

What Drives Anomaly Returns?

Lars A. Lochstoer and Paul C. Tetlock*

Columbia Business School

July 2016

Abstract

We provide novel evidence on which theories best explain stock return anomalies. Our estimates reveal whether anomaly returns arise from variation in the underlying firms' cash flows or their discount rates. For each of five well-known anomalies, we find that cash flow shocks explain more variation in anomaly returns than discount rate shocks. The cash flow and discount rate components of each anomaly's returns are negatively correlated. Most correlations between anomaly and market return components are small. Our evidence is inconsistent with theories of time-varying risk aversion and theories of common shocks to investor sentiment. It is most consistent with theories in which investors overextrapolate firm-specific cash flow news and those in which firm risk increases following negative cash flow news.

*Comments welcome. We thank seminar participants at the Chicago Asset Pricing Conference, Columbia University, Cornell University, London Business School, Swedish House of Finance conference, Francisco Gomes, and Alan Moreira for helpful comments. Contact information: Lochstoer: 405B Uris Hall, Columbia Business School, 3022 Broadway, New York, NY 10027. E-mail: LL2609@gsb.columbia.edu. Tetlock: 811 Uris Hall, Columbia Business School, 3022 Broadway, New York, NY 10027. E-mail: paul.tetlock@columbia.edu.

1 Introduction

Researchers in the past 30 years have uncovered robust patterns in stock returns that contradict classic asset pricing theories. A prominent example is that value stocks outperform growth stocks, even though these stocks are similarly risky by conventional measures. Myriad theories, behavioral and rational, attempt to explain such asset pricing anomalies. Yet widespread disagreement about the causes of these patterns remains because existing evidence is insufficient to differentiate competing explanations.

We contribute to this debate by providing novel evidence on the sources of anomaly returns. Rather than partitioning theories into those making behavioral or rational assumptions, we distinguish theories by their predictions of firms' cash flows and discount rates. Several theories predict that discount rate fluctuations drive variation in the returns of anomaly portfolios, whereas other theories predict that cash flow variation is more important. At one extreme, consider the model of noise trader risk proposed by De Long et al. (1990). Firm dividends (cash flows) are constant in this model, implying that all variation in returns arises from changes in discount rates. At the other extreme, consider the simplest form of the CAPM in which firm betas and the market risk premium are constant. Expected returns (discount rates) are constant in this setting, implying that all variation in returns arises from changes in expected cash flows. We introduce an empirical technique to decompose the variance in anomaly returns into cash flow and discount rate components, shedding new light on which theories can explain anomalies.

Our empirical work focuses on five well-known anomalies, based on value, size, profitability, investment, and issuance, and yields three sets of findings.¹ First, for all five anomalies, cash flows explain more variation in anomaly returns than discount rates. Second, for all five anomalies, shocks to cash flows and discount rates are strongly negatively correlated. Thus, firms with negative cash flow shocks tend to experience increases in discount rates.

¹We also find similar patterns in an unreported analysis of stock price momentum, as measured by Jegadeesh and Titman (1993).

This association contributes significantly to return variance in anomaly portfolios.

Third, anomaly cash flow and discount rate components exhibit weak correlations with market cash flow and discount rate components. In addition, there is little commonality in the cash flow and discount rate components of different anomaly returns beyond that arising from overlap in the sets of firms in anomaly portfolios. The correlations among the cash flow components of many anomalies' returns are insignificantly different from zero, and most correlations among the discount rate components are also low.

Our three sets of findings have important implications for theories of anomalies. First, theories in which discount rate variation is the primary source of anomaly returns, such as De Long et al. (1990), are inconsistent with the evidence on the importance of cash flow variation. Second, theories that emphasize commonality in discount rates, such as theories of time-varying risk aversion and those of common investor sentiment, are inconsistent with the low correlations among the discount rate components of anomaly returns. Third, theories in which anomaly cash flows are strongly correlated with market cash flows, such as Lettau and Wachter (2007), are inconsistent with the empirical correlations that are close to zero.

In contrast, theories of firm-specific biases in information processing and theories of firm-specific changes in risk are potentially consistent with these three findings. Such theories include behavioral models in which investors overextrapolate firms' cash flow news and rational models in which firm risk increases after negative cash flow shocks. In these theories, discount rate shocks amplify the effect of cash flow shocks on returns, consistent with the robustly negative empirical correlation between these shocks. These theories are also consistent with low correlations between anomaly return components and market return components.

Our approach builds on the seminal present-value decomposition introduced by Campbell and Shiller (1988) and applied to the firm-level by Vuolteenaho (2002). We exploit the present-value equation expressing each firm's book-to-market ratio in terms of expected returns and cash flows. To this end, we apply the clean surplus accounting relation of Ohlson (1995) to a log-linear approximation of book-to-market ratios, following Vuolteenaho

(2002) and Cohen, Polk, and Vuolteenaho (2003). We directly estimate firms' discount rates and expected cash flows using a vector autoregression (VAR) in which we impose the present-value relation. The VAR provides estimates of discount rates and expected cash flows at each horizon as a function of firm characteristics, such as profitability and investment, and aggregate variables, such as the risk-free rate. We differ from prior work in that we derive and analyze the implications of our firm-level estimates for interesting factor portfolios, such as the market, size, and value factors, to investigate the fundamental drivers of these factors' returns.

The premise of the approach is that investors use the characteristics in the VAR to form expectations of returns and cash flows, implying that their valuations are based on these characteristics. The characteristics could represent investors' perceptions of risk. We purposely select characteristics that serve as the basis for factor portfolios and betas in recent asset pricing models, such as the five-factor model of Fama and French (2015). Examples of such characteristics include book-to-market ratios, size, profitability, and investment. Firm characteristics could also represent investors' possibly mistaken beliefs about cash flows. In this spirit, we include characteristics such as share issuance that are featured in studies on asset pricing anomalies (see Daniel and Titman (2006) and Pontiff and Woodgate (2008)). Recognizing that these categories are not mutually exclusive, we design general tests that allow characteristics to forecast returns or earnings for any of the reasons above.

A key to our approach is that we aggregate firm-level VAR estimates rather than directly analyzing the returns and cash flows of anomaly portfolios in a VAR. Specifically, we consider five long-short quintile portfolios sorted by characteristics — book-to-market, investment, profitability, issuance, and size. We decompose the returns to these portfolios into cash flow and discount rate shocks based on the underlying firms' cash flow and discount rate shocks. This analysis allows us to test theories of anomalies as these typically apply to individual firms.

Moreover, the direct approach in which one analyzes the cash flows and returns of the

long-short portfolios obfuscates the cash flow and discount rate components of anomaly returns. Firms' weights in anomaly portfolios can change dramatically with the realization of stock returns and firms' changing characteristics. In the Appendix, we provide extreme examples in which firms' cash flows are constant but the direct VAR estimation suggests that all return variation in a rebalanced portfolio arises from cash flows. To our knowledge, our study is the first to recognize the pitfalls of the direct approach and offer a practical solution.

Complementing our main results on anomaly portfolios, the firm-level VAR yields insights into the sources of variation in individual firms' discount rates and cash flows. One notable finding is that the strongest predictors of long-run stock returns are book-to-market ratios and firm size, implying that large firms and those with high valuation ratios have significantly lower costs of equity capital. The high persistence of these variables helps to explain why they play important roles in long-run returns. Some patterns in short-run returns, such as the negative relation with investment, do not persist. Certain firms, such as those with low profitability and high investment rates, exhibit starkly different short-run and long-run discount rates. Thus, practitioners making capital budgeting decisions should exercise caution when applying short-run discount rates to long-run projects.

Our method and findings build on the growing body of research that exploits the present-value relation to investigate the relative importance of cash flows and discount rates in valuations. We contribute to this literature by characterizing the components of anomaly returns and relating them to each other as well as market return components. We build on the firm-level VAR introduced in Vuolteenaho's (2002) study of firm-level returns. Vuolteenaho (2002) finds that cash flow variation drives firm-level returns, but discount rate variation is important at the market level. The reconciliation to this tension is that there is more commonality in firms' discount rates than in their expected cash flows. Vuolteenaho (2002) does not consider anomaly portfolios, which are our primary focus.

Cohen, Polk, and Vuolteenaho (2003) use a portfolio approach to analyze the dynamics

of the value spread—i.e., the cross-sectional dispersion in book-to-market ratios. The study concludes that most of the spread comes from differences in expected cash flows. Our VAR approach allows us to use many characteristics beyond book-to-market ratios to forecast firms’ returns and cash flows. Because several of these characteristics predict returns and are correlated with book-to-market ratios, we infer that discount rate variation is more important than suggested by Cohen, Polk, and Vuolteenaho’s (2003) results. Our studies differ in that we analyze multiple anomalies and do so by aggregating estimates based on a firm-level VAR.

Lyle and Wang (2015) also apply the clean-surplus relation of Ohlson (1995) and log-linear techniques to relate book-to-market ratios to future cash flows and discount rates. They estimate the discount rate and cash flow components of firms’ book-to-market ratios by forecasting one-year returns using return on equity and book-to-market ratios. Lyle and Wang (2015) focus on stock return predictability at the firm level and do not analyze the sources of anomaly returns. Our work is related to studies that use the log-linear approximation of Campbell and Shiller (1988) for price-dividend ratios, typically applied to the market portfolio (see Campbell (1991), van Binsbergen and Koijen (2010), and Kelly and Pruitt (2013)). Lastly, our paper is related to the implied cost of capital literature (see, e.g., Claus and Thomas (2001) and Pastor, Sinha, and Swaminathan (2008)).

2 Theory

Empirical research identifies several asset pricing anomalies in which firm characteristics, such as firm profitability and investment, predict firms’ stock returns even after controlling for market beta. Theories of these anomalies propose that the properties of investor beliefs and firm cash flows vary with firm characteristics. Here we explain how decomposing anomaly returns into cash flow and discount rate shocks helps distinguish alternative explanations of anomalies.

The well-known value premium provides a useful setting for differentiating competing

theories. De Long et al. (1990) and Barberis, Shleifer and Vishny (1998) are examples of behavioral models that potentially explain this anomaly, while Zhang (2005) and Lettau and Wachter (2007) are examples of rational explanations.

To relate these models' predictions to our study, recall from Campbell (1991) that shocks to log stock returns can be decomposed approximately into shocks to expectations of future cash flows and returns:²

$$r_{i,t+1} - E_t[r_{i,t+1}] = CF_{i,t+1}^{shock} - DR_{i,t+1}^{shock}, \quad (1)$$

where

$$CF_{i,t+1}^{shock} = (E_{t+1} - E_t) \sum_{j=1}^{\infty} \kappa^{j-1} \Delta d_{i,t+j}, \quad (2)$$

$$DR_{i,t+1}^{shock} = (E_{t+1} - E_t) \sum_{j=2}^{\infty} \kappa^{j-1} r_{i,t+j}, \quad (3)$$

and where $\Delta d_{i,t+j}(r_{i,t+j})$ is the log of dividend growth (log of gross return) of firm i from time $t+j-1$ to time $t+j$, and κ is a log-linearization constant slightly less than 1. We define anomaly returns as the value-weighted returns of the stocks ranked in the highest quintile of a given characteristic minus the value-weighted returns of stocks ranked in the lowest quintile. We define anomaly cash flow shocks as the cash flow shocks to the top quintile portfolio minus the shocks to the bottom quintile portfolio. We similarly define anomaly discount rate shocks.

First, consider a multi-firm generalization of the De Long et al. (1990) model of noise trader risk. In this model, firm cash flows are constant but stock prices fluctuate because of random demand from noise traders. As the expectations in Equation (2) are rational, there are no cash flow shocks in this model. By Equation (1), all shocks to returns are due to discount rate shocks. Of course, the constant cash flow assumption is stylized and too

²The operator $(E_{t+1} - E_t)x$ is short-hand for $E_{t+1}[x] - E_t[x]$; the update in the expected value of x from time t to time $t+1$. The equation relies on a log-linear approximation of the price-dividend ratio around its sample average.

extreme. But, if one in the spirit of this model assumes that value and growth firms have similar cash flow exposures, the variance of net cash flow shocks to the long-short portfolio would be small relative to the variance of discount rate shocks. Thus, a finding that discount rate shocks only explain a small fraction of the return variance to the long-short portfolio would be inconsistent with this model. This theory does not make clear predictions about links between anomaly and market-level cash flows and discount rates.

Barberis, Shleifer, and Vishny (BSV, 1998) propose a model in which investors overextrapolate from long sequences of past firm earnings when forecasting future firm earnings. Thus, a firm that repeatedly experiences low earnings will be underpriced (a value firm) as investors are too pessimistic about its future earnings. The firm will have high expected returns as future earnings on average are better than investors expect. Growth firms will have low expected returns for analogous reasons. In this model, cash flow and discount rate shocks are intimately linked. Negative shocks to cash flows lead to low expected future cash flows. However, these irrationally low expectations manifest as positive discount rate shocks in Equations (2) and (3), as the econometrician estimates expected values under the objective probability measure. Thus, this theory predicts a strong negative correlation between cash flow and discount rate shocks at the firm and anomaly levels. This theory offers no clear guidance about the relation between anomaly and market return components.

Zhang (2005) provides a rational explanation for the value premium by modeling firms' production decisions. Persistent idiosyncratic productivity (earnings) shocks render firms, by chance, as either value or growth firms. Value firms, which have low productivity, have more capital than optimal because of adjustment costs. These firms' values are very sensitive to negative aggregate productivity shocks as they have little ability to smooth such shocks through disinvesting. Growth firms, on the other hand, have high productivity and suboptimally low capital stocks and therefore are not as exposed to negative aggregate shocks. Value (growth) firms' high (low) betas with respect to aggregate shocks justify their high (low) expected returns. Similar to BSV, this model predicts a negative relation between

firm cash flow and discount rate shocks. Different from BSV, the model predicts that the value anomaly portfolio has cash flow and discount rate shocks that are positively related to market cash flow and discount rate shocks on account of the high market beta of such a portfolio.

Lettau and Wachter (2007) propose a duration-based explanation of the value premium. In their model, growth firms are more exposed to shocks to market discount rates, which are not priced, and less exposed to market cash flow shocks, which are priced, than value firms. This model implies that cash flows shock to the long-short value portfolio are positively correlated with market cash flows and that discount rate shocks to the long-short portfolio are negatively correlated with market discount rates. It assumes low (actually, zero) correlation between discount rate and cash flow shocks.

In sum, models of anomaly returns have direct implications for the magnitudes and correlations of anomaly and market cash flow and discount rate shocks. We are unaware of any prior study that estimates these empirical moments. The anomaly theories apply to individual firms. Thus, one must analyze firm-level cash flow and discount rate shocks and then aggregate these into anomaly portfolio shocks. Extracting cash flow and discount rate shocks indirectly from dynamic trading strategies, such as the Fama-French value and growth portfolios, can lead to mistaken inferences as the trading itself confounds the underlying firms' cash flow and discount rate shocks. In the Appendix, we provide an example of a value-based trading strategy. The underlying firms only experience discount rate shocks, but the traded portfolio is driven solely by cash flow shocks as a result of rebalancing.

2.1 The Empirical Model

We begin our analysis by estimating a firm-level panel Vector Autoregression (VAR) as in Vuolteenaho (2002) to extract firm-level cash flow and discount rate shocks, relying on the following log-linear approximation for firms' book-to-market ratios:

$$bm_{i,t} \approx r_{i,t+1} - e_{i,t+1} + \kappa bm_{i,t+1}, \quad (4)$$

where $bm_{i,t}$ is the log of the book value of equity to the market value of equity of firm i in year t , $r_{i,t+1}$ and $e_{i,t+1}$ are the year $t + 1$ log return to equity and log accounting return on equity (ROE), respectively. The VAR imposes the present value relationship implied by the above approximation and it includes firm characteristics related to anomaly returns as described in the Data Section in addition to earnings, returns, and book-to-market ratios. Because the approach is standard in the literature, we relegate the description of the VAR to the Appendix.

We next analyze the sources of return variance for individual firms, the market portfolio, and anomaly portfolios, such as the long-short value minus growth portfolio. The VAR provides estimates of cash flow and discount rate shocks to firm-level returns: $CF_{i,t}^{shock}$ and $DR_{i,t}^{shock}$. We obtain portfolio-level variance decompositions by aggregating the portfolio constituents' $CF_{i,t}^{shock}$ and $DR_{i,t}^{shock}$. Because the firm-level variance decomposition applies to log returns, the portfolio cash flow and discount rate shocks are not simple weighted averages of the individual firms' cash flow and discount rate shocks. Therefore we approximate each firm's gross return using a second-order Taylor expansion around its current expected log return and then aggregate shocks to firms' gross returns using portfolio weights.

The first step in this process is to express gross returns in terms of the shocks to log returns using:

$$\begin{aligned} R_{i,t+1} &\equiv \exp(r_{i,t+1}) \\ &= \exp(E_t r_{i,t+1}) \exp(CF_{i,t+1}^{shock} - DR_{i,t+1}^{shock}), \end{aligned} \quad (5)$$

where $E_t r_{i,t+1}$ is the predicted value from the firm-level VAR regressions. A second-order

expansion at time t around a value of zero for both of the shocks yields:

$$R_{i,t+1} \approx \exp(E_t r_{i,t+1}) \left\{ 1 + CF_{i,t+1}^{shock} + \frac{1}{2} (CF_{i,t+1}^{shock})^2 - DR_{i,t+1}^{shock} + \frac{1}{2} (DR_{i,t+1}^{shock})^2 + CF_{i,t+1}^{shock} DR_{i,t+1}^{shock} \right\}. \quad (6)$$

We find that this approximation works very well in practice. Next we define:

$$CF_{i,t+1}^{level_shock} \equiv \exp(E_t r_{i,t+1}) \left\{ CF_{i,t+1}^{shock} + \frac{1}{2} (CF_{i,t+1}^{shock})^2 \right\}, \quad (7)$$

$$DR_{i,t+1}^{level_shock} \equiv \exp(E_t r_{i,t+1}) \left\{ DR_{i,t+1}^{shock} - \frac{1}{2} (DR_{i,t+1}^{shock})^2 \right\}, \quad (8)$$

$$CFDR_{i,t+1}^{cross} \equiv \exp(E_t r_{i,t+1}) CF_{i,t+1}^{shock} DR_{i,t+1}^{shock}. \quad (9)$$

as the cash flow and discount rate shocks to firm returns measured in levels. For a portfolio with weights $\omega_{i,t}^p$ on firms, we can approximate the portfolio return measured in levels using:

$$R_{p,t+1} - \sum_{i=1}^n \omega_{i,t}^p \exp(E_t r_{i,t+1}) \approx CF_{p,t+1}^{level_shock} - DR_{p,t+1}^{level_shock} + CFDR_{p,t+1}^{cross}, \quad (10)$$

where

$$CF_{p,t+1}^{level_shock} = \sum_{i=1}^n \omega_{i,t}^p CF_{i,t+1}^{level_shock}, \quad (11)$$

$$DR_{p,t+1}^{level_shock} = \sum_{i=1}^n \omega_{i,t}^p DR_{i,t+1}^{level_shock}, \quad (12)$$

$$CFDR_{p,t+1}^{cross} = \sum_{i=1}^n \omega_{i,t}^p CFDR_{i,t+1}^{cross}. \quad (13)$$

We decompose the variance of portfolio returns using

$$\begin{aligned} var(\tilde{R}_{p,t+1}) &\approx var\left(CF_{p,t+1}^{level_shock}\right) + var\left(DR_{p,t+1}^{level_shock}\right) \\ &\quad - 2cov\left(CF_{p,t+1}^{level_shock}, DR_{p,t+1}^{level_shock}\right) \\ &\quad + var\left(CFDR_{p,t+1}^{cross}\right), \end{aligned} \quad (14)$$

where $\tilde{R}_{p,t+1} \equiv R_{p,t+1} - \sum_{i=1}^n \omega_{i,t}^p \exp(E_t r_{i,t+1})$. We ignore covariance terms involving $CFDR_{p,t+1}^{cross}$ as these are very small in practice.

The VAR offers a parsimonious, reduced-form model of the cross-section of expected cash flows and discount rates at all horizons. In the Appendix, we show how the VAR specification is related to standard asset pricing models. In particular, the VAR specification concisely summarizes the dynamics of expected cash flows and returns, even when both consist of multiple components fluctuating at different frequencies. Fundamentally, shocks to a firm's discount rates arise from shocks to the product of the firm-specific quantity of risk and the aggregate price of risk, as well as shocks to the risk-free rate.

When analyzing cash flow and discount rate shocks to long-short portfolios, we obtain the anomaly cash flow (discount rate) shock as the difference in the cash flow (discount rate) shocks between the long and short portfolios. Taking the value anomaly as an example, suppose the long value portfolio and the short growth portfolio have the same betas with respect to all risk factors except the value factor. The VAR implies that discount rate shocks to this long-short portfolio can only arise from three sources: 1) shocks to the spread in the factor exposure between value and growth firms; 2) shocks to the price of risk of the value factor; or 3) shocks to the difference in return variance between the two portfolios. The third possibility arises because we analyze log returns. Similarly, cash flow shocks to this long-short portfolio only reflect these portfolios' differential exposure to cash flow factors.

3 Data

We use Compustat and CRSP data from 1962 through 2015 to estimate the components in the present-value equation. Our analysis requires panel data on firms' returns, book values, market values, earnings, and other accounting information, as well as time series data on factor returns, risk-free rates, and price indexes. Because computations of certain variables in the VAR require three years of historical accounting information, our estimation focuses on the period from 1964 through 2015.

We obtain all accounting data from Compustat, though we augment our book data with that from Davis, Fama, and French (2000). We obtain data on stock prices, returns, and

shares outstanding from the Center for Research on Securities Prices (CRSP). We obtain one-month and one-year risk-free rate data from one-month and one-year yields of US Treasury Bills, which are available on Kenneth French’s website and the Fama Files in the Monthly CRSP US Treasury Database, respectively. We obtain inflation data from the Consumer Price Index (CPI) series in CRSP.

We impose sample restrictions to ensure the availability of high-quality accounting and stock price information. We exclude firms with negative book values as we cannot compute the logarithms of their book-to-market ratios, which are key elements in the present-value equation. We include only firms with nonmissing market equity data at the end of the most recent calendar year. Firms also must have nonmissing stock return data for at least 225 days in the past year, which is necessary to accurately estimate stock return variance as discussed below. We exclude firms in the bottom quintile of the size distribution for the New York Stock Exchange to minimize concerns about illiquidity and survivorship bias. Lastly, we exclude firms in the finance and utility industries because accounting and regulatory practices distort these firms’ valuation ratios and cash flows. We impose these restrictions *ex ante* and compute subsequent book-to-market ratios, earnings, and returns as permitted by data availability. We use CRSP delisting returns and assume a delisting return is -90% in the rare cases in which the delisting return is missing.

When computing a firm’s book-to-market ratio, we adopt the convention of dividing its book equity by its market equity at the end of the June immediately after the calendar year of the book equity. With this convention, the timing of market equity coincides with the beginning of the stock return measurement period, allowing us to use the clean-surplus equation below. We compute book equity using Compustat data when available, supplementing it with hand-collected data from the Davis, Fama, and French (2000) study. We adopt the Fama and French (1992) procedure for computing book equity. Market equity is equal to shares outstanding times stock price per share. We sum market equity across firms that have more than one share class of stock. We define $\ln BM$ as the natural log of book-to-market

ratio.

We compute log stock returns in real terms to ensure consistency with lnBM and our log earnings measures below, which are denominated in real terms. We set real log annual stock returns equal to log returns minus the log of inflation, as measured by the log change in the CPI. Following the convention in asset pricing, we compute annual returns from the end of June to the following end of June. The benefit of this timing convention is that investors have access to December accounting data prior to the ensuing June-to-June period over which we measure returns.

Our primary measure of earnings is the log of clean-surplus return on equity, $\ln ROE^{CS}$, though we also compute log return on equity, $\ln ROE$, for comparison. We focus on clean-surplus earnings because our framework requires consistency between firms' book-to-market ratios, returns, and earnings. We define log clean-surplus earnings as in Ohlson (1995) and Vuolteenaho (2002), using log stock returns minus the change in log book-to-market ratios:

$$\ln ROE_{i,t+1}^{CS} \equiv r_{i,t+1} + \kappa bm_{i,t+1} - bm_{i,t}. \quad (15)$$

We extract this measure of clean-surplus earnings from the data as in Equation (15), thereby ensuring that the log-linear model holds for each firm at each time.

The log of return on equity is defined as log of one plus net income divided by last year's inferred book equity, where we substitute income before extraordinary items if net income is unavailable. We infer last year's book equity using current accounting information and the clean surplus relation—i.e., last year's book equity is this year's book equity plus dividends minus net income. We subtract the log inflation rate, based on the average CPI during the year, from log return on equity to obtain $\ln ROE$. We winsorize both earnings measures at $\ln(0.01)$ when earnings is less than -99%. We follow the same procedure for log returns and for log firm characteristics that represent percentages with minimum bounds of -100%. Alternative winsorizing or truncation procedures have little impact on our results.

Figures 1A and 1B compare clean-surplus earnings ($\ln ROE^{CS}$) with return on equity

($\ln ROE$) for two large, well-known firms, Apple and Caterpillar, in different industries. The figures show that the two earnings series closely track each other in most years. Large discrepancies occasionally arise from share issuance or merger events, which can cause violations of the clean surplus equation.

We compute several firm characteristics that predict short-term stock returns in historical samples. We compute each firm’s market equity (ME) or size as shares outstanding times share price. Following Fama and French (2015), we compute profitability (Prof) as annual revenues minus costs of goods sold, interest expense, and selling, general, and administrative expenses, all divided by book equity from the same fiscal year.³ Following Cooper, Gulen, and Schill (2008) and Fama and French (2015), we compute investment (Inv) as the annual percentage growth in total assets. Following Pontiff and Woodgate (2008), we compute share issuance (Issue) as the percentage change in adjusted shares outstanding over the past 36 months. We transform each of these four measures by adding one and taking its log, resulting in the following variables: $\ln ME$, $\ln Prof$, $\ln Inv$, and $\ln Issue$. We also subtract the log of gross domestic product from $\ln ME$ to ensure stationarity. We use an alternative stationary measure of firm size (SizeWt), equal to firm market capitalization divided by the total market capitalization of all firms in the sample, when applying value weights to firms’ returns for the purpose of forming portfolios.

We compute stocks’ annual return variances based on daily excess log returns, which are daily log stock returns minus the daily log return from the one-month risk-free rate as of the beginning of the month. A stock’s realized variance is simply the annualized average value of its squared daily excess log returns during the past year. We do not subtract each stock’s mean squared excess return to minimize estimation error in this calculation. We transform realized variance by adding one and taking its log, resulting in the variable $\ln RV$.

Table 1 presents summary statistics for the variables in our analysis. For ease of interpretation, we show statistics for nominal annual stock returns ($AnnRet$), nominal risk-free

³Novy-Marx (2013) uses a similar definition for profitability, except that the denominator is total assets instead of book equity.

rates (Rf), and inflation (Inflat) before we apply the log transformation. Similarly, we summarize stock return volatility (Volat) instead of log variance. We multiply all statistics by 100 to convert them to percentages, except the lnBM and lnME statistics, which retain their original scale.

Panel A displays the number of observations, means, standard deviations, and percentiles for each variable. The median firm has a log book-to-market ratio of -0.66 , which translates into a market-to-book ratio of $e^{0.66} = 1.94$. Valuation ratios range widely, as shown by the 10th and 90th percentiles of market-to-book ratios of 0.75 and 5.93. The variation in stock returns is substantial, ranging from -40% to 66% for the 10th and 90th percentiles.

Panel B shows that most correlations among the variables are modest. One exception is the mechanical correlation between the alternative size measures. The variables with the strongest correlations with book-to-market ratios are the firm return and size measures, which exhibit negative correlations ranging from -0.28 to -0.37 . The positive correlation of 0.39 between issuance and investment could be partly driven by mergers that trigger stock issuance and investment. Issuance and mergers cause deviations in clean-surplus accounting for the standard return on equity (lnROE) measure. Lastly, the substantial correlation of 0.55 between investment and clean-surplus return on equity is consistent with the well-known relationship between firm investment and cash flows.

4 VAR Estimation

We estimate the firm-specific and common predictors of firms' (log) returns and cash flows using a panel VAR system. Natural predictors of returns include characteristics that serve as proxies for firms' risk exposures or stock mispricing. As predictors of earnings, we use characteristics based on accounting metrics and market prices that forecast firm cash flows in theory and practice.

Our primary VAR specification includes eight firm-specific characteristics as predictors of firm returns and cash flows. Two firm characteristics are the lagged values of the dependent

variables ($\ln\text{Ret}$ and $\ln\text{ROE}^{\text{CS}}$). Five firm characteristics are those used in constructing the anomaly portfolios: $\ln\text{BM}$, $\ln\text{Prof}$, $\ln\text{Inv}$, $\ln\text{ME}$, and $\ln\text{Issue}$. The eighth firm characteristic is log realized variance ($\ln\text{RV}$), which captures potential differences between expected log returns and the log of expected returns as explained below. We standardize each independent variable by its full-sample standard deviation to facilitate interpretation of the regression coefficients. The only exceptions are $\ln\text{BM}$, which is not standardized to enable imposing the present-value relation in the VAR estimation, and the two lagged dependent variables. All log return and log earnings forecasting regressions include the log real risk-free rate ($\ln\text{Rf}$) to capture common time-series variation in firm valuations resulting from changes in market-wide discount rates.

Standard discount rates are based on expected returns, not expected log returns. Yet log returns must be the dependent variable in our regressions to be consistent with the log-linearization of book-to-market ratios. Including $\ln\text{RV}$ as a predictive variable in the VAR helps us isolate differences in expected log returns and the log of expected returns. Assuming annual stock returns are lognormally distributed, the expected difference between our dependent variable and standard discount rates is equal to half the variance of log returns, which is likely to be reflected in the predictive coefficient on $\ln\text{RV}$. Even if expected returns are unpredictable, we will find that stock return variance negatively forecasts log returns. However, the empirical results below indicate that $\ln\text{RV}$ is not a statistically significant predictor of either log returns or log earnings.

We estimate a first-order autoregressive system, allowing for one lag of each characteristic. A first-order VAR allows us to estimate the long-run dynamics of log returns and log earnings based on the short-run properties of a broad cross section of firms. We do not need to impose restrictions on which firms survive for multiple years, thereby mitigating statistical noise and survivorship concerns. As a robustness check, we investigate the second-order VAR specification and find very similar results as the second lags of the characteristics add little explanatory power.

The VAR system also includes regressions in which we forecast firm-specific and aggregate variables using a parsimonious specification. The only predictors of each firm characteristic are the firm’s own lagged value of its characteristic and the firm’s lagged log book-to-market ratio. For example, the only predictors of log investment are lagged log investment and lagged log book-to-market ratio. This restriction improves estimation efficiency without significantly reducing the explanatory power of the regressions. We model the real risk-free rate as a simple first-order autoregressive process.

The main concern with our panel VAR specification is that it omits an important common component in firms’ expected cash flows and discount rates. We address this issue in Section 7 by considering alternative VAR specifications in which we include the market-wide valuation ratio along with interactions with firm-level characteristics. Ultimately, our primary specification omits aggregate variables, except the risk-free rate, because they do not materially increase the explanatory power of the return and cash flow forecasting regressions and result in extremely high standard errors in return variance decompositions. Of course, it is possible that another not-yet-identified aggregate variable would materially improve on our forecasting regressions.

We conduct all tests using standard equal-weighted regressions, but we find that our findings are robust to applying value weights to each observation. Table 2 displays the coefficients for the regressions in which we forecast firms’ log returns and earnings in the first two columns. The third column in Table 2 shows the implied coefficients on firms’ log book-to-market ratios based on the clean surplus relation between log returns, log earnings, and log valuations. Standard errors are clustered by year and firm, following Petersen (2009), and appear in parentheses below the coefficients.

The findings in the log return regressions are related to the large literature on short-horizon forecasts of returns. We find that firms’ log book-to-market ratios and profitability are positive predictors of their log returns at the annual frequency, whereas log investment and share issuance are negative predictors of log returns. Log firm size and realized variance

weakly predict returns with the expected negative signs, though these coefficients are not statistically significant in this multivariate panel regression. The largest standardized coefficients are those for firm-specific log book-to-market ($0.037 = 0.83 \times 0.045$), profitability (0.043), and investment (-0.048). These predictors have standardized impacts of 3.7% to 4.8% on expected one-year log returns.

The second column of Table 2 shows the regressions predicting log earnings at the annual frequency. The main result is that log book-to-market ratio is by far the strongest predictor of log earnings. The coefficient on lagged lnBM is -0.143 , which is a standardized coefficient of -0.119 . The two other strong predictors of log earnings are the logs of firm-level returns and profitability, which have standardized coefficients of 0.060 (0.507×0.118) and 0.037. Other significant predictors of log earnings include the logs of firm-level issuance, size, and past earnings. Each of these variables exhibits a standardized impact of 1.3% to 1.4%.

The third column in Table 2 shows the implied coefficients of each lagged characteristic in a regression predicting log book-to-market ratios. Log book-to-market ratios are quite persistent as shown by the 0.846 coefficient on lagged log book-to-market. More interestingly, log investment and log issuance are significant positive predictors of log book-to-market, meaning that market-to-book ratios tend to decrease following high investment and issuance. These relations play a role in the long-run dynamics of expected log earnings and log returns of firms with high investment and issuance. Analogous reasoning applies to the positive coefficient on lagged log returns, which is statistically significant at the 5% level.

Table 3A shows regressions of firm characteristics on lagged characteristics and lagged book-to-market ratio. The most persistent characteristic is log firm size, which has a persistence coefficient of 0.973. We can, however, reject the hypothesis that this coefficient is 1.000, based on standard errors with firm and year clustering. The persistence coefficients on the logs of profitability, issuance, and realized variance range between 0.678 and 0.711. The persistence coefficients on the log of investment is just 0.154. All else equal, characteristics with high (low) persistence coefficients will be more important determinants of long-run cash

flows and discount rates. Lagged log book-to-market is a significant predictor of the logs of profitability, investment, issuance, and realized variance, but the incremental explanatory power from lagged valuations is modest in all regressions except the investment regression.

Table 3B shows that the aggregate variable, the lagged real risk-free rate (lnRf), is reasonably persistent, though not as persistent as firm size and valuation ratios. The persistence of the log real risk-free rate is 0.602. This estimate has little impact on expected long-run returns and cash flows simply because the risk-free rate is not a significant predictor of returns or cash flows, as shown in Table 2.

We now translate the VAR coefficients into estimates of the discount rate components of firms' log book-to-market ratios. Figure 2A plots the patterns in the implied cumulative coefficients for predicting log returns at horizons (N) ranging from 1 to 20 years. We compute the cumulative coefficients for predicting log returns by summing expected log returns across horizons, discounting by κ , enabling us to express the N -year discount rate component ($\widetilde{DR}_{i,t}^{(N)}$) as:

$$\widetilde{DR}_{i,t}^{(N)} = E_t \sum_{j=1}^N \kappa^{j-1} \tilde{r}_{i,t+j}, \quad (16)$$

where a tilde above a variable refers to its demeaned value. Similarly, Figure 2B plots the cumulative coefficients for predicting log earnings at horizons from 1 to 20 years. We obtain the N -year cash flow component of valuations ($\widetilde{CF}_{i,t}^{(N)}$) from the equation:

$$\widetilde{CF}_{i,t}^{(N)} = \sum_{j=1}^N \kappa^{j-1} E_t [\tilde{e}_{i,t+j}]. \quad (17)$$

These cumulative coefficients allow us to represent the discount rate and cash flow components in log book-to-market ratios from years 1 through 20 as affine functions of the characteristics in year 0. Appendix B explains how to compute $\widetilde{CF}_{i,t}^{(N)}$ and $\widetilde{DR}_{i,t}^{(N)}$ in terms of the VAR coefficients and firm characteristics.

Figure 2 shows that book-to-market and size are the most important predictors of long-run discount rates. The 20-year coefficient on log book-to-market is 25.8%, while the coefficient on log size is -13.6%. The high persistence of both variables implies that their

long-run impacts on valuation are much larger than their short-run impacts. In contrast, some effective predictors of short-run returns, such as log investment, have little long-run impact mainly because they are not highly persistent. In addition, investment positively predicts book-to-market ratios, which limits the extent to which its long-run impact can be negative. The long-run value and size coefficients imply that investors heavily discount the cash flows of value firms, whereas they pay more for the cash flows of large firms. Other notable predictors of 20-year cumulative log returns include log firm profitability and realized variance, which have coefficients of 12.1% and -8.6%, respectively. The negative effect of realized variance could arise because of the difference between expected log returns and log expected returns, or because realized variance negatively forecasts returns as found in Ang et al. (2006).

Figure 3 shows that book-to-market and size are also the most important predictors of long-run cash flows. The coefficients on log book-to-market and log size are -58.7% and -14.1%, respectively, for predicting cumulative log earnings at the 20-year horizon. Interestingly, log issuance, which positively predicts log earnings at the one-year horizon, is actually a negative predictor of long-run cash flows. This pattern is another consequence of the joint dynamics of issuance and book-to-market, as noted above. Imposing the present-value relation is essential for inferring the long-run dynamics of cash flows and returns.

To illustrate the importance of using long-run discount rates, we consider two alternative ways of computing a firm's discount rate. We contrast annualized infinite-horizon (i.e., long-run) coefficients based on the VAR with naive discount rates obtained by extrapolating short-horizon regressions. We compute the naive discount rate by extrapolating the one-year discount rate (expected log return), assuming expected log returns are constant at the one-year rate indefinitely. Thus, the naive rate is simply the one-year discount rate, $\widetilde{DR}^{(1)}$. The long-run discount rate is the annualized infinite-horizon discount rate component of firms' valuation ratios, $(1 - \kappa)\widetilde{DR}$, which takes into account the joint dynamics of firm and common characteristics and \widetilde{DR} is defined in Appendix B as the limit of $\widetilde{DR}^{(N)}$ as N

approaches infinity.

In Table 4, we present the short-run and long-run discount rates ($\widetilde{DR}^{(1)}$ and $(1 - \kappa)\widetilde{DR}$) along with standard errors in parentheses. The short-run standard errors are the same as those in the log return regression in the VAR. We compute the long-run standard errors by applying the delta method to the covariance matrix of the estimated A matrix coefficients. The last row in the table shows that the 9.53% volatility of short-run (i.e., naive long-run) expected returns vastly exceeds the 1.42% volatility of long-run expected returns. The long-run standard errors are much smaller than the short-run standard errors, with the exception of the long-run standard errors on the firm size coefficients, which are imprecisely estimated primarily because size is extremely persistent.

Figure 4 graphically compares the impact of each characteristic on the naive and long-run discount rates. The differential impacts on the two discount rates are stark for the investment, profitability, and book-to-market characteristics. For example, a one standard deviation increase in a firm's log investment is associated with a 0.16% lower long-run discount rate. However, if one naively extrapolates the one-year discount rate, a standardized increase in investment is associated with a 4.79% lower long-run discount rate. These magnitudes demonstrate that applying the wrong discount rate has severe consequences for firm and project valuation. Notably, the valuation error from extrapolation is small in the case of size because the extremely high persistence of firm size is reasonably consistent with the extrapolation assumption.

In summary, naive extrapolation of short-run discount rates produces erroneously high valuations for firms with high investment, low profitability, and low market-to-book ratios. For example, naive overvaluation is severe for unprofitable growth firms that invested aggressively during the technology boom of the late 1990s.

5 Firm-level Analysis

We now examine the decomposition of firms' log book-to-market ratios and returns implied by the regression results. We first analyze the correlations and covariances between total log book-to-market ($\ln BM$) and its two components (CF and DR). Panel A of Table 5 shows that DR and CF variation respectively account for 19.0% and 47.3% of variation in valuation ratios. Interestingly, covariation between DR and CF tends to amplify return variance, contributing a highly significant amount (33.8%) of variance. The last column shows that the correlation between the CF and DR components is negative and large at -0.564 . In economic terms, this correlation means that low expected cash flows are associated with high discount rates.

Panel B reveals a similar variance decomposition for firm returns. In particular, discount rate and cash flow shocks respectively account for 20.9% and 52.2% of return variance, and their covariance accounts for the remaining 27.0% of variance. The negative correlation between CF and DR shocks is pronounced at -0.409 . The negative correlation in cash flow and discount rate shocks could arise for behavioral or rational reasons. Investor overreaction to positive firm-specific cash flow shocks could lower effective firm discount rates (negative discount rate shocks). Alternatively firms with negative cash flow shocks could become more exposed to systematic risks, increasing their discount rates (positive discount rate shocks). Our decomposition indicates that discount rate variation is somewhat more important than suggested by prior studies, such as Cohen, Polk, and Vuolteenaho (2003).

A stylized example of an economy sheds light on the finding that discount rate variation contributes significantly to variation in valuations. Suppose the economy consists of four firms with cash flow (CF), discount rate (DR), and log book-to-market ratios ($bm = DR -$

CF) given by:

$$CF_1 = 0, DR_1 = 1, bm_1 = 1 \quad (18)$$

$$CF_2 = -1, DR_2 = 0, bm_2 = 1 \quad (19)$$

$$CF_3 = 2, DR_3 = 1, bm_3 = -1 \quad (20)$$

$$CF_4 = 1, DR_4 = 0, bm_4 = -1 \quad (21)$$

Applying the sorting method of Cohen, Polk, and Vuolteenaho (2003) to this economy, we group firms 1 and 2 together into a high bm portfolio and firms 3 and 4 together into a low bm portfolio. Grouping the firms and averaging their returns and earnings eliminates the variation in CF and DR within groups of firms with the same valuation. The high and low (H and L) bm portfolios have the following properties:

$$CF_H = -0.5, DR_H = 0.5, bm_H = 1 \quad (22)$$

$$CF_L = 1.5, DR_L = 0.5, bm_L = -1 \quad (23)$$

There is no discount rate variation at all across the two portfolios, which vary only in their cash flows. Based solely on this information, the natural but mistaken inference would be that cash flows account for 100% of variation in valuations.

In contrast, our regression approach considers each firm as a distinct observation and allows firms to differ along multiple dimensions, not just in their valuations. By controlling for bm in our regressions, we explicitly consider whether other firm characteristics capture variation in firms' cash flows and discount rates. For example, if firms with the same valuations in the economy above differ in their observed profitability, our method would correctly identify all cash flow and discount variation.

Prior research that sorts firms into bm portfolios cannot assess the correlation between the cash flow and discount rate components. This correlation is likely to be close to -1 across book-to-market sorted portfolios, assuming that cash flows and discount rates both

contribute at least somewhat to variation in valuations. One needs to analyze variation in firm characteristics other than book-to-market to evaluate the correlation between the components of valuations.

6 Portfolio-level Analysis

Now we analyze the implied discount rate (DR) and cash flow (CF) variation in returns to important portfolios, including the market portfolio and factor portfolios formed by cross-sectional sorts on value, size, profitability, investment, and issuance. We compute weighted averages of firm-level DR and CF estimates to obtain portfolio-level DR and CF estimates. We apply the approximation and aggregation procedure described in Section 2.

6.1 The Market Portfolio

We define the market portfolio as the value-weighted average of individual firms. We obtain firm-level expected log returns and log earnings from the VAR and apply the procedure in Section 2 to obtain the corresponding market-level discount rates and expected cash flows.

We compare the estimates from our aggregation approach to those from a standard aggregate-level VAR in the spirit of Campbell (1991). In the aggregate VAR, we use only the logs of (market-level) book-to-market ratio ($\ln BM_mkt$) and the real risk-free ($\ln Rf$) as predictors of the logs of market-level earnings and returns. Accordingly, this specification entails just three regressions in which market-level earnings, returns, and risk-free rates are the dependent variables and lagged book-to-market and risk-free rates are the independent variables.

We validate our panel VAR approach and compare it to the market-level VAR in Figure 5, which shows market cash flow and discount rate components from both VARs alongside realized market earnings and returns over the next 10 years. We construct the series of 10-year realized earnings (returns) based on firms' current market weights and their future 10-year earnings (returns). Thus, we forecast 10-year buy-and-hold returns to the market

portfolio, not the returns to an annually-rebalanced trading strategy. We do not rebalance the portfolio because the underlying discount rate estimates from the panel VAR are specific to firms. This distinction is important insofar as firm entry, exit, issuance, and repurchases occur.

The red and black lines in the top plot in Figure 5 are the predicted 10-year market earnings from our panel VAR and from the market-level VAR, respectively. Both predictions track realized 10-year market earnings very well, with an R^2 of 64.9% for the panel VAR and 55.4% for the market VAR. The bottom plot in Figure 5 shows that the predictions of 10-year returns from the two VARs are also similar, except that the panel VAR predicts lower returns around the 2000 period. Both sets of predictions exhibit positive relationships with realized 10-year returns. The R^2 of the panel VAR is 36.3%, whereas the R^2 of the market-level VAR is 19.1%. The plots in Figure 5 suggest that both VAR methods yield meaningful decompositions of valuations into CF and DR components. Even though the panel VAR does not directly analyze the market portfolio, aggregating the panel VAR's firm-level predictions results in forecasts of market cash flows and returns that slightly outperform forecasts based on the traditional approach.

Next we compare the implications of the two VARs for the sources of market returns. We compute the shocks to market cash flows and discount rates from both VARs, as in Equations (56) and (57) in Appendix B, and analyze the covariance matrix of these shocks. When calculating the aggregated panel VAR shock from time t to time $t + 1$, the updated expectation is based on the firms in the market portfolio at time t . Similarly, the shock from time $t + 1$ to $t + 2$ is based on the firms in the market portfolio at time $t + 1$.

Table 6 presents variance decompositions of market returns based on the panel VAR and the market-level VAR. The first four columns decompose the variance of predicted market returns from our approximation into four nearly exhaustive components: the variance of DR, variance of CF, variance of the cross term (CF*DR), and the covariance between CF and DR. We do not report the covariances between the cross term and the CF and DR terms

because these covariances are small. The fifth column reports the correlation between the DR and CF components of market returns. The last column reports the correlation between our approximation of market returns and actual market returns. This column shows that correlation is 0.986, indicating that our approximation is accurate. Standard errors based on the delta method appear in parentheses.

Table 6 shows that the panel and market-level VARs predict similar amounts of discount rate variation (18.3% and 28.1%, respectively), but the estimate from the panel VAR is more precise as measured by its standard error. Both estimates of DR variation are lower than those reported in prior studies. By restricting the sample of the market-level VAR to 1964 to 1990, we can reproduce the traditional finding that DR variation explains nearly all variation in market returns.

The estimates from the panel VAR imply that shocks to market cash flows account for 63.2% of market return variance, whereas the market-level VAR implies that CF shocks explain just 24.8% of return variance. The two VARs also differ in the implied correlations between the CF and DR components. The panel VAR indicates that the correlation is just -0.322 , whereas the market-level VAR implies a correlation of -0.892 .

One possible explanation for the difference in the two VARs is that the panel VAR relies on two log-linear approximations of market returns, which could introduce errors in the variance decomposition. However, we find that the predicted (log) market return based on the panel aggregation and approximations exhibits a correlation of 0.986 with the actual (log) market return. In addition, the cross term (CF*DR), which is unique to the approximate panel aggregation, accounts for less than 1% of market return variance.

A more likely reason for the discrepancy is that the panel VAR employs far more predictive variables than the market-level VAR, leading to a more accurate description of expected cash flows and discount rates. Another possible reason is that the market-level VAR suffers from two related biases induced by the reliance on market-level book-to-market ratios ($\ln\text{BM_mkt}$). The time-series properties of $\ln\text{BM_mkt}$ cause two problems: 1) this highly

persistent regressor causes a significant Stambaugh (1999) bias given the relatively short sample; and 2) an apparent structural break in $\ln\text{BM_mkt}$ occurs around 1990, as noted by Lettau and van Nieuwerburgh (2008) in the context of the market price-dividend ratio, implying that this regressor is actually non-stationary. We exclude $\ln\text{BM_mkt}$ from our primary panel VAR specification based on these considerations.

6.2 Anomaly Portfolios

We now turn to our analysis of the returns to long-short anomaly portfolios. Our goal is to bring new facts to the ongoing debate on the sources of anomalies. We estimate the cash flow and discount rate components of historical anomaly returns and analyze the covariance matrix of these shocks. We then evaluate whether theories that aim to explain anomaly returns make reasonable predictions about the cash flow and discount rate components of anomaly returns.

The anomaly portfolios represent trading strategies, where the underlying firms in the portfolio change every year based on firms' characteristic rankings in June. However, for any given year, the portfolio return is driven by the cash flow and discount rate shocks of the individual firms in the portfolio in that year. The firm-level VAR allows us to relate anomaly returns to underlying firm fundamentals. We aggregate the firm-level estimates using value weights within each quintile and then analyze portfolios with long positions in quintile 5 and short positions in quintile 1 according to firms' characteristic rankings. The aggregation procedure is otherwise analogous to that used for the market portfolio.

The plots in Figure 6 show that the cash flow and discount rate components of the value portfolio respectively forecast the 10-year earnings and returns for this portfolio. The predictor in the top plot in Figure 6 is the difference between the CF component of value and growth firms. Similarly, the realized cash flows in this plot represent the difference in the 10-year earnings of value and growth firms. The graphic shows that predicted earnings are correlated with future 10-year earnings, primarily in the second half of the sample. The

overall R^2 is modest at 18.6%. The bottom plot in Figure 6 depicts the relationship between the DR component of the value spread and future 10-year returns. This relationship is strong in both halves of the sample, and the overall R^2 is high at 47.1%.

Figure 7 presents the analogous R^2 statistics for the cash flow and discount rate components of the five long-short anomaly portfolios and the market portfolio. The DR component of the size anomaly portfolio forecasts its 10-year returns quite well ($R^2 = 61.9\%$), whereas the DR component of the issuance portfolio has more modest forecasting power for 10-year issuance anomaly returns ($R^2 = 22.5\%$). The R^2 values in Figure 7 range from 18.6% to 64.9%, implying correlations between the CF and DR components and their realized counterparts that range between 0.431 and 0.806. We conclude from this analysis that the aggregated cash flow and discount rate components plausibly reflect the long-short portfolios' cash flow and discount rate components.

Table 7 presents variance decompositions of anomaly returns for the five anomalies and is analogous to Table 6 for the market. Table 7 reveals consistent patterns across the five anomaly portfolios. Cash flow variation accounts for 37.1% to 51.8% of variation in anomaly returns, whereas discount rate variation by itself accounts for just 16.5% to 18.1% of anomaly return variance. The covariance between CF and DR is consistently negative, helping to explain why the covariance term accounts for 36.1% to 45.0% of anomaly return variance. The cross term (CF*DR) accounts for only 1.7% to 3.5% of anomaly return variance. The standard errors on the variance components never exceed 13.8% and are typically much lower, indicating the precision of these findings. The last column shows that the correlation between the approximation of anomaly returns and actual anomaly returns always exceeds 0.9, indicating that our approximation is accurate.

The relative importance of cash flows and the negative correlation between CF and DR are the most prominent effects. Theories of anomalies that rely heavily on independent variation in DR shocks, such as De Long et al. (1990), are inconsistent with the evidence in Table 7. In contrast, theories in which CF shocks are tightly linked with DR shocks have the

potential to explain the patterns in Table 7. Rational theories in which firm risk increases after negative cash flow realizations predict negative correlations between CF and DR shocks. Behavioral theories in which investors overreact to cash flow news are also consistent with this evidence.

Table 8 displays correlations between the components of market returns and those of anomaly returns. The four columns indicate the correlations between market cash flow and discount rate shocks and anomaly cash flow and discount rate shocks. Standard errors based on the delta method appear in parentheses.

The striking result in the first column of Table 8 is that none of the five anomaly cash flow shocks exhibits a large correlation with market cash flows. The correlations between market cash flows and the cash flows from the value, investment, and size anomalies range between -0.01 and 0.06 and are statistically indistinguishable from zero. The correlation between market cash flows and issuance cash flows is also statistically insignificant, though it is slightly larger at 0.23 . Only the correlation between profitability cash flows and market cash flows is statistically significant, though its economic magnitude is just -0.17 . These findings cast doubt on theories of anomalies that rely on cross-sectional differences in firms' sensitivities to aggregate cash flows. The evidence is ostensibly inconsistent with a broad category of risk-based explanations of anomalies, which includes Lettau and Wachter's (2007) theory of the value premium.

The fourth column in Table 8 reveals that discount rate shocks to the value, investment, and size anomalies are only weakly correlated with discount rate shocks to the market. However, two anomalies' DR shocks exhibit correlations with market DR shocks that are insignificantly different from zero at the 5% level. Market DR shocks are negatively correlated (-0.34) with DR shocks to the profitability anomaly and positively correlated (0.48) with DR shocks to the issuance anomaly. One interpretation is that firms with low profits and high equity issuance have a cost of capital that depends critically on the market-wide cost of equity capital. Such firms could be highly dependent on external equity finance. If so,

shocks to the market risk premium could drive variation in these firms' costs of capital.

Generalizing from the last column in Table 8, the weak correlation between most anomalies' DR shocks and market DR shocks is inconsistent with theories of common DR shocks. In theories such as Campbell and Cochrane (1999), commonality in DR shocks occurs because risk aversion varies over time. Similarly, theories in which anomalies are driven by common shocks to investor sentiment, such as Baker and Wurgler (2006), that affect groups of stocks and the market are at odds with the evidence on the lack of correlation in anomaly and market DR shocks.

The second and third columns in Table 8 indicate that there are few large cross-correlations between market DR and anomaly CF shocks or between market CF and anomaly DR shocks. All correlations are less than 0.4 in absolute value. However, five correlations exceed 0.3 in absolute value, and these five are statistically significant at the 5% level. The correlations between the market CF shock and the size and issuance DR shocks suggest that small firms and those with high equity issuance have lower costs of capital during good economic times. The correlations between the value, investment, and issuance CF shocks and the market DR shock are consistent with the idea that firms with high investment and equity issuance and low valuations have higher expected cash flows when market-wide discount rates fall. Because we are simultaneously testing many hypotheses, we are reluctant to overinterpret these cross-correlations. The low correlation between market and anomaly return components is consistent with theories in which idiosyncratic cash flow shocks affect firms' expected returns—e.g., Babenko, Boguth, and Tserlukevich (2016).

Tables 9A and 9B respectively report correlations among the anomaly CF shocks and among the anomaly DR shocks. Several of the correlations in both panels of Table 9 are statistically and economically significant. Notable negative correlations include those between investment and book-to-market, size and book-to-market, issuance and book-to-market, and issuance and profitability. Notable positive correlations include those between issuance and investment, size and investment, and profitability and investment.

However, nearly all of these correlations among return components simply reflect correlations among anomalies' total returns. For example, Table 9A shows that the CF shocks to the investment and book-to-market anomalies exhibit a strong negative correlation of -0.66 . Table 9B shows that the DR shocks to these anomalies exhibit a similarly strong negative correlation of -0.62 . The significant correlations in the two panels follow a strong pattern: the pairwise anomaly CF correlations are very similar in sign and magnitude to the pairwise anomaly DR correlations. Correlations in anomaly returns that apply to both the CF and DR shocks often arise because many of the same firms appear in multiple anomaly portfolios. Consistent with this interpretation, the notable correlations in Tables 9A and 9B typically have the same sign as the corresponding correlations in Table 1B, which reports the relationships between firm characteristics underlying the anomalies. We conclude that there is little commonality in the components of anomalies' returns beyond that arising from mechanical relationships.

7 Alternative Specifications

Here we consider two alternative VAR specifications in which we include the market-wide valuation ratio along with interactions with firm-level characteristics.⁴ Market valuations could capture common variation in firms' cash flows and discount rates and interactions with firm characteristics could capture firms' differential exposures to market-wide variation. The first alternative specification (*Spec2*) adds only the market-wide book-to-market ratio, as measured by the value-weighted average of sample firms' log book-to-market ratios, to our main specification (*Spec1*). The second alternative specification (*Spec3*) augments the first by including interaction terms between market-wide valuations and the five firm-level log characteristics as well as firm-level log realized variance.

The estimation of the key return and cash flow forecasting regressions in the VAR indi-

⁴In unreported tests, we explore specifications that include additional market-level and anomaly-level variables, such as aggregate versions of anomaly characteristics and spreads in valuations across anomaly portfolios.

cates that these additional regressors only modestly contribute to explanatory power. The adjusted R^2 in the return regression increases from 4.6% in *Spec1* to 5.5% in *Spec2*, and the coefficient on the added market-wide valuation variable is only marginally statistically significant ($p\text{-value} = 0.053$). In the earnings regression, the coefficient on market-wide valuation is robust statistically significant at the 1% level, but the adjusted R^2 barely increases from 24.3% in *Spec1* to 24.9% in *Spec2*. The findings for the second alternative specification, *Spec3*, suggest that the six interaction terms do not contribute incremental explanatory power beyond *Spec2*. Specifically, the adjusted R^2 for the return and earnings regressions are equal to or less than those for *Spec2* and the vast majority of the interaction coefficients are statistically insignificant. Overall, these two sets of regressions do not provide strong evidence that the more parsimonious primary specification, *Spec1*, is misspecified.

We now evaluate the implications of the alternative specifications for return variance decompositions. Table 10 shows the components of market return variance implied by *Spec2* and *Spec3*. The difference between Table 10 and Table 6A, which shows the results for *Spec1*, is striking. Whereas discount rate variation accounts for just 18.3% of return variance in *Spec1*, it accounts for 91.7% and 105.8% of variation in *Spec2* and *Spec3*, respectively. The main reason is that high market-wide book-to-market ratios apparently forecast higher returns and such valuations ratios are highly persistent, implying that their long-run impact is potentially large. However, this predictive relationship is quite weak statistically, so the standard errors on the variance decompositions are enormous in *Spec2* and *Spec3*. In fact, one cannot reject the hypothesis that DR variation accounts for 0% of variation in returns. Thus, the striking differences in point estimates across the specifications do not necessarily imply strikingly different conclusions.

Table 11 shows the components of anomaly return variance implied by *Spec2* and *Spec3*. Comparing Tables 11 and 7, we see that cash flow variation accounts for the bulk of anomaly return variance in all three VAR specifications. The finding that discount rates are negatively correlated with expected cash flows also generalizes from *Spec1* to *Spec2*, but it does not

obviously apply to *Spec3*, which allows for interaction terms between market-wide valuations and firm-level characteristics. The standard errors in *Spec3* are too large to draw reliable inferences about this correlation.

To assess which VAR specification provides the most meaningful decomposition of market and anomaly returns, we analyze the long-term forecasting power implied by each specification. Figure 8 shows the 10-year forecasts of market earnings and returns from *Spec2*, just as Figure 5 shows these forecasts for *Spec1*. Although adding market-wide valuations slightly improves the forecasting power in the one-year earnings regression, 10-year predictions based on the *Spec2* model are vastly inferior to those based on the more parsimonious *Spec1* model. The adjusted R^2 values of 65% for *Spec1* compared to just 5% for *Spec2* confirm the visual impression from the figures. The two specifications exhibit little difference in their ability to predict 10-year market returns ($R^2 = 36\%$ for *Spec1* vs. $R^2 = 33\%$ for *Spec2*).

Figure 9 shows the 10-year forecasting power of the three specifications for market earnings and returns as well as the earnings and returns of the five anomalies. The most notable difference arises in the forecasting power for market earnings. Both specifications that include market-wide valuations give rise to especially poor forecasts of 10-year market earnings. Apparent structural breaks in market-wide valuations, such as those proposed by Lettau and van Nieuwerburgh (2008), could help explain the poor long-term forecasting power of these two VAR specifications. There are few notable differences in the three specifications' abilities to predict long-term anomaly returns and earnings. This similarity is not surprising in light of the similar anomaly return decompositions predicted by the three specifications. We conclude that the more parsimonious *Spec1* not only gives rise to the most precise estimates of market and anomaly return components, but it also exhibits the most desirable long-term forecasting properties.

8 Conclusion

Despite decades of research on forecasting short-term stock returns, there is no widely accepted explanation for observed cross-sectional patterns in stock returns. We provide new evidence on the sources of anomaly returns by aggregating firm-level cash flow and discount rate estimates from a panel VAR system. Our aggregation approach enables researchers to study the components of portfolio returns, while avoiding the biases inherent in analyzing the cash flows and discount rates of rebalanced portfolios.

We contribute three novel findings to our understanding of stock return anomalies. First, cash flow variation is the primary driver of anomaly returns. Second, discount rate variation amplifies cash flow variation. Third, there is little commonality in the components of anomaly and market returns. Based on this evidence, the most promising theories of anomalies are those that emphasize the importance of firm-level cash flow variation as a driver of either changes in firm risk or errors in investors' expectations.

References

- Ang, Andrew, Robert Hodrick, Yuhang Xing, and Xioayan Zhang, 2006, The cross-section of volatility and expected returns, *Journal of Finance* 61, 259–299.
- Barberis, Nicholas, Andrei Shleifer, and Robert Vishny, 1998, A model of investor sentiment, *Journal of Financial Economics* 49, 307–343.
- Babenko, Ilona, Oliver Boguth, Yuri Tserlukevich, 2016, Idiosyncratic cash flows and systematic risk, *Journal of Finance* 81, 425–455.
- Baker, Malcolm, and Jeffrey Wurgler, 2006, Investor sentiment and the cross-section of stock returns, *Journal of Finance* 61, 1645–1680.
- van Binsbergen, Jules, and Ralph Koijen, 2010, Predictive regressions: A present value approach, *Journal of Finance* 65, 1439–1471.
- Campbell, John, 1991, A variance decomposition for stock returns, *Economic Journal* 101, 157–179.
- Campbell, John, and Robert Shiller, 1988, Stock prices, earnings, and expected dividends, *Journal of Finance* 43, 661–676.
- Campbell, John, and Luis Viceira, 2002, *Strategic Asset Allocation* (Clarendon Lectures in Economics, Cambridge, MA).
- Claus, J. and Jacob Thomas, 2001, Equity premia as low as three percent? Evidence from analysts' earnings forecasts from domestic and international stock markets, *Journal of Finance* 56, 1629 - 1666.
- Cohen, Randolph, Christopher Polk, and Tuomo Vuolteenaho, 2003, The value spread, *Journal of Finance* 58, 609–641.
- Cooper, Michael, Huseyin Gulen, and Michael Schill, 2008, Asset growth and the cross-section of stock returns, *Journal of Finance* 63, 1609–1651.

- Davis, James, Eugene Fama, and Kenneth French, 2000, Characteristics, covariance, and average returns: 1929-1997, *Journal of Finance* 55, 389–406.
- Daniel, Kent, and Sheridan Titman, 2006, Market reactions to tangible and intangible information, *Journal of Finance* 61, 1605–1643.
- De Long, Bradford, Andrei Shleifer, Lawrence Summers, and Robert Waldmann, 1990, Noise trader risk in financial markets, *Journal of Political Economy* 98, 703–738.
- Fama, Eugene, and Kenneth French, 1992, The cross-section of expected stock returns, *Journal of Finance* 47, 427–465.
- Fama, Eugene, and Kenneth French, 2015, A five-factor asset pricing model, *Journal of Financial Economics* 116, 1–22.
- Fama, Eugene, and James MacBeth, 1973, Risk, return, and equilibrium: Empirical tests, *Journal of Political Economy* 81, 607–636.
- Hansen, Lars, John Heaton, Junghoon Lee, and Nikolai Roussanov, 2007, Chapter 61: Intertemporal substitution and risk aversion, in James J. Heckman and Edward E. Learner, eds.: *Handbook of Econometrics*, Volume 6A (Elsevier, Amsterdam).
- Jegadeesh, Narasimhan, and Sheridan Titman, 1993, Returns to buying winners and selling losers: Implications for stock market efficiency, *Journal of Finance* 48, 65–91.
- Kelly, Bryan, and Seth Pruitt, 2013, Market expectations in the cross-section of present values, *Journal of Finance* 68, 1721–1756.
- Lettau, Martin, and Stijn van Nieuwerburgh, 2008, Reconciling the return predictability evidence, *Review of Financial Studies* 21, 1607–1652.
- Lettau, Martin, and Jessica Wachter, 2007, Why is long-horizon equity less risky? A duration-based explanation for the value premium, *Journal of Finance* 62, 55–92.

- Lyle, Matthew R., and Charles C.Y. Wang, 2015, The cross section of expected holding period returns and their dynamics: A present value approach, *Journal of Financial Economics* 116, 505–525.
- Novy-Marx, Robert, 2013, The other side of value: The gross profitability premium, *Journal of Financial Economics* 108, 1–28.
- Ohlson, James, 1995, Earnings, book values, and dividends in equity valuation, *Contemporary Accounting Research* 11, 661–687.
- Pastor, Lubos, M. Sinha, and B. Swaminathan, 2008, Estimating the intertemporal risk-return trade-off using the implied cost of capital, *Journal of Finance* 63, 2859–2897.
- Petersen, Mitchell A., 2009, Estimating standard errors in finance panel data sets: Comparing approaches, *Review of Financial Studies* 22, 435–480.
- Pontiff, Jeffrey, and Artemiza Woodgate, 2008, Share issuance and cross-sectional returns, *Journal of Finance* 63, 921–945.
- Shiller, Robert, 1981, Do stock prices move too much to be justified by subsequent changes in dividends? *American Economic Review* 71, 421–436.
- Stambaugh, Robert, 1999, Predictive regressions, *Journal of Financial Economics* 54, 375–421.
- Vuolteenaho, Tuomo, 2002, What drives firm-level stock returns? *Journal of Finance* 57, 233–264.
- Zhang, Lu, 2005, The value premium, *Journal of Finance* 60, 67–103.

Appendix A: Cash Flows vs. Discount Rates of Trading Strategies

Here we show that the cash flows and discount rates of rebalanced portfolios, such as anomaly portfolios, can differ substantially from those of the underlying firms in the portfolios. We provide examples below in which firms have constant cash flows, but all variation in returns to the rebalanced portfolio comes from cash flow shocks.

We first consider a stylized behavioral model of stock returns and cash flows. Assume that all firms pay constant dividends:

$$D_{i,t} = \bar{d}. \tag{24}$$

Assume that investors in each period erroneously believe that any given firm's dividend is permanently either $d_L < \bar{d}$ or $d_H > \bar{d}$. We define the firms associated with low (high) dividend beliefs to be value (growth) firms. The pricing of these firms satisfies:

$$P^{\text{value}} = \frac{d_L}{R - 1}, \tag{25}$$

$$P^{\text{growth}} = \frac{d_H}{R - 1}, \tag{26}$$

where R is the gross risk-free rate. Each period, with probability q , investors switch their beliefs about each stock's dividends either from d_L to d_H or from d_H to d_L . Investors believe their beliefs will last forever, whereas in reality they will switch with probability q in each period.

Now consider a value fund that invests only in stocks that investors currently believe will pay dividends of d_L . Further assume that there are only two firms in the economy—a firm that currently is a growth firm and a firm that currently is a value firm. When beliefs switch, the growth firm becomes a value firm and vice versa. This switch therefore induces trading in the value fund as the fund has to sell firms that become growth firms and buy the new value firms.

Such trading has a significant impact on the fund's dividends. Suppose that the fund

initially holds one share of the value stock, which implies that its initial wealth is $W_0 = P^{\text{value}}$. Assume investors do not switch beliefs in the next period. In this case, the fund's gross return is:

$$\begin{aligned} R_1^{\text{value}} &= \frac{P^{\text{value}} + \bar{d}}{P^{\text{value}}} \\ &= 1 + \frac{\bar{d}}{P^{\text{value}}}. \end{aligned} \quad (27)$$

Period 1 cum-dividend wealth is

$$W_1^{\text{cum}} = P^{\text{value}} + \bar{d}, \quad (28)$$

where ex-dividend wealth is P^{value} and dividend is $d_1 = \bar{d}$. Assume that beliefs switch in period 2. Then:

$$\begin{aligned} R_2^{\text{value}} &= \frac{(R-1) \left(\frac{d_H}{R-1} + \bar{d} \right)}{d_L} \\ &= \frac{d_H + (R-1)\bar{d}}{d_L} \\ &= \frac{d_H}{d_L} + \frac{\bar{d}}{P^{\text{value}}}. \end{aligned} \quad (29)$$

So fund wealth becomes:

$$W_2^{\text{cum}} = P^{\text{value}} \frac{d_H}{d_L} + \bar{d}.$$

Ex-dividend wealth is $W_2^{\text{ex}} = P^{\text{value}} \frac{d_H}{d_L}$, and the dividend is $d_2 = \bar{d}$ once again. The dividend price ratio of the strategy is now:

$$\frac{W_2^{\text{ex}}}{d_2} = \frac{P^{\text{value}}}{\bar{d}} \frac{d_H}{d_L} > \frac{P^{\text{value}}}{\bar{d}}. \quad (30)$$

The higher price-dividend ratio reflects high expected dividend growth next period.

Importantly, the fund now reinvests its capital gain into the current value stock and is able to purchase more than one share. Assuming beliefs do not switch in period 3, the fund's

wealth increases to:

$$\begin{aligned}
W_3^{cum} &= P^{\text{value}} \frac{d_H}{d_L} \left(1 + \frac{\bar{d}}{P^{\text{value}}} \right) \\
&= P^{\text{value}} \frac{d_H}{d_L} + \frac{d_H}{d_L} \bar{d}.
\end{aligned} \tag{31}$$

Now ex-dividend wealth is $W_3^{ex} = P^{\text{value}} \frac{d_H}{d_L}$ and $d_3 = \bar{d} \times d_H/d_L$, implying that dividend growth during this period is high as $d_3/d_2 = d_H/d_L > d_2/d_1 = 1$. The price-dividend ratio of the strategy is now:

$$\frac{W_3^{ex}}{d_3} = \frac{P^{\text{value}} \frac{d_H}{d_L}}{\bar{d} \times d_H/d_L} = \frac{P^{\text{value}}}{\bar{d}}, \tag{32}$$

meaning that the price-dividend ratio returns to its original value.

In summary, dividend growth of the dynamic value strategy varies over time, but expected returns to the strategy are constant and given by:

$$\begin{aligned}
E(R^{\text{value}}) &= 1 - q + q \frac{d_H}{d_L} + \frac{\bar{d}}{P^{\text{value}}} \\
&= 1 - q + q \frac{d_H}{d_L} + \frac{\bar{d}}{d_L} (R - 1).
\end{aligned} \tag{33}$$

A symmetric argument applies to the analogous growth strategy, which also has time-varying dividend growth and constant expected returns. We conclude that return variation in the dynamic trading strategies arises solely because of cash flow shocks even though all firms in the economy incur only discount rate shocks. Firm-level return variation is driven by changes in firms' expected returns, not their dividends—which are constant.

There are no discount rate shocks to the returns of these dynamic strategies when viewed from the perspective of an investor who invests in the value or growth funds. However, unexpected returns to such funds are in fact, under the objective measure of the econometrician, due to discount rate shocks to the underlying firms. The firms' actual expected returns vary, whereas their dividend growth does not.

This feature of rebalanced portfolios is not limited to the case of time-varying mispricing.

Consider a rational model in which value firms have riskier cash flows than growth firms. If time-variation in a firm's cash flow risk causes it to switch between being a value firm and a growth firm, the rational model delivers the same insights as the behavioral model discussed above.

In this example, we assume firms' log dividend growth is:

$$\Delta d_{i,t+1} = -\frac{1}{2}\sigma^2 + \sigma \left(\rho_{s_{i,t}} \varepsilon_{m,t+1} + \sqrt{1 - \rho_{s_{i,t}}^2} \varepsilon_{i,t+1} \right), \quad (34)$$

where $\varepsilon_{m,t+1}$ and $\{\varepsilon_{i,t+1}\}_i$ are uncorrelated standard normally distributed shocks representing aggregate and firm-specific dividend shocks, respectively. Firm exposure to aggregate dividend shocks is:

$$\rho_{s_{i,t}} = \begin{cases} \rho^H & \text{if } s_{i,t} = 1 \\ \rho^L & \text{if } s_{i,t} = 0 \end{cases}, \quad (35)$$

where $s_{i,t}$ follows a two-state Markov process where $\Pr\{s_{i,t+1} = 1 | s_{i,t} = 0\} = \Pr\{s_{i,t+1} = 0 | s_{i,t} = 1\} = \pi$. For ease of exposition, set $\rho^L = 0$ and $\rho^H = 1$. Initially, half of firms are in state 1, while the other half are in state 0. If a regime change occurs, all firms currently in state 1 switch to state 0, and vice versa.

The log stochastic discount factor is:

$$m_{t+1} = -\frac{1}{2}\gamma^2\sigma^2 - \gamma\sigma\varepsilon_{m,t+1}, \quad (36)$$

where we implicitly assume a zero risk-free rate and where $\gamma > 0$ represents risk aversion. These assumptions imply that the conditional mean and volatility of cash flow growth is constant. However, firm risk varies with $s_{i,t}$, which determines the covariance of cash flows with the pricing kernel, causing time-varying firm risk premiums.

Solving for the price-dividend ratio as a function of the state yields:

$$\begin{aligned} PD(s_{i,t}) &= E_t \left[e^{-\frac{1}{2}\gamma^2\sigma^2 - \gamma\sigma\varepsilon_{m,t+1} - \frac{1}{2}\sigma^2 + \sigma \left(\rho_{s_{i,t}} \varepsilon_{m,t+1} + \sqrt{1 - \rho_{s_{i,t}}^2} \varepsilon_{i,t+1} \right)} (1 + PD(s_{i,t+1})) \right] \\ &= e^{-\gamma\sigma^2\rho_{s_{i,t}}} (1 + \pi PD(s_{i,t+1} \neq s_{i,t}) + (1 - \pi) PD(s_{i,t+1} = s_{i,t})). \end{aligned} \quad (37)$$

Denote the price-dividend ratio in state j as PD_j . The price-dividend relation above is a system with two equations and two unknowns with the solution:

$$PD_1 = \frac{1}{e^{\gamma\sigma^2} - 1} \quad (38)$$

$$PD_0 = 1/\pi + \frac{1}{e^{\gamma\sigma^2} - 1} \quad (39)$$

These equations show that price-dividend ratios are higher in state 0 when dividend risk is low than in state 1 when dividend risk is high, implying that expected returns are higher in state 0 as expected dividend growth is constant across states. Firms' expected net returns are:

$$E_t [R_{i,t+1}|s_{i,t} = 0] - 1 = 0, \quad (40)$$

$$E_t [R_{i,t+1}|s_{i,t} = 1] - 1 = 2(e^{\gamma\sigma^2} - 1). \quad (41)$$

Since $PD_0 > PD_1$, we see that $E_t [R_{i,t+1}|s_{i,t} = 1] > E_t [R_{i,t+1}|s_{i,t} = 0]$. Thus, firms' price-dividend ratios fluctuate because of shocks to discount rates, not cash flows. Although there are cash flow shocks in returns arising from the contemporaneous dividend shock ($\sigma \left(\rho_{s_{i,t}} \varepsilon_{m,t+1} + \sqrt{1 - \rho_{s_{i,t}}^2} \varepsilon_{i,t+1} \right)$), dividends are unpredictable and therefore do not induce time-variation in the price-dividend ratio.

Now consider a value mutual fund that in each period buys firms that are currently in the low valuation state 1. With probability π , value firms held by the fund will switch to the high valuation state 0, meaning that they become growth firms. The fund sells all firms in each period and reinvests the proceeds in firms that are in the low valuation state 1. The fund pays out all firm dividends as they occur. The expected return to this strategy is constant and equal to $E_t [R_{i,t+1}|s_{i,t} = 1] - 1 = 2(e^{\gamma\sigma^2} - 1)$, even though all firms' expected returns vary over time.

We now analyze the growth of the value fund's dividends in each period. The first source

of fund dividend growth is growth in the underlying firms' dividends, which satisfy:

$$\frac{D_{i,t+1}}{D_{i,t}} = e^{-\frac{1}{2}\sigma^2 + \sigma\varepsilon_{m,t+1}}. \quad (42)$$

The second source of fund dividend growth is growth in the number of shares of value firms held by the fund. If value firms switch to growth firms, the fund will reap a capital gain and be able to buy more shares of the new value firms in the following period. Define the indicator variable $\mathbf{1}_{s_{i,t} \neq s_{i,t-1}}$ as equal to 1 if there was a regime shift from period $t - 1$ to period t and 0 otherwise. Accounting for both sources of growth, fund dividend growth is:

$$\frac{D_{t+1}^{Fund}}{D_t^{Fund}} = \mathbf{1}_{s_{i,t} \neq s_{i,t-1}} \frac{PD_0}{PD_1} e^{-\frac{1}{2}\sigma^2 + \sigma\varepsilon_{m,t+1}} + (1 - \mathbf{1}_{s_{i,t} \neq s_{i,t-1}}) e^{-\frac{1}{2}\sigma^2 + \sigma\varepsilon_{m,t+1}}, \quad (43)$$

where the term $\frac{PD_0}{PD_1} = 1 + \frac{e^{\gamma\sigma^2} - 1}{\pi}$ represents the capital gain from the prior period. Dividends are predictably high after high capital gains and low after low capital gains. The predictability in dividend growth leads to a time-varying price-dividend ratio for the mutual fund, even though its expected return is constant. Thus, discount rate shocks to the underlying value firms are cash flow shocks for the mutual fund implementing a value trading strategy.

Appendix B: The Panel VAR

Our goal is to decompose anomaly returns into discount rate and cash flow components and identify the determinants of these components. We first decompose firm returns and valuations, and then aggregate firm-level components to the portfolio level. We rely on the following log-linear approximation for firms' book-to-market ratios:

$$bm_{i,t} \approx r_{i,t+1} - e_{i,t+1} + \kappa bm_{i,t+1}, \quad (44)$$

where $bm_{i,t}$ is the log of the book value of equity to the market value of equity of firm i in year t , $r_{i,t+1}$ and $e_{i,t+1}$ are the year $t + 1$ log return to equity and log accounting return on

equity (ROE), respectively. The latter is defined as:

$$e_{i,t+1} \equiv \ln(1 + ROE_{i,t+1}), \quad (45)$$

where $ROE_{i,t+1}$ is firm i 's earnings in year $t + 1$ divided by book value of common equity in year t . The constant κ in Equation (4) is a log-linearization constant that we set to 0.96 in practice. Our main results are insensitive to small variations in kappa, such as 0.95 or 0.97, which span the values used in prior studies. Vuolteenaho (2002) derives Equation (4) by assuming the clean surplus accounting relation of Ohlson (1995) holds with equality:

$$D_{i,t} = E_{i,t} - \Delta BE_{i,t}, \quad (46)$$

where $E_{i,t}$ is earnings, $D_{i,t}$ is dividends, and $\Delta BE_{i,t}$ is the change in book equity from year $t - 1$ to year t .

Taking Equation (4) as an equality and recursively substituting n times results in the following relation between current book-to-market, the present values of earnings and returns, and book-to-market in n years:

$$bm_{i,t} = \sum_{j=1}^n \kappa^{j-1} (r_{i,t+j} - e_{i,t+j}) + \kappa^n bm_{i,t+n}. \quad (47)$$

Taking the limit as n approaches infinity under the transversality condition, we obtain the infinite-horizon present-value equation:

$$bm_{i,t} = \sum_{j=1}^{\infty} \kappa^{j-1} (r_{i,t+j} - e_{i,t+j}). \quad (48)$$

This relation shows that log book-to-market is approximately equal to the difference between cumulative future log returns and cumulative future log earnings.

We impose additional structure on the present-value equation by assuming that firms' expected log returns and log earnings are related to observable characteristics, $X_{i,t}$. We assume the first element in the $K \times 1$ characteristics vector, $X_{i,t}$, is firm i 's log book-to-market

ratio, while the remaining elements are firm-specific variables like the firm's profitability or investment, and aggregate variables such as the risk-free rate, aggregate profitability, and aggregate investment. Define the augmented and demeaned vector:

$$\tilde{X}_{i,t} = \begin{bmatrix} r_{i,t} - E[r_{i,t}] \\ e_{i,t} - E[e_{i,t}] \\ X_{i,t} - E[X_{i,t}] \end{bmatrix}, \quad (49)$$

and let

$$\tilde{X}_{i,t} = A\tilde{X}_{i,t-1} + \Sigma_X \varepsilon_{i,t}. \quad (50)$$

Here A is a $(K+2) \times (K+2)$ matrix, Σ_X is a $(K+2) \times (K+1)$ matrix of shock exposures and $\varepsilon_{i,t}$ is a $(K+1) \times 1$ vector of standard Normal shocks. Note that there is one less shock than there are variables in $\tilde{X}_{i,t}$. This is due to the present value constraint. In particular, let ι_j be a $(K+2) \times 1$ vector with a one in the j 'th row and zeros elsewhere. Now, the unconditionally demeaned log book-to-market ratios can be written:⁵

$$\tilde{b}m_{i,t} = (\iota'_1 - \iota'_2) E_t \sum_{j=0}^{\infty} \kappa^j \tilde{X}_{i,t+j+1} \quad (51)$$

$$= (\iota'_1 - \iota'_2) \frac{1}{\kappa} E_t \sum_{j=1}^{\infty} \kappa^j A^j \tilde{X}_{i,t} \quad (52)$$

$$= (\iota'_1 - \iota'_2) A (I - \kappa A)^{-1} \tilde{X}_{i,t} \quad (53)$$

Thus, the present value relation links the current book-to-market ratio with future expected returns and earnings in the form of $K+2$ present-value parameter restrictions for the A matrix and a stochastic singularity, which is why there are only $K+1$ shocks in Equation (50). In particular,

$$\iota'_3 \tilde{X}_{i,t} = (\iota'_1 - \iota'_2) A (I - \kappa A)^{-1} \tilde{X}_{i,t}. \quad (54)$$

⁵We demean all variables with their cross-sectional and time-series grand means for ease of exposition and without loss of generality.

and so

$$\iota'_3 = (\iota'_1 - \iota'_2) A (I - \kappa A)^{-1}. \quad (55)$$

We impose this restriction when estimating the coefficients in the A matrix using panel regressions based on the time-series and cross-section of stock returns and earnings.⁶ Following Campbell (1991), we decompose firm-level book-to-market ratio into a discount rate and a cash flow component. The discount rate component is:

$$\begin{aligned} \widetilde{DR}_{i,t} &= E_t \sum_{j=1}^{\infty} \kappa^{j-1} \tilde{r}_{i,t+j} \\ &= \iota'_1 A (I - \kappa A)^{-1} \tilde{X}_{i,t}, \end{aligned} \quad (56)$$

and the cash flow component is:

$$\begin{aligned} \widetilde{CF}_{i,t} &= E_t \sum_{j=1}^{\infty} \kappa^{j-1} \tilde{e}_{i,t+j} \\ &= \iota'_2 A (I - \kappa A)^{-1} \tilde{X}_{i,t}, \end{aligned} \quad (57)$$

such that $\widetilde{bm}_{i,t} = \widetilde{DR}_{i,t} - \widetilde{CF}_{i,t}$. We also define finite-horizon versions of the cash flow and discount rate components, which we later use in Figure 2 to forecast long-horizon log cash flows and log returns. The N -year discount rate component ($DR_{i,t}^{(N)}$) is:

$$\begin{aligned} \widetilde{DR}_{i,t}^{(N)} &= E_t \sum_{j=1}^N \kappa^{j-1} \tilde{r}_{i,t+j} \\ &= \iota'_1 (I - \kappa^N A^N) A (I - \kappa A)^{-1} \tilde{X}_{i,t}, \end{aligned} \quad (58)$$

The N -year cash flow component of valuations ($CF_{i,t}^{(N)}$) is:

$$\begin{aligned} \widetilde{CF}_{i,t}^{(N)} &= \sum_{j=1}^N \kappa^{j-1} E_t [\tilde{e}_{i,t+j}] \\ &= \iota'_2 (I - \kappa^N A^N) A (I - \kappa A)^{-1} \tilde{X}_{i,t}. \end{aligned} \quad (59)$$

⁶See, e.g., Hansen et al. (2007) for further details regarding the implications of a present value constraint in a VAR.

We decompose the variance of book-to-market ratios into fluctuations arising from discount rates and cash flows, respectively, by noting that:

$$var\left(\widetilde{bm}_{i,t}\right) = var\left(\widetilde{DR}_{i,t}\right) + var\left(\widetilde{CF}_{i,t}\right) - 2Cov\left(\widetilde{DR}_{i,t}, \widetilde{CF}_{i,t}\right). \quad (60)$$

We decompose shocks to realized returns into shocks to expectations about future and current cash flows and future discount rates:

$$r_{i,t+1} - E_t r_{i,t+1} = (E_{t+1} - E_t) \sum_{j=1}^{\infty} \kappa^{j-1} \tilde{e}_{i,t+j} - (E_{t+1} - E_t) \sum_{j=1}^{\infty} \kappa^{j-1} \tilde{r}_{i,t+j+1} \quad (61)$$

$$= CF_{i,t}^{shock} - DR_{i,t}^{shock}, \quad (62)$$

where $CF_{i,t}^{shock} \equiv (E_{t+1} - E_t) \sum_{j=1}^{\infty} \kappa^{j-1} \tilde{e}_{i,t+j}$, $DR_{i,t}^{shock} \equiv (E_{t+1} - E_t) \sum_{j=1}^{\infty} \kappa^{j-1} \tilde{r}_{i,t+j+1}$, and both shocks are simple functions of the VAR estimates. The variance decomposition of returns is then:

$$var\left(r_{i,t} - E_t r_{i,t+1}\right) = var\left(DR_{i,t}^{shock}\right) + var\left(CF_{i,t}^{shock}\right) - 2Cov\left(DR_{i,t}^{shock}, CF_{i,t}^{shock}\right). \quad (63)$$

Appendix C: Relation to Equilibrium Models

The VAR offers a parsimonious, reduced-form model of the cross-section of expected cash flows and discount rates at all horizons. Here we demonstrate that the VAR specification is related to standard asset pricing models. In well-known models such as Campbell and Cochrane's (1999) habit formation model and Bansal and Yaron's (2004) long-run risk model, the log stochastic discount factor is conditionally normally distributed and satisfies:

$$m_{t+1} = -r_{f,t} - \frac{1}{2} \|\lambda_t\|^2 + \lambda_t' \eta_{t+1}, \quad (64)$$

where λ_t is a $K \times 1$ vector of conditional risk prices, η_{t+1} is a $K \times 1$ vector of standard normal shocks, and $r_{f,t}$ is the risk-free rate. With conditionally normal log returns, applying the Law of One Price yields the following expression for the conditional expected log return

of firm i :

$$\begin{aligned} E_t[r_{i,t+1}] &= r_{f,t} - \frac{1}{2}v_{i,t} + cov_t(m_{t+1}, r_{i,t+1}) \\ &= r_{f,t} - \frac{1}{2}v_{i,t} + \beta'_{i,t}\lambda_t, \end{aligned} \quad (65)$$

where $v_{i,t} \equiv var_t(r_{i,t+1})$ is firm return variance, and $\beta_{i,t}^{(k)} = \frac{cov_t(\lambda_t^{(k)}\eta_{t+1}^{(k)}, r_{i,t+1})}{var_t(\lambda_t^{(k)}\eta_{t+1}^{(k)})}$ and $\beta_{i,t} = [\beta_{i,t}^{(1)} \ \beta_{i,t}^{(2)} \ \dots \ \beta_{i,t}^{(K)}]'$ represent firm betas.

We make simplifying assumptions to relate this setup to the VAR specification. Define firm risk premiums as $z_{i,t}^{(k)} \equiv \beta_{i,t}^{(k)}\lambda_t^{(k)}$ and $z_{i,t} = [z_{i,t}^{(1)} \ z_{i,t}^{(2)} \ \dots \ z_{i,t}^{(K)}]'$. Suppose that risk premiums, variances, and the risk-free rate evolve according to:

$$z_{i,t+1} = \bar{z} + A_z(z_{i,t} - \bar{z}) + \Sigma_{z,t}\varepsilon_{i,t+1}^z, \quad (66)$$

$$v_{i,t+1} = \bar{v} + A_v(v_{i,t} - \bar{v}) + \sigma_{v,t}\varepsilon_{i,t+1}^v, \quad (67)$$

$$r_{f,t+1} = \bar{r}_f + A_{r_f}(r_{f,t} - \bar{r}_f) + \sigma_{r,t}\varepsilon_{t+1}^{r_f}, \quad (68)$$

for all firms i . Assume firm log return on equity is also conditionally normal:

$$e_{i,t+1} = \mu + x_{i,t} + \sigma_{e,t}\varepsilon_{i,t+1}^e, \quad (69)$$

$$x_{i,t+1} = A_x x_{i,t} + \Sigma_{x,t}\varepsilon_{i,t+1}^x, \quad (70)$$

where $x_{i,t}$ is an $L \times 1$ vector of latent state variables determining expected return on equity. All shocks can be correlated.

Assuming the clean-surplus model described earlier, firm book-to-market ratios are given by:

$$bm_{i,t} = a_0 + a'_1 r_{f,t} + a'_2 z_{i,t} + a'_3 x_{i,t} + a_4 v_{i,t}. \quad (71)$$

Define the $(2K + L + 1) \times 1$ vector $s_{i,t} = [r'_{f,t} \ z'_{i,t} \ v_{i,t} \dots x'_{i,t}]'$ to consist of the stacked state

variables. We assume there exist $(2K + L + 1)$ observed characteristics, $\xi_{i,t}$, that span $s_{i,t}$:

$$\xi_{i,t} = A_1 + A_2 s_{i,t}, \tag{72}$$

where A_2 is invertible. With the characteristic spanning assumption, firms' book-to-market become a function of the observed characteristics, resulting in a VAR representation of the present-value relation. In sum, the VAR specification concisely summarizes the dynamics of expected cash flows and discount rates, even when both consist of multiple components fluctuating at different frequencies. The VAR yields consistent estimates even though there is heteroskedasticity across firms and time.

When analyzing long-short portfolios, we obtain the anomaly cash flow (discount rate) shock as the difference in the cash flow (discount rate) shocks between the long and short portfolios. Taking the value anomaly as an example, suppose the long value portfolio and short growth portfolio have the same betas with respect to all risk factors except the value factor (say, $\lambda_t^{(2)}$). According to Equation (65), discount rate shocks to this long-short portfolio can only arise from three sources: 1) shocks to the spread in the factor exposure between value and growth firms ($\beta_{\text{value},t}^{(2)} - \beta_{\text{growth},t}^{(2)}$); 2) shocks to the price of risk of the value factor ($\lambda_t^{(2)}$); or 3) shocks to the difference in return variance between the two portfolios. The third possibility arises because we analyze log returns. Similarly, cash flow shocks to this long-short portfolio only reflect these portfolios' differential exposure to cash flow factors.

Table 1 - Summary Statistics

Table 1: Panel A shows summary statistics for firm-level returns, cash flows, and characteristics. The first column lists the variables as defined in the text. The second column reports the number of firm-year observations, n . The remaining columns report the mean, standard deviation and various percentiles of the firm-year distribution for each variable. Panel B provides the correlation matrix for these variables. The sample spans the years 1964 through 2015.

Panel A:	N	Mean	SD	P1	P10	P50	P90	P99
AnnRet	68,642	12.66	50.70	-81.06	-39.82	7.69	65.60	175.00
Rf	68,642	5.35	3.12	0.12	0.31	5.55	8.61	13.96
Volat	63,564	38.69	16.94	15.34	21.83	35.11	58.75	102.45
SizeWt	68,642	0.31	1.06	0.00	0.01	0.06	0.55	5.38
lnROE ^{CS}	65,277	11.29	32.27	-86.64	-11.90	10.23	36.33	117.96
lnBM	67,299	-0.72	0.83	-3.04	-1.78	-0.66	0.29	0.98
lnME	68,593	4.89	1.31	2.90	3.37	4.65	6.73	8.62
lnProf	66,364	21.27	26.92	-72.50	6.39	23.33	39.14	77.92
lnInv	67,478	16.10	28.27	-31.35	-4.87	9.91	42.83	132.75
lnIssue	64,938	8.09	20.75	-29.03	-7.92	2.16	31.22	92.29

Panel B:	1	2	3	4	5	6	7	8	9
AnnRet (1)	1.00								
Rf (2)	0.03	1.00							
Volat (3)	-0.15	0.00	1.00						
SizeWt (4)	-0.02	-0.19	-0.11	1.00					
lnROE ^{CS} (5)	0.13	-0.01	-0.01	0.02	1.00				
lnBM (6)	-0.37	0.21	-0.03	-0.14	-0.20	1.00			
lnME (7)	0.23	-0.10	-0.35	0.54	0.08	-0.28	1.00		
lnProf (8)	0.09	0.04	-0.25	0.07	0.21	-0.11	0.19	1.00	
lnInv (9)	0.00	0.07	0.19	-0.03	0.55	-0.19	-0.03	0.02	1.00
lnIssue (10)	0.00	0.03	0.27	-0.08	0.11	-0.07	-0.10	-0.19	0.39

Table 2 - Return and Earnings Forecasting Regressions

Table 2: The table shows the regression coefficients for forecasting regressions of firms' annual real log returns (lnRet) and log annual real clean-surplus earnings (lnEarn) on the one-year-lagged value of the characteristics. The rightmost column shows the implied coefficients for a regression of firms' log book-to-market (lnBM) ratios on lagged characteristics. Beyond lnBM, the other characteristics are log profitability (lnProf), log asset growth (lnInv), log market equity (lnME), log three-year issuance (lnIssue), realized variance (lnRV), and the log one-year real risk-free rate (lnRf). The sample spans the years 1964 through 2015. Standard errors clustered by year and firm appear in parenthesis. N denotes the number of observations. The marks '+', '**', and '**' indicate significance at the 10, 5, and 1 percent levels, respectively.

	lnRet	lnROE ^{CS}	lnBM
Lag lnRet	-0.003 (0.056)	0.118** (0.014)	0.126* (0.055)
Lag lnROE ^{CS}	-0.021 (0.029)	-0.039* (0.016)	-0.019 (0.024)
Lag lnBM	0.045** (0.015)	-0.143** (0.010)	0.846** (0.019)
Lag lnProf	0.043** (0.014)	0.037** (0.009)	-0.007 (0.020)
Lag lnInv	-0.048** (0.012)	0.003 (0.005)	0.053** (0.010)
Lag lnME	-0.012 (0.012)	-0.013** (0.004)	-0.001 (0.011)
Lag lnIssue	-0.011+ (0.007)	0.014** (0.003)	0.027** (0.006)
Lag lnRV	-0.036 (0.025)	-0.007 (0.007)	0.030 (0.021)
Lag lnRf	0.000 (0.029)	0.012 (0.009)	0.011 (0.024)
Constant	-0.018 (0.059)	-0.027 (0.018)	-0.010 (0.051)
R^2	0.046	0.243	0.675
N	53,737	53,737	53,737

Table 3 - Characteristic Forecasting Regressions

Table 3: Panel A shows the regression coefficients for annual forecasting regressions of firm characteristics on their own lag as well as the firm's lagged book-to-market ratio. The characteristics are log profitability (lnProf), log asset growth (lnInv), log market equity (lnME), log three-year issuance (lnIssue), and realized variance (lnRV). Panel B reports the regression coefficients of the aggregate variable, the log one-year real risk-free rate (lnRf), which is regressed only on its own lag. The sample spans the years 1964 through 2015. Standard errors clustered by year and firm appear in parenthesis. N denotes the number of observations. The marks '+', '**', and '***' indicate significance at the 10, 5, and 1 percent levels, respectively.

Panel A:	Own Lag	Lag lnBM	R^2	N
lnProf	0.678** (0.045)	-0.084** (0.015)	45.6%	54,054
lnInv	0.154** (0.025)	-0.353** (0.038)	17.7%	54,068
lnME	0.973** (0.005)	0.020 (0.014)	91.0%	54,099
lnIssue	0.711** (0.010)	-0.029** (0.011)	61.1%	54,105
lnRV	0.696** (0.069)	-0.073* (0.035)	50.5%	53,695

Panel B:	Own Lag	R^2	N
Lag lnRf	0.602** (0.183)	36.3%	51

Table 4 - Short-run vs. Long-run Discount Rates

Table 4: The table shows the sensitivity of short-run and long-run expected returns to standardized increases in firm- and market-level characteristics. Short-run discount rates are one-year expected log returns. Long-run discount rates are the annualized discount rate component of log book-to-market ratios, obtained from the panel VAR. The sample spans the years 1964 through 2015. Standard errors clustered by year and firm appear in parenthesis. The last row indicates the total standard deviation of short-run and long-run discount rates. The marks '+', '*', and '**' indicate significance at the 10, 5, and 1 percent levels, respectively.

	Short-run Discount Rate (Annual)	Long-run Discount Rate (Annualized)
Lag lnRet	-0.11% (2.42%)	0.05% (0.08%)
Lag lnROE ^{CS}	-0.65% (0.90%)	-0.03% (0.03%)
Lag lnBM	3.81%** (1.27%)	1.02%** (0.21%)
Lag lnProf	4.38%** (1.41%)	0.48%** (0.13%)
Lag lnInv	-4.79%** (1.23%)	-0.16%** (0.04%)
Lag lnME	-1.17% (1.15%)	-0.74% (0.52%)
Lag lnIssue	-1.10%+ (0.67%)	-0.05% (0.06%)
Lag lnRV	-3.63% (2.48%)	-0.34% (0.23%)
Lag lnRf	0.10% (2.90%)	0.04% (0.20%)
St. Dev. of Discount Rate	9.53%	1.42%

Table 5 - Firm-level Variance Decompositions

Table 5: Panel A displays the decomposition of the variance of log book-to-market ratios into cash flow (CF) and discount rate (DR) components. Variances are computed over time and across firms and are measured as a fraction of the variance of book-to-market ratios. Panel B provides the variance decomposition of log returns into CF and DR components. The sample spans the years 1964 through 2015. Standard errors appear in parentheses. The marks '+', '**', and '***' indicate significance at the 10, 5, and 1 percent levels, respectively.

	$var(DR)$	$var(CF)$	$-2cov(DR, CF)$	$Corr(DR, CF)$
Panel A:				
Fraction of $var(\ln BM)$	0.190 ⁺ (0.110)	0.473 ^{**} (0.068)	0.338 ^{**} (0.094)	-0.564 ⁺ (0.295)
Panel B:				
Fraction of $var(r)$	0.209 ⁺ (0.117)	0.522 ^{**} (0.111)	0.270 ^{**} (0.064)	-0.409 [*] (0.160)

Table 6 - Market Variance Decompositions

Table 6: The table shows variance decompositions of market log returns into cash flow (CF) and discount rate (DR) components. Panel A shows the variance decomposition derived from the firm-level panel VAR, as explained in the text. Panel B shows the variance decomposition derived from the market-level VAR, using only the market's clean surplus earnings, returns, and book-to-market ratio. The sample spans the years 1964 through 2015. Standard errors appear in parentheses. The marks '+', '*', and '**' indicate significance at the 10, 5, and 1 percent levels, respectively.

	$var(DR)$	$var(CF)$	$var(Cross)$	$-2cov(DR, CF)$	$Corr(DR, CF)$	$Corr(Pred, Act)$
Panel A: Panel VAR						
Fraction of $var(r_m)$	0.183 (0.128)	0.632** (0.176)	0.009* (0.004)	0.219 (0.237)	-0.322 (0.466)	0.986** (0.001)
Panel B: Market VAR						
Fraction of $var(r_m)$	0.281 (0.226)	0.248 (0.181)		0.471** (0.052)	-0.892** (0.148)	

Table 7 - Anomaly Variance Decompositions

Table 7: The table shows decompositions of the variance of log anomaly returns into cash flow (CF) and discount rate (DR) components. The anomaly return is the difference between the log return of the top quintile portfolio and the log return of the bottom quintile portfolio, where the quintile sort is based on the relevant characteristic. The sample spans the years 1964 through 2015. Standard errors appear in parentheses. The marks '+', '**', and '***' indicate significance at the 10, 5, and 1 percent levels, respectively.

	$var(DR)$	$var(CF)$	$var(Cross)$	$-2cov(DR, CF)$	$Corr(DR, CF)$	$Corr(Pred, Act)$
<i>Book-to-market:</i>						
Fraction of $var(r_{bm})$	0.176 (0.098)	0.459** (0.121)	0.017* (0.007)	0.374** (0.046)	-0.658** (0.146)	0.967** (0.004)
<i>Profitability:</i>						
Fraction of $var(r_{prof})$	0.181+ (0.099)	0.466** (0.119)	0.032** (0.009)	0.361** (0.034)	-0.621** (0.101)	0.882** (0.026)
<i>Size:</i>						
Fraction of $var(r_{size})$	0.165 (0.109)	0.371** (0.138)	0.019** (0.007)	0.384** (0.038)	-0.777** (0.130)	0.938** (0.004)
<i>Issuance:</i>						
Fraction of $var(r_{issue})$	0.177+ (0.092)	0.483** (0.131)	0.035** (0.012)	0.450** (0.043)	-0.769** (0.091)	0.958** (0.008)
<i>Investment:</i>						
Fraction of $var(r_{inv})$	0.171* (0.083)	0.518** (0.124)	0.018** (0.004)	0.382** (0.045)	-0.641** (0.120)	0.950** (0.009)

Table 8 - Correlations between Anomaly and Market Return Components

Table 8: The table shows correlations between market cash flow and discount rate shocks and the anomaly cash flow and discount rate shocks. The sample spans the years 1964 through 2015. Standard errors appear in parentheses. The marks '+', '**', and '***' indicate significance at the 10, 5, and 1 percent levels, respectively.

	Market CF		Market DR	
	Anomaly CF	Anomaly DR	Anomaly CF	Anomaly DR
Book-to-market	0.06 (0.13)	-0.04 (0.07)	0.36** (0.10)	-0.17 (0.11)
Profitability	-0.17** (0.04)	0.17+ (0.11)	0.11 (0.16)	-0.34** (0.04)
Investment	-0.01 (0.16)	-0.07 (0.08)	-0.38** (0.08)	0.17+ (0.10)
Size	0.00 (0.14)	0.33** (0.11)	0.08 (0.08)	-0.04 (0.14)
Issuance	0.23 (0.14)	-0.32** (0.08)	-0.34* (0.16)	0.48** (0.07)

Table 9 - Correlations between Anomaly Return Components

Table 9: The table shows correlations between the cash flow shocks (Panel A) and discount rate shocks (Panel B) to various anomalies. The sample spans the years 1964 through 2015. Standard errors appear in parentheses. The marks '+', '**', and '**' indicate significance at the 10, 5, and 1 percent levels, respectively.

Panel A: Cash Flow Shocks	1	2	3	4
Book-to-market (1)	1.00			
Profitability (2)	-0.29** (0.03)	1.00		
Investment (3)	-0.66** (0.03)	0.25** (0.03)	1.00	
Size (4)	-0.18+ (0.11)	0.25** (0.05)	0.25** (0.06)	1.00
Issuance (5)	-0.27** (0.03)	-0.40** (0.03)	0.52** (0.03)	-0.14** (0.03)

Panel B: Discount Rate Shocks	1	2	3	4
Book-to-market (1)	1.00			
Profitability (2)	-0.30** (0.06)	1.00		
Investment (3)	-0.62** (0.04)	0.20* (0.08)	1.00	
Size (4)	-0.34** (0.02)	0.27** (0.04)	0.07 (0.10)	1.00
Issuance (5)	-0.25** (0.06)	-0.50** (0.04)	0.52** (0.04)	-0.24** (0.06)

Table 10 - Market Variance Decompositions in Alternative Specifications

Table 10: The table shows variance decompositions of market log returns into cash flow (CF) and discount rate (DR) components, derived from alternative specifications of the firm-level panel VAR, as explained in the text. Spec2 refers to the specification that includes the aggregate book-to-market ratio in the panel VAR, and Spec3 refers to the specification that in addition includes interaction terms as explained in the text. The sample spans the years 1964 through 2015. Standard errors appear in parentheses. The marks '+', '*', and '**' indicate significance at the 10, 5, and 1 percent levels, respectively.

		$var(DR)$	$var(CF)$	$var(Cross)$	$-2cov(DR, CF)$	$Corr(DR, CF)$	$Corr(Pred, Act)$
<i>Spec2</i>	Fraction of $var(r_m)$	0.917 (0.874)	0.262 (0.166)	0.005 (0.005)	-0.137 (0.906)	0.139 (0.843)	0.986** (0.002)
<i>Spec3</i>	Fraction of $var(r_m)$	1.058 (2.186)	0.331 (0.335)	0.019 (0.060)	-0.261 (2.231)	0.221 (1.586)	0.986** (0.002)

Table 11 - Anomaly Variance Decompositions in Alternative Specifications

Table 11: The table shows decompositions of the variance of log anomaly returns into cash flow (CF) and discount rate (DR) components. The anomaly return is the difference between the log return of the top quintile portfolio and the log return of the bottom quintile portfolio, where the quintile sort is based on the relevant characteristic. The variance-decompositions are derived from alternative specifications of the firm-level panel VAR, as explained in the text. Spec2 refers to the specification that includes the aggregate book-to-market ratio in the panel VAR, and Spec3 refers to the specification that in addition includes interaction terms as explained in the text. The sample spans the years 1964 through 2015. Standard errors appear in parentheses. The marks '+', '*', and '**' indicate significance at the 10, 5, and 1 percent levels, respectively.

		$var(DR)$	$var(CF)$	$var(Cross)$	$-2cov(DR, CF)$	$Corr(DR, CF)$	$Corr(Pred, Act)$
<i>Book-to-market:</i>							
<i>Spec2</i>	Fraction of $var(r_{bm})$	0.091 (0.080)	0.669** (0.220)	0.029 (0.027)	0.282** (0.082)	-0.571** (0.170)	0.967** (0.003)
<i>Spec3</i>	Fraction of $var(r_{bm})$	0.115 (0.117)	1.191 (0.824)	0.077 (0.221)	-0.107 (0.575)	0.144 (0.682)	0.967** (0.003)
<i>Profitability:</i>							
<i>Spec2</i>	Fraction of $var(r_{prof})$	0.114 (0.091)	0.619** (0.187)	0.043 (0.031)	0.302** (0.084)	-0.568** (0.132)	0.886** (0.022)
<i>Spec3</i>	Fraction of $var(r_{prof})$	0.376 (0.251)	0.713 (0.213)	0.038 (0.061)	-0.037 (0.338)	0.036 (0.314)	0.906** (0.026)
<i>Size:</i>							
<i>Spec2</i>	Fraction of $var(r_{size})$	0.079 (0.084)	0.608** (0.262)	0.028 (0.020)	0.299** (0.096)	-0.681** (0.158)	0.914** (0.008)
<i>Spec3</i>	Fraction of $var(r_{size})$	0.311 (0.528)	1.092 (0.870)	0.085 (0.186)	-0.221 (0.906)	0.190 (0.565)	0.919** (0.018)
<i>Issuance:</i>							
<i>Spec2</i>	Fraction of $var(r_{issue})$	0.084 (0.075)	0.838** (0.361)	0.067 (0.065)	0.362 (0.127)	-0.683** (0.151)	0.957** (0.006)
<i>Spec3</i>	Fraction of $var(r_{issue})$	0.218 (0.170)	1.066 (0.970)	0.133 (0.340)	0.214 (0.271)	-0.221 (0.321)	0.957** (0.006)
<i>Investment:</i>							
<i>Spec2</i>	Fraction of $var(r_{inv})$	0.083* (0.066)	0.842** (0.312)	0.045 (0.048)	0.240+ (0.135)	-0.454+ (0.274)	0.953** (0.006)
<i>Spec3</i>	Fraction of $var(r_{inv})$	0.103 (0.065)	0.912 (0.720)	0.054 (0.183)	0.134 (0.352)	-0.218 (0.658)	0.954** (0.008)

Figure 1 - Regular ROE vs. Clean-Surplus ROE

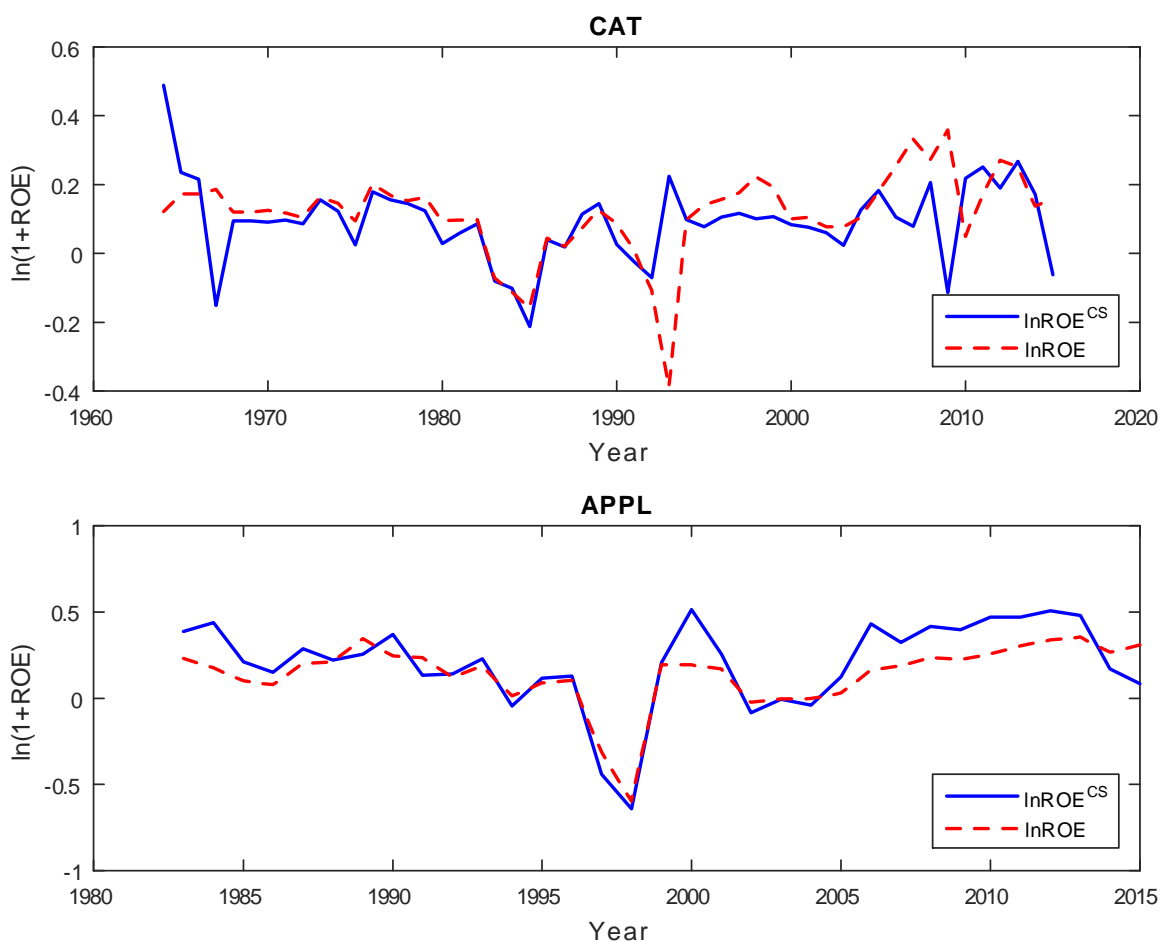


Figure 1: The top plot (plot A) shows the time-series of annual log ROE for Caterpillar (CAT). The dashed red line shows annual log ROE based on accounting data (net income and book equity). The solid blue line shows clean surplus ROE imputed as a residual from the log-linear valuation model, as explained in the text. The lower plot (plot B) shows the same for Apple (AAPL).

Figure 2 - Cumulative Return Forecasting Coefficients

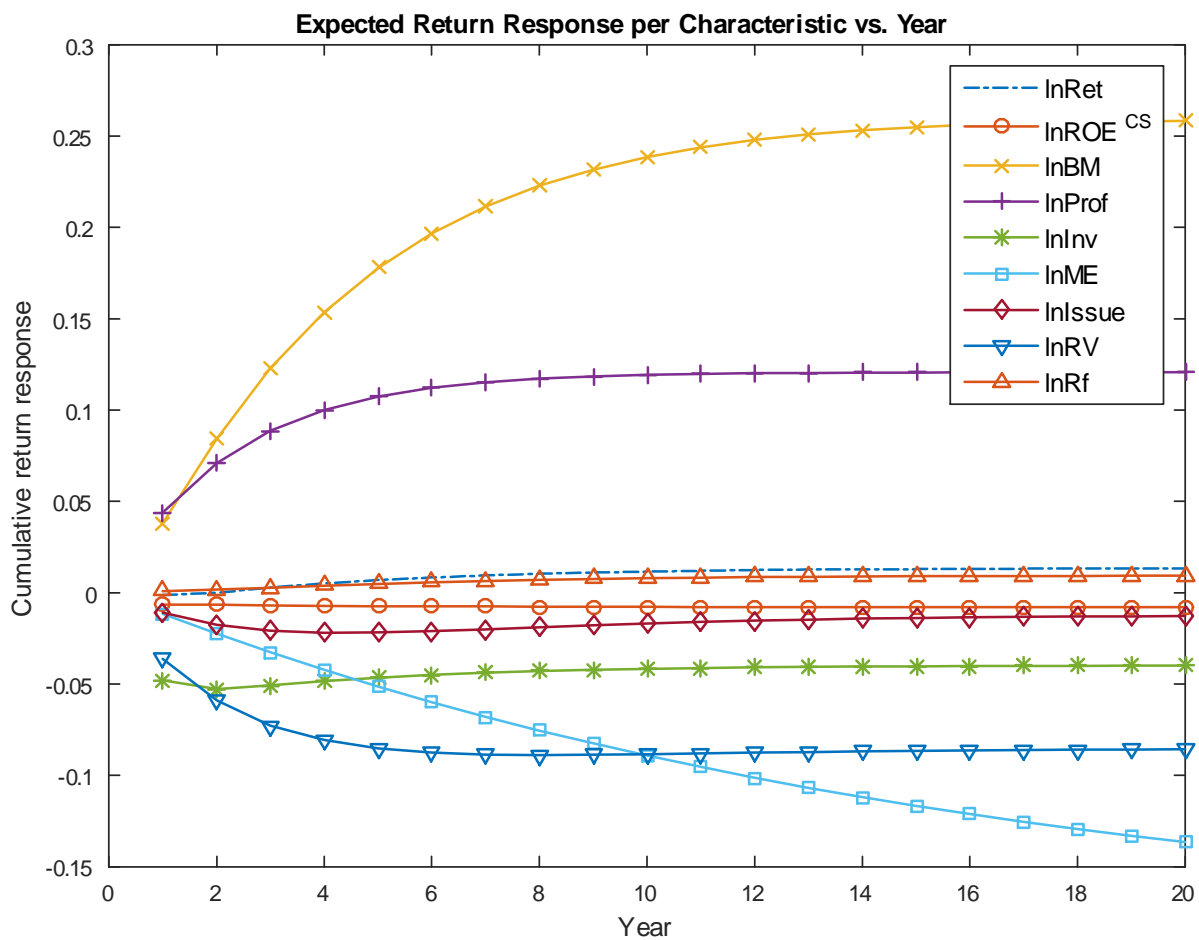


Figure 2: The plot shows the cumulative coefficients of real log return forecasts based on each characteristic (y-axis) for forecasting horizons of 1 to 20 years (x-axis), as implied by the panel VAR. When computing the cumulative coefficient, the coefficient for horizon j is multiplied by κ^j , where $\kappa = 0.96$ as in the text. Thus, the cumulative coefficient for each horizon represents the discount rate component of log book-to-market ratio for that horizon.

Figure 3 - Cumulative Earnings Forecasting Coefficients

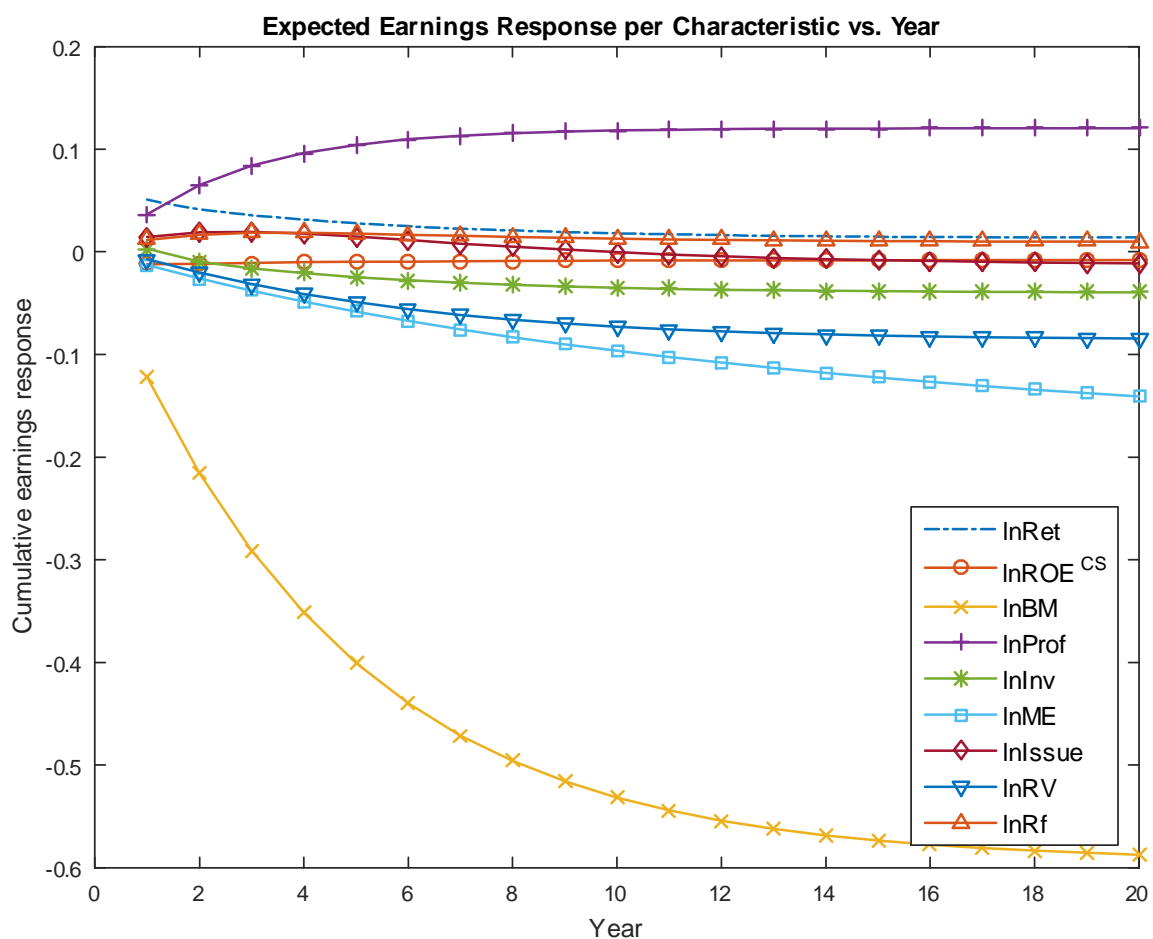


Figure 3: The plot shows the cumulative coefficients of real log earnings forecasts based on each characteristic (y-axis) for forecasting horizons of 1 to 20 years (x-axis), as implied by the panel VAR. When computing the cumulative coefficient, the coefficient for horizon j is multiplied by κ^j , where $\kappa = 0.96$ as in the text. Thus, the cumulative coefficient for each horizon represents the cash flow component of log book-to-market ratio for that horizon.

Figure 4 - Naive Extrapolation of Discount Rates

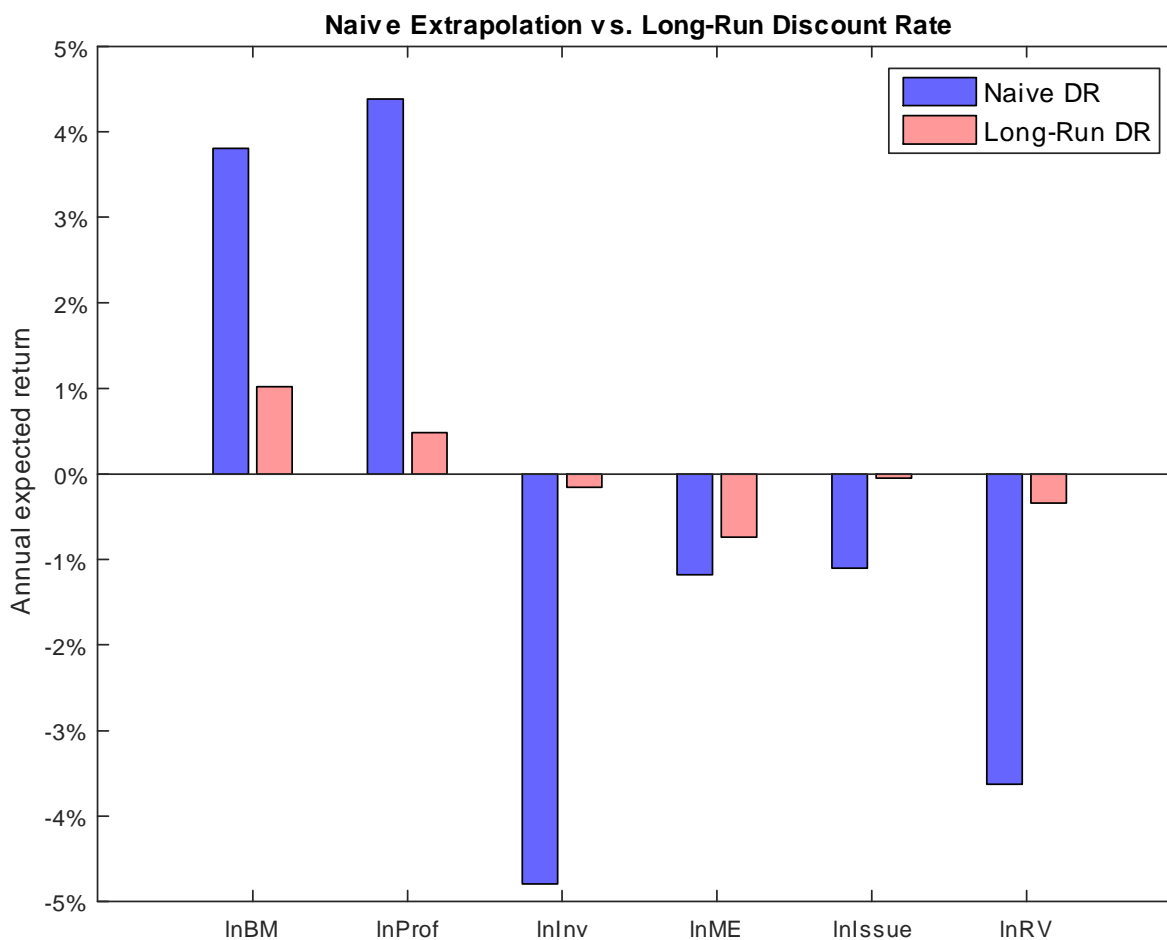


Figure 4: The figure compares a naïve estimate of the long-run discount rate to an estimate of the long-run discount rate based on the panel VAR. The naïve estimate is an extrapolation of the one-year discount rate. The long-run discount rate estimate is the annualized average of the discount rate applied to valuations implied by the panel VAR coefficients. The blue bars (leftmost in each category) are the naïve estimates, while the light red (rightmost in each category) bars are the long-run expected returns.

Figure 5 - Predicting 10-year Market Earnings and Returns

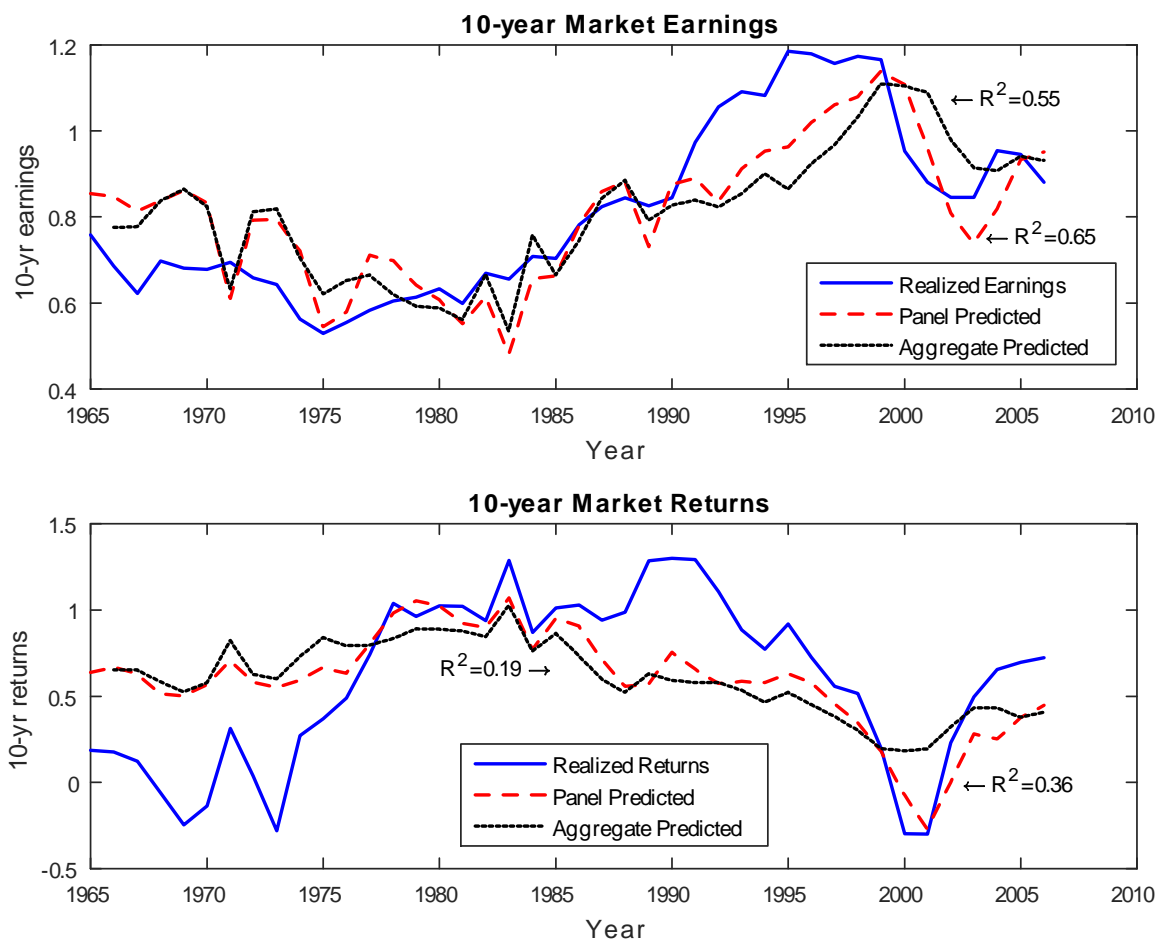


Figure 5: The top plot shows realized versus predicted 10-year log clean surplus earnings of the market portfolio. The solid blue line corresponds to realized earnings, while the dashed red and dotted black lines represent predicted earnings from the panel VAR and market-level VAR, respectively. The year on the x-axis is the year of the prediction -e.g., year 2005 corresponds to the 10-year realized earnings in 2006-2015. The bottom plot shows the corresponding for 10-year log real market returns.

Figure 6 - Predicting 10-year Value Anomaly Earnings and Returns

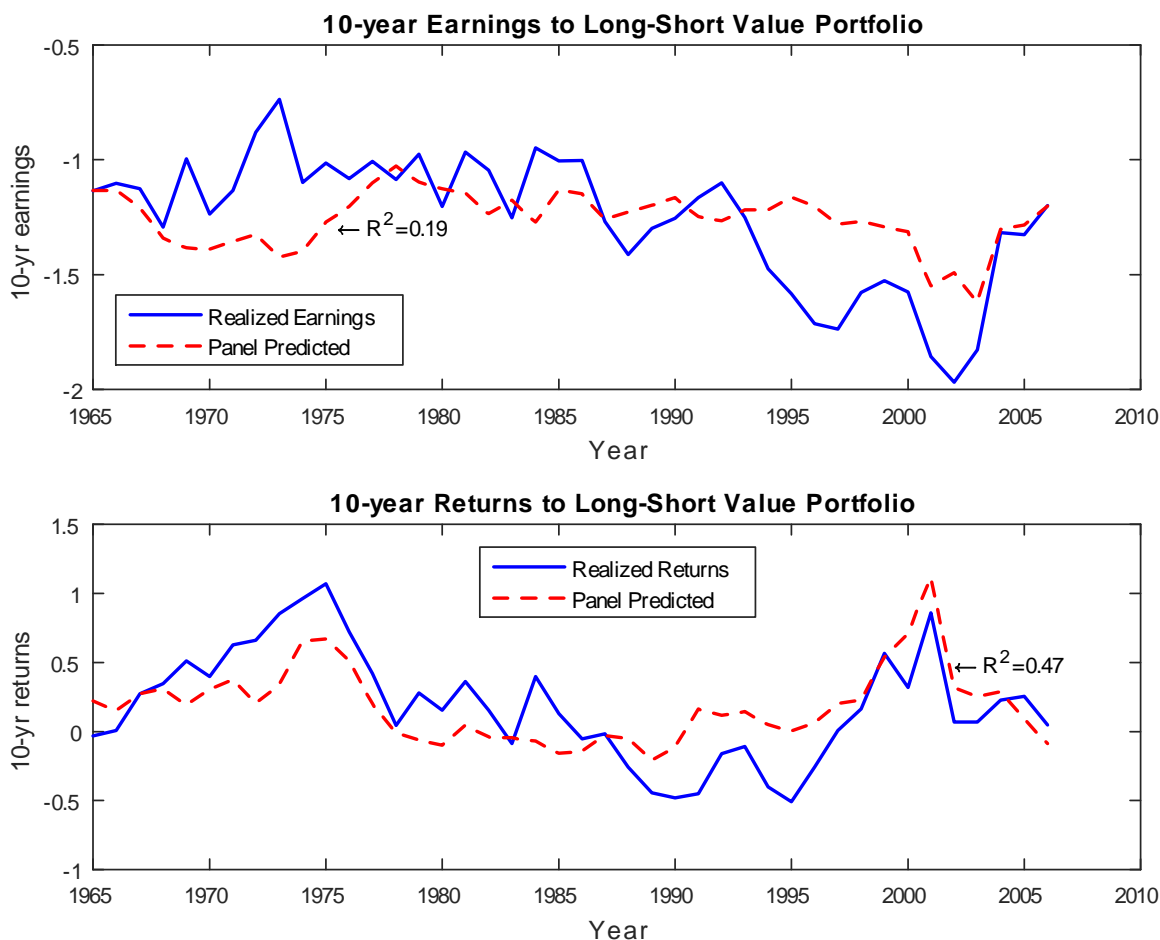


Figure 6: The top plot shows realized versus predicted 10-year log clean-surplus earnings of the long-short portfolio formed by sorting on book-to-market ratios. The solid blue line corresponds to realized earnings, while the dashed red line represent predicted earnings from the panel VAR. The year on the x-axis is the year of the prediction-e.g., year 2005 corresponds to the 10-year realized earnings in 2006-2015. The bottom plot shows the corresponding for 10-year real log returns to the long-short value portfolio.

Figure 7 - Predictive Power of Valuation Components

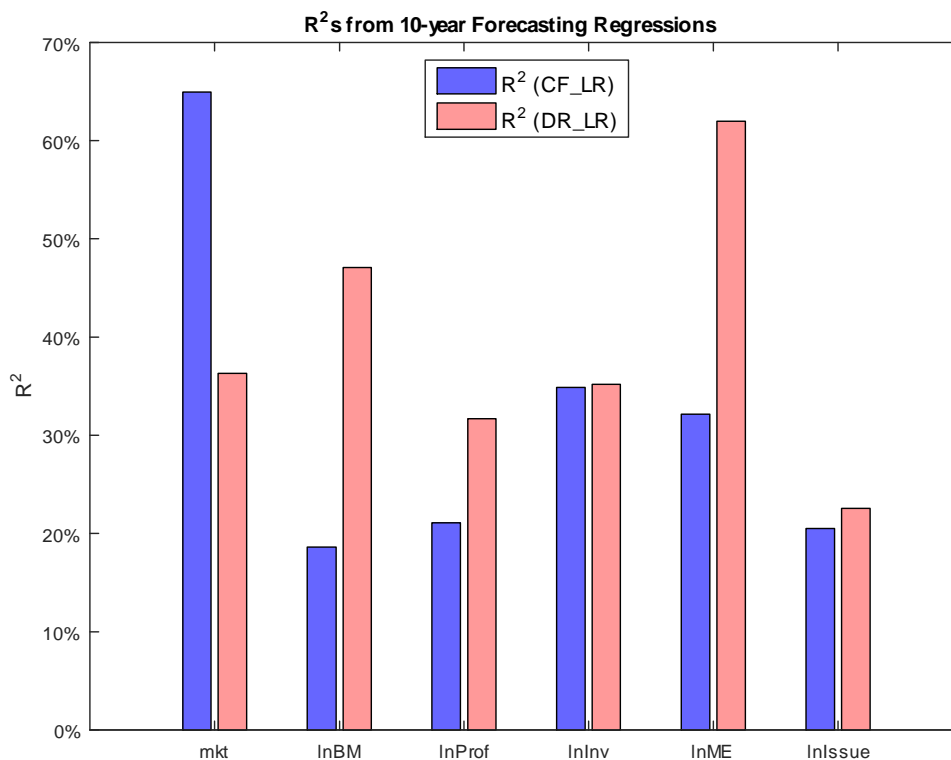


Figure 7: The figure shows the R2 statistics from regressions forecasting either 10-year earnings or returns of portfolios. The blue (left) bars represent the predictive power of regressions of 10-year log clean-surplus earnings on the cash flow components of firms' log book-to-market ratios (CF_LR) aggregated to the relevant portfolio level. The light red (right) bars represent the predictive power of regressions of 10-year log real returns on the discount rate components of firms' log book-to-market ratios (DR_LR). The portfolios are the market portfolio (Mkt), as well as top quintile minus bottom quintile portfolios sorted on book-to-market (B/M), profitability (Prof), investment (Inv), size (ME), and issuance (Issue). See the main text for details regarding the construction of the test portfolios and the corresponding cash flow and discount rate components. The sample spans the years 1964 through 2015.

Figure 8 - Predicting 10-year Market Earnings and Returns in Alternative Specifications

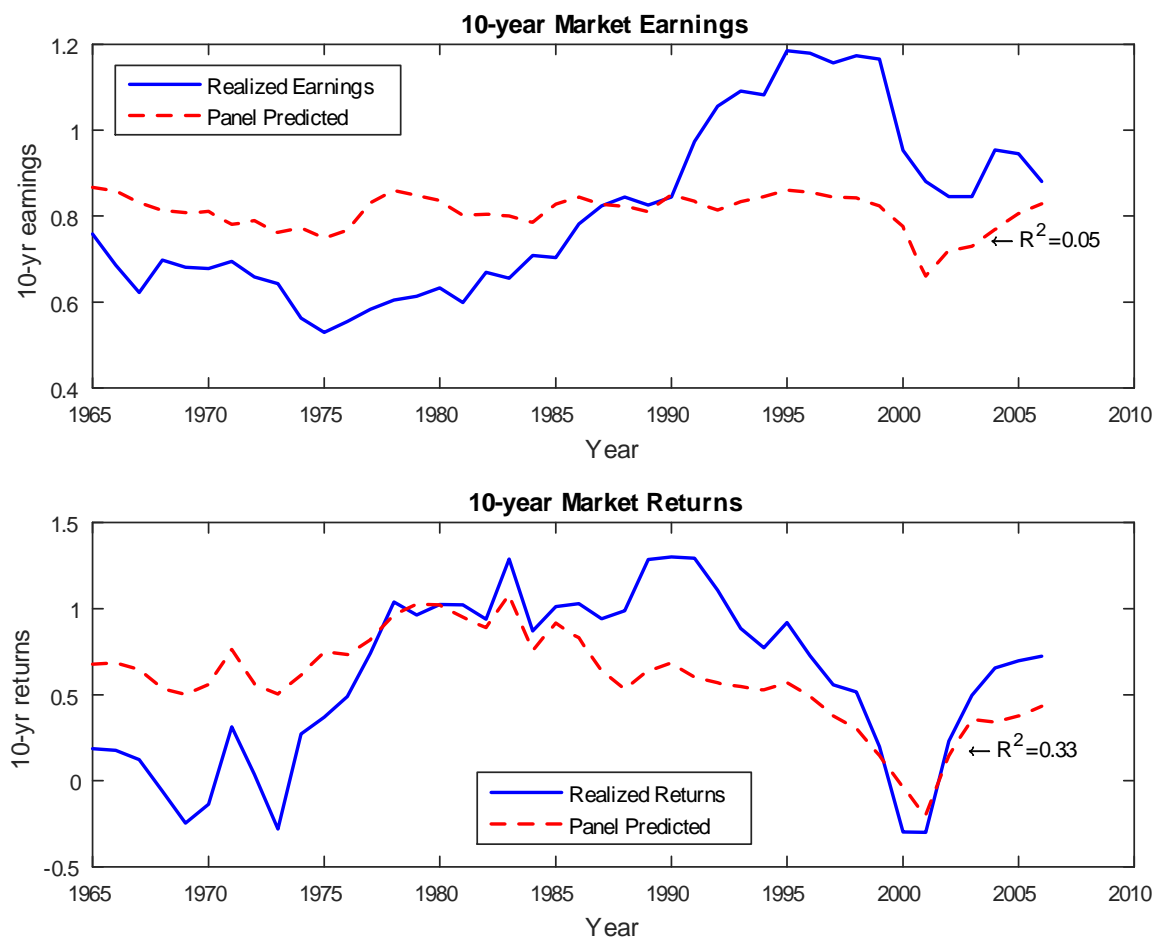


Figure 8: The top plot shows realized versus predicted 10-year log clean surplus earnings of the market portfolio. The solid blue line corresponds to realized earnings, while the red, dashed line represent predicted earnings from an alternative specification of the panel VAR (v2, where the aggregate book-to-market ratio is included in the VAR). The year on the x-axis is the year of the prediction -e.g., year 2005 corresponds to the 10-year realized earnings in 2006-2015. The bottom plot shows the corresponding for 10-year log real market returns.

Figure 9 - Predictive Power of Valuation Components in Alternative Specifications

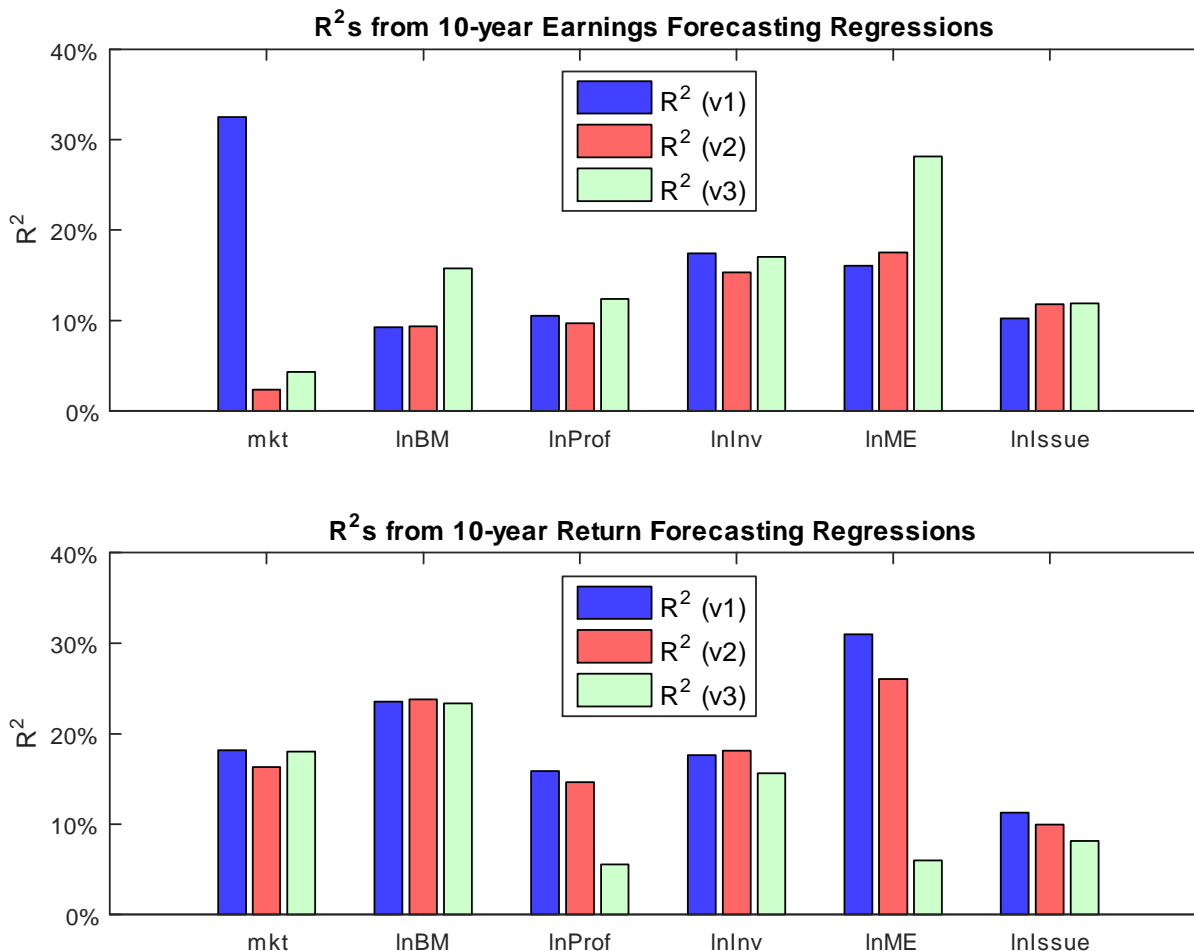


Figure 9: The figure shows the R^2 statistics from regressions forecasting either 10-year earnings (top plot) or returns (bottom plot) of portfolios. The dark blue (left) bars represent the predictive power of regressions using long-run cash flow or discount rate components of the log book-to-market ratios from the main specification of the panel VAR (specification v1). The red (middle) bars corresponds specification v2 (adding the aggregate book-to-market ratio to the panel VAR), while the light green (right) bars corresponds specification v3 (also adding interaction terms, as explained in the text). The portfolios are the market portfolio (Mkt), as well as top quintile minus bottom quintile portfolios sorted on book-to-market (B/M), profitability (Prof), investment (Inv), size (ME), and issuance (Issue). See the main text for details regarding the construction of the test portfolios and the corresponding cash flow and discount rate components. The sample spans the years 1964 through 2015.