Scale or Yield? A Present-Value Identity*

Thummim Cho
London School of Economics

Lukas Kremens
University of Washington

Dongryeol Lee
UCLA Anderson

Christopher Polk
London School of Economics

March 2023

*We thank two anonymous referees, John Campbell, Matthew Cedergren, Roberto Gomez Cram, Itay Goldstein Niels Gormsen, Dong Lou, Stig Lundeby, Ian Martin, Jeffrey Pontiff, Paul Tetlock, Tuomo Vuolteenaho, and seminar participants at Boston College, Baruch College, London School of Economics, University of Hong Kong, University of Washington, American Accounting Association 2021 Annual Meeting, and FIRN-Melbourne Asset Pricing Meeting, MFA 2022 Annual Meeting, SFS Cavalcade 2022, the FIRS 2022 Annual Meeting, and the EFA 2022 Annual Meeting for valuable comments.

Cho: Department of Finance, LSE, London WC2A 2AE, UK. Email: t.cho@lse.ac.uk. Kremens: Department of Finance and Business Economics, University of Washington, Seattle, WA, 98195, US. Email: lkremens@uw.edu. Lee: Department of Finance, UCLA Anderson School of Management, Los Angeles, CA 90095, US. Email: dongryeol.lee.phd@anderson.ucla.edu. Polk: Department of Finance, LSE, London WC2A 2AE, UK and CEPR. Email: c.polk@lse.ac.uk.

Scale or Yield? A Present-Value Identity

Abstract

We propose a loglinear present-value identity in which investment ("scale"), profitability ("yield"), and discount rates determine a firm's market-to-book ratio. Our identity reconciles existing influential market-to-book decompositions and facilitates novel insights from three empirical applications: (i) Both investment and profitability are important contributors to the value spread and stock return news variance. (ii) Any cross-sectional return predictability has a mirror image in cash-flow fundamentals, providing asset-pricing theories with additional moments to match. (iii) The investment spread significantly improves the predictability of time-series variation in the value premium and justifies the poor performance of value in recent years.

JEL classification: G11, G12

1 Introduction

Cash flows to equity holders depend on the scale of the firm's equity capital and its profitability. Hence, investment ("scale"), profitability ("yield"), and discount rates on cash flows are the three fundamental determinants of a stock's market value.

This paper develops a loglinear, firm-level present-value identity in which investment, profitability, and discount rates determine the firm's market-to-book equity ratio (M/B). Conceptually, the identity improves on the the influential M/B decompositions of Vuolteenaho (2002) and Fama and French (2006) and reconciles their apparent inconsistency. Empirically, the identity (i) produces a richer decomposition of the cross-section of valuations and return news that enables us to evaluate recent asset pricing theories, (ii) casts any cross-sectional return predictability to firm fundamentals, and (iii) improves on the identity-based time-series return forecast model of Cohen et al. (2003).

Our identity decomposes the log of M/B as follows:

$$mb_t \approx \sum_{j=1}^{\infty}\rho^{j-1}E_t\left[roe_{t+j} - r_{t+j} + iva_{t+j}\right], \tag{1}$$

where roe for return on equity measures profitability, r is the discount rate, and iva for "investment value added" measures the valuation effect of investment in book equity. Importantly, iva does not simply measure the *direction* or *amount* of book equity growth but how it *interacts* with the market-to-book ratio of retained equity. Furthermore, while firms could adjust book equity either through the share issuance/repurchase ("net issuance") channel or through the plowback/payout ("payout") channel, iva focuses on the net issuance channel, and the empirically-stable payout channel drops out as a parameter in the linearization.[1]

To understand why iva is important, suppose that a firm with $roe > r$ in all periods finds a future project with the same $roe > r$ as the existing capital. For simplicity, assume that the firm plans to

[1]We find that the stickiness of dividend policy in the data makes the plowback/payout decision a much smaller driver of the investment channel. For this reason, we simply call the effect of net-issuance-driven book equity investment "investment value added." For more details, see Appendix B.2.

raise the necessary funds by issuing new shares.[2] Since the project has a positive NPV but does not affect today's book equity, the news should raise the firm's M/B today: $mb_t \uparrow$. However, absent iva—as is the case with Vuolteenaho's (2002) decomposition—, equation (1) cannot capture this rise in mb, since the news leaves the expected roe and r unchanged. The iva term captures such a valuation effect on the scale margin. By issuing new equity shares at price above the book value per share, existing shareholders see their book-equity per share rise, earning the unchanged ROE on a larger equity per share.[3]

Our identity improves on the two influential M/B decompositions. Vuolteenaho (2002) approximates the log M/B as a spread between future profitability and returns:

$$mb_t \approx \sum_{j=1}^{\infty} \rho^{j-1} E_t \left[roe_{t+j} - r_{t+j} \right], \tag{2}$$

where ρ is a constant (around 0.96 in annual data). Nevertheless, equation (2) misses the role of investment emphasized in Fama and French (2006, 2015) and captured by our iva term: "A caveat is that [equation (2)] does not explain the role of an investment factor" (Campbell (2018), p. 226). Hence, to make equation (2) hold in practice, the literature uses a "clean-surplus" roe that conflates profitability and investment, two sources of cash-flow fundamentals with very different economic interpretations. Staying true to equation (2) by using the accounting roe in applications, another common practice, can lead to large estimation errors. The iva variable explicitly adds the missing investment channel to Vuolteenaho, bridging its gap with Fama and French's decomposition below. Introducing iva also obviates the need for the clean-surplus roe that some researchers use in practice when applying Vuolteenaho, closing the gap between theory and practice.

Fama and French (2006) restate the dividend discount model as

$$\frac{M_t}{B_t} = \frac{\sum_{j=1}^{\infty} E_t \left(Y_{t+j} - \Delta B_{t+j} \right) / \left(1 + R \right)^j}{B_t}, \tag{3}$$

[2] If the firm initially issues debt but later reverts to a target capital structure by retaining earnings or issuing equity, the logic of the example is qualitatively unchanged. Note also that the current book equity should remain unchanged.

[3] In other words, existing shareholders appropriate the positive NPV of the project and new shareholders get a fair deal (an expected return of r).

which they use to relate expected stock returns (R) to current M/B, future "profitability" ($E_t Y_{t+j}/B_t$), and future "investment" ($E_t \Delta B_{t+j}/B_t$). However, unlike the loglinear models of Campbell and Shiller (1988) and Vuolteenaho (2002), equation (3) is nonlinear in nonstationary variables, limiting its applicability. Our identity's loglinearity allows for a more precise empirical analysis of how future profitability and investment relate to valuations and returns in a present-value framework.

Past research certainly anticipated a loglinear decomposition of M/B with an investment term. For example, Lochstoer and Tetlock (2020, Appendix E) use a numerical example to interpret the clean-surplus adjustment done to Vuolteenaho's roe as investment value added through net issuance. However, it is not obvious how to analytically introduce an investment variable that internalizes the complementary between the direction of net book equity investment and the sign of the roe-discount-rates spread, since the variable encoding the interaction of these two must be strictly positive to fit the loglinear framework. Our contribution is to formally derive, from first principles, equation (1) with an investment term and to apply it to draw out novel empirical insights.

Applying the identity, we document three new empirical facts about stock valuations and returns and relate them to recent asset pricing theories (Section 4). First, both profitability (roe) and investment (iva) are important determinants of firm valuations and returns—accounting for approximately 40% vs. 17% of the cross-sectional M/B variations and 67% vs. 18% of the return news variance—and the correlation between roe news and iva news is robustly negative. Second, Lochstoer and Tetlock's (2020) finding that for key anomaly portfolios, cash-flow news correlates negatively with discount-rate news arises through both roe news and iva news. Last, high expected iva (e.g., growth options)—as opposed to high expected profitability—has become an increasingly important feature of today's growth firms.

The three empirical findings based on the granular decompositions of valuations and return news variance have important implications for asset pricing theories. The first finding on the importance of iva for valuations and returns as well as the negative correlation between roe news and iva news cast doubt on theories of the value premium that rely on productivity shock as a single source of cash-flow risk. For example, in Zhang (2005), firms with a higher idiosyncratic productivity have a higher market-to-book ratio (i.e., are "growth firms") and less risk (less vulnerable to costly divestment in bad times). For these firms, positive iva news (i.e., an expansion of capital, given high market-to-book) arises from positive productivity shocks, thus making the correlation of profitability and investment news counterfactually positive. Models which decouple productivity from investment-specific shocks (e.g., Belo et al., 2014; Kogan and Papanikolaou, 2013, 2014) to generate a value spread offer more flexibility to match this negative correlation.

The second finding, combined with the first, raises the bar for matching the negative correlation between CF news and DR news found in Lochstoer and Tetlock (2020). Kogan and Papanikolaou (2013), for instance, study investment-specific technology (IST) shocks as drivers of the value of growth options. Belo et al. (2019) use exposures to equity issuance cost shocks (ICS). In both cases, the mechanisms are rich enough or could be extended to generate three pairwise negative correlations of discount-rate news, profitability news, and iva-news, bringing the models closer to reality, but the current calibrations do not match these new facts. Kogan et al. (2022) provides a more flexible investment-based framework with which to match all three news correlations.

Finally, the third finding can help rationalize the rising valuations of some growth firms in recent years (e.g., in the tech sector) despite low profitability: these firms do not necessarily need to reach the same profitability as more established firms to justify similar valuations, as long as their expected profitability exceeds the discount rate and they are able to aggressively scale the excess profitability. If indeed firms time the market for their own stock (Baker and Wurgler, 2002), the rising use of share repurchases could be one specific incidence that raised valuation dispersion through the iva channel.[4] Relatedly, the high valuations of "meme" stocks could be partly sustained through high expected future iva—the expectation that the firm would issue new shares at high prices.

We consider two additional applications of our identity. The identity implies that any return predictor must also predict roe, iva, and/or changes in the valuation ratio. That is, each finding of

[4] Fama and French (2001) were one of the first to comment on the declining incidence of dividend-paying firms, and Skinner (2008) and Grullon et al. (2011) explain that firms have substituted to using share repurchases to gain flexibility with which to time investment opportunities.

return predictability comes with its own 'anatomy' in terms of the other variables in the identity. Section 5 applies this anatomy to return forecasts associated with asset growth, profitability, and cash-flow duration. In each case, the return predictor is associated with cash-flow predictions that are economically and statistically larger, as well as more persistent over longer horizons than the return prediction. Each return predictor further forecasts the two cash-flow components in opposite directions. The composite cash-flow measure from clean-surplus ROE therefore conceals patterns in firm fundamentals associated with these return predictors. These patterns are important for understanding the sources of price and return variation, regardless of whether these fundamentals are tied to risk premia or to systematic errors in investor expectations.[5]

Our third and last application augments the time-series portfolio return predictability model of Cohen et al. (2003) to include the iva spread as an additional predictor using all of the variables in our identity to predict HML returns (Section 6). The value spread—previously shown to predict HML returns in the time series (Cohen et al., 2003)—is a weak *univariate* predictor of returns to value-minus-growth bets in recent data. Restoring the time-series predictability of HML returns requires an additional predictor that accounts for the long-run cash-flow information embedded in the value spread. roe predicts the two cash-flow components in opposite directions and therefore, by itself, fails to improve the HML forecast. In contrast, our new iva variable predicts long-run cash flows better, notably through a fall in profitability. Indeed, iva raises the forecasting power of the value spread and the profitability spread, in addition to it being a negative predictor of HML returns. While we are not the first to argue that cash-flow predictors improve return predictions (Cohen et al., 2003; Asness et al., 2000), our delineation of two fundamentally different cash-flow channels highlights why long-run cash-flow predictability is multifaceted and why familiar, persistent cash-flow characteristics like profitability fall short.

To summarize, we generalize Vuolteenaho (2002)'s framework to restore an explicit investment channel in a loglinear M/B decomposition, bringing coherence to research that relates firm characteristics to stock returns through an identity. Alongside Campbell and Shiller (1988)'s price-

[5] De La O and Myers (2021) study survey data and find that variation in the price-dividend ratio of the S&P 500 can be largely explained by variation in subjective cash-flow expectations expressed in surveys. Franzoni et al. (2022) analyze the ability of the scale margin to explain the cross-section of expected short-horizon stock returns.

dividend ratio decomposition for the aggregate market, our firm-level identity can potentially serve as a workhorse present-value identity in asset pricing research. Indeed, separating profitability and investment-based drivers of cash-flow growth provides a diagnostic tool for asset pricing models and generates novel insights on the sources of cash-flow growth differentials between growth and value firms, the sources of cross-sectional return predictability of asset growth and profitability, and the time-series predictability of value-minus-growth portfolio returns.

2 A loglinear present-value identity with investment

Like the present value of any scalable project, the market value of a firm's equity depends on

(i) the spread between the return on capital inside (return on equity) and outside (discount rate) the firm; and on

(ii) the amount of capital on which this spread is earned, determined by movement of capital in and out of the firm through issuance and repurchases (investment value added).

The existing influential present-value identities struggle to map this economic intuition to tractable notions of firm-level fundamentals that capture both of these channels.

We clarify these economic links with our new firm-level identity, which bridges the gap between the existing M/B decompositions and generates novel empirical findings. We derive our identity in a loglinear framework in the spirit of Campbell and Shiller (1988), since the linearity allows empiricists to measure the relative importance of the three valuation components and theorists to seek an approximate analytical solution or simplify their numerical calibration.

2.1 The identity

Begin with the definition of return:

$$P_t = \frac{1}{1 + R_{t+1}} \left(D_{t+1} + P_{t+1} \right), \tag{4}$$

where P is price, R is return, and D is dividend per share. As shown in detail in Appendix A.1, multiplying both sides by N_t/B_t and then multiplying and dividing the right-hand side by $(B_t + Y_{t+1})/B_t$, that is, $1 + ROE_{t+1}$, and by the cum-dividend book value per share, $D_{t+1} + B_{t+1}/N_{t+1}$, implies

$$\begin{aligned}\frac{P_t N_t}{B_t} &= \frac{1}{1+R_{t+1}} \times \left(\frac{B_t + Y_{t+1}}{B_t}\right) \times \left(\frac{D_{t+1} + B_{t+1}/N_{t+1}}{(B_t + Y_{t+1})/N_t}\right) \\ &\quad \times \left(\frac{D_{t+1}}{D_{t+1} + B_{t+1}/N_{t+1}} + \frac{P_{t+1}N_{t+1}}{B_{t+1}} \times \frac{B_{t+1}/N_{t+1}}{D_{t+1} + B_{t+1}/N_{t+1}}\right),\end{aligned} \tag{5}$$

where N is the number of shares, B is total (as opposed to per-share) book value of equity, and Y is total earnings. Equation (5) leads to our exact nonlinear identity, an intermediate step to our approximate loglinear identity.

Remark 1 ***(Exact nonlinear identity).*** *The market-to-book ratio can be stated as*

$$\frac{M_t}{B_t} = \underbrace{\frac{1}{1+R_{t+1}}}_{\text{discount rate}} \times \underbrace{(1 + ROE_{t+1})}_{\text{profitability}} \times \underbrace{(1 + IVA_{t+1})}_{\text{investment value added}} \times \left(1 + \left(\frac{M_{t+1}}{B_{t+1}} - 1\right) \underbrace{\Lambda_{t+1}}_{\text{plowback}}\right), \tag{6}$$

where $M_t = P_t N_t$ is the market value of equity, $ROE_t = Y_t/B_{t-1}$ is the return on equity, $IVA_t = (D_t + BPS_t)/(BPS_{t-1} + EPS_t) - 1$ is the investment value added which encodes the value added or lost—taking ROE and R as given—from net investment in equity capital each period, and $\Lambda_t = BPS_t/(D_t + BPS_t)$ is the plowback ratio with $BPS_t = B_t/N_t$ denoting the book value per share and $EPS_t = Y_t/N_{t-1}$ denoting earnings per share.

Remark 1 is intuitive. Ceteris paribus, a firm with a higher market-to-book ratio must have, in future periods, lower stock returns (R), higher return on (book) equity (ROE), or more valuable growth or downsize opportunities as captured by both the investment value added (IVA) and the expression involving the plowback ratio (Λ). The last two expressions capture how a firm's net investment in book equity can change its market value without necessarily affecting ROE and R.

IVA measures the change in the cum-dividend book value *per share* through share issuance or repurchase ("net issuance"). Intuitively, without net issuance, $D_t + BPS_t = BPS_{t-1} + EPS_t \implies$

$IVA_t = 0$, since dividends and new book equity must be financed entirely by old book equity and earnings. To see this formally, rewrite IVA using $N_{t-1}D_t + B_t = B_{t-1} + Y_t + (N_t - N_{t-1})\, P_t^*$, where P_t^* is the average net issuance price adjusted by net issuance costs.[6]

Remark 2 ***(Investment value added).*** *The investment value added measures the change in future cash flows per share owing to changes in equity capital employed (through share issuance or repurchases), rather than changes in the return on a unit of equity (ROE):*

$$IVA_t = \frac{D_t + BPS_t}{BPS_{t-1} + EPS_t} - 1 = \left(\frac{N_t - N_{t-1}}{N_{t-1}}\right)\left(\frac{P_t^*}{BPS_t} - 1\right)\frac{BPS_t}{BPS_{t-1} + EPS_t}. \tag{7}$$

Furthermore, in the absence of net issuance costs, the sign of $P_t^*/BPS_t - 1$ depends on the spread between returns on equity and discount rates in the future.[7]

IVA is positive if the firm issues shares at a price (net of issuance costs) above the book value per share or repurchases shares at a price (plus repurchase costs) below the book value per share. Holding returns on equity and the plowback ratio fixed, such an increase in the cum-dividend book value per share unambiguously increases future cash flows to existing shareholders. Importantly, a positive IVA does not mean an increase in cum-dividend *total* book equity, which could be driven by the creation of new shares and can therefore have an ambiguous effect on the cash flows to existing shareholders.[8] The name "investment value added" reflects that even a decrease in total book equity due to share repurchases could add "value" to the existing shareholders if it expands the (cum-dividend) per-share book equity. Since IVA captures the valuation consequences of net issuance, it explicitly addresses the limitations of clean-surplus approaches like those of Vuolteenaho (2002) or Ohlson (1995, 2000).

The Internet Appendix plots the average industry-level $iva = \log(1+IVA)$ against log market-

[6]See the derivation in Appendix A.2

[7]Equation (6) implies $P_t/BPS_t = M_t/B_t = X_{t+1} + \sum_{j=1}^{\infty}\left\{\Pi_{s=1}^{j} X_{t+s}\Lambda_{t+s}\left(X_{t+s+1} - 1\right)\right\}$ where $X_t \equiv (1 + ROE_t)\,(1 + IVA_t)\,/\,(1 + R_t)$. Setting all future IVA to be zero as a benchmark, $P_t/BPS_t - 1 > 0$ if $ROE_{t+j} > R_{t+j}$ for all $j \geq 1$, or in words, when the ROEs exceed the discount rates.

[8]Its ambiguous effect on firm value suggests that book equity (or similarly asset) growth, the investment variable in Fama and French's (2006; 2015) identity, cannot be a stand-alone variable in a loglinear decomposition of the market-to-book ratio.

to-book for the 10 Fama-French industries, along with the time-series volatilities and correlation between the two variables (Figure E.1). Firms in higher-valuation industries tend to have higher iva both across industries and over time. However, since IVA measures the interaction between the share issuance decision and the ratio of net issuance price and the book value per share, it has a weak correlation of only 0.136 with a conventional net issuance variable such as the one used by Pontiff and Woodgate (2008).

Investment in book equity can also take place through the payout-versus-plowback decision. Holding R_{t+1}, ROE_{t+1}, and IVA_{t+1} fixed, paying out dividends when $M_{t+1} < B_{t+1}$ or plowing back to book equity when $M_{t+1} > B_{t+1}$ increases the market value at time t. However, since dividend payouts tend to be sticky, we find this channel of net investment in book equity to be less important empirically. Our loglinearization therefore drops the interaction of Λ with the market-to-book ratio as an approximation error, preserving only the net issuance channel for net investment in book equity through IVA. Appendix B.1 develops more intuition on how the issuance and plowback channels aggregate into investment in book equity in an exact, nonlinear identity.

To obtain our approximate loglinear identity, take the log of both sides of (6) to write

$$mb_t = -r_{t+1} + roe_{t+1} + iva_{t+1} + \log\left(1 + \left(\exp\left(mb_{t+1}\right) - 1\right)\exp\left(\lambda_{t+1}\right)\right), \tag{8}$$

where $mb_t = \log\left(\frac{M_t}{B_t}\right)$, $r_t = \log\left(1 + R_t\right)$, $roe_t = \log\left(1 + ROE_t\right)$, $iva_t = \log\left(1 + IVA_t\right)$, and $\lambda_t = \log\left(\Lambda_t\right)$. Next, approximate the nonlinear quantity $\log\left(1 + \left(\exp\left(mb_{t+1}\right) - 1\right)\exp\left(\lambda_{t+1}\right)\right)$ using a multivariate Taylor approximation around $mb_t = 0$ and $\lambda_t = \log(\rho)$ to obtain our approximate loglinear relation, where $\log\left(\rho\right)$ corresponds to the long-run average value of $\log\left(\frac{P_t}{D_t+P_t}\right)$ in Campbell and Shiller (1988). This value is an appropriate choice for the long-run average of λ_t, since it measures the log of the ratio of book value per share to cum-dividend book value per share, which we expect to equal the ratio of price to cum-dividend price in the long run, when price equals the book value per share. The approximation around $mb_t = 0$ moves the dividend-policy term to the approximation error, since a dollar inside the firm is just as valuable as outside around $mb_t = 0$. Crucially, introducing the iva variable limits this source of approximation error to dollars moved out of the firm through dividends as opposed to net repurchases.

Remark 3 ***(A loglinear present-value identity with investment).*** *In the presence of net issuance, a firm's log market-to-book ratio is approximately linear in its stock return, return on equity, investment value added, and log market-to-book ratio in the next period:*

$$mb_t \approx -r_{t+1} + roe_{t+1} + iva_{t+1} + \rho mb_{t+1}. \tag{9}$$

The approximation tends to be highly accurate: the time-series average of the yearly cross-sectional R^2s of the fit is 99.4%. Figure 1 visualizes this by plotting, for Apple and IBM stocks, the actual log annual returns with the approximated log returns defined as $\hat{r}_{t+1} = roe_{t+1} + iva_{t+1} + \rho mb_{t+1} - mb_t$. The small approximation error suggests that the decision to drop the payout-versus-plowback channel as an approximation error is relatively harmless. Appendix B.2 also shows that the payout-versus-plowback channel contributes relatively little to the cross-sectional variance in the market-to-book ratio and in return news.

Iterating forward and imposing $\lim_{j\to\infty} \rho^j mb_{t+j} = 0$, we decompose the log market-to-book ratio into contributions from future discount rates, returns on equity, and the value added from future value-enhancing investment in or divestment from book equity.

$$mb_t \approx -\sum_{j=1}^{\infty} \rho^{j-1} r_{t+j} + \sum_{j=1}^{\infty} \rho^{j-1} roe_{t+j} + \sum_{j=1}^{\infty} \rho^{j-1} iva_{t+j}. \tag{10}$$

Since the relationship holds ex-post, it also holds ex-ante in expectation.

Any firm decision that affects valuation—including a change in debt—is reflected in equation (10), with the exception of the earnings plowback/payout decision that is captured by the approximation error. For example, if the firm issues debt, the higher leverage could increase the profitability per unit of equity (roe). The higher leverage could also increase systematic risk and raise the discount rates, changing the r term, but it would not directly affect iva, as the transaction does not change the level of book equity.[9] The identity also remains valid whether or not book equity has

[9]Of course, future iva may change if the debt issuance signals that the firm will issue more equity in the future.

been adjusted for intangibles (see Appendix C).

Writing equation (10) from the perspective of time $t-1$ and taking a difference between its conditional expectations at $t-1$ and at t implies that news about a stock's time-t return, $N_{r,t}$, is due to news about the firm's current and future ROEs, investment value added, or discount rates. (From hereon, we drop the approximation sign for simplicity.)

$$\begin{aligned} N_{r,t} = r_t - E_{t-1} r_t = (E_t - E_{t-1}) \sum_{j=0}^{\infty} \rho^j roe_{t+j} + (E_t - E_{t-1}) \sum_{j=0}^{\infty} \rho^j iva_{t+j} \\ - (E_t - E_{t-1}) \sum_{j=1}^{\infty} \rho^j r_{t+j} = \underbrace{N_{roe,t} + N_{iva,t}}_{N_{CF,t}} - N_{DR,t}, \end{aligned} \tag{11}$$

where $N_{x,t}$ denotes the component of return news driven by variable x, and $N_{CF,t} = N_{roe,t} + N_{iva,t}$ and $N_{DR,t}$ respectively denote the composite cash-flow news and discount-rate news. This equation forms the basis of our analysis in Section 4, where we document the quantitative importance of the news about future investment value added. A detailed description of how to estimate these quantities from a VAR can be found in Campbell (1991).

Restating equation (10) to have long-horizon returns on one side, $\sum_{j=1}^{\infty} \rho^{j-1} r_{t+j} \approx -mb_t + \sum_{j=1}^{\infty} \rho^{j-1} roe_{t+j} + \sum_{j=1}^{\infty} \rho^{j-1} iva_{t+j}$, shows that anything that predicts long-horizon stock returns (left-hand side) must predict either long-horizon ROEs or long-horizon investment value added, unless its predictive power comes solely from its contemporaneous correlation with M/B. We use this decomposition in Section 5 to restate any long-horizon return predictability linked to a characteristic in terms of its ability to predict future ROEs and/or IVAs.

2.2 Comparison to Vuolteenaho (2002)

Setting investment value added (iva) to zero in all periods reduces equation (9) to Vuolteenaho's equation,

$$mb_t \approx -r_{t+1} + roe_{t+1} + \rho mb_{t+1}. \tag{12}$$

In this case, the long-run expression in equation (10) also simplifies to make future ROEs the sole determinant of infinite-horizon cash-flow fundamentals, leading to the following remark:

> "Whether one chooses to think about infinite-horizon cash-flow fundamentals in terms of dividend growth or ROE is a matter of taste, however" (Vuolteenaho (2002), p.235).

This argument has permeated large parts of the asset pricing literature. However, it only applies to *clean-surplus ROE*, which aggregates cash flows to equity from operations (conventional earnings) with those from net issuance, thereby conflating operating profitability with net payout policy.

While some authors are careful to use clean-surplus ROE in their applications, many do not.[10] Furthermore, following Chen and Zhao (2009)'s critique that VAR-based results depend on the choice of the state variables, many authors have explicitly modeled future cash flows using simple accounting ROEs either in the VAR framework or in another, exacerbating the issue. We show in Section 4 that more than one-quarter of cash-flow news is news about investment value added. As a consequence, equating simple ROE news with composite cash-flow news is far from innocuous.

2.3 Comparison to Campbell and Shiller (1988)

Our loglinear identity is closely related to that of Campbell and Shiller (1988) but expresses cash-flow fundamentals using future ROEs and IVAs rather than future dividends. Our approach has the advantage of relating a firm's valuation ratio directly to the two fundamental and distinct sources of cash flows and tends to generate empirically smaller approximation errors (Panel C of Table 1). The next remark summarizes the relation to the loglinear dividend-price ratio decomposition of Campbell and Shiller.

Remark 4 ***(Relation to the Campbell-Shiller decomposition).*** *Our model in* (9) *can be exactly*

[10]Note that this is not a critique of the practice to back out cash-flow news as the residual return-news component in a VAR that explicitly estimates discount-rate news. We emphasize that the argument—made by Chen and Zhao (2009)—that the choice of which news term to back out as the residual affects the return news decomposition, does not apply to close approximations like ours, the one in Campbell (1991), or the one in Vuolteenaho (2002), as long as one uses his approximation in conjunction with clean-surplus ROE.

restated as

$$p_t \approx -r_{t+1} + \log\left(1 + \exp\left(d_{t+1} - \log\left(BPS_{t+1}\right)\right)\right) + (1-\rho)\log\left(BPS_{t+1}\right) + \rho p_{t+1}. \qquad (13)$$

Approximating the nonlinear term $\log\left(1 + \exp\left(d_{t+1} - \log\left(BPS_{t+1}\right)\right)\right)$ *around* $d_{t+1} - \log\left(BPS_{t+1}\right) = \log\left(1/\rho - 1\right)$*, the long-run average log dividend-price ratio, leads to Campbell and Shiller (1988):*

$$pd_t \approx k - r_{t+1} + \Delta d_{t+1} + \rho pd_{t+1}, \qquad (14)$$

where pd *is the log price-dividend ratio,* Δd *is the log dividend growth, and* $k = -\log\rho - (1-\rho)\log(1/\rho - 1)$ *is a constant arising from the Taylor approximation.*

The constant, k, is missing from the identities in the market-to-book ratio. Loosely speaking, if book returns (that is, ROE) equal discount rates, book values equal market values such that the log market-to-book ratio is zero. This does not hold for the comparison of dividend growth and discount rates, such that the Campbell-Shiller identity in the price-dividend ratio includes additional terms collected in k. These additional terms are functions of the non-zero expansion point for the Taylor approximation, and therefore depend on the average log price-dividend ratio.

2.4 Comparison to Fama and French (2006)

Fama and French (2006) restate the dividend discount model as

$$\frac{M_t}{B_t} = \frac{\sum_{j=1}^{\infty} E_t\left(Y_{t+j} - \Delta B_{t+j}\right)/\left(1+R\right)^j}{B_t}. \qquad (3)$$

Based on this identity, they offer a motivation for the influential five-factor model (Fama and French (2015)): expected future level of earnings (Y_{t+j} or "profitabiliity") and increase in book equity (ΔB_{t+j} or "investment") have positive and negative ceteris-parius relations with stock returns (R), respectively. They also use the identity to clarify that the return predictability tied to any variable other than M/B, profitability, and investment must come from its ability to improve the forecasts of $\{Y_{t+j}\}$ and $\{\Delta B_{t+j}\}$.[11] Our identity has both practical and conceptual advantages in drawing out these types of insights.

First, equation (3) features nonstationary variables (earnings and changes in book equity) interacting in a nonlinear way, making it less suited for direct empirical implementation; as a result, the long-horizon analysis in Fama and French (2006) only considers horizons up to 3 years. In contrast, the linearity of our identity and the stationarity of the profitability and investment components of our market-to-book decomposition allows one to cleanly attribute the return predictability of any variable to its ability to forecast the profitability vs. investment components of future fundamentals. Furthermore, our identity allows expected returns to vary through time, eliminating the need to use the internal rate of return (R) to measure expected returns.

Second, interpreting Y as profitability and ΔB as investment could lead to conceptual issues. The negative ceteris-paribus relation between "investment" and returns requires holding the future *levels* of earnings fixed. Requiring more investment to reach the same level of earnings is, of course, value-reducing and must be offset by lower discount rates to hold M/B constant. However, controlling for future levels of (unscaled) earnings over a long horizon as one varies investment is a less intuitive exercise than controlling for scaled profitability measures such as roe (or gross profitability (Novy-Marx, 2013)). The negative ceteris-paribus relation may also give the appearance of defying the capital budgeting principle that increasing book equity could add *or* destroy value. In contrast, the investment variable in our expression of the same identity, iva, takes into account the valuation ratio at which investment or disinvestment through issuance and repurchases occurs and therefore produces an unambiguously *positive* ceteris-paribus relation between future iva and future returns.[12]

[11] "If variables not explicitly linked to this decomposition, such as *Size* and momentum, help forecast returns, they must do so by implicitly improving forecasts of profitability and investment, or by capturing horizon effects in the term structure of returns" (Fama and French 2015, p.2).

[12] IVA measures value-enhancing investment/divestment, and, consistent with our identity, we find that returns load positively on iva news, controlling for profitability news and discount-rate news. However, we do not necessarily expect a positive coefficient on iva when forecasting next-period returns in the VAR as the coefficient more broadly reflects the information in iva about future scale, yield, and valuation, as equation (19) in Section 5 shows must generally be the case for any variable forecasting returns. On a separate note, our iva variable achieves the unambiguously positive ceteris-paribus relation to returns by measuring investment in book equity *per share*. However, its construction and interpretation have little relation to Aharoni et al. (2013), who point out that the empirical implementation in

3 Data

We combine monthly stock price data from the Center for Research in Security Prices (CRSP) and annual accounting data from CRSP/Compustat Merged (CCM) to create a merged annual dataset for the period 1965–2022. We make a number of adjustments to the CRSP dataset before merging. We require all stocks to be domestically incorporated (CRSP share code of 10 or 11) and listed on one of the three major exchanges (CRSP exchange code 1 through 3). We include firms with valid market equity data for the current and prior month.

We convert the CRSP dataset from monthly to annual frequency by calculating annual compounded returns from the beginning of July in year $t-1$ to the end of June in year t. If a stock is delisted during the 12 months leading to June of year t, we use compounded returns until the time of delisting. However, to ensure that there is no bias to our results because of firms leaving the sample, we include delisting returns using the approach suggested by Shumway (1997) and Shumway and Warther (1999). We compute annualized dividends such that the annualized variables satisfy the basic return identity, $P_{t-1} = (P_t + D_t) / (1 + R_t)$. The appendix details how to use CRSP and Compustat to construct annualized variables that satisfy our exact nonlinear identity. Although rare (less than 1% of firms), there are instances in which one CCM firm has multiple common shares in the CRSP dataset. We aggregate all CRSP variables at the firm level before merging with CCM.

We follow the Fama-French convention to ensure that accounting data are publicly available for forecasts made the following June. Accordingly, we match the CCM observations in calendar year $t-1$ with CRSP data in June of year t and use the merged data to construct our final variables. We adjust the quantities for any changes in the number of shares between December and the following June. Since market equity observations as of the fiscal year end in December sometimes differ substantially between CRSP and Compustat (e.g., due to inconsistencies in the timing of stock splits or issuance), we adjust all Compustat variables by the ratio of the two market equity values to ensure that the ratio of an accounting value to a market value is not distorted by this data issue.

Fama and French (2006) must use change in total book equity rather than per-share book equity to be consistent with equation (3). Furthermore, like Fama and French (2006), Aharoni et al. (2013) are constrained by the nonstationarity of their variables and limit their analysis to horizons of up to two years.

To ensure that the natural logs of variables such as return or return on equity are well-defined for all firms, we follow Vuolteenaho (2002) to construct all variables for a composite portfolio, formed each June and investing 90% in the firm's equity and 10% in the one-month Treasury Bill, whose returns we accumulate over the subsequent year. Specifically, we compute the log book-to-market value of equity ratio (bm_t), return (r_t), return on equity (roe_t), and investment value added, computed as $iva_t = \log\left(\frac{D_t + BPS_t}{BPS_{t-1} + EPS_t}\right)$, for the composite portfolio as they are the main variables of interest. We also compute additional variables used in the analysis of the composite portfolio. We report summary statistics of the key variables of interest in Table 1. Cohen et al. (2003) and Lochstoer and Tetlock (2020) point out that portfolio formation can lead to differences between firm-level and portfolio-level variables. When we perform decompositions for portfolios (e.g., Section 4.1), we therefore construct portfolio-level variables that respect the present-value identity. See Appendix G for details.

For each firm, we further require valid observations for all relevant variables in years t and $t-1$. For the firm-level VAR, we exclude firms with book-to-market more than 100 or less than 1/100 in either year t or $t-1$, again in line with Vuolteenaho (2002). To further rule out that the VAR estimation is driven by microcaps, we exclude firms with market equity in the lowest decile of the NYSE size distribution at $t-1$. For portfolio-based exercises, such as our HML forecast in Section 6, we retain these firms in order to ensure consistency with the portfolio-based literature in cross-sectional asset pricing.

Following Cochrane (2005) and Campbell (2018), we set $\rho = 0.96$. In the original Campbell-Shiller approximation, the ρ parameter is a function of the average dividend-price ratio. We note that our key results and interpretations are not sensitive to the exact choice of this parameter. One factor that has affected the calibration of ρ over time is the widespread use of share issuance and repurchases. Our new approximation (9), which features the iva term, therefore improves upon the return decomposition of Vuolteenaho (2002) by explicitly accounting for this increasingly prevalent corporate behavior. We now turn to the core of our empirical analysis.

4 Profitability vs. investment in valuations and returns

How much of the variation in the cash-flow component of firm values is driven by profitability vs. scale? Our first application takes a step further than Vuolteenaho (2002) and Cohen et al. (2003) to decompose the cross-sectional variation in the market-to-book ratio and in unexpected stock returns into the three components implied by, respectively, equations (10) and (11).

4.1 Valuations

Present-value identities imply that a high valuation ratio must either forecast low returns or high cash-flow growth (Cochrane, 2008). Asset pricing theory has often sought to combine the two elements by linking the value premium to differences in cash-flow duration (e.g., Campbell and Vuolteenaho (2004), Lettau and Wachter (2007), or more recently Gormsen and Lazarus (2022)).

In recent years, a low realized value premium has led to a lively discussion among academics and practitioners about the 'death of value' (e.g., Arnott et al., 2021). Just like many of the theoretical (e.g., duration-based) explanations for a positive expected value premium, the negative realization suggests a particularly large cash-flow growth differential between high and low M/B firms. Yet, the empirical evidence for this growth differential appears mixed (e.g., Chen (2017) finds that dividend growth of value firms actually *outpaces* that of growth firms).

Rather than looking for growth differentials or return spreads in isolation, equation (10) lets us address both questions in a coherent manner and decompose the cross-sectional variation in market-to-book ratios into expected returns, expected profitability (roe), and expected changes in scale (iva). To see this, take the cross-sectional covariance of both sides of (10) with mb_t and divide both sides by the variance of mb_t:

$$1 \approx -\frac{Cov(\sum_{j=1}^{\infty} \rho^{j-1} r_{t+j}, mb_t)}{Var(mb_t)} + \frac{Cov(\sum_{j=1}^{\infty} \rho^{j-1} roe_{t+j}, mb_t)}{Var(mb_t)} + \frac{Cov(\sum_{j=1}^{\infty} \rho^{j-1} iva_{t+j}, mb_t)}{Var(mb_t)}. \tag{15}$$

In doing so, we refine the exercise of Cohen et al. (2003), who use Vuolteenaho's approximation (2) and clean-surplus ROE to arrive at a composite cash-flow component.

We compute the discounted sums using realized variables over 5, 10, and 15-year horizons. In order to deal with delisting, we form 25 portfolios, double-sorted by size and book-to-market quintiles, and compute the discounted sum of the relevant variables at the portfolio level using a discount factor $\rho = 0.96$. We then regress these sums, as well as the discounted market-to-book ratio on the portfolio's current market-to-book ratio (with year fixed effects). We estimate

$$\sum_{j}^{J} \rho^{j-1} y_{i,t+j} = \alpha_t + \beta \, mb_{i,t} + \varepsilon_{i,t}, \tag{16}$$

for $y \in \{roe, iva, r\}$ so that β estimates each of the three components of the decomposition in equation (15).

Table 2 shows that future roe accounts for around 40% of market-to-book variation, iva for around 17%, and future returns for up to 13%. Since the sums are truncated, the future market-to-book ratio is responsible for a large share of the cross-sectional variation in current market-to-book ratios, around 23% at the fifteen-year horizon.[13]

In line with the findings of Cohen et al. (2003), but seemingly contrary to those of Chen (2017), growth firms do have higher realized cash-flow growth driven by both profitability (roe) and investment (iva), justifying their higher valuations. Over the whole sample, the contribution of higher future profitability outweighs that of future changes in firm scale. An important distinction between our results and some of those in Chen (2017) is that we consider firm fundamentals (earnings) rather than dividends as the relevant cash-flow measure, and do so over a longer horizon in growth realizations (fifteen years).

Panel B reports the results from the same exercise, using VAR-implied, firm-level, infinite-horizon expectations, i.e. $E_t \sum_j^J \rho^{j-1} x_{i,t+j}$, on the left-hand side of regression (16). We describe the underlying VAR in the following subsection. Relative to using realized quantities at truncated horizons, the VAR-based results attribute 17% of variation to iva expectations (slightly smaller than in Panel A, given the lack of truncation), and larger shares to discount rates (29%) and prof-

[13] Hence, $1 \approx -\frac{Cov(\sum_{j=1}^{J} \rho^{j-1} r_{t+j}, mb_t)}{Var(mb_t)} + \frac{Cov(\sum_{j=1}^{J} \rho^{j-1} roe_{t+j}, mb_t)}{Var(mb_t)} + \frac{Cov(\sum_{j=1}^{J} \rho^{j-1} iva_{t+j}, mb_t)}{Var(mb_t)} + \frac{Cov(\rho^{J} mb_{t+J}, mb_t)}{Var(mb_t)}$ is the decomposition we work with.

itability (48%). The remainder is due to the impact of future payout decisions, i.e., the approximation error incurred by our new identity (1) (see Appendix B.2).[14]

Figure 2 plots the decomposition over time and shows that the relative contributions of profitability and investment have changed remarkably over time and flipped in recent years. Per-share cash-flow growth in high market-to-book firms is increasingly driven by equity issuance at high valuations and the resulting increases in scale per share. Over the whole sample, we confirm the well-established result that value firms have higher returns, although with some important time-variation in the magnitude of this result, e.g., around the build-up of the dotcom bubble and its subsequent crash.

The literature has proposed various ad hoc market-to-book decompositions that improve return predictability. For instance, Gerakos and Linnainmaa (2018) find that the component of market-to-book that correlates with recent increases in size (market capitalization) is a better return predictor than its orthogonal complement component. In Table E.11, we report the results of an exercise analogous to that in Table 2. Indeed, we find that discount-rate variation is a more important driver of the size-related component of market-to-book than it is of the orthogonal component. We further find that these differences actually persist over longer return horizons. The orthogonal component instead predominantly reflects future profitability. Interestingly, the size-related component identified by Gerakos and Linnainmaa (2018) also predicts variation in the iva-component of cash flows much more strongly than the orthogonal part of the valuation ratio. This finding once again points to the importance of accounting for both sources of cash flows when examining return predictability in the context of cash-flow predictability.

In Appendix C, we adjust the identity to include intangible capital in market-to-book. We find that over the one-year horizon, the resulting metric is less cross-sectionally correlated with adjusted profitability and iva. In turn, the adjusted market-to-book ratio is (and must be) a better predictor of cross-sectional return differences, thus supporting the findings of Eisfeldt et al. (2021) and

[14]The approaches in Panels A and B differ along various dimensions, which limit the precision of quantitative comparisons: for instance, Panel B decomposes variation across firms rather than across portfolios, estimates variation in infinite-horizon quantities rather than observing finite-horizon realizations, and it excludes microcaps. We confirm, however, that allowing the VAR coefficients to vary by size quintile leaves the results in Panel B essentially unchanged.

embedding them within a comprehensive identity-based framework of the market-to-book ratio.

4.2 Return news

4.2.1 Firm level

To understand the drivers of stock return news, we run the following parsimonious vector autoregression (VAR) in the panel of firm-year observations:

$$z_{i,t+1} = \mu_{t+1} + Bz_{i,t} + u_{i,t+1}, \tag{17}$$

where $z_{i,t} = [r_{i,t}, bm_{i,t}, iva_{i,t}, roe_{i,t}]$. To treat each yearly cross-section equally, we estimate the system using weighted-least-squares regressions that deflate each firm-year observation by the number of firms in the year as in Vuolteenaho (2002) and Lochstoer and Tetlock (2020). As our VAR includes year fixed effects, the firm-level analysis—including that on return news—should be interpreted as purely cross-sectional.

Panel A of Table 3 uses the estimated system to decompose the return news variance as in equation (11).[15] Among the two cash-flow news components, the ratio of the contributions is roughly 3:1 between profitability news and investment news. Lochstoer and Tetlock (2020) find a correlation between firm-level cash-flow news and discount-rate news of -0.42. In our sample and specification, this correlation is -0.46, and our decomposition shows that it is negative for *both* cash-flow components: A positive shock to either iva-news or roe-news is associated with lower subsequent returns. Consistent with the results in Vuolteenaho (2002), the variance contribution from discount-rate news is relatively small. These results are based on the direct estimate of all three news terms and do not rely on backing out one as a residual from the others, although doing so generates similar results thanks to the accuracy of our approximation.[16] Importantly, we find that the correlation between the two components of cash-flow news is negative, meaning that news

[15]We report the estimated coefficient matrix B in Panel A of Table E.1.

[16]In Appendix, Table E.6, we also calculate decompositions where we back out each of the cash-flow news terms as a residual from the present-value relationship implied by equation (11). Those results closely resemble those in Table 3.A, confirming that these estimates are not meaningfully affected by the approximation error.

about higher future investment tend to coincide with lower expectations of future profitability. This observation based on our novel decomposition points to a richer set of cash-flow dynamics than those contemplated by many models of the value premium which rely on a single source of risk. Following Vuolteenaho (2002), Table E.5 also reports the variance decomposition by size quintile: all three covariances between N_{DR}. N_{roe}, and N_{iva} are negative in each size quintile; the share of iva news in total news variance is flat at 17–19%.

As a separate but related quantification of how much each news component contributes to the cross-sectional variance of unexpected stock returns, we regress the VAR-implied return news $r_{t+1} - E_t r_{t+1} = N_{roe,t+1} + N_{iva,t+1} - N_{DR,t+1}$ on combinations of its components, with results reported in Table 4. Our interest in this exercise is in the R^2; we report both the R^2 from the regression and the resulting R^2 when the coefficients are constrained to their theoretical values. The first three columns each omit one of the components, and the R^2 only comes close to one in column (4) with all three regressors. The last two columns decompose cash-flow news, $N_{CF} = N_{roe}+N_{iva}$, into its profitability and investment components. The R^2-differential—in line with the numbers in Table 3 Panel A—suggests that investment news accounts for roughly one quarter of the volatility of cash-flow news.

We assess the cross-sectional heterogeneity of the investment news contribution by running the same exercise within characteristic quintiles. We find that the contribution of investment news to cash-flow news variance is largest among growth firms (bottom book-to-market quintile) and low-profitability firms (bottom ROE quintile) at around 30% . In comparison, the variance contribution among value firms and high-roe firms is similar to that of the overall sample at around 25%. While iva generally captures the present value of a firm's scale decisions, this heterogeneity result suggests that the rise in the iva-share of cross-sectional M/B-dispersion reflects the valuation of growth options in highly priced, but not necessarily highly profitable, firms. We do not find a meaningful difference between the top and bottom size quintiles, indicating that the importance of our novel cash-flow component is not concentrated in very small firms. We note that the relative contribution of investment news to the variance of cash-flow news has increased in recent years–from under 20% in 1965–1989 to around 30% since 1990.

To assess the robustness of these VAR-based result, we compare news decompositions across subsamples. Table E.4 reports analogs of Table 3 for the subsamples in which we find iva news to account for a larger share of cash-flow news. In each case, we compute news decompositions based on (i) the transition matrix of the baseline VAR and (ii) the transition matrix estimated only within the respective subsample. In each case, we find very similar results, suggesting that our key takeaways from the negative news correlations are robust to VAR specification, and that differences across subsamples arise from different shocks to different firms or at different times, rather than from variation of our VAR-coefficient estimates across samples. In Appendix F, we further augment the baseline VAR for the full sample by adding log asset growth and additional lags to the state variables, and find consistent results (Table E.2).

4.2.2 Portfolio level

How does the above split of cash-flow news into profitability and investment components look for diversified portfolios? We run similar variance decompositions for return news on the aggregate stock market and the mean-variance efficient portfolio of the market and popular "anomaly" strategies.

We run a time-series VAR for the market portfolio. Following Campbell and Vuolteenaho (2004) and Campbell et al. (2018), we include additional state variables in the annual VAR system beyond those variables required by our loglinear model of the book-to-market ratio: (i) the term yield spread (TY) between ten-year and one-year treasury yields, (ii) the default yield spread (DEF) between between corporate bonds rated Baa and Aaa by Moody's, and (iii) the small stock value spread (VS), that is, the difference in log book-to-market between small stocks in the top and bottom terciles from Kenneth French's website.

We report the aggregate return news variance decomposition in Table 3 Panel B.[17] In line with the findings of a large literature, aggregate return variance is predominantly driven by discount-rate news. Within cash-flow news, the contribution from investment news is slightly lower, but on a similar order of magnitude as that of aggregate profitability news.

[17]The VAR coefficient matrix and residual variance-covariance matrix are shown in in Panels A and B of Table E.3

We then decompose the return news of five well-known anomalies, using long-short quintile portfolios (high minus low) sorted on value (book-to-market), size (market equity), profitability (ROE), investment (asset growth), and the momentum characteristic. We follow Lochstoer and Tetlock (2020) to compute the variance decomposition for the mean-variance efficient combination of the five anomaly portfolios and the market. For the purposes of comparability, we use the "cum market" MVE portfolio weights of Lochstoer and Tetlock (2020): 0.80, -0.21, -0.35, 0.73, -1.87, and 1.35 for the market, value, size, profitability, investment, and momentum, respectively.

The results for the MVE portfolio in Table 3 Panel C are similar to Lochstoer and Tetlock's regarding the split between discount-rate and cash-flow news, with the latter accounting for the majority of the variation. Additionally, we find that—lower than in the cross-section, but higher than in the aggregate time-series—around one tenth of the variation in the mean-variance efficient portfolio return stems from variation in investment news.

While discount-rate news predominates in market returns, systematic cash-flow news drives the returns of anomaly portfolios (in line with Lochstoer and Tetlock (2020)). In most cases, profitability news variance accounts for roughly half of total return variation. The individual variance contribution from investment news is smaller, but exceeds one fifth for the profitability anomaly. An important insight from this analysis based on our novel decomposition is the correlation between cash-flow news terms and discount-rate news: in line with the results of Lochstoer and Tetlock (2020), cash-flow news is negatively correlated with discount-rate news, and this negative correlation arises through both sources of cash-flow news, but with a larger magnitude of the correlation for investment-driven cash-flow news at the firm level and in the MVE portfolio. As for individual firms, profitability news and investment news are negatively correlated for the MVE portfolio, once again pointing to multidimensional sources of priced cash-flow risk. We decompose the return news variance for the individual anomalies (Table E.7). All anomalies exhibit falling expected returns when expected cash flows of the anomaly portfolio rise, regardless of whether this is driven by profitability or investment news. This covariance component contributes around 25% of the total anomaly return news variances. If expressed in terms of the composite cash-flow news measure as in Lochstoer and Tetlock (2020), rather than profitability, this channel of variance contribution appears smaller, particularly for the profitability and investment anomalies.

4.3 Implications for cross-sectional asset pricing

Decomposing the cash-flow component in the present-value identity into profitability and investment has produced three key findings: (i) profitability news and investment news both contribute importantly to firm-level return news variance and are negatively correlated (Table 3), (ii) Lochstoer and Tetlock's (2020) finding that cash-flow news correlate negatively with discount-rate news arises through both cash-flow channels (Table 3), and (iii) The investment component accounts for an increasing share of market-to-book dispersion (Figure 2).

The first challenges theories of the value premium that rely on productivity shocks as a single source of cash-flow risk (e.g., Zhang, 2005). Quantitatively large iva news (18% of return news variance) require a high sensitivity of investment to productivity shocks. Critically, capital expansion at market-to-book above one in response to a positive future productivity shock generates a *positive* correlation between iva news and roe news, contrary to what we find.[18] Instead, theories that contemplate shocks to the arrival of investment opportunities separately from shocks to overall profitability are better suited to match this new stylized fact (e.g., Belo et al., 2014, 2019; Kogan and Papanikolaou, 2014). These models are not only more likely to generate quantitatively important iva news, but also allow in principle for a negative correlation with roe news if shocks that lower the cost of future investment expand equity capital (positive iva news) may lower the profitability of the average unit of capital (negative roe news).

The second finding provides a more granular perspective on Lochstoer and Tetlock's (2020) finding that cash-flow news and discount-rate news are negatively correlated at the firm/portfolio level, which they argue are consistent with "behavioral models in which investors overreact to news about firms' long-run CFs (e.g., Daniel et al., 2001) and risk-based models in which firm risk

[18] Models of this type do not distinguish between debt and equity financing, but the distinction is less relevant if firms pursue a leverage target. Theoretically, the negative correlation could come from extreme value firms with a market-to-book ratio below one shrinking their book equity (resulting in a positive iva shock) in response to a negative profitability shock. However, we find that the negative correlation is not just a feature of the average firm, but also robustly present among growth firms (see Table E.4).

increases after negative news about long-run CFs (e.g., Kogan and Papanikolaou, 2013)" (p.1420). To the extent that overreaction is responsible for the negative correlation between cash-flow news and discount-rate news, our finding implies that investors must be overreacting to news about a firm's future investment opportunities as well as news about its future profitability. Furthermore, our finding presents a more granular challenge to the risk-based explanation, since the existing literature on profitability and investment (broadly defined) in cross-sectional asset pricing typically assigns risk premia of opposite signs to the two characteristics.

In Kogan and Papanikolaou (2013) (KP) and Belo et al. (2019), for example, a favorable investment-cost shock makes firms less risky, generating a negative correlation between iva news (N_{iva}) and discount-rate news (N_{DR}) as in the data. However, the same shock makes the firm less profitable, pushing the correlation between roe news (N_{roe}) and N_{DR} upward. On the other hand, for tractability, KP uses a knife-edge restriction that makes an aggregate *productivity* shock have no effect on firm risk. As a result, the net correlation between N_{roe} and N_{DR} is counterfactually positive. Relaxing the knife-edge restriction to allow value functions to depend nonlinearly on aggregate productivity, as done in Kogan et al. (2022), can restore a negative N_{roe}-N_{DR} correlation. In Belo et al. (2019), a positive (idiosyncratic) productivity shock makes the firm less risky, contributing to a negative correlation between N_{roe} and N_{DR}. However, the same shock increases investment and could make the correlation between N_{roe} and N_{iva} counterfactually positive. Overall, generating the three pairwise negative correlations among N_{roe}, N_{iva}, and N_{DR} to make the model more realistic requires a careful calibration of the relative volatilities of investment-specific shocks and productivity shocks and could help discipline the model further.

The last result goes to the age-old question of what makes growth firms expensive. Traditionally, growth firms have been more profitable than value firms (and generally had lower discount rates). The emergence of highly-valued yet less profitable firms, for instance in the tech sector, has challenged this view. One potential explanation is that these firms are simply overpriced and accordingly will have low return realizations. Another is that their profitability will rise sharply in the future to justify high ex ante valuations. Our new variable, iva, delineates a third explanation: it codifies how expectations of a rapid future (profitability-neutral) expansion raise expected shareholder cash flows and a firm's present value. Figure 2 suggests that this third explanation has carried an increasing weight in rationalizing valuation dispersion across firms, particularly since the mid 1990s. As an aside, iva as a valuation component also provides a useful framework to think about the tangible value of firms timing their own stock market (Baker and Wurgler, 2002), or—in an extreme form—of being a "meme stock", that is, the ability to issue new equity at temporarily inflated share prices.

5 Sources of cross-sectional return predictability

Our second application of the identity shows that for every empirical pattern of return predictability, there must be a predictability pattern for cash-flow fundamentals (roe and iva) and/or the market-to-book ratio. The strength of this identity-based approach to return predictability is that any internally consistent theoretical model or explanation of return predictability—be it risk-based or behavioral—must produce *some* predictability pattern in roe, iva, or market-to-book.

To see this, rewrite (9) to have one-period return on the left-hand side:

$$r_{t+1} = roe_{t+1} + iva_{t+1} + \rho mb_{t+1} - mb_t. \tag{18}$$

Hence, any short-horizon return predictor x_t must also predict roe, iva, or a change in mb. To be more concrete, on both sides, take a cross-sectional covariance with x_t and divide by the cross-sectional variance of x_t to get

$$\frac{cov(r_{t+1}, x_t)}{var(x_t)} = \frac{cov(roe_{t+1}, x_t)}{var(x_t)} + \frac{cov(iva_{t+1}, x_t)}{var(x_t)} + \frac{cov(\rho mb_{t+1} - mb_t, x_t)}{var(x_t)}, \tag{19}$$

which states that the univariate predictive coefficient of r_{t+1} on x_t equals the sum of the univariate predictive coefficients on the right-hand side. A similar logic applies to a multivariate setting with several return predictors as well as to predicting longer-horizon returns, for which the present-value identity in equation (10) rearranges to

$$\sum_{j=1}^{J}\rho^{J-1}r_{t+j}=\sum_{j=1}^{J}\rho^{J-1}roe_{t+j}+\sum_{j=1}^{J}\rho^{J-1}iva_{t+j}+\rho^{J}mb_{t+J}-mb_{t}. \tag{20}$$

That is, a predictor of long-horizon returns must predict roe, iva, or a long-horizon change in mb.

To illustrate our approach, we characterize in what sense, and over what time-frame, a known cross-sectional return predictor forecasts these cash-flow drivers of returns as delineated by our identity.

5.1 An anatomy of return predictability

We run cross-sectional bivariate regressions of the future realizations of each variable in equation (20) on several potential predictors of returns (each time controlling for market-to-book):

$$\sum_{j=1}^{J}\rho^{j-1}y_{i,t+j}=\alpha_t+\beta_1\,mb_{i,t}+\beta_2\,x_{i,t}+\varepsilon_{i,t}, \tag{21}$$

where $y\in\{r,roe,iva\}$ and $J\in\{1,5,15\}$ years.[19] As for the predictor x, we examine asset growth (ag), return on equity (roe), and cash-flow duration (Dur). We also report the results for using $x=iva$, as we reference the result in Section 6. These regressions use realized values of y rather than VAR-implied expectations.

Table 5 reports the estimated $\widehat{\beta}_2$ for different combinations of the choice of y, x and J. Looking at return prediction first ($y=r$), longer-horizon return predictability tends to be stronger than one-year return predictability, although the sign of the effect is in general consistent with the literature on monthly return predictability. High-profitability firms have higher returns, whereas high-asset-growth and high-duration firms have (insignificantly) lower returns. Our new variable iva is associated with lower subsequent returns, with stronger results than for asset growth, the more common investment-based predictor, particularly at medium to long horizons.

[19] We use our identity variables plus asset growth (the most common investment-based predictor) and duration (e.g., Weber, 2018; Gormsen and Lazarus, 2022), but this analysis can be done for any arbitrary return predictor.

Next, we study the mirror image of return predictability in cash flows, that is, the right-hand side of Equation (20). Starting with asset growth (first row of Table 5), the top-right panel of Figure 3 illustrates that, controlling for current market-to-book, asset growth (weakly) predicts lower returns which accumulate up until a horizon of seven to eight years, and then revert partially. This result is complementary to the short-run predictability results documented in several other papers (e.g., Fama and French (2015)). In the context of equation (20), the negative return realization is composed of steadily falling shareholder cash flows: profitability per unit of equity capital (roe) declines significantly and persistently, and this decline is only partially offset by steadily rising book equity per share (iva). But while the cash-flow decline is steady and persistent, the return realization is not. The valuation ratio of firms with initially high asset growth declines slightly over short-to-medium horizons (adding to negative return realizations) followed by a steady rise, thus halting the decline in cumulative returns.

Of course, the directional finding that asset growth negatively predicts profitability is not new. Fama and French (2006) show that asset growth negatively predicts short-run profitability once they control for other characteristics such as market-to-book. Similarly, an extensive literature in accounting has shown that growth in accruals and other net operating assets negatively predicts return on assets (e.g., Fairfield et al., 2003). What is novel is that the decline in fundamentals (profitability) associated with today's asset growth is persistent while the decline in returns and the market-to-book ratio is temporary and reverts afterwards; plausible explanations of the "investment anomaly" need to match these patterns in fundamentals and the valuation ratio. For instance, a model positing a positive unconditional relation between profitability and risk premia and a negative unconditional relation between market-to-book and risk premia should be able to explain why low profitability and high market-to-book over a long horizon *conditional* on initially high asset growth, do not translate into low returns over the same long horizon.

For comparison with the asset growth case, we visualize the return predictability using profitability (roe) and its anatomy. The middle panels of Figure 3 suggest that the cross-sectional return prediction from roe is more persistent than that from asset growth (and strongly significant at medium-to-long horizons). The higher returns on high-roe firms manifest themselves through a profitability-driven rise in long-run cash flows per share (roe). Again, the investment-based component of these long-run cash flows (iva) moves in the opposite direction, with negative long-run iva attenuating the effect of rising profitability. Market-to-book rises initially but then falls and offsets some of this rise in cash flows over longer horizons.

Lastly, consider the cross-sectional return predictions from cash-flow duration (Weber, 2018) (bottom row of Table 5; bottom panels of Figure 3). High-duration firms tend to earn lower returns at all horizons (up to 15 years), consistent with the short-run return predictability documented by Weber. They tend to have higher future roe and lower iva. However, none of these patterns are statistically significant in our sample.

What do we learn from these anatomy exercises? We see two common features: (i) Predicted cross-sectional differences in cash flows tend to be larger and more persistent than the predicted differences in returns. (ii) Predicted differences in the two cash-flow components go in opposite directions. The first validates using cash-flow fundamentals as a starting point to explain return predictability: return predictors proxy for cash-flow differences which are either associated with risk, or subject to investor biases. The second finding highlights the importance of understanding scale-based, profitability-neutral cash-flow drivers captured by iva separately from profitability.

Concretely, the first two anatomies (asset growth and roe) are consistent with the IST-view of the investment discount and the profitability premium (e.g., Kogan and Papanikolaou, 2013), as long as growth options are less profitable than assets in place: high-investment firms have lower ROE, gradually realize growth options (high future iva, low future roe), and earn lower risk premia.[20] However, while the findings of Gormsen and Lazarus (2022) suggest that these anomalies share a common duration-based explanation, the anatomy of the duration-based predictor (constructed following Weber (2018)) is hard to square with the other two: high-duration firms earn lower returns, but have *higher* future profitability and *lower* iva. The duration-coefficients for roe and iva are indistinguishable from zero, but they are statistically different from the ones for asset growth and profitability. Duration-based return predictability is often seen as a monolith, irre-

[20] The formulation of this view in Kogan and Papanikolaou (2013), however, fails to match the negative correlation of roe news and discount-rate news (see Section 4 for more details and discussion).

spective of whether duration is extrapolated from historical cash flows (Weber, 2018) or obtained from dividend strips (Gormsen and Lazarus, 2022), but Figure 3 points to important distinctions between the results unveiled via the different methodologies.

5.2 Sorting on investment value added

Incidentally, the information in cross-sectional variation in iva is not subsumed by asset growth. While iva negatively predicts returns, Table 6 shows that a long-short portfolio based on iva generates positive *abnormal* returns, controlling for the five Fama and French (2015) factors and momentum (Carhart, 1997). The sign-flip between returns and alphas is driven by large negative loadings on the CMA and RMW factors: while the high-minus-low-iva portfolio correlates strongly with high-investment and low-profitability firms, its returns are less negative than those implied by these factor loadings. Double sorts on iva and book-to-market, operating profitability, or size reveal that this pattern is particularly strong among growth firms and large firms and does not vary with firm profitability. The purpose of this exercise is to show that the role of the iva variable—derived from a rigorous present-value identity—in the VAR or other forecasting settings does not simply re-package known relationships between investment-based variables and returns or other fundamentals. As a result, this variable also highlights a particular part of the cross-section of average returns that is not adequately explained by the preeminent existing factors.

6 Forecasting the time series of HML returns

The final application is the *time-series* forecast of value-minus-growth (HML) portfolio returns. Applying the long-run identity in equation (10) to the HML portfolio, we obtain an adapted version of the equation Cohen et al. (2003) use to motivate forecasting HML returns with the value spread:

$$\underbrace{\sum_{j=1}^{\infty} \rho^{j-1} r_{t+j}^{HML}}_{\equiv HML_t^{LR}} \approx \underbrace{(bm_t^H - bm_t^L)}_{\equiv VS_t} + \underbrace{\sum_{j=1}^{\infty} \rho^{j-1}(roe_{t+j}^H - roe_{t+j}^L) + \sum_{j=1}^{\infty} \rho^{j-1}(iva_{t+j}^H - iva_{t+j}^L)}_{\equiv CF_t^{LR}} \quad (22)$$

where HML_t^{LR}, VS_t, and CF_t^{LR} respectively denote the spread in long-run returns, in book-to-market ("value spread"), and in long-run cash flows between value and growth portfolios. The equation shows that the value spread should predict either future HML returns or/and future spread in cash flows, and Cohen et al. found that the value spread does predict (short-run) HML returns with a significant positive coefficient in a pre-2000 sample.

In the last two decades, however, the value spread has grown while HML returns have been low. This pattern does not show up in Figure 2, since long-run cash flows and returns associated with these recent market-to-book ratios have not yet materialized. Hence, to understand the value spread in the more recent period, we conduct a cross-sectional analog to the exercise in Cochrane (2008), extrapolating long-run forecasting coefficients from the corresponding short-run regressions as well as the persistence of the market-to-book ratio.[21] The results in Figure 4 show that in recent years, short- or long-run discount-rate variation is no longer reflected in the market-to-book variation, leaving expected long-run cash flows to drive the entire variation in the market-to-book. This is consistent with the value spread not lining up well with HML returns in recent years.

Given this background, what else does equation (22) say about timing HML returns? First, predictors of long-run cash flows forecast HML returns, even when the value spread does not:[22]

$$Cov(HML_t^{LR}, x_t \mid VS_t) = Cov(CF_t^{LR}, x_t \mid VS_t). \tag{23}$$

That is, controlling for the value spread, a characteristic x_t that best predicts long-run HML returns is one that best forecasts long-run cash flows. Second, such a cash-flow-based predictor x_t may help restore part of the value spread's predictive power:[23]

$$\begin{aligned} Cov(HML_t^{LR}, VS_t \mid x_t) = & Var(HML_t^{LR} \mid x_t) - Var(CF_t^{LR} \mid x_t) \\ & + Cov(CF_t^{LR}, -VS_t \mid x_t) \end{aligned} \tag{24}$$

[21] See Appendix D.1 for further details on the exercise.

[22] Taking a covariance of both sides of equation (22) with x_t conditional on VS_t, $Cov(HML_t^{LR}, x_t \mid VS_t) = Cov(VS_t + CF_t^{LR}, x_t \mid VS_t) = Cov(CF_t^{LR}, x_t \mid VS_t)$.

[23] From equation (22), $Cov(HML_t^{LR}, VS_t \mid x_t) = Cov(HML_t^{LR}, HML_t^{LR} - CF_t^{LR} \mid x_t) = Var(HML_t^{LR} \mid x_t) - Cov(HML_t^{LR}, CF_t^{LR} \mid x_t) = Var(HML_t^{LR} \mid x_t) - Var(CF_t^{LR} \mid x_t) + Cov(CF_t^{LR}, -VS_t \mid x_t)$.

That is, controlling for x_t helps VS_t predict HML returns better, particularly if x_t captures variation in future cash flows that correlates with the value spread ($Cov(CF_t^{LR}, -VS_t \mid x_t)$ is large), holding fixed the overall predictability of long-run HML returns, i.e., $Var(HML_t^{LR} \mid x_t)$.

Precisely for these reasons, the original time-series forecast model of Cohen et al. included clean-surplus (CS) ROE spread as a cash-flow predictor. However, because their work found a point estimate on the CS-ROE spread that was statistically indistinguishable from zero, subsequent work on time-series return predictability largely ignored cash-flow controls, focusing only on the value spread to time HML returns.

We find that including cash-flow controls is now critical for timing HML returns. Table 7 shows that cash-flow controls such as the profitability spread (spread in roe between value and growth portfolios) and iva spread (corresponding spread in iva) not only predict HML directly, but also restores part of the value spread's predictive power. The iva spread seems particularly important, almost doubling the t-statistic on the value spread, turning the profitability spread into a significant predictor, and noticeably raising the R^2.

The iva spread helps time HML returns mainly because it negatively predicts the roe component of future cash flows (Table 5). In fact, this not only makes the iva spread a significant (negative) return predictor, but also enhances the explanatory power of both the value spread and the profitability spread. When the value spread increases because value firms have a particularly poor long-run profitability outlook compared to growth firms, the iva spread between value and growth firms widens to reveal this expectation, leaving the residual movement in the value spread to reflect changes in the discount rate. While a rise in current profitability (roe) spread between value and growth firms could mean a rise in future profitability spread or a fall in future iva spread, the former effect gets captured by a fall in current iva spread, leaving the residual rise in profitability spread to forecast a *lower* future cash flow spread and lower long-run HML returns.[24]

Appendix D.2 shows that the findings of Table 7 also apply to out-of-sample forecasts. In the

[24] At first sight, the negative sign on the profitability spread in Table 7 is puzzling in light of equation (22). However, the equation says that future HML returns must have a positive ceteris-paribus relation to the *future* long-run spreads in profitability (roe) and iva and is silent about how *current* profitability relates to future HML returns, given it predicts one cash-flow component (roe) positively, and the other (iva) negatively (Table 5).

years following the CPV sample, cash-flow and discount-rate shares in cross-sectional market-to-book dispersion have shifted. As a result, the value spread alone no longer predicts value-minus-growth returns. Once supplemented with a successful set of predictors of long-run cash flows—such as the iva spread and profitability spread—forecasting power improves.

Finally, the cash-flow predictor x_t in equations (23) and (24) could be any characteristic that forecasts future cash flows, not just the profitability spread or the iva spread. Appendix D.3 shows that the issuer-repurchaser spread of Greenwood and Hanson (2012) is a good alternative cash-flow predictor with which to supplement the value spread in timing HML returns and does this by forecasting lower long-run profitability. The same appendix section analyzes other HML predictors proposed in the literature.[25] We find some predictability of monthly HML returns from the lagged HML return (Ilmanen et al., 2021) and the inverse of lagged three-year HML-volatility (Moreira and Muir, 2017). These predictors do not interact with the value spread, indicating that their power arises from variation in the *term structure* of HML returns, rather than variation in the level of cumulative long-run returns.

7 Conclusion

We derive a loglinear present-value identity that links today's market-to-book equity ratio to future investment ("scale"), profitability ("yield"), and discount rates. By explicitly allowing net investment in book equity (through net issuance) to influence the firm's present value, our identity departs from Vuolteenaho's (2002) framework and helps quantify the relative importance of scale vs. yield in firm valuations and returns. It also allows empiricists to relate stock return predictability to the predictability of future profitability and investment at any horizon, generating new insights into return predictability.

In sum, we offer a potentially indispensable tool not only for empiricists wishing to study stock prices and returns through the lens of an identity but also for theorists seeking to match the joint

[25] While several papers propose factor-timing strategies, the overall evidence is mixed. Ilmanen et al. (2021) conclude that there is only "modest predictability that likely fails to overcome implementation frictions."

term structures of risk premia and cash-flow fundamentals. Indeed, linking long-horizon returns to both long-horizon profitability and book equity investment would help facilitate the analysis of stock price levels.

References

Aharoni, G., B. Grundy, and Q. Zeng (2013). Stock returns and the Miller Modigliani valuation formula: Revisiting the Fama French analysis. *Journal of Financial Economics 110*(2), 347–357.

Arnott, R. D., C. R. Harvey, V. Kalesnik, and J. T. Linnainmaa (2021). Reports of value's death may be greatly exaggerated. *Financial Analysts Journal 77*(1), 44–67.

Asness, C. S., J. A. Friedman, R. J. Krail, and J. M. Liew (2000). Style timing: Value versus Growth. *Journal of Portfolio Management 26*(3), 50–60.

Baker, M. and J. Wurgler (2002). Market timing and capital structure. *The Journal of Finance 57*(1), 1–32.

Belo, F., X. Lin, and S. Bazdresch (2014). Labor hiring, investment, and stock return predictability in the cross section. *Journal of Political Economy 122*(1), 129–177.

Belo, F., X. Lin, and F. Yang (2019). External equity financing shocks, financial flows, and asset prices. *Review of Financial Studies 32*(9), 3500–3543.

Campbell, J. Y. (1991). A variance decomposition for stock returns. *Economic Journal 101*, 151–179.

Campbell, J. Y. (2018). *Financial Decisions and Markets: A Course in Asset Pricing*. Princeton University Press.

Campbell, J. Y., S. Giglio, C. Polk, and R. Turley (2018). An intertemporal CAPM with stochastic volatility. *Journal of Financial Economics 128*(2), 207–233.

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

Campbell, J. Y. and T. Vuolteenaho (2004). Bad beta, good beta. *American Economic Review 94*(5), 1249–1275.

Carhart, M. M. (1997). On persistence in mutual fund performance. *Journal of Finance 52*(1), 57–82.

Chen, H. (2017). Do cash flows of growth stocks really grow faster? *Journal of Finance 72*(5), 2279–2330.

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

Cochrane, J. H. (2005). *Asset Pricing*. Princeton University Press.

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

Daniel, K., D. Hirshleifer, and A. Subrahmanyam (2001). Overconfidence, arbitrage, and equilibrium asset pricing. *Journal of Finance 56*, 921–965.

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

Diebold, F. X. and R. S. Mariano (1995). Comparing predictive accuracy. *Journal of Business and Economic Statistics 13*(3), 253–263.

Eisfeldt, A., E. Kim, and D. Papanikolaou (2021). Intangible value. *Critical Finance Review, Forthcoming*.

Eisfeldt, A. and D. Papanikolaou (2013). Organizational capital and the cross-section of expected returns. *Journal of Finance 68*(4).

Fairfield, P. M., J. S. Whisenant, and T. L. Yohn (2003). Accrued earnings and growth: Implications for future profitability and market mispricing. *Accounting Review 78*(1), 353–371.

Fama, E. F. and K. R. French (1995). Size and book-to-market factors in earnings and returns. *Journal of Finance 50*(1), 131–155.

Fama, E. F. and K. R. French (2001). Disappearing dividends: changing firm characteristics or lower propensity to pay? *Journal of Financial Economics 60*(1), 3–43.

Fama, E. F. and K. R. French (2006). Profitability, investment, and average returns. *Journal of Financial Economics 82*, 491–518.

Fama, E. F. and K. R. French (2015). A five-factor asset pricing model. *Journal of Financial Economics 116*(1), 1–22.

Franzoni, F., D. Obricky, and R. Resendes (2022). The wealth creation effect in stock returns. *Working Paper*.

Gerakos, J. and J. T. Linnainmaa (2018). Decomposing value. *Review of Financial Studies 31*(5), 1825–1854.

Gormsen, N. and E. Lazarus (2022). Duration-driven returns. *Journal of Finance Forthcoming*.

Goyal, A. and I. Welch (2008). A comprehensive look at the empirical performance of equity premium prediction. *Review of Financial Studies 21*(4), 1455–1508.

Greenwood, R. and S. G. Hanson (2012). Share issuance and factor timing. *Journal of Finance 67*(2), 761–798.

Grullon, G., B. Paye, S. Underwood, and J. P. Weston (2011). Has the propensity to pay out declined? *Journal of Financial and Quantitative Analysis 46*(1), 1–24.

Hansen, L. P. and R. J. Hodrick (1980). Forward exchange rates as optimal predictors of future spot rates: An econometric analysis. *Journal of Political Economy 88*(5), 829–853.

Ilmanen, A., R. Israel, R. Lee, T. Moskowitz, and A. Thapar (2021). How do factor premia vary over time? a century of evidence. *Journal of Investment Management Forthcoming*.

Kogan, L., J. Li, and H. Zhang (2022). Operating hedge and gross profitability premium. *Working Paper*.

Kogan, L. and D. Papanikolaou (2013). Firm characteristics and stock returns: The role of investment-specific shocks. *Review of Financial Studies 26*(11), 2718–2759.

Kogan, L. and D. Papanikolaou (2014). Growth opportunities, technology shocks, and asset prices. *Journal of Finance 69*(2), 657–718.

Lettau, M. and J. A. Wachter (2007). Why is long-horizon equity less risky? A duration-based explanation of the value premium. *Journal of Finance 62*(1), 55–92.

Lochstoer, L. A. and P. C. Tetlock (2020). What drives anomaly returns. *Journal of Finance 75*(3), 1517–1455.

Moreira, A. and T. Muir (2017). Volatility-managed portfolios. *Journal of Finance 72*(4), 1611–1644.

Novy-Marx, R. (2013). The other side of value: The gross profitability premium. *Journal of Financial Economics 108*(1), 1–28.

Ohlson, J. A. (1995). Earnings, book values, and dividends in equity valuation. *Contemporary Accounting Research 11*(2), 661–687.

Ohlson, J. A. (2000). Residual income valuation: The problems. *Working Paper*.

Pontiff, J. and A. Woodgate (2008). Share issuance and cross-sectional returns. *Journal of Finance 63*(2), 921–945.

Shumway, T. (1997). The delisting bias in CRSP data. *Journal of Finance 52*(1), 327–340.

Shumway, T. and V. Warther (1999). The delisting bias in CRSP's Nasdaq data and its implications for the size effect. *Journal of Finance 54*(6), 2361–2379.

Skinner, D. J. (2008). The evolving relation between earnings, dividends, and stock repurchases. *Journal of Financial Economics 87*(3), 582–609.

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

Weber, M. (2018). Cash flow duration and the term structure of equity returns. *Journal of Financial Economics 128*(3), 486–503.

Zhang, L. (2005). The value premium. *Journal of Finance 50*(1), 67–103.

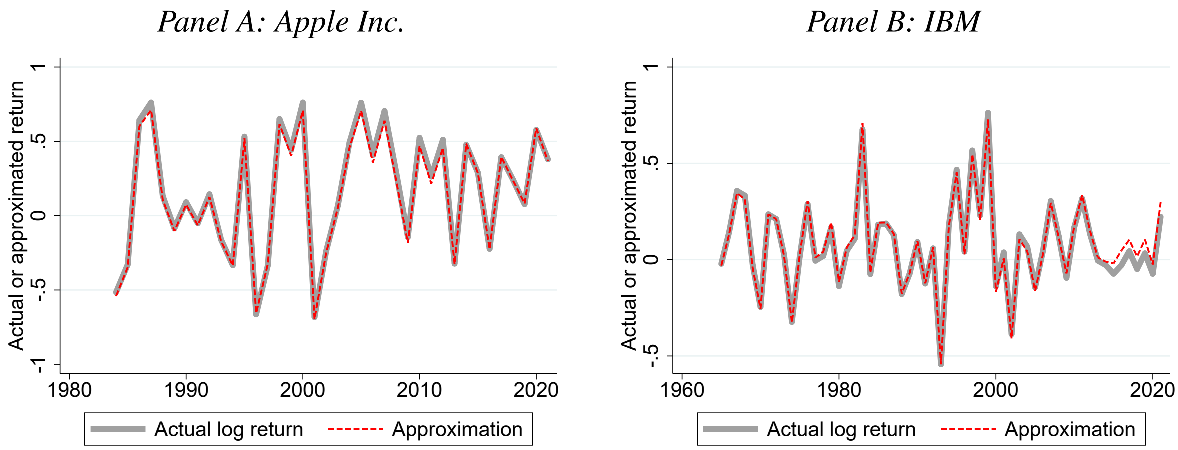

Figure 1: **Actual versus approximated log returns.** This figure plots the actual versus approximated log annual returns on Apple Inc. and IBM stocks to illustrate the accuracy of our loglinear identity. The approximated log annual return is defined as $\hat{r}_t = roe_t + iva_t + \rho mb_t - mb_{t-1}$.

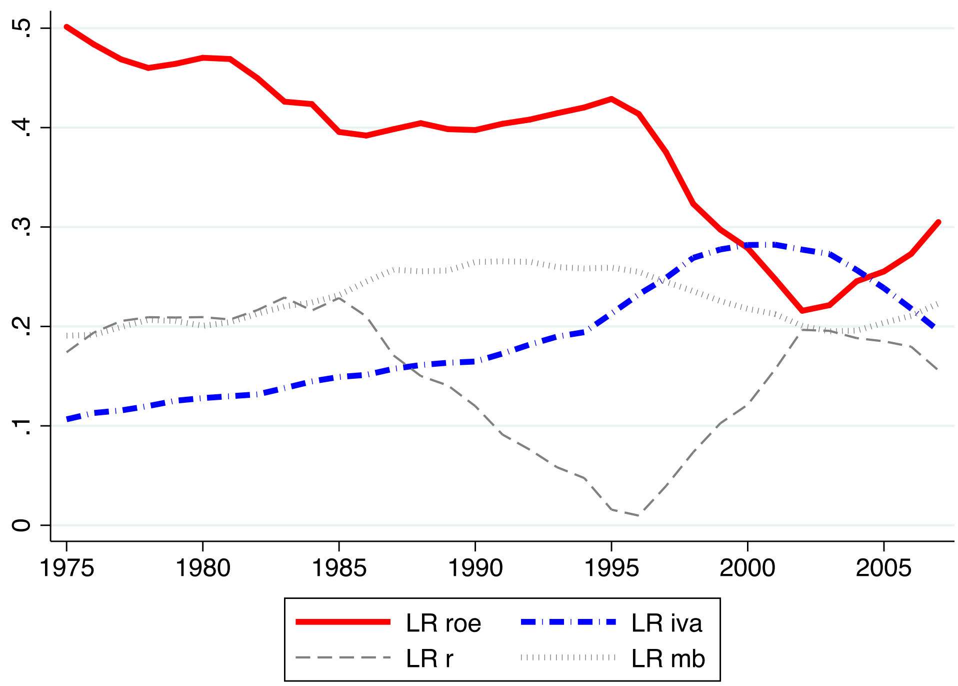

Figure 2: **Cross-sectional variance decomposition of the market-to-book ratio.** This figure plots the ratios of the cross-sectional variance in the market-to-book ratio explained by future long-run returns (r), roe, iva, and persistence in market-to-book (mb). We track realizations of returns, profitability, and iva over a 15-year horizon and run rolling-window regressions of the discounted sums and the future market-to-book ratio on today's market-to-book ratio. Each coefficient is obtained from a trailing 10-year window.

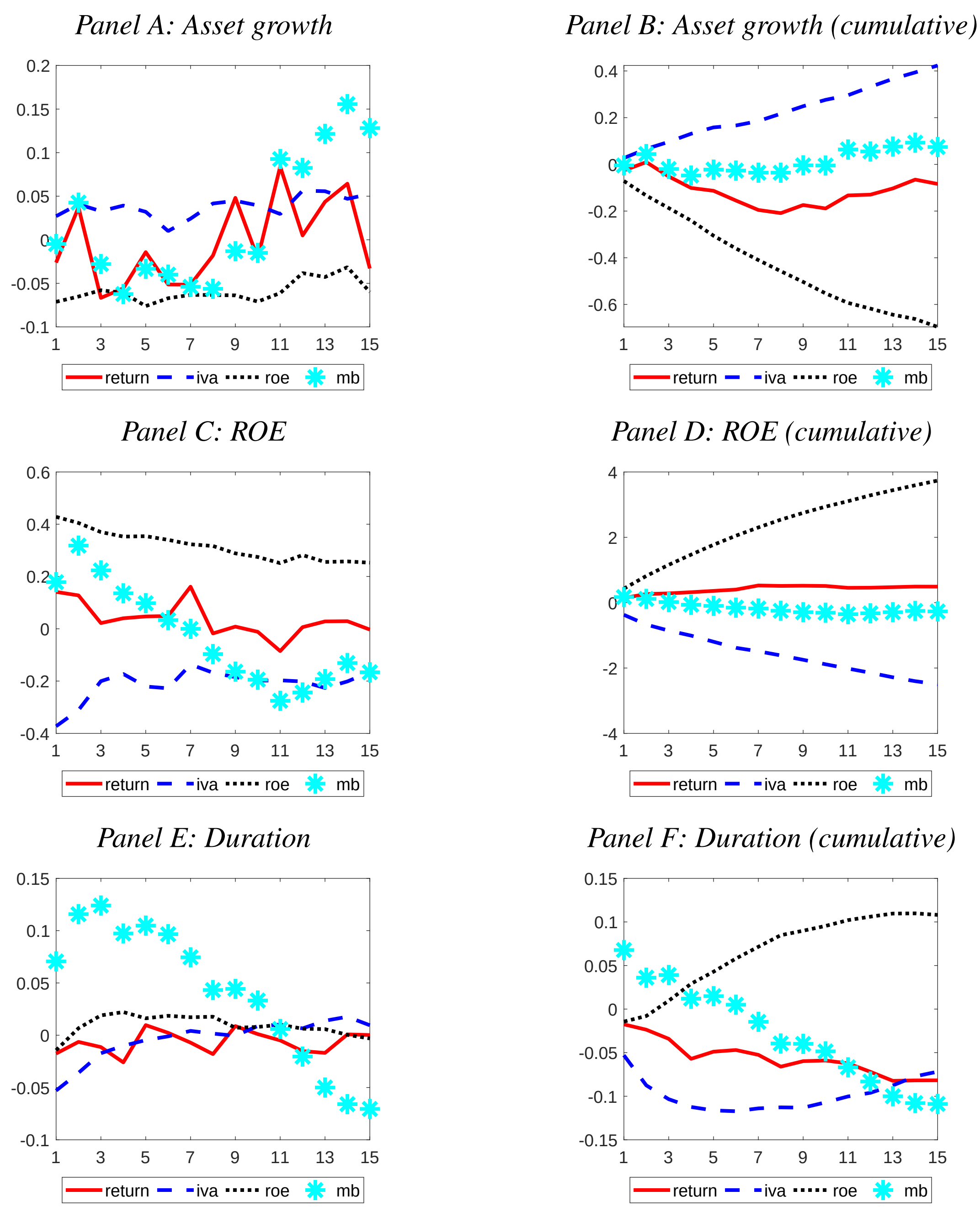

Figure 3: **Return predictability and its anatomy: asset growth (top) or profitability (bottom).** This figure plots horizon-dependent forecasting coefficients from cross-sectional long-horizon regressions (21). We use realized variables on 25 portfolios double-sorted on size and market-to-book. The regressions are bivariate, using market-to-book (mb) and the respective characteristic as predictors of multi-horizon realizations of the identity variables. The left panels use one-year realizations of the dependent variables at the respective horizon; the right panels use cumulative realized variables discounted by ρ as shown in equation (21). We plot coefficients on asset growth, ROE, and duration.

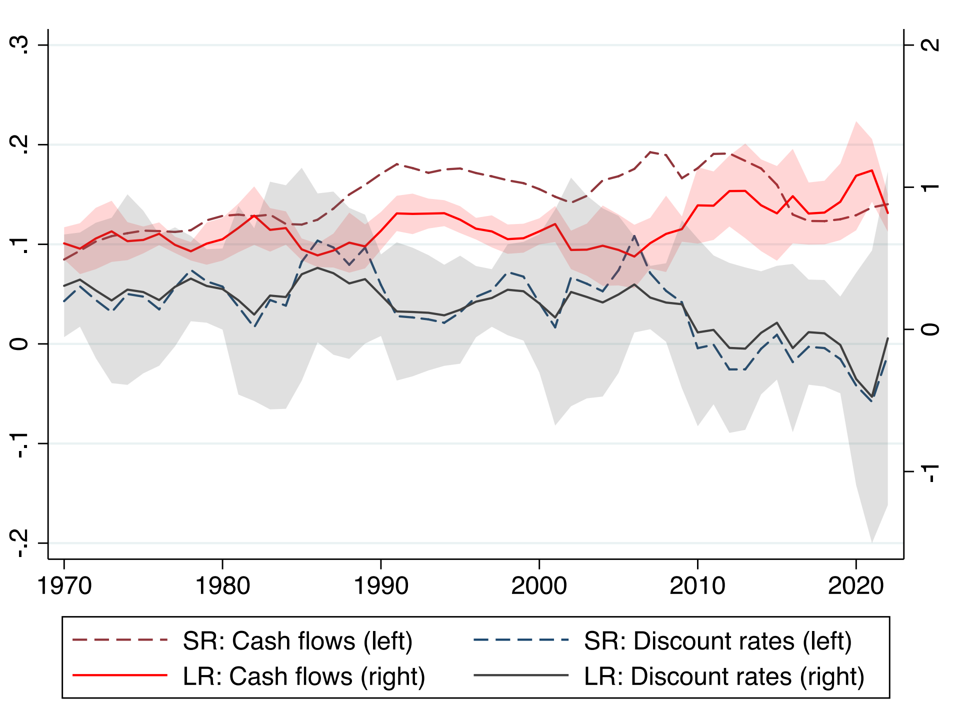

Figure 4: **Short-run and long-run forecasts of cash flows and returns.** This figure visualizes the time-variation in the composition of cross-sectional return- and cash-flow predictability from the market-to-book ratio. Over rolling five-year windows, we run cross-sectional regressions of one-year stock returns and cash-flows ($cf_{i,t} = roe_{i,t} + iva_{i,t}$) on the lagged market-to-book ratio. The dashed lines show the slope coefficients $\hat{\beta}^{SR}$ from these regressions (left vertical axis). We then compute the long-run coefficients as $\hat{\beta}_t^{LR} = \hat{\beta}_t^{SR}/(1-\rho\hat{\phi}_t)$, where $\rho = 0.96$ and $\hat{\phi}_t$ is the estimated coefficient from a cross-sectional regression of market-to-book on its lag over the same rolling window. The solid lines plot the long-run coefficients (right vertical axis). Short-run standard errors are clustered by firm and year, and long-run standard errors obtained using the delta method. We plot 95% confidence intervals for the long-run coefficients and omit the short-run confidence intervals for readability.

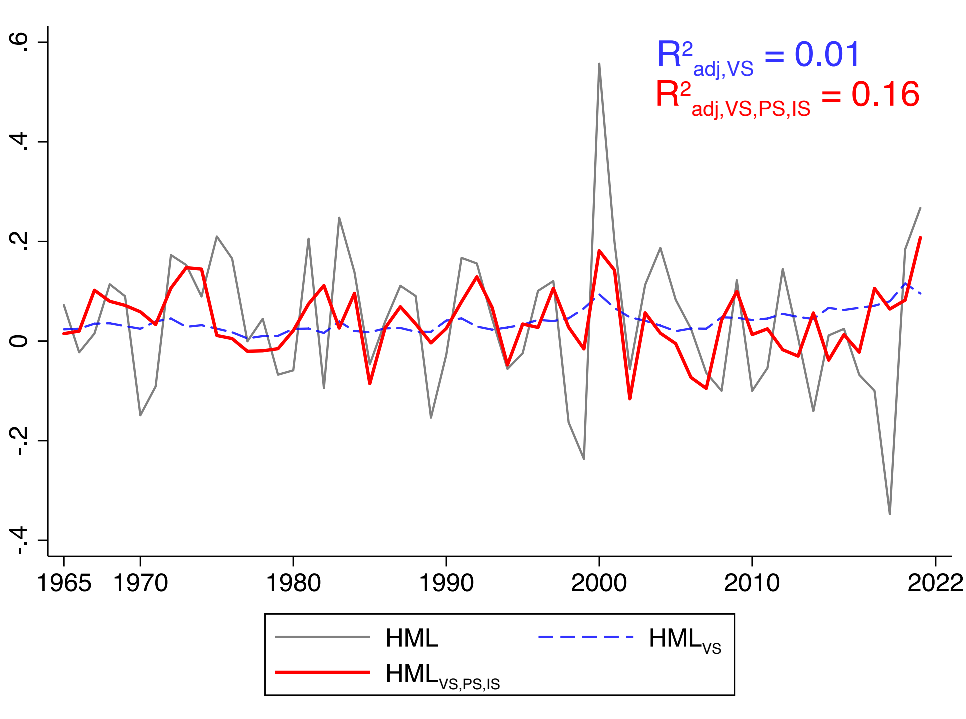

Figure 5: **Forecasts of HML.** This figure visualizes the forecasting performance of the value spread (VS), profitability spread (PS), and iva spread (IS) for annual HML returns. The respective fitted values are obtained from the annual regressions in Table 7 Panel A columns (1) and (3).

Table 1: **Descriptive statistics**

This table presents summary statistics for our key variables. The sample consists of firm-year observations from CRSP and Compustat between 1965 and 2022. r_t and roe_t are the log annual stock return and return on equity, respectively, between June of year $t-1$ and June of year t, mb_t is the market-to-book ratio defined as the log of the market value of equity as of June of year t over the book value of equity in the last fiscal year of $t-1$, and $iva_t = \log\left(\frac{D_t + BPS_t}{BPS_{t-1} + EPS_t}\right)$ is "investment value added" that measures the present value consequence of adjusting book equity through net issuance, where D, BPS, and EPS denote dividend, book equity, and earnings per share. For the aggregate market, TY_t is the term spread, DEF_t is the default spread, PE_t is the price-earnings ratio, and VS_t is the value spread. Panel C presents key statistics of approximation errors from the three versions of identities: a new identity we propose in this paper (*CKLP*), Campbell and Shiller (1988) identity (*Campbell-Shiller*), and Vuolteenaho (2002) identity without the clean-surplus adjustment (*Vuolteenaho w/o CS adj*). The two far-right columns compare the magnitudes for all stocks, and stocks in the top size decile (*Large*).

Panel A. Individual stocks

Variable	N	Mean	St. Dev.	Min	1%	25%	Median	75%	99%	Max
r_t	97,295	0.060	0.308	-1.477	-0.821	-0.100	0.074	0.238	0.801	1.411
roe_t	97,295	0.087	0.136	-1.568	-0.420	0.061	0.104	0.141	0.352	0.905
iva_t	97,295	0.010	0.112	-1.093	-0.268	-0.019	0.004	0.027	0.406	1.316
mb_t	97,295	0.489	0.590	-2.063	-0.854	0.095	0.448	0.864	1.942	2.647

Panel B. Aggregate stock market

Variable	N	Mean	St. Dev.	Min	1%	25%	Median	75%	99%	Max
r_t	57	0.101	0.152	-0.288	-0.288	0.026	0.121	0.178	0.494	0.494
roe_t	57	0.115	0.022	0.047	0.047	0.101	0.119	0.131	0.153	0.153
iva_t	57	0.003	0.021	-0.080	-0.080	-0.003	0.005	0.014	0.047	0.047
mb_t	57	0.683	0.361	-0.119	-0.119	0.464	0.703	0.951	1.420	1.420
TY_t	57	1.025	1.100	-1.230	-1.230	0.240	1.030	1.700	3.090	3.090
DEF_t	57	1.028	0.406	0.390	0.390	0.760	0.920	1.210	2.130	2.130
VS_t	57	1.599	0.161	1.311	1.311	1.496	1.579	1.687	2.068	2.068

Panel C. Present value models: Comparison

	Approximation error e_t									$\lvert e_t\rvert > \lvert e_t^{CKLP}\rvert$	
Model	Mean	St. Dev.	Min	1%	25%	Median	75%	99%	Max	All stocks	Top size decile
CKLP	-0.007	0.023	-0.157	-0.057	-0.016	-0.005	0.002	0.057	0.266		
Campbell-Shiller	-0.023	0.027	-0.428	-0.082	-0.038	-0.016	-0.003	-0.000	0.003	76.6%	62.0%
Vuolteenaho w/o CS adj	-0.017	0.115	-1.361	-0.425	-0.037	-0.010	0.016	0.270	1.097	81.0%	81.7%

Table 2: **Market-to-book decomposition**

This table reports the results from decompositions of cross-sectional variation in valuation ratios. In Panel A, we compute the discounted sums of future roe, future iva, and future log returns based on realizations over a multi-year horizon. To deal with delisting firms without introducing look-ahead bias, we form 25 size- and BM-sorted portfolios using independent NYSE size and BM quintile breakpoints. We then compute the relevant variables at the portfolio level for the subsequent J years and run cross-sectional regressions of each of the components on the market-to-book ratio. We report point estimates and two-way clustered standard errors (by portfolio and year) of β-coefficients from

$$\sum_{j=1}^{J} \rho^{j-1} y_{i,t+j} = \alpha_t + \beta\, mb_{i,t} + \varepsilon_{i,t}. \tag{16}$$

In Panel B, we compute the discounted sums of expected future returns, expected future roe, and expected future iva at the firm level from the VAR described in more detail in Section 4.2.1. Standard errors in Panel B are obtained from a bootstrap reflecting uncertainty in the estimation of the dependent variable via the underlying VAR. All variables are observed annually between 1965 and 2022.

Panel A. Baseline

J	$-\sum_{j=1}^{J} \rho^{j-1} r_{t+j}$	$\sum_{j=1}^{J} \rho^{j-1} roe_{t+j}$	$\sum_{j=1}^{J} \rho^{j-1} iva_{t+j}$	$\rho^J mb_{t+J}$
1	0.023	0.089	0.044	0.798
	(0.013)	(0.011)	(0.013)	(0.018)
5	0.070	0.253	0.125	0.502
	(0.025)	(0.045)	(0.040)	(0.033)
10	0.115	0.338	0.161	0.324
	(0.027)	(0.073)	(0.058)	(0.027)
15	0.125	0.402	0.172	0.233
	(0.029)	(0.089)	(0.067)	(0.022)

Panel B. VAR-based, firm-level decomposition

J	$-\sum_{j=1}^{J} \rho^{j-1} E_t\left[r_{t+j}\right]$	$\sum_{j=1}^{J} \rho^{j-1} E_t\left[roe_{t+j}\right]$	$\sum_{j=1}^{J} \rho^{j-1} E_t\left[iva_{t+j}\right]$
∞	0.286	0.475	0.170
	(0.067)	(0.066)	(0.022)

Table 3: **Variance decomposition of return news**

This table reports the covariance matrix of the decomposition of return news into discount-rate news (N_{DR}), investment news (N_{iva}), and profitability news (N_{roe}), where all of the news terms are computed explicitly. For ease of interpretation, the left half of each Panel reports standard deviations on the diagonal, and correlation coefficients for the off-diagonal elements. The right half reports each covariance term as a fraction of the total return news variance. The off-diagonal terms multiply the respective covariance by -2 in the N_{DR}-column and by 2 in the N_{roe}-column. Panel A is based on the firm-level VAR reported in Table E.1, Panel B on an aggregate VAR (Table E.3), and Panel C aggregates the firm-level news terms into the mean-variance efficient portfolio of Lochstoer and Tetlock (2020). We assign an equal weight to each cross-section. All variables are observed annually between 1965 and 2022.

Panel A. Firm-level news variance

	σ (diag), ρ (off-diag)			Contribution to $\sigma^2_{N_r}$		
	N_{DR}	N_{roe}	N_{iva}	$-N_{DR}$	N_{roe}	N_{iva}
N_{DR}	0.088			0.091		
N_{roe}	-0.235	0.239		0.116	0.671	
N_{iva}	-0.432	-0.244	0.124	0.111	-0.171	0.182

Panel B. Market-level news variance

N_{DR}	0.130			1.145		
N_{roe}	0.205	0.033		-0.119	0.074	
N_{iva}	0.469	0.590	0.026	-0.214	0.069	0.046

Panel C. MVE portfolio-level news variance

N_{DR}	0.058			0.071		
N_{roe}	-0.367	0.164		0.149	0.579	
N_{iva}	-0.598	-0.118	0.083	0.122	-0.069	0.147

Table 4: **How much does investment contribute to return news variance?**

This table decomposes return news (N_R) and cash-flow news (N_{CF}) from VAR in Table E.1. We regress return news on discount-rate news (N_{DR}), profitability news (N_{roe}), and investment news (N_{iva}) at the firm-level. The last two column regress cash-flow news on profitability and investment news. We assign an equal weight to each cross-section. All variables are observed annually between 1965 and 2022. In parentheses we report two-way clustered standard errors by firm and year. The row labeled $\tilde{R}^2$ reports the R^2 for a regression that fixes the slope coefficients at their theoretically implied values of one and minus one, respectively. Panel B repeats the exercise for subsamples: $Growth/Value$, $Large/Small$, $Profitable/Unprofitable$ indicate firms in the top/bottom market-to-book, market capitalization, and roe quintiles, respectively. All cut-offs are based on the distribution among NYSE stocks in the year preceding the observed stock return. In brackets, we report bootstrapped p-values intervals for the difference in R^2 and $\tilde{R}^2$ between the top and bottom quintiles from the regressions of N_{CF} on N_{roe}. The $Recent$ sample spans years 1990 through 2022.

Panel A. Full sample

	N_r	N_r	N_{CF}	N_{CF}
N_{DR}	-2.033	-1.300		
	(0.024)	(0.004)		
N_{roe}	0.831	1.021	0.920	1.061
	(0.011)	(0.001)	(0.022)	(0.004)
N_{iva}		0.998		1.109
		(0.003)		(0.006)
Observations	97295	97295	97295	97295
R^2	0.888	0.995	0.719	0.985
$\tilde{R}^2$	0.805	0.987	0.713	0.981

Panel B. Subsamples

	Growth		*Large*		*Profitable*		*High duration*		*Recent*	
N_{roe}	0.887	1.067	0.962	1.062	0.956	1.060	0.825	1.057	0.870	1.063
N_{iva}		1.086		1.071		1.076		1.117		1.107
R^2	0.686	0.979	0.716	0.982	0.729	0.979	0.683	0.985	0.690	0.984
$\tilde{R}^2$	0.671	0.975	0.703	0.978	0.724	0.975	0.649	0.981	0.675	0.979

	Value		*Small*		*Unprofitable*		*Low duration*		*Early*	
N_{roe}	0.814	1.041	0.870	1.057	0.779	1.050	0.921	1.051	1.057	1.054
N_{iva}		1.118		1.133		1.124		1.095		1.117
R^2	0.716	0.988	0.721	0.988	0.675	0.987	0.737	0.985	0.798	0.989
$\tilde{R}^2$	0.674	0.984	0.703	0.983	0.615	0.982	0.729	0.981	0.796	0.985
(ΔR^2)	[0.028]		[0.602]		[0.013]		[0.000]		[0.000]	
$(\Delta \tilde{R}^2)$	[0.431]		[0.479]		[0.006]		[0.000]		[0.000]	

Table 5: **Predicting long-horizon returns and their cash-flow drivers**

This table seeks to understand the mirror image of long-horizon return predictability based on the long-horizon cash flow counterparts implied by the identity (Section 5). We run the following regressions,

$$\sum_{j=1}^{J} \rho^{j-1} y_{i,t+j} = \alpha_t + \beta_1 \, mb_{i,t} + \beta_2 \, x_{i,t} + \varepsilon_{i,t}, \tag{21}$$

for $y \in \{r, roe, iva\}$ (reported in different column groups), $x = \{ag, roe, iva, Dur\}$ (reported in different rows), and $J \in \{1, 5, 15\}$ years (reported in different sub-columns), where ag is asset growth and Dur is duration. Hence, each reported coefficient is the estimated $\widehat{\beta}_2$ from a separate regression. The regression controls for current market-to-book (mb_t) to analyze return and cash-flow predictability beyond what mb_t predicts. The estimates are based on realized variables of 25 portfolios double-sorted on size and market-to-book. All variables are observed annually between 1965 and 2022. In parentheses, we report standard errors clustered by portfolio and year for $J = 1$, and Hansen and Hodrick (1980) standard errors with $J - 1$ lags for $J = 5$ and 15.

	$\sum_{j=1}^{J} \rho^{J-1} r_{t+j}$			$\sum_{j=1}^{J} \rho^{J-1} roe_{t+j}$			$\sum_{j=1}^{J} \rho^{J-1} iva_{t+j}$		
J	1	5	15	1	5	15	1	5	15
ag_t	-0.026	-0.113	-0.084	-0.071	-0.306	-0.696	0.027	0.158	0.423
	(0.067)	(0.090)	(0.096)	(0.019)	(0.062)	(0.106)	(0.026)	(0.036)	(0.064)
roe_t	0.141	0.360	0.490	0.428	1.771	3.735	-0.373	-1.197	-2.497
	(0.139)	(0.089)	(0.097)	(0.082)	(0.062)	(0.113)	(0.051)	(0.036)	(0.067)
iva_t	-0.283	-0.441	-0.628	-0.149	-0.616	-1.362	0.023	0.242	0.815
	(0.163)	(0.089)	(0.095)	(0.117)	(0.061)	(0.109)	(0.112)	(0.036)	(0.066)
Dur_t	-0.018	-0.049	-0.082	-0.014	0.043	0.108	-0.053	-0.116	-0.072
	(0.037)	(0.088)	(0.092)	(0.018)	(0.059)	(0.110)	(0.022)	(0.034)	(0.068)
Observations	1100	1100	1100	1100	1100	1100	1100	1100	1100

Table 6: **Six-factor alphas of iva-sorted portfolios**

This table reports monthly alphas relative to the Fama-French-Carhart six-factor model for iva quintile portfolios (Panel A) as well as portfolios based on double sorts (Panel B) on iva and either book-to-market, operating profitability, and size. We report Newey-West standard errors in parentheses with automatically selected lag-length. The sample spans years 1965 to 2022.

Panel A. iva quintiles

iva quintile	1	2	3	4	5	5-1
α	-0.034	-0.062	-0.082	-0.055	0.170	0.205
	(0.051)	(0.035)	(0.043)	(0.034)	(0.049)	(0.085)

Panel B. Double-sorts: BM, OP, and Size

BM \ iva	1	2	3	4	5	5-1
1	-0.071	0.021	-0.069	0.059	0.377	0.448
	(0.065)	(0.065)	(0.073)	(0.065)	(0.081)	(0.106)
5	0.053	-0.079	0.087	0.006	0.060	0.007
	(0.168)	(0.083)	(0.089)	(0.071)	(0.099)	(0.197)
OP \ iva	1	2	3	4	5	5-1
1	-0.296	-0.166	-0.046	0.015	0.122	0.418
	(0.127)	(0.095)	(0.099)	(0.086)	(0.089)	(0.162)
5	0.015	-0.030	-0.092	-0.031	0.257	0.242
	(0.069)	(0.081)	(0.100)	(0.083)	(0.088)	(0.113)
Size \ iva	1	2	3	4	5	5-1
1	-0.072	0.003	0.168	0.169	0.009	0.081
	(0.096)	(0.062)	(0.078)	(0.071)	(0.076)	(0.107)
5	0.006	-0.084	-0.097	-0.058	0.208	0.202
	(0.053)	(0.041)	(0.047)	(0.044)	(0.064)	(0.092)

Table 7: **Predicting HML**

This table reports the result from monthly and annual time-series forecasting regressions of HML returns—simple net returns rather than log returns—using the value spread, profitability (roe) spread, and iva spread as return predictors. In parentheses, we report Newey-West standard errors with a lag of 12 months. In brackets, we report the corresponding t-stats for convenience.

	Annual			Monthly		
	(1)	(2)	(3)	(4)	(5)	(6)
Value spread	0.096	0.133	0.150	0.007	0.009	0.009
	(0.124)	(0.109)	(0.101)	(0.010)	(0.008)	(0.008)
	[0.78]	[1.21]	[1.48]	[0.69]	[1.12]	[1.16]
Profitability spread		-0.293	-0.680		-0.023	-0.046
		(0.282)	(0.256)		(0.027)	(0.022)
		[-1.04]	[-2.66]		[-0.88]	[-2.05]
iva spread			-1.089			-0.082
			(0.311)			(0.023)
			[-3.50]			[-3.57]
Observations	57	57	57	684	684	684
Adjusted R^2	0.006	-0.004	0.165	0.001	0.001	0.021

A Derivations

A.1 Exact nonlinear identity (Remark 1)

Begin with the definition of return:

$$P_t = \frac{1}{1 + R_{t+1}} \left(D_{t+1} + P_{t+1} \right),$$

where P is price, R is return, and D is dividend per share. The firm uses the previous year's book equity to realize earnings, engages in net issuance, and distributes dividends to old shareholders; hence, total dividend payment at $t + 1$ is $D_{t+1}N_t$, where N_t is the number of shares at the end of time t.

Multiply both sides by N_t/B_t, where N is the number of shares and B is total book equity, to obtain

$$\frac{P_t N_t}{B_t} = \frac{1}{1 + R_{t+1}} \times \frac{1}{B_t/N_t} \times (D_{t+1} + P_{t+1})$$

On the right-hand side, multiply and divide by $BPS_t + EPS_{t+1}$ and $D_{t+1} + BPS_{t+1}$ where $BPS_t \equiv B_t/N_t$ is book value per share and $EPS_{t+1} \equiv Y_{t+1}/N_t$ for total earnings Y_{t+1} is earnings per the previous year's shares:

$$\frac{M_t}{B_t} = \frac{1}{1 + R_{t+1}} \times \frac{BPS_t + EPS_{t+1}}{BPS_t} \times \frac{D_{t+1} + BPS_{t+1}}{BPS_t + EPS_{t+1}} \times \frac{D_{t+1} + P_{t+1}}{D_{t+1} + BPS_{t+1}},$$

where $M_t = P_t N_t$ on the left-hand side is the total market value. Since the last term can be written as

$$\frac{D_{t+1} + P_{t+1}}{D_{t+1} + BPS_{t+1}} = \frac{D_{t+1}}{D_{t+1} + BPS_{t+1}} + \frac{P_{t+1}}{BPS_{t+1}} \times \frac{BPS_{t+1}}{D_{t+1} + BPS_{t+1}},$$

we can rewrite the last term as an interaction of the plowback ratio and the new market-to-book ratio:

$$\frac{M_t}{B_t} = \frac{1}{1+R_{t+1}} \times \underbrace{\frac{B_t + Y_{t+1}}{B_t}}_{1+ROE_{t+1}} \times \underbrace{\frac{D_{t+1} + BPS_{t+1}}{BPS_t + EPS_{t+1}}}_{1+IVA_{t+1}} \times \left(1 + \left(\frac{M_{t+1}}{B_{t+1}} - 1\right) \underbrace{\frac{BPS_{t+1}}{D_{t+1} + BPS_{t+1}}}_{\Lambda_{t+1}}\right),$$

where ROE is now expressed in terms of firm-level quantities rather than per-share quantities (by multiplying both the numerator and denominator by N_t).

A.2 Investment value added (Remark 2)

To express IVA in terms of net issuance, begin with the expression for IVA and add and subtract B_t/N_{t-1} in the numerator to obtain

$$1 + IVA_t = \frac{D_t + BPS_t}{BPS_{t-1} + EPS_t} = \frac{D_t + B_t/N_{t-1} - B_t/N_{t-1} + B_t/N_t}{BPS_{t-1} + EPS_t}.$$

The law of motion for equity capital is $N_{t-1}D_t + B_t = B_{t-1} + Y_t + (N_t - N_{t-1})\, P_t^*$, since dividend payment and new book equity are financed by the old book equity and new earnings net of the dollar raised or spent on equity issuance or repurchases. Dividing the law of motion by N_{t-1} to have $D_t + B_t/N_{t-1}$ on the left-hand side and substituting it into the last expression for $1 + IVA_t$,

$$IVA_t = \frac{BPS_t - BPS_t N_t/N_{t-1}}{BPS_{t-1} + EPS_t} + \frac{N_t - N_{t-1}}{N_{t-1}} \frac{P_t^*}{BPS_{t-1} + EPS_t}.$$

Finally, a simple rearrangement gives

$$IVA_t = \frac{N_t - N_{t-1}}{N_{t-1}} \left(\frac{P_t^*}{BPS_t} - 1\right) \left(\frac{BPS_t}{BPS_{t-1} + EPS_t}\right).$$

B Investment via issuance and plowback

Throughout the paper, we use the term investment to refer to increases in book equity, since this is what drives the relevant channels in our present-value framework. Net investment in book equity is composed of issuance, repurchases, earnings retention (plowback) and dividend payouts. For reasons outlined in the main text, we focus on the net issuance portion of this channel (iva) as the empirical object of interest.

To build some intuition of how the different investment pieces fit together, and how our choice to carve out the issuance-based channel affects the approximation error in the loglinearization, this section first lays out the key economic intuition in the form of an exact nonlinear identity that aggregates the investment channel (B.1). We then delineate why the approximation error in our chosen loglinear framework can be interpreted as incorporating the net plowback channel, and how this channel quantitatively contributes to variation in valuation ratios (B.2).

B.1 An alternative exact identity for intuition

To illustrate the joint role of issuance and plowback in the present-value identity, we restate the exact nonlinear identity as follows:

$$\frac{M_t}{B_t} = \underbrace{\frac{1}{1+R_{t+1}}}_{\text{discount rate}} \times \underbrace{(1+ROE_{t+1})}_{\text{profitability}} \times \overbrace{[1 \; + \; (\frac{M_{t+1}}{B_{t+1}} - 1) \; \times \; \underbrace{\frac{B_{t+1}}{B_t + Y_{t+1}}}_{\text{investment in book equity}}]}^{\text{value added from investment in book equity}}, \tag{B.1}$$

where $\frac{B_{t+1}}{B_t+Y_{t+1}}$ is (net) investment in book equity through retained earnings, net share issuance, and dividend payout. Clearly, holding all else fixed, a drop in discount rates (R) or a rise in profitability (ROE) raises today's M/B unambiguously. However, since investment in book equity ($\frac{B_{t+1}}{B_t+Y_{t+1}}$) is made at the expense of cash flows to shareholders today, more investment in book equity adds value only when the market value of retained capital exceeds the book value—i.e., if the future M/B is greater than one. In addition, future M/B unambiguously raises today's M/B, meaning that when equation (B.1) is iterated forward, discount rates, profitability, and investment value added in all future periods affect today's M/B in the same direction.

To see how to derive equation (B.1), begin with $P_t = \frac{D_{t+1}+P_{t+1}}{1+R_{t+1}}$. Multiply both sides by N_t/B_t to obtain $\frac{M_t}{B_t} = \frac{1}{1+R_{t+1}}\left[\frac{D_{t+1}N_t}{B_t} + \frac{P_{t+1}N_t}{B_t}\right]$. Noting that $D_{t+1}N_t + B_{t+1} = B_t + Y_{t+1} +$

$(N_{t+1} - N_t) P_{t+1}$, rewrite $D_{t+1}N_t$ in terms of the book equity change, earnings, and cash flows from net issuance: $\frac{M_t}{B_t} = \frac{1}{1+R_{t+1}} \left[-\frac{B_{t+1}}{B_t} + 1 + \frac{Y_{t+1}}{B_t} + \frac{P_{t+1}N_{t+1} - P_{t+1}N_t}{B_t} + \frac{P_{t+1}N_t}{B_t} \right]$. Defining terms appropriately and rearranging, $\frac{M_t}{B_t} = \frac{1}{1+R_{t+1}} \left[1 + ROE_{t+1} + \left(\frac{M_{t+1}}{B_{t+1}} - 1 \right) \frac{B_{t+1}}{B_t} \right]$, which we can again rearrange to be $\frac{M_t}{B_t} = \frac{1}{1+R_{t+1}} (1 + ROE_{t+1}) \left[1 + \left(\frac{M_{t+1}}{B_{t+1}} - 1 \right) \frac{B_{t+1}}{B_t + Y_{t+1}} \right]$. Note that this derivation assumes that net issuance price equals end-of-period market price.

B.2 Plowback value-added

As outlined in the main text, our expression of the exact non-linear identity contains two channels for valuations to reflect future investment decisions: via the issuance-repurchase decision and via the payout-plowback decision. In the baseline linearization, we focus on the former, which we encode our novel iva variable, and defer the latter to the approximation error. However, the exact nonlinear identity (6) can also be linearized without any approximation. Simply define *plowback value added* as $PVA_t = \left(\frac{M_t}{B_t} - 1 \right) \Lambda_t$ and $pva_t = \log(1 + PVA_t) - \rho mb_t$. Taking logs of (6), the log market-to-book ratio can be exactly written as

$$\begin{aligned} mb_t &= -r_{t+1} + roe_{t+1} + iva_{t+1} + pva_{t+1} + \rho mb_{t+1} \\ &= -\sum_{j=1}^{\infty} \rho^{j-1} r_{t+j} + \sum_{j=1}^{\infty} \rho^{j-1} roe_{t+j} + \sum_{j=1}^{\infty} \rho^{j-1} iva_{t+j} + \sum_{j=1}^{\infty} \rho^{j-1} pva_{t+j}, \end{aligned} \tag{B.2}$$

where the second equation once again imposes the transversality condition on infinite-horizon market-to-book. The above definition of PVA closely mirrors that of IVA in equation (7). Just like IVA interacts net issuance with the market-to-book ratio, PVA captures the interaction of the plowback ratio Λ with market-to-book. Its presence in the exact linear identity (B.2) reflects the economic intuition that retaining earnings rather than paying dividends increases value when the marginal dollar is more valuable inside the firm than outside.

In the main text, we ignore plowback value-added by grouping it with the approximation error since we find that iva is empirically a much more relevant investment channel. Equation (B.2) allows us to do this more quantitatively. Tables E.8 and E.9 repeat the analyses in Tables 2 and 3, respectively. The value contribution of future pva accounts for almost 10% of cross-sectional dispersion in market-to-book, but close to none of the dispersion in return news. These results are consistent with the conventional wisdom that dividend policy is sticky within-firm but varies more meaningfully across firms.

C Intangible value

The original goal of 'value investing' is to identify stocks that are 'cheap', in the sense of offering high future returns. But of course, Fama and French (1995) and Cohen et al. (2003) show that some stocks are 'cheap' because they have low expected cash-flow growth, and we confirm that this can be due to both low profitability gains or low expansion opportunities. In more recent work, Eisfeldt and Papanikolaou (2013) and others have stressed that scaling prices by book equity to determine 'cheapness' misses cross-sectional differences in 'intangible' capital. Eisfeldt et al. (2021) then show that a value factor that sorts on intangibles-adjusted market-to-book outperforms the standard HML factor.

In this section, we show that the present-value identity can be easily adjusted to decompose a market-to-book ratio that accounts for intangible capital. We then decompose the adjusted market-to-book ratio into future profitability and future issuance-driven book-equity growth, each adjusted for intangibles.

C.1 Adjusting the identity

We denote intangible capital by $I_{i,t}$ and follow Eisfeldt et al. (2021) in specifying its law of motion as $I_{i,t} = (1-\delta)I_{i,t-1} + \theta \text{SG\&A}_{i,t}$, with initial condition $I_{i,s} = \text{SG\&A}_{i,s}/(g+\delta)$ in firm i's initial year s. We also follow their parameter choices of $\theta = 1$ for the capitalization rate of SG&A expenses and $\delta = 0.2$ for the depreciation rate of intangible capital. We can then define $B_t^* = B_t + I_t$, and $IVA_t^* = \left(\frac{N_t - N_{t-1}}{N_t}\right)\left(\frac{M_t}{B_t^*} - 1\right)\frac{B_t^*}{B_{t-1}^* + Y_t}$ to rewrite our identity (10) as

$$mb_{t-1}^* \approx -\sum_{j=0}^{\infty} \rho^j r_{t+j} + \sum_{j=0}^{\infty} \rho^j roe_{t+j}^* + \sum_{j=0}^{\infty} \rho^j iva_{t+j}^*, \tag{C.1}$$

where $mb_t^* = \log(M_t/B_t^*)$, $roe_t^* = \log(1 + Y/B^*)$, and $iva^* = \log(1 + IVA_t^*)$. The resulting profitability variable roe^* adds back SG&A expenses and is, in that sense, a step closer to a gross-profitability figure of the type used to establish the profitability anomaly in Novy-Marx (2013).

C.2 The intangible-adjusted value spread

If the 'cheapness' of value stocks is meant to capture high returns, then—as outlined in Section 4.1 and consistent with the findings of Cohen et al. (2003)—many value stocks are cheap for a reason: low cash-flow growth. Relative to Cohen et al. (2003), we show that these differences in cash-flow growth are driven to a similar extent by future gains in profitability, as they are by future expansions of equity capital. But what do differences in intangible-adjusted market-to-book tell us about differences in future returns and fundamentals? We repeat the portfolio-based cross-secitonal decomposition from Section 4.1 for M/B^*. Table E.10 reports the full-sample results for different horizons. Compared to the conventional market-to-book ratio, we find that the adjusted valuation ratio the fraction of cross-sectional variance that accounts for variation in short-horizon returns rises slightly, while that of roe^* falls relative to that of roe in Table 2.

So is intangible-adjusted book-to-market a better measure of 'value' in the sense of expected short-run returns? Eisfeldt et al. (2021) find that adjusting the conventional HML factor for intangibles produces higher returns. They also find that the long leg of the adjusted HML portfolio, relative to the conventional one, contains firms with higher (current) productivity or gross profitability. They conclude that this is "likely due to the intangible value factor sorting more effectively on productivity, profitability, financial soundness, and on other valuation ratios..." (p. 1). The present-value identity allows us to put more structure on this explanation. Adjusted market-to-book is indeed more informative about future return differences, because it is less informative about differences in future cash-flow growth. Of course, this explanation is not inconsistent with value firms also having higher *current* productivity or profitability. Instead, our exercise embeds these intuitively sensible conditions on these characteristics within the identity-based present-value framework that ties valuation ratios to returns and cash-flow fundamentals more rigorously.

D HML forecasting

D.1 The cross-sectional analog of Cochrane (2008)

We generate Figure 4 by applying a method used in Cochrane (2008) to the cross-section. Using rolling five-year windows, we run cross-sectional firm-level regressions of annual (i) cash-flows (constructed as $cf_{i,t} = roe_{i,t} + iva_{i,t}$) and (ii) stock returns on the lagged market-to-book ratio. Denoting these short-run coefficients by $\hat{\beta}_t^{SR}$, we then construct the implied *long-run* coefficients as $\hat{\beta}_t^{LR} = \hat{\beta}_t^{SR}/(1-\rho\hat{\phi}_t)$, where $\rho = 0.96$ and $\hat{\phi}_t$ is the estimated coefficient from a cross-sectional regression of market-to-book on its lag over the same rolling window. Short-run standard errors are clustered by firm and year; long-run standard errors are obtained from short-run standard errors using the Delta method. We omit confidence intervals on short-run coefficients for readability.

D.2 Out-of-sample forecasting

We assess the out-of-sample performance of our model that augments the CPV-forecast of HML returns based on the value spread with the profitability spread and the iva spread. As in Table 4, we split the sample in January 1990 and use the first 25 years to estimate different forecasting models. We then compute expanding-window out-of-sample R^2 for HML return forecasts following Goyal and Welch (2008) for the augmented model versus the univariate prediction using the value spread.

Figure E.2 plots the cumulative out-of-sample performance of these expanding-window forecasts over time. During the early nineties, the univariate forecast performs similarly out-of-sample as the trivariate one, in line with the CPV-motivation: as long as the discount-rate share in market-to-book dispersion is relatively stable, the value spread predicts HML returns. We see in Figure 2, however, that the discount-rate share falls in the 90s as the iva-share keeps rising. With shifts in the decomposition of market-to-book dispersion, the value spread must be supplemented by variables that predict long-horizon cash-flows and/or short-horizon changes in market-to-book. The iva spread and roe spread (jointly) do this, so the trivariate specification outperforms the univariate from the run-up to the dotcom bubble onwards. To assess the statistical significance of these results, we compute Diebold-Mariano tests (Diebold and Mariano, 1995). The relevant full-sample

p-value vis-à-vis the univariate model is 0.075.

D.3 Other predictors

The value spread has to forecast long-run differences in returns and/or cash flows between value and growth firms. The basic motivation in using the profitability spreand and the iva spread to forecast HML is as a control for cash-flow differences contained in the value spread. Return predictors that complement the value spread (in the sense of reinforcing each other's forecasting power) will often do so because—like the profitability spread and the iva spread—the complementary predictor forecasts cash flows. In short-run return forecasts, however, a given variable may also be a successful predictor if it correlates with changes in the term structure of HML returns. We therefore consider a number of additional predictors proposed in previous studies to see if and how they interact with the value spread.[26]

We consider the aggregate market-to-book ratio and the Aaa-Baa default spread (Cohen et al., 2003), the lagged HML return (Ilmanen et al., 2021), the inverse of lagged three-year HML-volatility (Moreira and Muir, 2017), and the issuer-repurchaser spread of the book-to-market characteristic (Greenwood and Hanson, 2012). The results are in Table E.12. Except for the last of these variables, none alter predictability coming from the roe spread and iva spread, nor do they complement the value spread (the t-statistic on the value spread falls relative to Table 7, as does adjusted R^2 for annual returns). None of them are individually significant predictors of annual returns. Only the lagged HML return and its historical three-year inverse volatility predict monthly returns. These results suggest that the inverse volatility of Moreira and Muir (2017) and HML momentum (Ilmanen et al., 2021) predict variations in the term structure of HML returns, but are unlikely to predict long-run factor returns.

The issuer-repurchaser spread (Greenwood and Hanson, 2012), instead, behaves similarly to the iva spread: it negatively predicts both monthly and annual returns and it strongly interacts with the value spread and the profitability spread: it raises the coefficient and the t-statistic on the value

[26]While several papers propose factor-timing strategies, the overall evidence is mixed. Ilmanen et al. (2021) conclude that there is only "modest predictability that likely fails to overcome implementation frictions."

spread (very close to the original results in CPV's older sample), but it loses its own significance once the value spread is dropped from the set of predictors. Just like the iva spread it also interacts with the profitability spread in HML predictions. These findings are new relative to Greenwood and Hanson (2012) and suggest that some of the predictive power stems from its ability to forecast long-run cash flows. Indeed, we find in regressions analogous to those in Table 5 that—similar to iva—an issuer-repurchaser dummy like the one used to construct the issue-repurchaser spread negatively predicts long-run profitability, and positively predicts long-run iva (Table E.13).

We also conduct a similar out-of-sample exercise to the one in Figure E.2 for a trivariate model that replaces the iva spread with the issuer-repurchaser spread. We find that (i) this model similarly outperforms the univariate value-spread forecast, (ii) this outperformance is similarly driven by the post-dotcom-bubble period, and (iii) the out-of-sample performances of the two trivariate models are not statistically distinguishable (Diebold-Mariano p-value of 0.40).

E Supplementary Figures and Tables

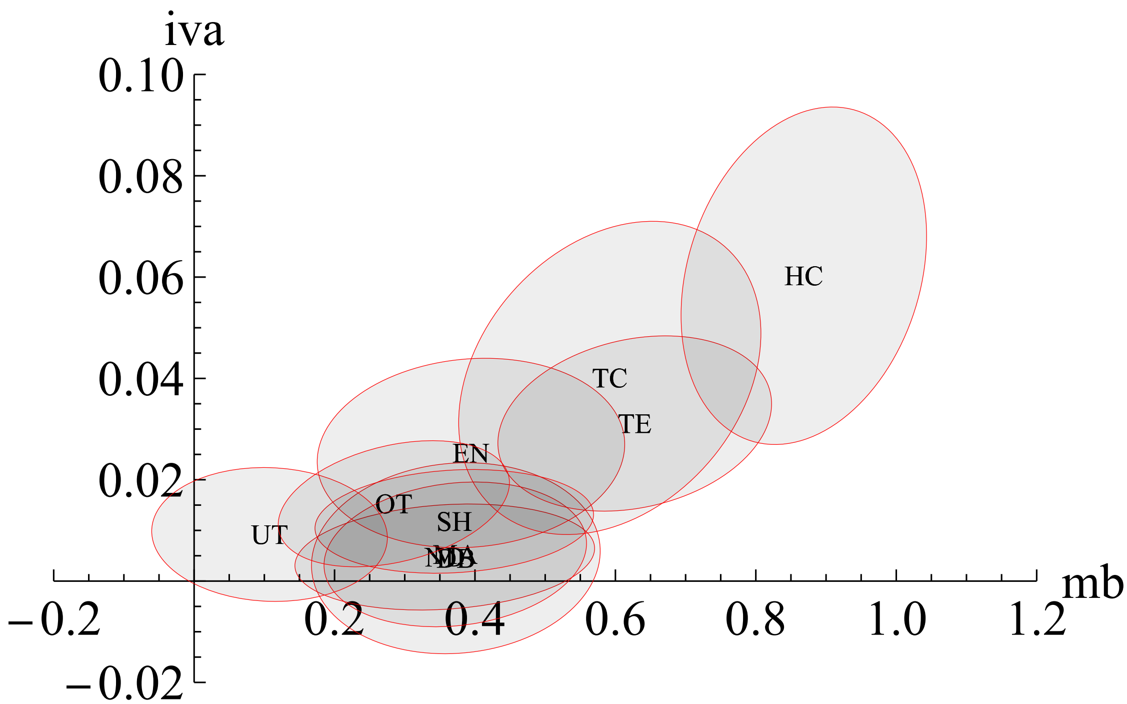

Figure E.1: **Market-to-book and iva across industries.** This figure plots the log market-to-book ratio (mb_t) and investment value added (iva_t) for the ten Fama-French industries. All years receive equal weight. For each industry, the averages are surrounded by a confidence ellipse, whose orientation reflects the time-series correlation of industry-level iva and mb and whose size reflects their volatilities. Under joint normality, each ellipse would contain 20% of the observations for the given industry. The industries are: Consumer durables (DB), Consumer non-durables (ND), Energy (EN), Healthcare (HC), Manufacturing (MA), Other (OT), Shopping (SH), Technology (TE), Telecommunications (TC), Utilities (UT). See Kenneth French's website for detailed definitions and composition.

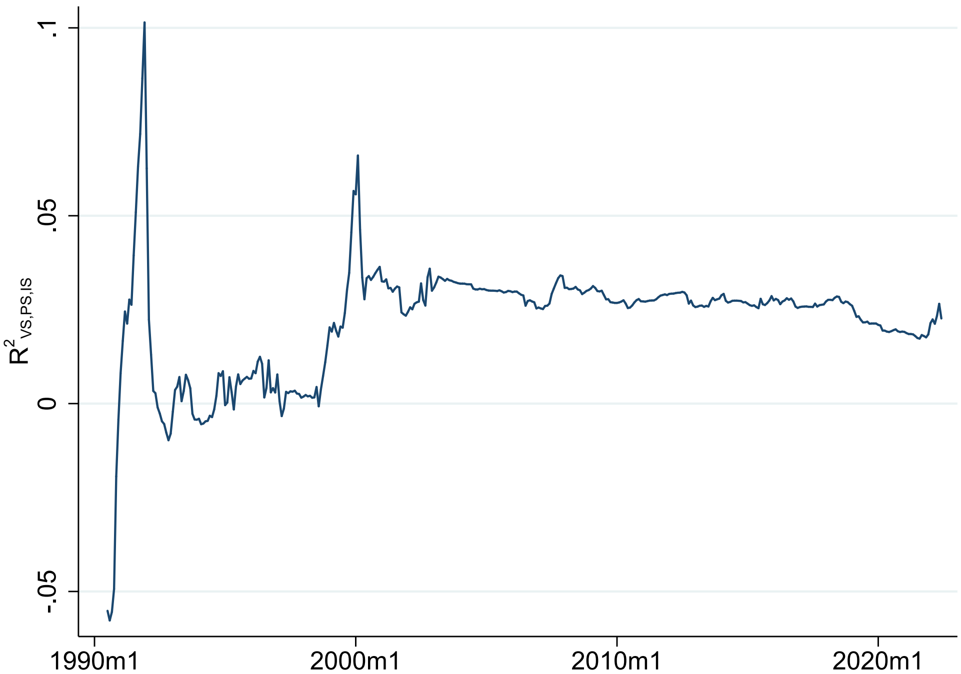

Figure E.2: **Out-of-sample performance of augmented forecast models relative to univariate value spread.** This figure visualizes the out-of-sample forecasting performance of a trivariate model using the value spread (VS), the profitability spread (PS), and the IVA-spread (IS) relative to the univariate model using only the value spread. Out-of-sample forecasts are based on expanding-window regressions and the OOS-R^2 is computed following Goyal and Welch (2008) as $R^2_{t,OOS} = 1 - \frac{\sum_{s \leq t}(r_s - \hat{r}_s)^2}{\sum_{s \leq t}(r_s - r^*_s)^2}$. where $\hat{r}$ is obtained from the respective trivariate forecast and r^* is a forecast based on the value spread.

Table E.1: **VAR results – Cross-section of firms**

This table reports the results from a VAR of log return, log book-to-market ratio, log return on equity (profitability), and investment value added with time fixed effects. Panel A reports VAR coefficients and Panel B reports variance-covariance matrix of the VAR residuals. We assign an equal weight to each cross-section. All variables are observed annually between 1965 and 2022. We report in parentheses two-way clustered standard errors by firm and year.

Panel A. Transition matrix

	r_{t-1}	roe_{t-1}	iva_{t-1}	bm_{t-1}	R^2
r_t	0.035	0.101	-0.114	0.046	0.010
	(0.034)	(0.026)	(0.029)	(0.014)	
roe_t	0.079	0.328	-0.093	-0.048	0.249
	(0.007)	(0.018)	(0.013)	(0.003)	
iva_t	0.015	-0.118	0.012	-0.033	0.045
	(0.004)	(0.014)	(0.017)	(0.003)	
bm_t	0.067	0.093	0.041	0.897	0.685
	(0.033)	(0.031)	(0.029)	(0.016)	

Panel B. Variance-covariance of residuals

	r_t	roe_t	iva_t	bm_t
r_t	0.099	0.007	0.000	-0.093
roe_t	0.007	0.015	-0.003	0.005
iva_t	0.000	-0.003	0.013	0.010
bm_t	-0.093	0.005	0.010	0.111

Table E.2: **Augmented VAR results – Cross-section of firms**

This table repeats the analysis from Table E.1 after adding log asset growth to the VAR and adding extra lags to appropriate state variables. All variables are observed annually between 1965 and 2022. We report two-way clustered standard errors by firm and year in parentheses. Panel C reports the news variance decomposition analogous to those in Table 3.

Panel A. Transition matrix

	r_{t-1}	roe_{t-1}	iva_{t-1}	bm_{t-1}	ag_{t-1}	r_{t-2}	roe_{t-2}	iva_{t-2}	ag_{t-3}	r_{t-3}	ag_{t-3}	r_{t-4}	ag_{t-4}	R^2
r_t	0.010	0.109	-0.020	0.025	-0.074	-0.030	0.090	-0.004	-0.061	0.017	-0.026	0.003	-0.024	0.013
	(0.032)	(0.027)	(0.026)	(0.012)	(0.014)	(0.022)	(0.022)	(0.019)	(0.011)	(0.018)	(0.011)	(0.014)	(0.012)	
roe_t	0.087	0.230	-0.069	-0.048	-0.027	0.044	0.104	-0.071	-0.023	0.017	-0.023	0.011	-0.012	0.268
	(0.007)	(0.016)	(0.014)	(0.003)	(0.004)	(0.005)	(0.009)	(0.009)	(0.003)	(0.004)	(0.003)	(0.004)	(0.004)	
iva_t	0.028	-0.087	-0.047	-0.017	0.037	0.022	-0.086	0.029	0.010	0.016	0.011	0.007	0.010	0.049
	(0.004)	(0.018)	(0.022)	(0.002)	(0.006)	(0.004)	(0.010)	(0.010)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	
bm_t	0.119	0.009	-0.108	0.942	0.098	0.106	-0.096	-0.046	0.061	0.023	0.022	0.019	0.030	0.703
	(0.031)	(0.030)	(0.026)	(0.013)	(0.015)	(0.022)	(0.025)	(0.022)	(0.012)	(0.017)	(0.012)	(0.014)	(0.012)	
ag_t	0.138	0.005	-0.047	-0.049	0.082	0.111	-0.028	-0.002	0.040	0.062	0.042	0.031	0.042	0.167
	(0.008)	(0.015)	(0.012)	(0.003)	(0.008)	(0.006)	(0.009)	(0.009)	(0.007)	(0.004)	(0.006)	(0.005)	(0.005)	

Panel B. Variance-covariance of residuals

	r_t	roe_t	iva_t	bm_t	ag_t
r_t	0.071	0.005	0.001	-0.067	0.002
roe_t	0.005	0.011	-0.002	0.004	0.002
iva_t	0.001	-0.002	0.009	0.007	0.005
bm_t	-0.067	0.004	0.007	0.080	0.006
ag_t	0.002	0.002	0.005	0.006	0.025

Panel C. Variance decomposition of return news

	σ (diag), ρ (off-diag)			Contribution to $\sigma^2_{N_r}$		
	N_{DR}	N_{roe}	N_{iva}	$-N_{DR}$	N_{roe}	N_{iva}
N_{DR}	0.071			0.080		
N_{roe}	-0.144	0.223		0.073	0.797	
N_{iva}	-0.460	-0.326	0.109	0.113	-0.253	0.189

Table E.3: **VAR results – Aggregate stock market**

This table reports the results from a VAR of market-level log return, log book-to-market ratio, log return on equity (profitability), and iva, as well as four additional state variables, term yield spread (TY), default spread (DEF), and value spread (VS). Panel A reports VAR coefficients and Panel B reports variance-covariance matrix of the VAR residuals. All variables are observed annually between 1965 and 2022. We report standard errors in parentheses and p-values from the F-test in the last column.

Panel A. Transition matrix

	r_{t-1}	roe_{t-1}	iva_{t-1}	bm_{t-1}	TY_{t-1}	DEF_{t-1}	VS_{t-1}	R^2	$p(F)$
r_t	0.052	1.137	-0.852	0.179	0.027	-0.046	0.216	0.109	0.355
	(0.149)	(1.211)	(1.076)	(0.084)	(0.023)	(0.065)	(0.169)		
roe_t	0.018	0.192	0.096	-0.005	-0.005	-0.006	-0.048	0.441	0.022
	(0.017)	(0.141)	(0.125)	(0.010)	(0.003)	(0.008)	(0.020)		
iva_t	0.005	-0.065	0.093	0.007	-0.001	0.009	0.022	0.078	0.442
	(0.021)	(0.172)	(0.152)	(0.012)	(0.003)	(0.009)	(0.024)		
bm_t	-0.033	-0.999	1.102	0.836	-0.034	0.056	-0.251	0.842	0.000
	(0.148)	(1.206)	(1.071)	(0.084)	(0.023)	(0.065)	(0.168)		
TY_t	-0.669	1.182	-3.337	-0.119	0.467	0.492	1.448	0.386	0.040
	(0.892)	(7.252)	(6.445)	(0.506)	(0.140)	(0.392)	(1.012)		
DEF_t	-0.323	5.051	0.821	0.196	-0.011	0.537	-0.196	0.580	0.003
	(0.272)	(2.214)	(1.968)	(0.154)	(0.043)	(0.120)	(0.309)		
VS_t	-0.140	0.827	-2.638	-0.140	0.001	0.083	0.602	0.678	0.000
	(0.095)	(0.769)	(0.683)	(0.054)	(0.015)	(0.042)	(0.107)		

Panel B. Variance-covariance of residuals

	r_t	roe_t	iva_t	bm_t	TY_t	DEF_t	VS_t
r_t	0.020	0.001	0.001	-0.020	-0.009	-0.013	-0.002
roe_t	0.001	0.000	0.000	-0.000	-0.004	-0.001	-0.000
iva_t	0.001	0.000	0.000	-0.000	-0.000	-0.001	-0.000
bm_t	-0.020	-0.000	-0.000	0.020	0.006	0.012	0.002
TY_t	-0.009	-0.004	-0.000	0.006	0.729	0.064	0.010
DEF_t	-0.013	-0.001	-0.001	0.012	0.064	0.068	0.004
VS_t	-0.002	-0.000	-0.000	0.002	0.010	0.004	0.008

Table E.4: **News decomposition – Subsample analysis**

This table mirrors the analysis in Table 3, but reports the decomposition for different subsamples for which we find a particularly high share of iva news in the composite cash-flow news in Table 4. In each case, we report a decomposition based on the baseline VAR (estimated across all observations, Panels A, C, and E), and based on a subsample VAR estimated only within the respective subsample (Panels B, D, and F).

Panel A. Recent (1990–2022), baseline VAR

	σ (diag), ρ (off-diag)			Contribution to $\sigma^2_{N_r}$		
	N_{DR}	N_{roe}	N_{iva}	$-N_{DR}$	N_{roe}	N_{iva}
N_{DR}	0.098			0.093		
N_{roe}	-0.194	0.275		0.101	0.725	
N_{iva}	-0.418	-0.314	0.148	0.117	-0.245	0.209

Panel B. Recent (1990–2022), subsample VAR

N_{DR}	0.062			0.040		
N_{roe}	-0.064	0.289		0.023	0.852	
N_{iva}	-0.640	-0.306	0.152	0.124	-0.274	0.236

Panel C. Growth (top MB-quintile), baseline VAR

N_{DR}	0.100			0.085		
N_{roe}	-0.392	0.277		0.184	0.647	
N_{iva}	-0.322	-0.285	0.157	0.086	-0.209	0.208

Panel D. Growth (top MB-quintile), subsample VAR

N_{DR}	0.124			0.144		
N_{roe}	-0.241	0.228		0.127	0.483	
N_{iva}	-0.537	-0.357	0.178	0.221	-0.269	0.294

Panel E. Unprofitable (bottom ROE-quintile), baseline VAR

N_{DR}	0.132			0.117		
N_{roe}	0.083	0.374		-0.055	0.936	
N_{iva}	-0.508	-0.447	0.202	0.181	-0.452	0.272

Panel F. Unprofitable (bottom ROE-quintile), subsample VAR

N_{DR}	0.109			0.102		
N_{roe}	0.241	0.275		-0.124	0.651	
N_{iva}	-0.747	-0.325	0.216	0.302	-0.333	0.402

Table E.5: **News decomposition – Size Quintiles**

This table mirrors the analysis in Table 3, but reports the decomposition for different size quintiles. All of the decompositions are based on the baseline VAR, that is, use the same transition matrix consistent with Table IV in Vuolteenaho (2002). We use NYSE breakpoints to construct size quintiles.

Panel A. Size Bin 1 (Small)

	σ (diag), ρ (off-diag)			Contribution to $\sigma^2_{N_r}$		
	N_{DR}	N_{roe}	N_{iva}	$-N_{DR}$	N_{roe}	N_{iva}
N_{DR}	0.112			0.098		
N_{roe}	-0.079	0.316		0.043	0.773	
N_{iva}	-0.522	-0.325	0.158	0.144	-0.251	0.193

Panel B. Size Bin 2

N_{DR}	0.094			0.091		
N_{roe}	-0.213	0.256		0.105	0.671	
N_{iva}	-0.454	-0.233	0.132	0.116	-0.161	0.178

Panel C. Size Bin 3

N_{DR}	0.082			0.087		
N_{roe}	-0.304	0.220		0.142	0.627	
N_{iva}	-0.414	-0.201	0.116	0.102	-0.133	0.174

Panel D. Size Bin 4

N_{DR}	0.073			0.085		
N_{roe}	-0.416	0.190		0.185	0.579	
N_{iva}	-0.323	-0.166	0.106	0.080	-0.107	0.178

Panel E. Size Bin 5 (Big)

N_{DR}	0.064			0.084		
N_{roe}	-0.527	0.160		0.224	0.534	
N_{iva}	-0.255	-0.135	0.092	0.063	-0.083	0.178

Table E.6: **News decomposition: Approximation errors**

This table mirrors the analysis in Table 3, but computes investment news and profitability news, respectively, as a residual from the identity relationship.

Panel A. Backing out N_{iva} as a residual

	σ (diag), ρ (off-diag)			Contribution to $\sigma^2_{N_r}$		
	N_{DR}	N_{roe}	N_{iva}	$-N_{DR}$	N_{roe}	N_{iva}
N_{DR}	0.088			0.078		
N_{roe}	-0.331	0.252		0.148	0.643	
N_{iva}	-0.432	-0.191	0.124	0.096	-0.121	0.157

Panel B. Backing out N_{roe} as a residual

N_{DR}	0.088			0.078		
N_{roe}	-0.235	0.239		0.099	0.576	
N_{iva}	-0.585	-0.138	0.139	0.144	-0.093	0.195

Table E.7: **News decomposition – Anomalies**

This table reports news variance decompositions for anomaly portfolios, formed as value-weighted long-short portfolios of top and bottom NYSE characteristic quintiles. We report standard deviations on the diagonals, and correlation coefficients for the off-diagonal elements in the left half of each panel. The right halves report the respective covariance term as a fraction of total return news variance. The off-diagonal terms multiply the respective covariance by -2 in the N_{DR}-column and by 2 in the N_{roe}-column. Firm-level news terms are obtained from a panel VAR with cross-sectional medians of all four identity variables.

Panel A. Book-to-market

	σ (diag), ρ (off-diag)			Contribution to $\sigma^2_{N_r}$		
	N_{DR}	N_{roe}	N_{iva}	$-N_{DR}$	N_{roe}	N_{iva}
N_{DR}	0.043			0.089		
N_{roe}	-0.619	0.099		0.250	0.461	
N_{iva}	-0.419	0.279	0.034	0.058	0.088	0.054

Panel B. Size

N_{DR}	0.036			0.070		
N_{roe}	-0.732	0.090		0.259	0.445	
N_{iva}	-0.638	0.281	0.032	0.080	0.089	0.056

Panel C. Profitability

N_{DR}	0.031			0.118		
N_{roe}	-0.209	0.077		0.123	0.725	
N_{iva}	-0.256	-0.331	0.042	0.081	-0.261	0.213

Panel D. Investment

N_{DR}	0.030			0.092		
N_{roe}	-0.657	0.067		0.274	0.473	
N_{iva}	-0.307	0.033	0.030	0.056	0.014	0.092

Panel E. Momentum

N_{DR}	0.041			0.085		
N_{roe}	-0.538	0.093		0.205	0.430	
N_{iva}	-0.752	0.324	0.036	0.110	0.107	0.063

Table E.8: **PVA — Market-to-book decomposition**

This table repeats the analysis in Table 2 for the four-way decomposition described in Appendix B.2. Again, all variables are observed annually between 1965 and 2022. In parentheses, we report two-way clustered standard errors by firm and year.

J	$-\sum_{j=1}^{J} \rho^{j-1} r_{t+j}$	$\sum_{j=1}^{J} \rho^{j-1} roe_{t+j}$	$\sum_{j=1}^{J} \rho^{j-1} iva_{t+j}$	$\sum_{j=1}^{J} \rho^{j-1} pva_{t+j}$	$\rho^J mb_{t+J}$
1	0.023	0.089	0.044	0.013	0.798
	(0.013)	(0.011)	(0.013)	(0.003)	(0.018)
5	0.070	0.253	0.125	0.049	0.502
	(0.025)	(0.045)	(0.040)	(0.014)	(0.033)
10	0.115	0.338	0.161	0.078	0.324
	(0.027)	(0.073)	(0.058)	(0.023)	(0.027)
15	0.125	0.402	0.172	0.097	0.233
	(0.029)	(0.089)	(0.067)	(0.031)	(0.022)

Table E.9: **PVA —Variance decomposition of return news**

This table repeats the analysis in Table 3.A for the four-way decomposition described in Appendix B.2. Again, all variables are observed annually between 1965 and 2022.

	σ (diag), ρ (off-diag)				Contribution to $\sigma^2_{N_r}$			
	N_{DR}	N_{roe}	N_{iva}	N_{pva}	$-N_{DR}$	N_{roe}	N_{iva}	N_{pva}
N_{DR}	0.088				0.078			
N_{roe}	-0.235	0.239			0.099	0.576		
N_{iva}	-0.432	-0.244	0.124		0.096	-0.147	0.157	
N_{pva}	-0.768	0.313	0.280	0.036	0.049	0.054	0.025	0.013

Table E.10: **Intangible adjusted market-to-book decomposition**

This table reports the results from decompositions of cross-sectional variation in intangible-adjusted market-to-book. We compute the discounted sums based on realizations over a multi-year horizon. To deal with delisting firms without introducing look-ahead bias, we form 25 portfolios, double-sorted by size and BM-quintiles each year. We then run cross-sectional regressions of each of the components on the market-to-book ratio. We then compute the relevant variables at the portfolio level for the subsequent T years. All variables are observed annually between 1975 and 2022. In parentheses, we report two-way clustered standard errors by firm and year.

J	$-\sum_{j=1}^{J} \rho^{j-1} r_{t+j}$	$\sum_{j=1}^{J} \rho^{j-1} roe_{t+j}$	$\sum_{j=1}^{J} \rho^{j-1} iva_{t+j}$	$\rho^{J} mb_{t+J}$
1	0.0284	0.0493	0.0406	0.8256
	(0.017)	(0.006)	(0.007)	(0.016)
5	0.1259	0.1335	0.1970	0.4433
	(0.035)	(0.022)	(0.024)	(0.027)
10	0.1685	0.1572	0.3181	0.2242
	(0.038)	(0.037)	(0.045)	(0.022)
15	0.1817	0.1587	0.3786	0.1334
	(0.042)	(0.047)	(0.061)	(0.014)

Table E.11: **Market-to-book decomposition: Gerakos and Linnainmaa (2018)**

This table repeats the analysis in Table 2 using a Gerakos and Linnainmaa (2018)'s market-to-book ratio decomposition. Gerakos and Linnainmaa (2018) decompose market-to-book ratio into a size effect (mb^s) and its orthogonal component (mb^o). Panel A and B report point-estimates of β-coefficient from

$$\sum_{j=1}^{J} \rho^{j-1} y_{i,t+j} = \alpha_t + \beta\, x_{i,t} + \varepsilon_{i,t}, \tag{E.2}$$

where $x \in \{mb^s, mb^o\}$. By our identity, the coefficients in the same row approximately sum to zero:

$$\begin{aligned} 0 \quad \approx \quad & \frac{Cov(-\sum_{j=1}^{J} \rho^{j-1} r_{t+j}, x_t)}{Var(x_t)} + \frac{Cov(\sum_{j=1}^{J} \rho^{j-1} roe_{t+j}, x_t)}{Var(x_t)} + \frac{Cov(\sum_{j=1}^{J} \rho^{j-1} iva_{t+j}, x_t)}{Var(x_t)} \\ & + \frac{Cov(\rho^J mb_{t+J} - mb_t, x_t)}{Var(x_t)} \end{aligned} \tag{E.3}$$

All variables are observed annually between 1965 and 2022. In parentheses, we report two-way clustered standard errors by portfolio and year.

Panel A. Size effect

J	$-\sum_{j=1}^{J} \rho^{j-1} r_{t+j}$	$\sum_{j=1}^{J} \rho^{j-1} roe_{t+j}$	$\sum_{j=1}^{J} \rho^{j-1} iva_{t+j}$	$\rho^J mb_{t+J} - mb_t$
1	0.0285	0.0886	0.0572	-0.1907
	(0.015)	(0.017)	(0.009)	(0.021)
5	0.0856	0.2169	0.1612	-0.5273
	(0.028)	(0.052)	(0.026)	(0.046)
10	0.1274	0.2559	0.2070	-0.6689
	(0.025)	(0.077)	(0.039)	(0.049)
15	0.1256	0.2845	0.2243	-0.7218
	(0.026)	(0.092)	(0.044)	(0.059)

Panel B. Orthogonal

J	$-\sum_{j=1}^{J} \rho^{j-1} r_{t+j}$	$\sum_{j=1}^{J} \rho^{j-1} roe_{t+j}$	$\sum_{j=1}^{J} \rho^{j-1} iva_{t+j}$	$\rho^J mb_{t+J} - mb_t$
1	0.0079	0.0903	0.0125	-0.1129
	(0.018)	(0.024)	(0.013)	(0.027)
5	0.0194	0.3090	0.0176	-0.3560
	(0.042)	(0.095)	(0.037)	(0.080)
10	0.0449	0.4607	-0.0152	-0.4905
	(0.044)	(0.140)	(0.047)	(0.130)
15	0.0685	0.5568	-0.0534	-0.5527
	(0.047)	(0.164)	(0.053)	(0.160)

Table E.12: **Predicting HML: Other predictors**

This table reports the result from monthly and annual forecasting regressions of HML returns—simple net returns rather than log returns—using the value spread, the profitability spread, the iva spread, and several other predictors proposed by previous research. The returns—the left-hand variable—are measured in percentage and all predictors are standardized for an easier comparison. In parentheses, we report Newey-West standard errors with a lag of 12 months (annual) and clustered by year (monthly, to reflect annual refreshing of accounting variables) respectively.

	Annual			Monthly		
	(1)	(2)	(3)	(4)	(5)	(6)
Value spread	2.884	6.069		0.168	0.472	
	(2.897)	(2.778)		(0.167)	(0.201)	
ROE spread	-3.728	-3.467	-2.478	-0.232	-0.220	-0.154
	(1.748)	(1.672)	(1.731)	(0.103)	(0.100)	(0.116)
IVA spread	-5.821	-2.935	-4.075	-0.320	-0.117	-0.205
	(2.107)	(2.913)	(2.923)	(0.122)	(0.154)	(0.162)
Market M/B	0.381	-0.297	2.191	-0.129	-0.186	0.013
	(2.941)	(3.021)	(2.287)	(0.205)	(0.193)	(0.193)
Default spread	1.212	1.161	1.824	-0.223	-0.198	-0.208
	(2.189)	(2.259)	(2.112)	(0.131)	(0.120)	(0.126)
Return spread	-0.108	-0.031	-1.707	0.380	0.342	0.368
	(2.609)	(2.294)	(2.449)	(0.139)	(0.147)	(0.138)
Inverse vol. (3y)	-2.862	-2.086	-2.163	-0.282	-0.179	-0.267
	(1.997)	(1.753)	(1.714)	(0.139)	(0.126)	(0.127)
Issuer-Rep. spread		-5.438	-3.023		-0.469	-0.252
		(3.007)	(2.859)		(0.189)	(0.161)
Observations	55	55	55	661	661	661
Adjusted R^2	0.148	0.203	0.159	0.041	0.051	0.044

Table E.13: **Predicting long-horizon returns and their cash-flow drivers**

In analogy to Table 5, this table seeks to understand the mirror image of return predictability from issuance based on the long-horizon cash-flow counterparts implied by the identity. We run the following regressions for $y \in \{r, roe, iva\}$:

$$\sum_{j=1}^{J} \rho^{j-1} y_{i,t+j} = \alpha_t + \beta_1 \, mb_{i,t} + \beta_2 \, r_{i,t} + \beta_3 \, roe_{i,t} + \beta_4 \, iva_{i,t} + \beta_5 \, IssRep_{i,t} + \varepsilon_{i,t}. \tag{21}$$

The variable $IssRep_{i,t} = 1$ if log share growth for portfolio i in year t exceeds 10% and $GH_{i,t} = -1$ if log share growth is below -0.5%. These thresholds correspond to those used by Greenwood and Hanson (2012) in the construction of the issuer-repurchaser spread. The estimates are based on realized variables of 25 portfolios double-sorted on size and market-to-book. All variables are observed annually between 1965 and 2022. In parentheses, we report standard errors clustered by portfolio and year for $J = 1$, and Hansen and Hodrick (1980) standard errors with $J - 1$ lags for $J = 5$ and 15.

	$\sum_{j=1}^{J} \rho^{J-1} r_{t+j}$			$\sum_{j=1}^{J} \rho^{J-1} roe_{t+j}$			$\sum_{j=1}^{J} \rho^{J-1} iva_{t+j}$		
J	1	5	15	1	5	15	1	5	15
$IssRep_t$	0.001	-0.002	0.036	-0.012	-0.056	-0.117	0.009	0.030	0.080
	(0.012)	(0.019)	(0.024)	(0.004)	(0.012)	(0.031)	(0.004)	(0.009)	(0.021)
mb_t	-0.037	-0.095	-0.122	0.067	0.218	0.360	0.039	0.121	0.206
	(0.016)	(0.035)	(0.039)	(0.008)	(0.034)	(0.086)	(0.006)	(0.024)	(0.057)
r_t	-0.001	-0.132	-0.192	0.035	-0.170	-0.489	0.091	0.258	0.293
	(0.102)	(0.088)	(0.092)	(0.026)	(0.059)	(0.099)	(0.017)	(0.035)	(0.058)
roe_t	0.136	0.283	0.488	0.408	1.520	3.144	-0.305	-0.986	-2.114
	(0.100)	(0.191)	(0.211)	(0.078)	(0.338)	(0.897)	(0.043)	(0.271)	(0.635)
iva_t	-0.277	-0.428	-0.653	-0.111	-0.470	-1.062	-0.001	0.160	0.610
	(0.163)	(0.359)	(0.359)	(0.082)	(0.310)	(0.673)	(0.075)	(0.255)	(0.445)
Observations	1100	1100	1100	1100	1100	1100	1100	1100	1100
R^2	0.049	0.129	0.158	0.822	0.763	0.711	0.575	0.618	0.579

F Robustness: Augmenting the VAR

As a test of the robustness of our findings to different specifications of the VAR, we run an augmented specification which adds additional lags to variables as in Vuolteenaho (2002)'s "long" VAR specification and adds log asset growth as an additional state variable. Adding these variables allows us to understand which potential drivers of iva might be responsible for the key aspects of the estimate of the transition matrix in the baseline VAR. We report the transition matrix and return variance decomposition in Table E.2.

Asset growth is a significant predictor of all other state variables, but the gains in adjusted R^2 relative to our baseline VAR system are minimal. More importantly, our key results continue to hold in the augmented VAR with qualitatively unchanged economic magnitudes.

G Variable construction

Computing implied annual dividend. For each security i, we want the return identity to hold for annual returns:

$$P_{i,t-1} = \frac{1}{1+R_{i,t}} \left(D_{i,t} + P_{i,t}\right), \tag{G.1}$$

where $t-1$ is June of year $t-1$ and t is June of year t. The problem is that the CRSP data has monthly returns and dividends. That is, prices, returns, and dividends in the monthly CRSP data satisfy the following relation instead (here, we drop the subscript i for brevity):

$$\begin{aligned}
P_{t-1} &= \tfrac{1}{1+R_{t-1+1/12}} \left(D_{t-1+1/12} + P_{t-1+1/12}\right) \\
P_{t-1+1/12} &= \tfrac{1}{1+R_{t-1+2/12}} \left(D_{t-1+2/12} + P_{t-1+2/12}\right) \\
&\vdots \\
P_{t-1+11/12} &= \tfrac{1}{1+R_{t-1+11/12}} \left(D_{t} + P_{t}\right),
\end{aligned}$$

which implies the following identity for annual returns:

$$P_{t-1} = \frac{D_{t-1+1/12}}{1+R_{t-1+1/12}} + \frac{D_{t-1+2/12}}{\left(1+R_{t-1+1/12}\right)\left(1+R_{t-1+2/12}\right)} + ... + \frac{D_t}{\left(1+R_{t-1+1/12}\right) \times ... \times \left(1+R_t\right)} + \frac{P_t}{1+R_t}.$$

In other words, with the annual return identity, we are imaging an investor who would reinvest all dividends back into the firm, so the effective annual dividends that make equation (G.1) hold is

$$D_t = \Pi_{s=2}^{12}\left(1 + R_{t-1+s/12}\right) D_{t-1+1/12} + \Pi_{s=3}^{12}\left(1 + R_{t-1+s/12}\right) D_{t-1+2/12} + ... + D_t.$$

In practice, calculating this quantity is tricky, as one needs to adjust for not just the returns within a year, but also the splits, etc.

Hence, we first compute annual returns and back out the implied dividend using equation (G.1). To do so, first compute the split-adjusted share prices:

$$\begin{aligned} P_{i,t-1} &= |PRC_{i,t-1}| \div CFACPR_{i,t-1} \\ P_{i,t} &= |PRC_{i,t}| \div CFACPR_{i,t} \end{aligned} \tag{G.2}$$

(It is important to adjust the delisting price for splits as well, since it is used to account for delisting returns for delisting securities.) Also compute cumulative one-year return as

$$R_{i,t} = \Pi_{s=1}^{12}\left(1 + R_{i,t-1+s/12}\right) - 1. \tag{G.3}$$

Then, compute the implied dividend as

$$D_{i,t} = (1 + R_t)\, P_{i,t-1} - P_{i,t}. \tag{G.4}$$

Since the $CFACPR$ is a rounded number, this formula could generate negative dividends. To prevent this from generating large negative implied dividends, set the implied dividend to be zero if the absolute dividend amount is within the 0.2% rounding error region.

An important question is whether or not the share adjustment in (G.2) account for share repurchases or issuances. It does not, meaning that the implied dividend also does not include share repurchases.[27] That is, $P_{i,t}$ is the June t price of one share purchased in June $t-1$ and held for a

[27] To allow the return identity in (G.1) to account for the effect of share repurchases or issuances, we would no longer be able to use cumulative 1-year returns as R_t and D_t could be negative (when issuance exceeds repurchases and dividends).

year, undeterred by share repurchases or issuances.

Market equity dynamics and firm-level market values. Equation (G.1) is in terms of the price of single share. To state this in terms of the market value of equity for each security, first ensure that the shares outstanding are adjusted for splits:

$$\begin{aligned} N_{i,t-1} &= SHROUT_{i,t-1} \times CFACSHR_{i,t-1} \\ N_{i,t} &= SHROUT_{i,t} \times CFACSHR_{i,t} \end{aligned} \tag{G.5}$$

Multiply both sides of equation (G.1) by $N_{i,t-1}$ and rearrange to get

$$M_{i,t-1}\left(1+R_{i,t}\right) = D_{i,t}^{total} + \frac{N_{i,t-1}}{N_{i,t}} M_{i,t},$$

where $M_{i,t} = P_{i,t}N_{i,t}$ and $D_{i,t}^{total} = D_{i,t}N_{i,t-1}$. Next, to compute the firm-level market equity and dividend, compute the sum of each side overall all share classes issued by the same firm:

$$\sum_i M_{i,t-1}\left(1+R_{i,t}\right) = D_t^{firm} + \sum_i \frac{N_{i,t-1}}{N_{i,t}} M_{i,t},$$

where $D_t^{firm} = \sum_i D_{i,t}^{total}$. Defining the firm-level market equity to be $M_t = \sum_i M_{i,t}$ and the value-weighted firm-level return to be $1 + R_t = \sum_i M_{i,t-1}\left(1+R_{i,t}\right)/M_t$, we obtain

$$\left(1+R_t\right) M_{t-1} = D_t^{firm} + \frac{N_{t-1}}{N_t} M_t, \tag{G.6}$$

where

$$\frac{N_{t-1}}{N_t} = \frac{1}{M_t}\sum_i \frac{N_{i,t-1}}{N_{i,t}} M_{i,t}. \tag{G.7}$$

Firm-level accounting values. Before merging the firm-level market values with accounting values, we adjust for the discrepancy in the market value of equity in Compustat and CRSP:

$$B_t = B_{t-1}^{dec} \times \frac{M_{CRSP,t-1}^{dec}}{M_{Comp,t-1}^{dec}} \times \left(\frac{N_{t-1}^{dec}}{N_t}\right)^{-1}$$
$$Y_t = NI_{t-1}^{dec} \times \frac{M_{CRSP,t-1}^{dec}}{M_{Comp,t-1}^{dec}} \times \left(\frac{N_{t-1}^{dec}}{N_t}\right)^{-1}$$

where

$$\frac{N_{t-1}^{dec}}{N_t} = \frac{1}{M_t} \sum_i \frac{N_{i,t-1}^{dec}}{N_{i,t}} M_{i,t}. \tag{G.8}$$

The firm-level identity. Take equation (G.6) and write it as

$$M_{t-1} = \frac{1}{1+R_t}\left(D_t^{firm} + \frac{N_{t-1}}{N_t} M_t\right). \tag{G.9}$$

Multiply and divide quantities to get

$$\begin{aligned} \frac{M_{t-1}}{B_{t-1}} \quad = \quad & \frac{1}{1+R_t} \times \frac{B_{t-1}+Y_t}{B_{t-1}} \times \frac{D_t^{firm} + \frac{N_{t-1}}{N_t} B_t}{B_{t-1}+Y_t} \\ & \times \left(\frac{D_t^{firm}}{D_t^{firm} + \frac{N_{t-1}}{N_t} B_t} + \frac{M_t}{B_t} \times \frac{\frac{N_{t-1}}{N_t} B_t}{D_t^{firm} + \frac{N_{t-1}}{N_t} B_t}\right). \end{aligned} \tag{G.10}$$

However, note that the dividend policy variables drop out of the approximation.

Firm-level accounting values. Suppose we define

$$N_{t-1} = \frac{\sum_i ME_{i,t-1} N_{i,t-1}}{\sum_i ME_{i,t-1}}$$

multiply both sides by the number of shares outstanding as of year $t-1$ such that

$$ME_{i,t-1} = \frac{1}{1+R_{i,t}}\left(D_{i,t}^{total} + ME_{i,t}\right), \tag{G.11}$$

where

$$ME_{i,t-1} = P_{i,t-1} \times SHROUT_{i,t-1} \tag{G.12}$$

$$ME_{i,t} = P_{i,t} \times SHROUT_{i,t-1} = |PRC_{i,t}| \times SHROUT_{i,t-1} \times \frac{CFACPR_{i,t}}{CFACPR_{i,t-1}} \tag{G.13}$$

$$D_{i,t}^{total} = D_{i,t} \times SHROUT_{i,t-1}. \tag{G.14}$$

Theory does not give a clear guidance on how to treat the delisting cash outflow. To prevent diluting the iva variable with delisting payouts, which are very different in nature from regular payouts, we treat delisting cash flows as part of market equity. That is, we add delisting dividends to market equity and set the delisting dividend to zero.

Composite portfolios. In our VAR analysis, we form composite portfolios of stocks by mixing it with the Treasury bill to mitigate the influence of outliers. To do this, we would like to compute composite portfolio values such that equation (6) holds. To do this, define

$$\begin{aligned}
ME_{c,t-1} &= ME_{t-1} \\
1 + R_{c,t} &= 0.9\left(1 + R_t\right) + 0.1\left(1 + R_{f,t}\right) \\
D_{c,t} &= 0.9D_t^{firm} + 0.1\left(1 - \rho\right)\left(1 + R_{f,t}\right) ME_{t-1} \\
ME_{c,t} &= 0.9ME_t + 0.1\rho\left(1 + R_{f,t}\right) ME_{t-1}
\end{aligned}$$

Use $\rho = 0.96$ to be consistent with Campbell 2017 bottom of p. 134 ("...a reasonable value for the loglinearization parameter ρ is in the range 0.95–0.96 if the time period is one year."). Note that the market equity dynamics holds with these composite variables:

$$ME_{c,t-1} = \frac{D_{c,t} + ME_{c,t}}{1 + R_{c,t}}$$

To compute the composite variable ratios, additionally define the following:

$$\begin{aligned}
BE_{c,t-1} &= 0.9BE_{t-1} + 0.1ME_{t-1} \\
Y_{c,t} &= 0.9Y_t + 0.1R_{f,t}ME_{t-1} \\
BE_{c,t} &= 0.9BE_t + 0.1\rho\left(1 + R_{f,t}\right) ME_{t-1}
\end{aligned}$$

Then, the ratio variables can be computed as usual based on the composite values defined before:

$$\frac{ME_{c,t-1}}{BE_{c,t-1}} = \frac{ME_{t-1}}{0.9BE_{t-1} + 0.1ME_{t-1}}$$

$$\frac{BE_{c,t-1}}{ME_{c,t-1}} = \frac{0.9BE_{t-1} + 0.1ME_{t-1}}{ME_{t-1}}$$

$$\frac{ME_{c,t}}{BE_{c,t}} = \frac{0.9ME_t + 0.1\rho\left(1 + R_{f,t}\right)ME_{t-1}}{0.9BE_t + 0.1\rho\left(1 + R_{f,t}\right)ME_{t-1}}$$

$$\frac{BE_{c,t}}{ME_{c,t}} = \frac{0.9BE_t + 0.1\rho\left(1 + R_{f,t}\right)ME_{t-1}}{0.9ME_t + 0.1\rho\left(1 + R_{f,t}\right)ME_{t-1}}$$

$$1 + ROE_{c,t} = 1 + \frac{Y_{c,t}}{BE_{c,t-1}} = 1 + \frac{0.9Y_t + 0.1R_{f,t}ME_{t-1}}{0.9BE_{t-1} + 0.1ME_{t-1}}$$

$$\frac{N_{c,t-1}}{N_{c,t}} = \frac{0.9ME_t\frac{N_{t-1}}{N_t} + 0.1\rho\left(1 + R_{f,t}\right)ME_{t-1}}{0.9ME_t + 0.1\rho\left(1 + R_{f,t}\right)ME_{t-1}}$$

$$G_{t+1} = \frac{D_{c,t} + \frac{N_{c,t-1}}{N_{c,t}}BE_{c,t}}{BE_{c,t-1} + Y_{c,t}}$$

Identity for a multiple-stock portfolio. The firm-specific return definition in equation (G.9) implies,

$$M_{k,t-1}\left(1 + R_{k,t}\right) = D_{k,t}^{firm} + \frac{N_{k,t-1}}{N_{k,t}}M_{k,t},$$

where k denotes a firm and variables with no subscript will now denote portfolio-level quantities. Taking a sum across firms on both sides,

$$\sum_k M_{k,t-1}\left(1 + R_{k,t}\right) = D_t^{agg} + L_tM_t,$$

where $D_t^{agg} \equiv \sum_k D_{k,t}^{firm}$ is the sum of all firm-level dividends, $L_t \equiv \frac{1}{M_t}\sum_k M_{k,t}\frac{N_{k,t-1}}{N_{k,t}}$ is the value-weighted change in the number of shares, and $M_t = \sum_k M_{k,t}$ is the sum of all market equities of firms in the portfolio. Defining the value-weighted return as $R_t \equiv \frac{1}{M_{t-1}} \sum_k M_{k,t-1} R_{k,t}$,

$$M_{t-1}\left(1+R_t\right) = D_t^{agg} + L_t M_t.$$

Multiply and divide portfolio-level accounting quantities defined as the sum of the firm-level quantities ($B_t \equiv \sum_k B_{k,t}$ and $Y_t \equiv \sum_k Y_{k,t}$) to get

$$\begin{aligned} \frac{M_{t-1}}{B_{t-1}} = & \quad \frac{1}{1+R_t} \times \underbrace{\frac{B_{t-1}+Y_t}{B_{t-1}}}_{1+ROE_t} \times \underbrace{\frac{D_t^{agg}+L_t B_t}{B_{t-1}+Y_t}}_{1+IVA_t} \\ & \times \left(\frac{D_t^{agg}}{D_t^{agg}+L_t B_t} + \frac{M_t}{B_t} \times \frac{L_t B_t}{D_t^{agg}+L_t B_t} \right), \end{aligned} \tag{G.15}$$

which gives us the portfolio-level measures of the variables in the identity.