to this evidence is that a lot of people are making dumb mistakes. That could be. But it also might not be unreasonable to deviate from present-value logic. Consider the housing market. If you own a home, which discount rate did you use when deciding how much to pay for it? We are guessing you did not calculate a discounted cash flow model. Almost no one does. Homebuyers look at comparable sales, recent transaction prices for similar homes in similar neighborhoods. The pricing rule is backward-looking: what did this kind of house sell for recently? Not what is the discounted present value of the rents this house will generate over the next thirty years?

This backward-looking approach is so deeply embedded that the dominant methodology for measuring home prices, the Case-Shiller repeat-sales index, is explicitly built on it (Case and Shiller, 1987, 1989). This does not make you a bad economist. It just means that not every market operates according to present-value principles. Bond markets do. Housing markets clearly do not. The evidence above suggests that, in many cases, neither do equity markets.

Gao and Martin (2021) calls time-varying R and G "the empirically relevant case." That is a claim about Step 5. The question is whether Steps 1-4 are empirically relevant, too. The examples above demonstrate that Steps 1 and 2 are not always justified. Without them it is not possible to perform the forward iteration in Step 3, rendering any notion of transversality in Step 4 a moot point. In these scenarios, the Gordon model does not describe how stocks are priced, and the reason has nothing to do with time-varying parameters.

2.3 Required Precision

Suppose, for the sake of argument, that Steps 1 through 4 hold. Investors reliably apply one-period-ahead present-value logic at t and at every future date. The forward iteration runs. The transversality condition clears the tail. On the surface, a time-varying extension of the Gordon model now looks reasonable. But could an investor actually use such a model to set prices? The answer is no. Even before allowing R and G to vary over time, the Gordon formula requires investors to know both parameters with an implausible level of precision.

We use the "accounting identity" language to quantify how much precision is needed. For something to qualify as an accounting identity, you have to be able to reliably count each term. In the Dividend Discount Model (Equation 12), the single largest term is next year's expected dividend, $\frac{\mathbb{E}_t[\mathrm{Div}_{t+1}]}{1+\mathrm{R}}$. Every subsequent term is smaller by a factor of $\left(\frac{1+\mathrm{G}}{1+\mathrm{R}}\right) < 1$. This puts an upper bound on the amount of noise the Gordon model can tolerate. If errors have a larger price impact than the present value of next year's dividend forecast, we cannot reliably tell whether the larger term has been "counted" in the price.

Proposition 2.3. *Suppose prices obey the Gordon model in Equation (2) with* $\mathrm{G} \geq 0\%$. *If noise distorts the cap rate by* $|\Delta[\mathrm{R}{-}\mathrm{G}]| = \delta > 0$, *then the resulting price change will exceed the present value of next year's dividend forecast whenever*

$$|\delta| > \mathrm{DY}^2 \qquad \rightsquigarrow \qquad |\Delta\mathrm{Price}_t| > \frac{\mathbb{E}_t[\mathrm{Div}_{t+1}]}{1+\mathrm{R}} \tag{16}$$

The S&P 500 has an average dividend yield of $\overline{\mathrm{DY}} \approx 2\%$, so the threshold is $\overline{\mathrm{DY}}^2 \approx 0.04\%$. This is preposterously small. Standard estimates of the equity risk premium range from 4% to 8%, a spread that is 100× larger. The variance decomposition literature treats Equation (1) as an accounting identity. Researchers attribute all fluctuations in log PD to revisions in expected future returns or dividend growth. This is wrong. If investors were actually using the kind of model that Campbell-Shiller is meant to approximate, then log PD would have to move for lots of other random reasons. Allowing R and G to change over time does not produce a better model. It makes the problem worse.

3 Subjective Beliefs

We started by asking: what would it take for Campbell-Shiller to hold in-sample? The answer put restrictions on where the formula could be used. Next, we assumed these conditions were met and asked: what would it take to extend the formula from realized data to expected outcomes under correct expectations? The answer further narrowed the effective range. This section assumes that Campbell-Shiller holds both in-sample and under correct expectations. The question is now: when will the formula remain valid under subjective beliefs?

The conventional wisdom: always. The correct answer: sometimes. Subsection 3.1 spells out what is required. The resulting forecast errors must satisfy a precise adding-up condition. Subsection 3.2 shows that leading models of biased beliefs can be both consistent and inconsistent with Campbell-Shiller. It depends on the amount of additional modeling overhead. Researchers currently view Campbell-Shiller variance decompositions as a model-free way to study subjective beliefs. Subsection 3.3 shows the exact opposite. The resulting calculations have the same form and logic as the Gordon pricing rule. Far from being model-free, these variance decompositions rely on the specific modeling assumptions that Campbell-Shiller was meant to relax.

3.1 Adding-Up Condition

Suppose that Campbell-Shiller holds under correct expectations. Equation (1) describes how investors price a stock. A researcher runs a survey and obtains investors' subjective beliefs about this stock's future dividend growth and returns, $\widetilde{\mathbb{F}}_t[\Delta \log \mathrm{Div}_{t+h}]$ and $\widetilde{\mathbb{F}}_t[\mathrm{R}_{t+h}]$. She then plugs these numbers into Equation (1) in place of the objectively correct predictions, $\mathbb{E}_t[\Delta \log \mathrm{Div}_{t+h}]$ and $\mathbb{E}_t[\mathrm{R}_{t+h}]$. The resulting formula will not always remain valid. Equality only persists if the resulting forecast errors satisfy a precise adding-up condition. This subsection characterizes that condition.

Researchers in this area usually think in terms of subjective expectations, $\widetilde{\mathbb{E}}_t[\cdot]$. They treat "expectations" as a synonym for any form of beliefs. The survey chapter by Adam and Nagel (2023) opens by noting that the market price of assets “reflects investors’ price and payout expectations.” The chapter itself is titled “Expectations data in asset pricing.” Survey responses about future events get recorded as expected outcomes. Everything gets interpreted in this way.

$\widetilde{\mathbb{F}}_t[\cdot]$ is something different. It is a subjective forecast, representing an arbitrary mapping from random variables to real numbers. It need not be linear. It need not correspond to any probability measure. It may be specified independently across variables and horizons. A subjective forecast, $\widetilde{\mathbb{F}}_t[\cdot]$, is a more permissive object than a subjective expectation, $\widetilde{\mathbb{E}}_t[\cdot]$. Every $\widetilde{\mathbb{E}}_t[\cdot]$ is a legal $\widetilde{\mathbb{F}}_t[\cdot]$. The converse is not true.

By framing everything in terms of expectations, researchers have been imposing lots of additional structure on investors’ beliefs. $\widetilde{\mathbb{E}}_t[\cdot]$ allows investors to be wrong about the data-generating process, but their views must still be coherent. Linearity, iterated expectations, and respect for almost-sure identities all follow from being a well-defined expectation.

To get a better feel for the extra structure in subjective expectations, $\widetilde{\mathbb{E}}_t[\cdot]$, let’s look at an example. Consider an investor whose beliefs are anchored in Gordon logic. Her model says that the required return and the dividend growth rate are both constants. Rearranging Equation (2) gives a single pricing constraint on her primitives, $(\mathrm{R}-\mathrm{G}) = \mathrm{DY}$. Thus, she really only has one degree of freedom, not two. If the investor’s return forecast is +1%pt too high, then her anticipated dividend-growth rate must be +1%pt too high as well. The two errors are locked by the structure of her own model. The adding-up condition in Equation (3) follows automatically.

This internal-consistency requirement disappears when we replace the subjective expectations operator, $\widetilde{\mathbb{E}}_t[\cdot]$, with an arbitrary subjective forecasting rule, $\widetilde{\mathbb{F}}_t[\cdot]$. The investor can now believe next year’s return will be +1%pt higher without needing to make any revisions to her forecast for dividend growth. There is no longer any requirement that the investor reconcile her views about the future with the current price level. The adding-up condition becomes a single linear restriction on an infinite-dimensional set of forecasts. It is generically violated under $\widetilde{\mathbb{F}}_t[\cdot]$. Satisfying instances are knife-edge cases.

Proposition 3.1. *Suppose that Campbell-Shiller holds under correct expectations as shown in Equation (1). If we replace $\mathbb{E}_t[\cdot]$ with $\widetilde{\mathbb{F}}_t[\cdot]$, the formula continues to hold if and only if the following adding-up condition is satisfied*

$$\underbrace{\sum_{h=1}^{\infty} \frac{(\widetilde{\mathbb{F}}_t - \mathbb{E}_t)[\Delta \log \mathrm{Div}_{t+h}]}{(1+\overline{\mathrm{DY}})^{h-1}}}_{\substack{\text{Present value of errors about future} \\ \text{dividend growth discounted at } \overline{\mathrm{DY}}}} = \underbrace{\sum_{h=1}^{\infty} \frac{(\widetilde{\mathbb{F}}_t - \mathbb{E}_t)[\mathrm{R}_{t+h}]}{(1+\overline{\mathrm{DY}})^{h-1}}}_{\substack{\text{Present value of return} \\ \text{errors discounted at } \overline{\mathrm{DY}}}} \tag{3}$$

The adding-up condition says that forecast errors about future returns and future dividend growth must have the same present value. But the discount rate in this calculation is not R. It is $\overline{\mathrm{DY}}$. The time-cost of money is R, but the cost of mistakes gets discounted at $\overline{\mathrm{DY}}$. This is a direct consequence of the log-linearization. Campbell-Shiller approximates around the average log PD ratio. $\overline{\mathrm{DY}}$ winds up taking the place of $(\mathrm{R}-\mathrm{G})$ in the discounting because $\rho = \left(\frac{1}{1+\overline{\mathrm{DY}}}\right)$.

A simple counterexample. Suppose investors unilaterally increase their S&P 500 return forecast for next year, $\widetilde{\mathbb{F}}_t[\mathrm{R}_{t+1}] = \mathbb{E}_t[\mathrm{R}_{t+1}] + 1\%\mathrm{pt}$. All other forecasts are objectively correct. Investors' approach to forecasting next year's return has no impact on realized data. So the left-hand side of Equation (1) stays the same. By contrast, the right-hand side has moved by 1%pt. Equality has been violated. The proposition above answers a harder question: Which subjective beliefs are consistent with Campbell-Shiller? We now work through three examples to provide a better sense of the implications.

Perpetually Wrong. Suppose investors overestimate every future return by the same amount

$$\widetilde{\mathbb{F}}_t[\mathrm{R}_{t+h}] = \mathbb{E}_t[\mathrm{R}_{t+h}] + \delta \qquad \text{for all } h = 1, 2, 3, \ldots \tag{17}$$

$\delta > 0$ is the magnitude of the mistake. Assume dividend-growth forecasts are correct, $\widetilde{\mathbb{F}}_t[\Delta \log \mathrm{Div}_{t+h}] = \mathbb{E}_t[\Delta \log \mathrm{Div}_{t+h}]$ for all $h \geq 1$. Hence, we are looking at a unilateral mistake about returns. The dividend-growth side of Equation (3) is zero. There are no errors to discount at $\overline{\mathrm{DY}}$ per year.

The return side is different. There we get a geometric sum

$$\sum_{h=1}^{\infty} \frac{\delta}{(1+\overline{\mathrm{DY}})^{h-1}} \;\approx\; \delta \times \left(\frac{1}{\overline{\mathrm{DY}}}\right) \tag{18}$$

The adding-up condition fails by $\delta \times \left(\frac{1}{\overline{\mathrm{DY}}}\right)$. For $\overline{\mathrm{DY}} \approx 2\%$ and $\delta = 1\%$pt, the violation is 0.50. We write approximately because we are ignoring the difference between $(1+\overline{\mathrm{DY}})^{h-1}$ and $(1+\overline{\mathrm{DY}})^{h}$ in the denominator.

The multiplier $\left(\frac{1}{\overline{\mathrm{DY}}}\right)$ is the Gordon multiple. In the Gordon model, $\overline{\mathrm{DY}} = (\mathrm{R}–\mathrm{G})$, and the current price of a stock is $\mathbb{E}_t[\mathrm{Div}_{t+1}] \times \left(\frac{1}{\mathrm{R}-\mathrm{G}}\right)$. A perpetual δ error on future returns is capitalized at exactly the rate Gordon uses to capitalize cash flows. A small bias turns into a large inconsistency because $\overline{\mathrm{DY}}$ is small.

Moving Target. Now suppose the investor's return-forecast error decays geometrically with horizon at rate $\phi \in [0,1)$

$$\widetilde{\mathbb{F}}_t[\mathrm{R}_{t+h}] = \mathbb{E}_t[\mathrm{R}_{t+h}] + \phi^{h-1} \cdot \delta \qquad \text{for all } h = 1, 2, 3, \ldots \tag{19}$$

$\delta > 0$ is the magnitude of the period-one error. Assume that dividend-growth forecasts are correct, just like before. The ϕ^{h-1} on the right-hand side reflects AR(1) dynamics with persistence ϕ.

The adding-up violation in Equation (3) now corresponds to a different geometric sum

$$\sum_{h=1}^{\infty} \frac{\phi^{h-1} \cdot \delta}{(1+\overline{\mathrm{DY}})^{h-1}} \;\approx\; \delta \times \left(\frac{1}{1-\rho \cdot \phi}\right) \tag{20}$$

This multiplier $\left(\frac{1}{1-\rho\cdot\phi}\right)$ recurs throughout the Campbell-Shiller literature. The approximation comes from treating $(1+\overline{\mathrm{DY}})^{h-1}$ like $(1+\overline{\mathrm{DY}})^{h}$ in the denominator.

For small $\overline{\mathrm{DY}}$, we have $\left(\frac{1}{1-\rho\cdot\phi}\right) \approx \frac{1}{\overline{\mathrm{DY}}+(1-\phi)}$. This is the Gordon multiple $\left(\frac{1}{\overline{\mathrm{DY}}}\right)$ with a mean-reversion correction $(1-\phi)$ added to the denominator. When $\phi = 1$ the correction vanishes, and we return to the perpetually wrong case from before. When $\phi = 0$, the multiplier is approximately 1 and a $\delta = 1\%$pt error produces a 0.01 violation, no amplification. For an interior case with $\overline{\mathrm{DY}} \approx 2\%$ and $\phi = 0.6$, the multiplier is roughly 2.4×.

Measurement Error. Both of the preceding examples involved biased beliefs. The same gap between $\widetilde{\mathbb{F}}_t[\cdot]$ and $\mathbb{E}_t[\cdot]$ can also arise from a more mundane source: estimation error. A researcher who plugs sample estimates of $\mathbb{E}_t[\Delta \log \text{Div}_{t+h}]$ and $\mathbb{E}_t[\text{R}_{t+h}]$ into Equation (1) is replacing the true conditional expectations with noisy estimates. Suppose these estimation errors are each IID white noise with mean zero, mutually independent across all horizons and across variables, with at least one having positive variance.

The difference between the left- and right-hand side of Equation (3) would be a mean-zero random variable with variance

$$\frac{\sigma_{\text{Div}}^2 + \sigma_{\text{R}}^2}{1 - \rho^2} \tag{21}$$

σ_{Div}^2 and σ_{R}^2 are the variances of the dividend-growth and return errors. The probability that this random variable equals exactly zero is zero.

Equation (3) is violated almost surely. The errors do not need to be biased. They do not need to be persistent. Even unbiased, independent, mean-zero measurement noise violates the adding-up condition. This rhymes with Proposition 2.3, which showed that forward-looking Campbell-Shiller under correct expectations is fragile to small deviations in its inputs. Here we land on the same conclusion from the subjective-beliefs side. Two routes, one verdict: Campbell-Shiller does not tolerate ordinary noise.

Several papers mention that subjective beliefs must satisfy a transversality condition for Campbell-Shiller to hold. For example, Adam and Nagel (2023) notes that "for the terminal price to disappear, it must also be common knowledge that the expected discounted terminal price is equal to zero under all agents' beliefs." This discussion misses the mark. Subjective transversality is a condition on the tail of an infinite sum obtained by iterating the one-period Campbell-Shiller identity forward under $\widetilde{\mathbb{F}}_t[\cdot]$. Forward iteration requires $\widetilde{\mathbb{F}}_t[\cdot]$ to be linear, which requires $\widetilde{\mathbb{F}}_t[\cdot]$ to be an expectation under some probability measure. By the time subjective transversality is well-defined, $\widetilde{\mathbb{F}}_t[\cdot]$ has already been promoted to $\widetilde{\mathbb{E}}_t[\cdot]$, and all the structure needed for adding-up is in place.

3.2 Popular Models

Imagine that investors unilaterally increase their return forecast for the S&P 500 by +1%pt. We saw in the introduction that this choice of $\widetilde{\mathbb{F}}_t[\cdot]$ violates Campbell-Shiller. We also saw how it could be brought back in line by assuming investors also overestimate the S&P 500's dividend growth next year by +1%pt. Given the S&P 500's average dividend yield, $\overline{\mathrm{DY}} \approx 2\%$, investors could also overestimate dividend growth three years out by $(1+\overline{\mathrm{DY}})^2 \cdot 1\%\mathrm{pt} \approx 1.04\%\mathrm{pt}$.

If a set of subjective beliefs violates Campbell-Shiller, it is always possible to restore consistency by layering on additional assumptions. We now look at several leading models of biased subjective beliefs from the behavioral-finance literature and find the same pattern. The simple, single-variable version of each bias is inconsistent with Campbell-Shiller. But it is possible to fix things by adding more structure.

Return extrapolation. Return extrapolation assumes that investors overweight recent return surprises when forecasting future returns. Suppose demeaned returns follow an AR(1), $\mathrm{R}_t = \phi \cdot \mathrm{R}_{t-1} + \varepsilon_t$, and let $K_t = \varepsilon_t$ denote the most recent surprise. The subjective forecast at horizon h is

$$\widetilde{\mathbb{F}}_t[\mathrm{R}_{t+h}] \;=\; \mathbb{E}_t[\mathrm{R}_{t+h}] \;+\; \phi^h \cdot (\theta \cdot K_t) \tag{22}$$

$\theta > 0$ controls the strength of the bias. ϕ^h comes from the rational forecast revision. New information at t changes the h-step rational return forecast by $\phi^h \cdot \varepsilon_t$. Investors with extrapolative beliefs overweight this revision by θ. The adding-up violation sums to $\phi \cdot (\theta \cdot K_t) \times \left(\frac{1}{1-\rho\cdot\phi}\right)$.

Restoring consistency requires a general-equilibrium model. Barberis et al. (2015) embeds the bias in a two-agent economy where extrapolators trade against rational investors. Prices emerge from market clearing, which simultaneously determines expected returns, expected dividend growth, and valuations. Jin and Sui (2022) does the same with a representative agent and a subjective Euler equation. In both cases, the equilibrium pins down the joint distribution. The beliefs become $\widetilde{\mathbb{E}}_t[\cdot]$, and the adding-up condition is satisfied by construction. But it is the equilibrium apparatus, not the extrapolation, that delivers this result. Return extrapolation on its own violates Campbell-Shiller.

Diagnostic expectations. Bordalo et al. (2018) proposes that investors overweight states that have become relatively more likely, generating overreaction to recent news. The mathematics is analogous to dividend-growth extrapolation. Suppose demeaned dividend growth follows an AR(1), $\Delta \log \mathrm{Div}_t = \phi \cdot \Delta \log \mathrm{Div}_{t-1} + \varepsilon_t$, and let $K_t = \varepsilon_t$ denote the most recent dividend-growth shock. When new information arrives at time t, the rational forecast revision at horizon h is $\mathbb{E}_t[\Delta \log \mathrm{Div}_{t+h}] - \mathbb{E}_{t-1}[\Delta \log \mathrm{Div}_{t+h}] = \phi^h \cdot K_t$. Diagnostic expectations amplify this revision by $\theta > 0$

$$\widetilde{\mathbb{F}}_t[\Delta \log \mathrm{Div}_{t+h}] \;=\; \mathbb{E}_t[\Delta \log \mathrm{Div}_{t+h}] \;+\; \theta \cdot (\phi^h \cdot K_t) \tag{23}$$

The ϕ^h is inherited from the forecast revision, which decays at rate ϕ because dividend growth has persistence ϕ. The adding-up violation is $\theta \cdot (\phi \cdot K_t) \times \left(\frac{1}{1-\rho \cdot \phi}\right)$.

Because the bias applies to dividend growth, a simpler fix is available. Bordalo, Gennaioli, La Porta, and Shleifer (2019) derives the price from the discounted sum of diagnostically expected dividends, then define the return forecast as the residual from the Campbell-Shiller identity. This forces the return side of Equation (3) to match the dividend-growth side by construction. The consistency comes from treating one forecast as a residual of the other, not from diagnostic expectations per se.

Bordalo et al. (2024) does not take this step. They apply the diagnostic distortion only to expected dividend growth while setting required returns to a constant unrelated to the belief distortion. The price is computed by plugging the distorted cash-flow forecasts into the forward-looking Campbell-Shiller formula. There is no offsetting return error. This is inconsistent with the adding-up condition, and the resulting variance decomposition inherits the violation.

Personal experience. Malmendier and Nagel (2011) proposes that agents overweight personally experienced returns when forming expectations. Returns are approximately unpredictable, $\mathrm{R}_t = \mu + \varepsilon_t$. The anchor K_t is the investor's lifetime experienced return, which updates slowly, $K_t = \theta \cdot K_{t-1} + (1-\theta) \cdot \varepsilon_t$. As $\theta \longrightarrow 1$, old experiences fade more and more slowly. The subjective forecast is

$$\widetilde{\mathbb{F}}_t[\mathrm{R}_{t+h}] \;=\; \mathbb{E}_t[\mathrm{R}_{t+h}] \;+\; K_t \tag{24}$$

The bias is the same at every horizon. The investor simply thinks returns are K_t higher, indefinitely. This is exactly the perpetual error from Equation (17), giving an adding-up violation of $K_t \times \left(\frac{1}{1-\rho}\right)$.

Malmendier and Nagel (2011) measures these experience effects from the UBS/Gallup survey, where cohort-level return expectations are correlated with cohort-specific experienced returns. Those survey-measured expectations are $\widetilde{\mathbb{F}}_t[\cdot]$. A researcher who combined them with separately sourced dividend-growth expectations in the forward-looking Campbell-Shiller formula would violate adding-up. Restoring consistency requires an equilibrium model.

Natural expectations. Fuster et al. (2010) proposes that agents fit a parsimonious model to data generated by a higher-order process. Suppose dividend growth follows an AR(2), $\Delta \log \mathrm{Div}_t = \phi_1 \cdot \Delta \log \mathrm{Div}_{t-1} + \phi_2 \cdot \Delta \log \mathrm{Div}_{t-2} + \varepsilon_t$, but investors fit an AR(1), $\Delta \log \mathrm{Div}_t = \hat{\phi}_1 \cdot \Delta \log \mathrm{Div}_{t-1} + \hat{\varepsilon}_t$. The agent's subjective forecast is

$$\widetilde{\mathbb{F}}_t[\Delta \log \mathrm{Div}_{t+h}] \;=\; \hat{\phi}_1^h \cdot \Delta \log \mathrm{Div}_t \tag{25}$$

Because $\hat{\phi}_1$ exceeds the true AR(2) eigenvalues, the agent overestimates persistence. Short-run momentum is captured but long-run mean reversion is missed. Unlike the other applications, there is no separate anchor K_t. The forecast error at horizon h involves three exponential terms, and the adding-up violation is generically nonzero.

Fuster, Hebert, and Laibson (2012) solves a consumption-based asset-pricing model in which the agent prices assets via an Euler equation under the misspecified model. The Euler equation simultaneously pins down expected returns and expected dividend growth under the agent's coherent (but wrong) probability measure. The beliefs become $\widetilde{\mathbb{E}}_t[\cdot]$, and the adding-up condition is satisfied by construction. But a researcher who took the natural-expectations forecast for dividend growth and combined it with separately measured return expectations would be working with $\widetilde{\mathbb{F}}_t[\cdot]$, and the adding-up condition would not hold.

Every major class of belief distortion in the asset-pricing literature follows the same pattern. The reduced-form, single-variable version of the bias violates the adding-up condition. It can be brought back in line by adding more modeling infrastructure: equilibrium conditions, market clearing, Euler equations, calibrated parameters. But this raises a question. What is the underlying logic behind what all these different sets of extra assumptions are doing?

The answer is that the assumptions close the asset-pricing model enough to cross-validate the investor's forecasts. Given her forecasts of any subset of the observables, the closed model recovers the rest. This invertibility is what makes the model consistent with Campbell-Shiller.

Proposition 3.2. *Suppose investors have beliefs that can be cross-validated. That is, for every $n \in \{1,...,N\}$, the subjective forecast of the nth variable equals the value implied by their asset-pricing model when given the other $(N-1)$ forecasts.*

(a) $\widetilde{\mathbb{F}}_t[\cdot]$ *represents a conditional expectation operator.*

(b) The adding-up condition in Equation (3) *is satisfied.*

Property (a) is Pearl (1988)'s recursive factorization theorem. A system of structural equations that admits a unique solution for any subset of variables determines a probability measure under which the variables are jointly distributed. The asset-pricing model the investor uses is such a system. Cross-validation is the invertibility property that lets the theorem apply. The investor's forecasts are her point predictions for each observable, and under cross-validation they coincide with the conditional expectations under the implied measure.

Property (b) follows from (a) immediately. Once it has been determined that $\widetilde{\mathbb{F}}_t[\cdot]$ is actually a conditional expectation, we immediately know it is linear. Apply it to the realized-values Campbell-Shiller identity. Apply $\mathbb{E}_t[\cdot]$ to the same identity. Both produce equations with the same observed left-hand side, $\log \text{PD}_t$. Subtracting gives the adding-up condition.

What Proposition 3.2 reveals is that the modeling overhead the theorist layers on top of a behavioral bias has a single unifying purpose: to make the model invertible. Once invertibility is achieved, the predictions of the model are no longer specific to the bias the theorist started with. They are properties of the closed asset-pricing model itself. The bias chooses where to start. The closure determines where to end up.

For return biases (return extrapolation, personal experience), a theorist needs to find some way of generating offsetting dividend-growth errors. But the bias is a story about returns, not dividends, so there is no natural way to produce these errors by hand. The only fix is a general-equilibrium model. You need an Euler equation that simultaneously determines expected returns, expected dividend growth, and valuations under a single coherent measure.

A simpler fix is available for dividend-growth biases (diagnostic and natural expectations). Prices follow directly from dividend forecasts. Once you have a biased price, returns are mechanically determined. A researcher can compute the price from discounted biased dividends and define the return forecast as the residual from the Campbell-Shiller identity. This forces the return errors to offset the dividend-growth errors by construction.

In academic models, adding-up holds because the theorist has chosen the closure that makes it hold. An investor cannot do the same on her own. To enforce adding-up, she would need detailed knowledge of her own forecast errors at every horizon, and the discipline to offset her return errors with dividend-growth errors of exactly the right magnitude. Anyone who understood the structure of her own errors that well would just stop making them.

3.3 Variance Decomposition

Researchers currently view the Campbell-Shiller formula as a relatively model-free way to study subjective beliefs. It is not. For Campbell-Shiller to hold under subjective beliefs, the resulting forecast errors must obey a precise adding-up condition. Beliefs can be wrong, but they cannot be arbitrarily wrong. Hence, papers that use Equation (1) are smuggling a lot of structure in through the back door. Variance decompositions are a particularly clear example of the problems this creates.

Campbell-Shiller variance decompositions are designed to answer the following question: Which variable is the main driver of market fluctuations? Researchers currently assume that the formula always holds: in-sample, in expectation, and under arbitrary subjective beliefs. Hence, they take it for granted that all variation in the S&P 500's log PD ratio must be due to changes in future dividend growth or returns. Variance decompositions operationalize this assumption, giving researchers a way to infer which input matters more.

Here is the logic. At each time t, we can observe the S&P 500's log PD ratio. Suppose we could also compute both of the infinite sums on the right-hand side of Equation (1). In that case, we could regress each infinite sum on the S&P 500's current log PD ratio. For dividend growth, the time-series regression would look like

$$\underbrace{\textstyle\sum_{h=1}^{\infty} \rho^{h-1} \cdot \Delta \log \mathrm{Div}_{t+h}}_{\text{One number at each time } t} \overset{\text{OLS}}{\sim} \alpha_{\mathrm{Div}} + \beta_{\mathrm{Div}}(\infty) \cdot \log \mathrm{PD}_t \tag{26}$$

We could run the same sort of regression for returns to get an analogous infinite-horizon slope coefficient, $\beta_{\mathrm{R}}(\infty)$.

To see why this is helpful, note that both slope coefficients take the form $\frac{\mathbb{C}\mathrm{ov}[X,\, \log \mathrm{PD}]}{\mathbb{V}\mathrm{ar}[\log \mathrm{PD}]}$ where X represents an infinite sum involving either dividend growth or returns. Hence, if we assume that all variation in the S&P 500's log PD ratio must come from one of two sources, we can conclude that

$$100\% \;=\; \beta_{\mathrm{Div}}(\infty) \;+\; \beta_{\mathrm{R}}(\infty) \tag{27}$$

This restriction is the key implication of a Campbell-Shiller variance decomposition. The whole approach boils down to this equation.

Notice how the logic behind this calculation differs from the way that researchers use other asset-pricing models. For example, when you deploy the CAPM, you fit the model to the data and then evaluate the size of the resulting errors. A researcher does the opposite when performing a variance decomposition. She assumes Equation (27) holds and uses it to estimate one of the two components. Violations of Campbell-Shiller do not show up as errors. They get absorbed into the parameter being estimated.

Moreover, it is not physically possible to run the time-series regression in Equation (26). Computing the dependent variable involves summing an infinite series at each point in time. Researchers use theory to get around this problem. First, they estimate a one-period analog to the coefficient of interest, $\beta_{\text{Div}}(1)$. Researchers replace the infinite sum on the left-hand side of Equation (26) with data on the S&P 500 the following year. Then, they scale up this short-term estimate using a theory-implied multiple to get $\beta_{\text{Div}}(\infty)$. All remaining variation in the S&P 500's log PD ratio gets labeled as $\beta_{\text{R}}(\infty)$.

Cochrane (2008) performs this exercise using realized data. First, he regresses next year's S&P 500 dividend growth on the index's current log PD ratio. This regression yields $\beta_{\text{Div}}(1) \approx 0.8\%$. He then estimates a persistence of $\phi \approx 0.92$, giving a multiplier of roughly 10×. The full-horizon estimate is thus

$$\beta_{\text{Div}}(\infty) \;=\; \overbrace{\beta_{\text{Div}}(1)}^{0.8\%} \times \overbrace{\left(\frac{1}{1 - \underbrace{\rho}_{0.98} \cdot \underbrace{\phi}_{0.92}}\right)}^{\sim 10\times} \;\approx\; 8\% \tag{4}$$

The paper attributes the remaining 92% of variance to discount-rate changes.

De la O and Myers (2021) runs the same calculations using data from surveys. This data yields a much larger short-term slope coefficient, $\beta_{\text{Div}}(1) \approx 39\%$. Survey respondents also expect future dividend growth to be less persistent than it actually is, $\phi \approx 0.60$, leading to a smaller multiplier, 2.4×. Putting the pieces together, the authors arrive at

$$\beta_{\text{Div}}(\infty) \;=\; \overbrace{\beta_{\text{Div}}(1)}^{39\%} \times \overbrace{\left(\frac{1}{1 - \underbrace{\rho}_{0.98} \cdot \underbrace{\phi}_{0.60}}\right)}^{\sim 2.4\times} \;\approx\; 93\% \tag{28}$$

Apparently, subjective return expectations only account for the remaining 7%.

Which number is correct? Both cannot be right. Yet both papers claim to be performing a model-free exercise based on an accounting identity that always holds. Cochrane (2008) describes the regression coefficients as "linked by an identity" implied by the Campbell-Shiller linearization. De la O and Myers

(2021) writes that they "can calculate the variance decomposition for the price-dividend ratio without making any assumptions about the objective distribution for earnings and dividends."

These two papers are performing calculations that are structurally identical to the Gordon pricing rule. Gordon solves an infinite present-value sum by assuming constant dividend growth, which collapses the infinite sum to a more manageable calculation: a short-term estimate, $\mathbb{E}_t[\mathrm{Div}_{t+1}]$, times a theory-implied multiple, $\left(\frac{1}{\mathrm{R}-\mathrm{G}}\right)$. Cochrane (2008) and De la O and Myers (2021) side-step the infinite sum required by the variance decomposition in Equation (27) by assuming AR(1) persistence in the predictor. This choice is what collapses the infinite sum to a more manageable calculation: a short-term estimate, $\beta_{\mathrm{Div}}(1)$, times a theory-implied multiple, $\left(\frac{1}{1-\rho\cdot\phi}\right)$. Same operational problem. Same structural solution.

Campbell-Shiller is a model in all but name. It is imperative that researchers treat it as such. Calling Equation (1) an "identity" is assumption laundering. The variance-decomposition exercise in Cochrane (2008) and De la O and Myers (2021) is not model free. It comes from imposing the exact constant-parameter restrictions that Campbell-Shiller was originally designed to relax. The original result got published because everyone understood that the Gordon model was a model. By contrast, the variance decomposition above treats Equation (1) as a fact, making it harder to decipher what is actually going on.

Conclusion

Our goal is not to discredit Campbell and Shiller (1988). There is nothing incorrect about the original paper. The formula does sometimes hold, and it is a useful extension of the Gordon model. The log-linear approximation is exactly the right tool for the job in situations where prices come from a time-varying Gordon model. We are making a specific claim about how the literature has come to view Campbell-Shiller.

Researchers currently believe that Equation (1) is an accounting identity that always holds no matter what. Not true. It fails in realized data whenever per-share dividends diverge from total payouts. It fails under correct expectations whenever investors do not apply present-value logic, or when they cannot estimate the cap rate to basis-point precision. And it fails under subjective beliefs whenever the resulting forecast errors do not satisfy the adding-up condition in Equation (3).

We did not write this paper to scold researchers for past mistakes. Every example and application we use comes from leading researchers that do careful work. The takeaway is not: "these people did something wrong." It is: "wow, there is a lot more going on inside Campbell-Shiller than I realized."

The Gordon model does not describe how every asset gets priced. Often, the reason has nothing to do with constant parameters. A homebuyer who looks at comparable recent sales in the same neighborhood is not approximating a time-varying Gordon model. A sell-side analyst who sets a price target by multiplying their EPS forecast for next year times a trailing PE is not approximating Gordon logic, either (Ben-David and Chinco, 2025). Both of these pricing rules anchor on a recently observed number rather than discounting a future cash-flow stream. When researchers insist that it is possible to interpret every asset's price through the lens of Campbell-Shiller, they blind themselves to the true pricing mechanism in these scenarios.

Our goal is to open up a broader range of modeling possibilities. Time-varying Gordon is such a narrow lens. Markets can be so much more interesting. Prices can anchor on comparable transactions, as in the housing market. Price targets can come from a backward-looking multiple times a short-term earnings forecast, as in sell-side equity research. Valuations can reflect payout-policy choices that Campbell-Shiller does not see. None of these possibilities are available to a researcher who treats Equation (1) as an accounting identity. To appreciate them, you have to be open to the possibility that the formula does not hold. This paper is an argument for that openness.

References

Adam, K., A. Marcet, and J. Beutel (2017). Stock price booms and expected capital gains. American Economic Review 107(8), 2352–2408.

Adam, K., A. Marcet, and J. Nicolini (2016). Stock-market volatility and learning. Journal of Finance 71(1), 33–82.

Adam, K. and S. Nagel (2023). Expectations data in asset pricing. In Handbook of Economic Expectations, pp. 477–506. Elsevier.

Barberis, N., R. Greenwood, L. Jin, and A. Shleifer (2015). X-capm: An extrapolative capital asset pricing model. Journal of Financial Economics 115(1), 1–24.

Ben-David, I. and A. Chinco (2025). Expected EPS × Trailing PE: Pricing without discounting. Working paper.

Bernanke, B. and K. Kuttner (2005). What explains the stock market's reaction to Federal Reserve policy? Journal of Finance 60(3), 1221–1257.

Bordalo, P., N. Gennaioli, R. La Porta, and A. Shleifer (2019). Diagnostic expectations and stock returns. Journal of Finance 74(6), 2839–2874.

Bordalo, P., N. Gennaioli, R. La Porta, and A. Shleifer (2024). Finance without (exotic) risk. Working paper.

Bordalo, P., N. Gennaioli, and A. Shleifer (2018). Diagnostic expectations and credit cycles. Journal of Finance 73(1), 199–227.

Boudoukh, J., R. Michaely, M. Richardson, and M. Roberts (2007). On the importance of measuring payout yield: Implications for empirical asset pricing. Journal of Finance 62(2), 877–915.

Campbell, J. (1991). A variance decomposition for stock returns. The Economic Journal 101(405), 157–179.

Campbell, J. (2017). Financial decisions and markets: a course in asset pricing. Princeton University Press.

Campbell, J. and R. Shiller (1988). The dividend-price ratio and expectations of future dividends and discount factors. Review of Financial Studies 1(3), 195–228.

Case, K. E. and R. J. Shiller (1987). Prices of single-family homes since 1970: New indexes for four cities. New England Economic Review, 45–56.

Case, K. E. and R. J. Shiller (1989). The efficiency of the market for single-family homes. American Economic Review 79(1), 125–137.

Chen, L. and X. Zhao (2009). Return decomposition. Review of Financial Studies 22(12), 5213–5249.

Chinco, A., S. Hartzmark, and A. Sussman (2022). A new test of risk factor relevance. Journal of Finance 77(4), 2183–2238.

Cho, T., L. Kremens, D. Lee, and C. Polk (2024). Scale or yield? A present-value identity. Review of Financial Studies 37(3), 950–988.

Cochrane, J. (2001). Asset Pricing. Princeton University Press.

Cochrane, J. (2011). Presidential address: Discount rates. Journal of finance 66(4), 1047–1108.

Cochrane, J. H. (2008). The dog that did not bark: A defense of return predictability. Review of Financial Studies 21(4), 1533–1575.

Cohen, R., C. Polk, and T. Vuolteenaho (2003). The value spread. Journal of Finance 58(2), 609–641.

De la O, R. and S. Myers (2021). Subjective cash flow and discount rate expectations. Journal of Finance 76(3), 1339–1387.

De la O, R. and S. Myers (2024). Which subjective expectations explain asset prices? Review of Financial Studies 37(6), 1829–1867.

Doyle, A. C. (1890). The Sign of the Four. London: Spencer Blackett.

Engsted, T., T. Q. Pedersen, and C. Tanggaard (2012). The log-linear return approximation, bubbles, and predictability. Journal of Financial and Quantitative Analysis 47(3), 643–665.

Fuster, A., B. Hebert, and D. Laibson (2012). Natural expectations, macroeconomic dynamics, and asset pricing. NBER Macroeconomics Annual 26(1), 1–48.

Fuster, A., D. Laibson, and B. Mendel (2010). Natural expectations and macroeconomic fluctuations. Journal of Economic Perspectives 24(4), 67–84.

Gao, C. and I. Martin (2021). Volatility, valuation ratios, and bubbles: An empirical measure of market sentiment. Journal of Finance 76(6), 3211–3254.

Gordon, M. (1962). The investment, financing, and valuation of the corporation. RD Irwin.

Gourinchas, P. and H. Rey (2007). International financial adjustment. Journal of Political Economy 115(4), 665–703.

Greenwood, R. and A. Shleifer (2014). Expectations of returns and expected returns. Review of Financial Studies 27(3), 714–746.

Hartzmark, S. and A. Sussman (2025). Price agnostic demand. Working paper.

Jin, L. J. and P. Sui (2022). Asset pricing with return extrapolation. Journal of Financial Economics 145(2), 273–295.

Larrain, B. and M. Yogo (2008). Does firm value move too much to be justified by subsequent changes in cash flow? Journal of Financial Economics 87(1), 200–226.

Lettau, M. and S. Ludvigson (2001). Consumption, aggregate wealth, and expected stock returns. Journal of Finance 56(3), 815–849.

Lettau, M., S. Ludvigson, and P. Manoel (2018). Characteristics of mutual-fund portfolios: Where are the value funds? Working paper.

Lintner, J. (1956). Distribution of incomes of corporations among dividends, retained earnings, and taxes. American Economic Review 46(2), 97–113.

Malmendier, U. and S. Nagel (2011). Depression babies: Do macroeconomic experiences affect risk taking? Quarterly Journal of Economics 126(1), 373–416.

Nagel, S. and Z. Xu (2022). Asset pricing with fading memory. Review of Financial Studies 35(5), 2190–2245.

Pearl, J. (1988). Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. Morgan Kaufmann.

Vuolteenaho, T. (2002). What drives firm-level stock returns? Journal of Finance 57(1), 233–264.

A Technical Appendix

Proof. **(Proposition 1.2)**

By definition, the per-share dividend is $\text{Div}_{t+h} = \frac{\text{TotDiv}_{t+h}}{\#\text{Shares}_{t+h}}$. *Taking logs and first-differencing gives*

$$\Delta \log \text{Div}_{t+h} \;=\; \Delta \log \text{TotDiv}_{t+h} \;-\; \Delta \log \#\text{Shares}_{t+h} \tag{A.1}$$

Subtracting $\Delta \log \text{TotPmt}_{t+h}$ *from both sides gives Equation* (7).

Remark. The proposition assumes R_{t+1} *remains constant. This leaves only one channel through which the per-share substitution can distort Campbell-Shiller:* $\Delta \log \text{TotPmt}_{t+1}$. *Allowing* R_{t+1} *to change would introduce additional distortions, making it even harder to defend the claim that Campbell-Shiller is a mere accounting identity. Holding* R_{t+1} *fixed thus amounts to studying the best-case scenario for the conventional wisdom.* □

Proof. **(Proposition 2.3)**

Under the Gordon model in Equation (2), *the forward dividend yield equals the cap rate,* $\text{DY} = (\text{R}-\text{G})$.

Cap-rate decrease. Suppose noise lowers the cap rate by $\delta > 0$. *The price moves from* $\text{Price}_t = \mathbb{E}_t[\text{Div}_{t+1}] \times \left(\frac{1}{\text{R}-\text{G}}\right)$ *to* $\mathbb{E}_t[\text{Div}_{t+1}] \times \left(\frac{1}{[\text{R}-\text{G}]-\delta}\right)$, *so*

$$\Delta \text{Price}_t \;=\; \mathbb{E}_t[\text{Div}_{t+1}] \times \left(\frac{1}{\text{R}-\text{G}}\right) \cdot \left[\frac{\delta}{(\text{R}-\text{G})-\delta}\right] \tag{A.2}$$

Requiring $|\Delta \text{Price}_t| > \frac{\mathbb{E}_t[\text{Div}_{t+1}]}{1+\text{R}}$ *and cancelling out* $\mathbb{E}_t[\text{Div}_{t+1}]$ *gives*

$$\delta \;>\; \frac{(\text{R}-\text{G})^2}{(1+\text{R})+(\text{R}-\text{G})} \;=\; \text{DY}^2 \cdot \left(\frac{1}{1+\text{R}+\text{DY}}\right) \tag{A.3}$$

The second equality uses $\text{DY} = (\text{R}-\text{G})$.

Cap-rate increase. Suppose instead that the cap rate increases by $\delta > 0$. *The analogous calculation yields*

$$\delta \;>\; \frac{(\text{R}-\text{G})^2}{(1+\text{R})-(\text{R}-\text{G})} \;=\; \text{DY}^2 \times \left(\frac{1}{1+\text{G}}\right) \tag{A.4}$$

The two thresholds differ because $\left(\frac{1}{\mathrm{R}-\mathrm{G}}\right)$ *is convex. A drop in the cap rate moves the price by more than an equal-sized rise, so the decrease case requires a smaller* δ *to exceed a given price change.*

Sufficient condition. Both exact thresholds in Equations (A.3) *and* (A.4) *lie strictly below* DY^2 *whenever* $\mathrm{R} > 0$ *and* $\mathrm{G} > 0$*. Hence* $|\delta| > \mathrm{DY}^2$ *is sufficient regardless of the direction of the perturbation. For the S&P 500,* $\mathrm{DY} \approx 2\%$*, so the threshold is* $\mathrm{DY}^2 \approx 0.04\%\mathrm{pt}$. □

Proof. **(Proposition 3.1)**

Campbell-Shiller holds under correct expectations by assumption. Thus, replacing $\mathbb{E}_t[\cdot]$ *with* $\widetilde{\mathbb{F}}_t[\cdot]$ *will preserve equality if and only if the following difference vanishes*

$$\sum_{h=1}^{\infty} \rho^{h-1} \cdot \left\{ (\widetilde{\mathbb{F}}_t - \mathbb{E}_t)[\Delta \log \mathrm{Div}_{t+h}] \right\} = \sum_{h=1}^{\infty} \rho^{h-1} \cdot \left\{ (\widetilde{\mathbb{F}}_t - \mathbb{E}_t)[\mathrm{R}_{t+h}] \right\} \tag{A.5}$$

Rewriting $\rho = \left(\frac{1}{1+\overline{\mathrm{DY}}}\right)$ *gives the desired result.*

Remark. Transversality plays no role in this argument. The infinite-horizon Campbell-Shiller formula under $\mathbb{E}_t[\cdot]$ *is taken as the starting assumption, so the transversality condition is already baked in. A separate "subjective transversality" condition on* $\widetilde{\mathbb{F}}_t[\cdot]$ *is not well-defined.* $\widetilde{\mathbb{F}}_t[\cdot]$ *need not correspond to any probability measure, so the tail expectation* $\lim_{H\to\infty} \rho^H \cdot \widetilde{\mathbb{F}}_t[\log \mathrm{PD}_{t+H}]$ *has no coherent interpretation.*

Remark. The proposition asks when the observed $\log \mathrm{PD}_t$ *coincides with the subjective-forecast-based right-hand side of Equation* (1)*. Both sides of the subtraction above share the same observed left-hand side. Neither* $\mathbb{E}_t[\cdot]$ *nor* $\widetilde{\mathbb{F}}_t[\cdot]$ *affect* $\log \mathrm{PD}_t = \log\left(\frac{\mathrm{Price}_t}{\mathrm{Div}_t}\right)$*. Campbell-Shiller approximates the multiple when writing Gordon as* $\mathrm{Div}_t \times \left(\frac{1+\mathrm{G}}{\mathrm{R}-\mathrm{G}}\right)$ *rather than* $\mathbb{E}_t[\mathrm{Div}_{t+1}] \times \left(\frac{1}{\mathrm{R}-\mathrm{G}}\right)$. □

Proof. **(Proposition 3.2)**

The investor's asset-pricing model is a system of structural equations relating N *observables. This is known as a structural causal model (SCM). The cross-validation hypothesis says this system is invertible. It is possible to recover the nth variable given investor's forecasts for the remaining* $(N-1)$ *observables.*

Property (a). *The argument follows Pearl (1988)'s recursive factorization theorem. Here are the key steps:*

1. *Write the asset-pricing model as a system of structural equations*

$$X_n \;=\; f_n(\mathrm{Pa}_n, U_n), \qquad n = 1, \ldots, N \tag{A.6}$$

   Pa_n *denotes the parents of* X_n *in the implied directed graph.* $(U_1, \ldots, U_N)$ *are exogenous noise variables.*

2. *Each structural equation induces a conditional distribution,* $\widetilde{\mathbb{Q}}(X_n \mid \mathrm{Pa}_n)$*. When* $f_n(\cdot)$ *is deterministic, this conditional is a point mass. With noise, it is non-degenerate.*
3. *Pearl (1988)'s recursive factorization theorem states that the joint distribution can be written as*

$$\widetilde{\mathbb{Q}}(X_1, \ldots, X_N) \;=\; \prod_{n=1}^{N} \widetilde{\mathbb{Q}}(X_n \mid \mathrm{Pa}_n) \tag{A.7}$$

4. *Cross-validation guarantees the SCM is well-posed. Recovering the* N*-th forecast from any* $(N-1)$ *others requires both existence (the structural equations admit a consistent joint solution) and uniqueness (that solution is a function of the* $(N-1)$ *inputs, not a correspondence). Without existence, no joint distribution would exist; without uniqueness, multiple joint distributions would be consistent with the same structural equations. The cross-validation hypothesis rules out both cases.*
5. *Given a valid SCM, an investor's forecast for each variable must coincide with its conditional expectation under* $\widetilde{\mathbb{Q}}$*. Hence* $\widetilde{\mathbb{F}}_t[\cdot]$ *corresponds to the subjective expectation operator defined by* $\widetilde{\mathbb{Q}}$*.*

Property (b). *From property (a), we know that* $\widetilde{\mathbb{F}}_t[\cdot]$ *is an expectation operator. Hence, it is linear. We can then apply* $\widetilde{\mathbb{F}}_t[\cdot]$ *and* $\mathbb{E}_t[\cdot]$ *separately to the realized-values Campbell-Shiller identity, which is assumed to hold in-sample. The left-hand side,* $\log \mathrm{PD}_t$*, is observed at* t *and so is unaffected. The right-hand side distributes through the infinite sum by linearity. Differencing the two equations gives*

$$0 \;=\; \sum_{h=1}^{\infty} \rho^{h-1} \cdot \left\{ (\widetilde{\mathbb{F}}_t - \mathbb{E}_t)[\Delta \log \mathrm{Div}_{t+h}] \;-\; (\widetilde{\mathbb{F}}_t - \mathbb{E}_t)[\mathrm{R}_{t+h}] \right\} \tag{A.8}$$

Hence, the adding-up condition in Equation (3) is satisfied. □