

NBER WORKING PAPER SERIES

WHAT DRIVES FIRM-LEVEL STOCK RETURNS?

Tuomo Vuolteenaho

Working Paper 8240
<http://www.nber.org/papers/w8240>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 2001

I would like to thank Tom Berglund, Jonathan Berk, John Campbell, John Cochrane, Randy Cohen, Tim Doede, Eugene Fama, Richard Green, J. B. Heaton, John Heaton, Peter Hecht, Anita Kelly, Matti Keloharju, Robert Kimmel, Owen Lamont, Aaron Levovitz, Robert Merton, Tobias Moskowitz, Christopher Polk, David Robinson, Katherine Schipper, Erik Stafford, Per Strömberg, Richard Thaler, Samuel Thompson, and Pietro Veronesi for helpful discussions. I also received useful comments from an anonymous referee and the seminar participants at the Anderson School, Columbia Business School, EFA2000 Meeting, Fuqua School of Business, Haas School of Business, Harvard Business School, Harvard University Department of Economics, Johnson Graduate School of Management, Kellogg Graduate School of Management, London Business School, MIT Sloan School of Management, Simon Graduate School of Business Administration, and Wharton School. I am grateful for the financial support of the Foundation for Economic Education, the Emil Aaltonen Foundation, the Oscar Mayer Foundation, and the John Leusner Fund at the University of Chicago Graduate School of Business. The views expressed herein are those of the author and not necessarily those of the National Bureau of Economic Research.

© 2001 by Tuomo Vuolteenaho. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

What Drives Firm-Level Stock Returns?

Tuomo Vuolteenaho

NBER Working Paper No. 8240

April 2001

JEL No. G120, G140

ABSTRACT

I use a vector autoregressive model (VAR) to decompose an individual firm's stock return into two components: changes in cash-flow expectations (i.e., cash-flow news) and changes in discount rates (i.e., expected-return news). The VAR yields three main results. First, firm-level stock returns are mainly driven by cash-flow news. For a typical stock, the variance of cash-flow news is more than twice that of expected-return news. Second, shocks to expected returns and cash flows are positively correlated for a typical small stock. Third, expected-return-news series are highly correlated across firms, while cash-flow news can largely be diversified away in aggregate portfolios.

Tuomo Vuolteenaho

Harvard University

Department of Economics

Littauer Center 312

1875 Cambridge Street

Cambridge, MA 02138, USA

and NBER

t_vuolteenaho@harvard.edu

By definition, a firm's stock returns are driven by shocks to expected cash flows (i.e., cash-flow news) and/or shocks to discount rates (i.e., expected-return news). There is a substantial body of research measuring the relative importance of cash-flow and expected-return news for aggregate portfolio returns (e.g., Campbell [1991] and Campbell and Ammer [1993]), but virtually no evidence on the relative importance of these components at the firm level.

In this paper, I estimate a vector autoregression (VAR) from a large firm-level panel (the 1954-1996 CRSP-COMPUSTAT intersection). The VAR and Campbell's (1991) return-decomposition framework enable me to decompose the firm-level stock return into cash-flow and expected-return news and to estimate how important these two sources of stock-return variation are for an individual firm. In addition, I measure whether positive cash-flow news is typically associated with an increase or decrease in expected returns.

The VAR is designed to capture the following empirical return-predictability results. Historically, small firms have earned higher average stock returns than large firms (Banz's [1981] "size effect"). Past long-term losers have outperformed past long-term winners ("long-term reversal," DeBondt and Thaler [1985]), while past short-term winners have outperformed past short-term losers ("momentum," Jegadeesh and Titman [1993]). High book-to-market-equity firms have earned higher average stock returns than low book-to-market-equity firms ("book-to-market anomaly," Rosenberg, Reid, and Lanstein [1985]). Controlling for other characteristics, firms with higher profitability have earned higher average stock returns (Haugen and Baker [1996]). Also, high-leverage firms have historically outperformed low-leverage firms (Bhandari's [1988] "leverage effect"). I avoid modeling corporate dividend policy by using an accounting-based present-value model and excluding any dividend-based variables from the VAR.¹

My first result concerns the relative importance of cash-flow and expected-return news for firm-level stock returns. For excess returns (log stock return less log risk-free return), the variance of expected-return news (0.0645 or standard deviation 22%) is approximately one half of the variance of cash-flow news (0.1002 or standard deviation 32%). For market-adjusted returns (log return less cross-sectional average log return), the variance of expected-return news (0.0161 or standard deviation 13%) is one fifth that of cash-flow-news (0.0801 or standard deviation 28%). Thus, information about future cash flows is the dominant factor driving firm-level stock returns.

The variance-decomposition differs by firm size, i.e., the market capitalization of equity. Both cash-flow-news and expected-return-news variance decrease as a function of firm size. The ratio of expected-

return-news to total-return variance is also higher for small firms. The cross-sectional differences in the cash-flow-news variances make sense if large firms are better-diversified portfolios of investment projects than small firms. This diversification argument is unlikely to completely explain the patterns in expected-return-news variances, however.

I also find that cash-flow news is positively correlated with shocks to expected returns for a typical stock. This correlation appears to be the largest for the smallest stocks, declining (nearly) monotonically in size. These correlations are used to interpret previous evidence relating to market over- or underreaction to news about future cash flows.

Finally, I reconcile the firm-level-return-variance decomposition with the aggregate-return-variance decompositions in the previous literature. Campbell and Shiller (1988a,b), Campbell (1991), and Campbell and Ammer (1993) decompose the variance of aggregate stock returns. In post-war data, the expected-return-news variance dominates the cash-flow-news variance. I show that the highly variable firm-level cash-flow-news component is largely diversified away in aggregate portfolios. I calculate the two news series for individual firms and then aggregate them into an equal-weight index portfolio. While the variance of cash-flow news is twice that of expected-return news for firm-level excess returns, the cash-flow-news variance is only three quarters of the expected-return-news variance for an equal-weight portfolio. It appears that while cash-flow information is largely firm specific, expected-return information is predominantly driven by systematic, macroeconomic components.

The rest of the paper is organized as follows. Section I outlines the stock-return-decomposition methodology. Section II describes the sample. Section III discusses the results. Section IV concludes.

I. Decomposing the stock return

Campbell (1991) and Campbell and Ammer (1993) use the dividend-growth model of Campbell and Shiller (1988a) to decompose aggregate stock returns. In this paper, I use an accounting-based present-value formula derived by Vuolteenaho (2000). This formula uses ROE (earnings over book equity) instead of dividend growth as the basic cash-flow fundamental and provides a natural basis for variable selection in my empirical analysis. Whether one chooses to think about infinite-horizon cash-flow fundamentals in terms of dividend growth or ROE is a matter of taste, however.

Three main assumptions are made in order to derive the ROE-based version of the approximate present-value model. First, book equity (B), dividend (D), and market equity (M) are assumed to be strictly positive. Second, the difference between log book equity (b) and log market equity (m) and the difference between log dividend (d) and log book equity are assumed to be stationary, even though the series individually have an integrated component. Third, earnings (X), dividends, and book equity must satisfy the clean-surplus identity:

$$B_t = B_{t-1} + X_t - D_t \quad (1)$$

– book equity this year equals book equity last year plus earnings less dividends. Armed with these assumptions, Vuolteenaho (2000) derives a model for the log book-to-market ratio (denoted by θ):

$$\theta_{t-1} = k_{t-1} + \sum_{j=0}^{\infty} \rho^j r_{t+j} - \sum_{j=0}^{\infty} \rho^j (e_{t+j} - f_{t+j}), \quad (2)$$

where ROE is denoted by $e_t = \log(1 + X_t / B_{t-1})$ and the excess log stock return by $r_t = \log(1 + R_t + F_t) - f_t$. R_t is the simple excess stock return, F_t is the interest rate, f_t is log one plus the interest rate, and k is a constant plus the approximation error. Analogously to the Campbell-Shiller model, the book-to-market ratio can be (temporarily) low if future cash flows are high and/or future excess stock returns are low.

As long as some dividends are paid, the discount coefficient satisfies $\rho < 1$; the optimal value in my sample is 0.967. (The details of the derivation and the choice of ρ are reproduced in Appendix 1.) The results of this paper are insensitive to small changes in the discount coefficient: All of the main empirical results can be reproduced with ρ set to 0.95 or 1.00.

Equation (2) allows one to decompose the unexpected stock return into an expected-return component and a cash-flow component, along the lines of Campbell (1991). I take the change in expectations of (2) from $t-1$ to t and reorganize:

$$r_t - E_{t-1}r_t = \Delta E_t \sum_{j=0}^{\infty} \rho^j (e_{t+j} - f_{t+j}) - \Delta E_t \sum_{j=1}^{\infty} \rho^j r_{t+j} + \kappa_t, \quad (3)$$

where ΔE_t denotes the change in expectations from $t-1$ to t (i.e., $E_t(\cdot) - E_{t-1}(\cdot)$). Defining the two return components as cash-flow news (N_{cf}) and expected-return news (N_r) yields:

$$N_{cf,t} \equiv \Delta E_t \sum_{j=0}^{\infty} \rho^j (e_{t+j} - f_{t+j}) + \kappa_t, \quad N_{r,t} \equiv \Delta E_t \sum_{j=1}^{\infty} \rho^j r_{t+j} \quad (4)$$

Since $r_t - E_{t-1}r_t = N_{cf,t} - N_{r,t}$, the unexpected excess stock return can be high if either expected future excess returns decrease and/or expected future excess ROEs (i.e., ROE less interest rate) increase.²

The approximation error of the return-news equation is denoted by $\kappa_t \equiv \Delta E_t k_{t-1}$. The error is ascribed to either the cash-flow or expected-return term, depending on how the model is implemented. Both cases are considered below, and the results are robust to the choice.

The unexpected-return variance is decomposed into three components using equations (3) and (4):

$$\text{var}(r_t - E_{t-1}r_t) = \text{var}(N_{r,t}) + \text{var}(N_{cf,t}) - 2\text{cov}(N_{r,t}, N_{cf,t}) \quad (5)$$

This variance decomposition in equation (5) is used to assess the importance of expected-return and cash-flow news as drivers of stock returns.

II. Data

A. Basic data

The basic data come from the CRSP-COMPUSTAT intersection, 1954-1996. The Center for Research in Securities Prices (CRSP) monthly stock file contains monthly prices, shares outstanding, dividends and returns for NYSE, AMEX, and NASDAQ stocks. The COMPUSTAT annual research file contains the relevant accounting information for most publicly traded US stocks. In addition, I use rolled-over one-month Treasury-bill return from Ibbotson Associates as the risk-free rate. All variables are of annual frequency.

B. Data requirements

In order to be included in my sample, a firm-year must satisfy the following COMPUSTAT data requirements. First, I require all firms to have a December fiscal-year end of $t-1$, in order to align accounting variables across firms. Second, a firm must have $t-1$, $t-2$, and $t-3$ book equity available, where t denotes time in years. Third, I require $t-1$ and $t-2$ net income and long-term debt data.

A number of CRSP data requirements must also be satisfied. A valid market-equity figure must be available for $t-1$, $t-2$, and $t-3$. I require that there is a valid trade during the month immediately preceding the period t return. This requirement ensures that the return predictability is not spuriously induced by stale prices or other similar market micro-structure issues. I also require at least one monthly return observation during each of the preceding five years, from $t-1$ to $t-5$. In addition, I screen out clear data errors and mismatches by excluding firms with $t-1$ market equity less than \$10 million and book-to-market more than

100 or less than 1/100. I carefully avoid imposing any COMPUSTAT or CRSP requirements on year t data, because these data are used in the dependent variables of my regressions.

C. Variable definitions

The stock returns are calculated as follows. The simple stock return is an annual value-weight return on a firm's common stock issues (typically one). If no return data are available, I substitute zeros for both returns and dividends. Annual returns are compounded from monthly returns, recorded from the beginning of June to the end of May. Delisting returns are included when available in CRSP. If a firm is delisted but the delisting return is missing, I investigate the reason for disappearance. If the delisting is performance-related, I assume a -30% delisting return. Otherwise, I assume a zero delisting return.³

Market equity (combined value of all common stock classes outstanding) is taken from CRSP as of the end of May. If the year t market equity is missing, I compound the lagged market equity with return without dividends.

For book equity, I prefer COMPUSTAT data item 60, but if it is unavailable I use item 235. Also, if short- and/or long-term deferred taxes are available (data items 35 and 71), I add them to book equity. If both data items 60 and 235 are unavailable, I proxy book equity by last periods book equity plus earnings less dividends. If neither earnings nor book equity is available, I assume that the book-to-market ratio has not changed and compute the book equity proxy from last periods book-to-market and this periods market equity. I treat negative or zero book equity values as missing.

GAAP ROE is the earnings over last period's book equity, measured according to the US Generally Accepted Accounting Principles. I use the COMPUSTAT data item 172, earnings available for common. When earnings are missing, I use the clean-surplus formula to proxy for earnings: the change in book equity plus dividends. In every case, I do not allow the firm to lose more than their book equity. That is, I define the net income as maximum of reported net income (or clean-surplus net income, if earnings are not reported) and negative of the beginning of the period book equity. Hence, the minimum GAAP ROE is truncated to -100%.

I calculate leverage as book equity over the sum of book equity and book debt. The book debt is the sum of debt in current liabilities (34), total long-term debt (9), and preferred stock (130).

D. Transformations

The log transformations may cause problems if some stock returns and/or ROEs are close to -100% or if some of the book-to-market ratios are close to zero or infinity. I solve this complication by redefining the firm as a portfolio of 90% common stock and 10% Treasury-bills using market values. Every period, the portfolio is rebalanced to these weights. This affects not only stock return and accounting return on equity, but also the book-to-market equity, pulling this ratio slightly towards one. After adding this risk-free investment, the ratios and returns are sufficiently well behaved for log transformations. Simple market and accounting returns on this portfolio closely approximate simple returns on the firm's common stock only. The accounting identities hold for the transformed quantities. Furthermore, this transformation method is superior to purely statistical transformations (such as the Box-Cox transformation), because the transformed quantities still correspond to an investment strategy. The results are robust to moderate perturbations (± 0.025) of the T-bill weight.

Each year, market equity is log transformed, cross-sectionally demeaned, and divided by cross-sectional standard deviation to form my size variable. I also log transform leverage. Because these variables are merely indicators whose scale does not matter, I do not make further leverage adjustments.

E. Descriptive statistics

Table I shows descriptive statistics for the variables. Panel A reports means, standard deviations, and percentiles for excess log returns and corresponding variables. Panel B shows the standard deviations and percentiles for market-adjusted variables. I estimate the descriptive statistics from pooled data. Thus, the dispersion metrics capture both cross-sectional and time-series variation in the variables. A notable feature of the descriptive statistics is that firm-level excess ROE is almost as variable as firm-level excess returns (standard deviation of 0.27 vs. 0.32). Market adjusting the data reduces the return standard deviation considerably but has little effect on the ROE standard deviation.

III. Results and discussion

A. Firm-level variance decomposition of market-adjusted returns

A vector autoregressive model provides a convenient way to implement the return and return variance decompositions. Let $z_{i,t}$ be a vector of firm-specific state variables describing a firm i at time t . In

particular, let the first element of $z_{i,t}$ be the firm's stock return, defined as market-adjusted log return. An individual firm's state vector is assumed to follow a linear law:

$$z_{i,t} = \Gamma z_{i,t-1} + u_{i,t} \quad (6)$$

The VAR coefficient matrix Γ is assumed to be constant, both over time and across firms. The error term $u_{i,t}$ is assumed to have a covariance matrix Σ and to be independent of everything known at $t-1$. At this stage, I make no assumptions on how the errors are correlated across firms. The model is homogeneous over firms – a firm is expected to behave similarly to others with the same values of the state variables. Because the error terms are not necessarily perfectly correlated across firms, however, two firms that are equal today do not have to be equal tomorrow.

The VAR implies a return decomposition. Define $e1' \equiv [1 \ 0 \ \dots \ 0]$ and

$$\lambda' \equiv e1' \rho \Gamma (I - \rho \Gamma)^{-1} \quad (7)$$

The definition (7) introduced by Campbell (1991) simplifies the expressions considerably: Expected-return news can then be conveniently expressed as $\lambda' u_{i,t}$ and cash-flow news as $(e1' + \lambda') u_{i,t}$. If returns are unpredictable, i.e., the first row of Γ is zeros, expected-return news is identically zero and the entire return is due to cash-flow news. Effectively, this method first computes expected-return news directly and then backs out the cash-flow news as unexpected return plus expected-return news. It may seem that this indirect method of calculating cash-flow news as a residual relies on heavier assumptions than directly calculating the change in the discounted sum of clean-surplus ROEs does. Section E.1 uses ROE directly and shows that the results are robust to this choice.

For the variance decomposition of unexpected returns, the innovation covariance matrix Σ is required, in addition to the Γ matrix. Equation (8) shows the formulas for the elements of the news covariance matrix:

$$\begin{aligned} \text{var}(N_r) &= \lambda' \Sigma \lambda \\ \text{var}(N_{cf}) &= (e1' + \lambda') \Sigma (e1 + \lambda) \\ \text{cov}(N_r, N_{cf}) &= \lambda' \Sigma (e1 + \lambda) \end{aligned} \quad (8)$$

In estimating the VAR coefficient matrix, I trade off efficiency for robustness and simplicity. I estimate the VAR from the panel using the weighted least-squares (WLS) approach and one pooled prediction regression per state variable. Instead of using the optimal but unknown GLS weights or unit OLS weights, I weigh each cross-section equally, much like the Fama-MacBeth (1973) procedure does. In

practice this means deflating the data for each firm-year by the number of firms in the corresponding cross-section. The OLS and WLS point estimates are similar, and the main findings of the paper are not sensitive to the choice between OLS and WLS. I use two methods to calculate cross-correlation consistent standard errors. The standard-error estimates obtained using Rogers' (1983, 1993) robust standard-error method agree closely with ones obtained using Shao and Rao's (1993) jackknife method. (Appendix 2 contains the formulas.)

First, I consider a parsimonious VAR specification. This specification uses market-adjusted log stock return, log book-to-market, and log GAAP ROE as the state variables. Only one lag of each is used to predict the state vector evolution. I dub this specification "the short VAR," because it uses only three predictive coefficients per equation. Parameter estimates (presented in Table II) imply that expected returns are high when past one-year return, the book-to-market ratio and profitability are high. Expected profitability is high when past stock return and past profitability are high and the book-to-market ratio low. The expected future book-to-market ratio is mostly affected by the past book-to-market ratio. As expected, unexpected profitability and stock return covary positively (approximately 30% correlation).

Figure 1 illustrates how the short VAR captures the main stylized return-predictability facts in the previous literature. The top graph of Figure 1 shows the cumulative response of returns to a 25% return shock. Initially, the price continues rise for one year, then flat-lines for two years, and then slowly decays for decades. In the end, 23 percentage points are permanent and 2 percentage points temporary. These price-momentum and price-reversal patterns are roughly consistent with the results of DeBondt and Thaler (1985) and Jegadeesh and Titman (1993, 1999).

The bottom graph of Figure 1 shows the cumulative response of returns to a 25% cash-flow shock. If expected returns were constant, this shock would result in exactly 25% realized return. Instead, the initial response is only 20%. According to one interpretation, a cash-flow shock typically coincides with a temporary increase in risk, and hence in equilibrium expected returns. Another interpretation is that it takes the market two years to fully incorporate the cash-flow shock into prices. Irrespective of the interpretation, these patterns are consistent with the results summarized in Bernard and Thomas (1989, 1990).

The variance decomposition implied by the short VAR is shown in Panel A of Table III, and cash-flow news is the main driver of firm-level stock returns. The expected-return-news standard deviation is 13% (variance 0.0161 with 0.0069 standard error) and the cash-flow-news standard deviation 28% (variance

0.0801 with 0.0130 standard error). The ratio of expected-return-news variance to total-unexpected-return variance is approximately 0.25 (with 0.10 standard error). The correlation between the two news series is 0.41 and 3.4 standard errors from zero.

I also confirm these basic results by estimating a richer, “long” VAR. The predictive variables include four lags of past stock returns, the book-to-market ratio, two lags of profitability, two lags of leverage, and one lag of size. I motivate this lag order as a second order cointegrating VAR, with an additional lag of returns to capture possible longer horizon return autocorrelation that is not captured by the book-to-market ratio. The long VAR variance-decomposition results are shown in Panel B of Table III. The results are similar to the short VAR results. The estimated expected-return variance is slightly larger and the correlation between the two news terms is slightly higher. Due to the additional free parameters, the standard errors are somewhat larger, however. Many other elaborate specifications give similar results – the qualitative results are not sensitive to the lag order.

These results suggest that the firm-level returns are mainly driven by cash-flow news, but this does not necessarily imply that the expected-return variation is unimportant for firm-level stock prices. For example, there may be stable and significant low-frequency variation in expected returns that has little effect on one-period unexpected returns but causes large, near-permanent swings in prices.

To examine price-level effects, I define “atypical discount,” p_r , as

$$p_{r,t-1} \equiv \sum_{j=0}^{\infty} \rho^j (E_{t-1} r_{t+j} - \bar{r}) \quad (9)$$

Above, $\bar{r} \equiv E(r)$ denotes the normal or typical log risk premium or discount rate.⁴ As seen from equation (9), atypical discount is the cumulative effect of return predictability on prices. The log market value of a firm equals with the firm’s log book equity plus the future expected log ROEs discounted at typical discount rates less the firm’s atypical discount:

$$m_{t-1} \approx b_{t-1} + \sum_{j=0}^{\infty} [\rho^j E_{t-1}(e_{t+j} - \bar{r})] - p_{r,t-1} \quad (10)$$

Substituting the definition (9) to equation (10) recovers equation (2). (Cochrane [1991,1992] discusses similar decompositions in the context of the aggregate stock market.)

If a stock’s atypical discount is high, the expected future returns on this stock are also high and the stock price low (relative to the expected cash flows). At any point of time, stocks with large positive

atypical discount could be classified as “value” stocks, and stocks with large negative atypical discount as “glamour” stocks.

One can use the homogenous VAR to assess the variability of market-adjusted atypical discounts, $\tilde{p}_{r,t} \equiv \sum_{j=0}^{\infty} \rho^j E_{t-1} \tilde{r}_{t+j}$. Within the VAR, market-adjusted atypical discount is calculated as $\rho^{-1} \lambda' (z_t - \bar{z})$ and the variance of atypical discounts as $\rho^{-2} \lambda' \text{cov}(z) \lambda$, where \bar{z} denotes the mean state vector. The variance estimate computed using the VAR parameter estimates is 0.0673 with 0.0278 standard error. I also approximate the distribution of $\tilde{p}_{r,t}$ with the histogram of the fitted values of $\tilde{p}_{r,t}$ computed from the VAR, shown as Figure 2. As seen from Figure 2, it is slightly more likely that a stock’s atypical discount is extremely low (i.e., expected returns are extremely low) than extremely high (i.e., expected returns are extremely high), where extreme is taken to mean $\exp(\tilde{p}_{r,t})$ more than $\frac{1}{2}$ the market price or twice the market price. Atypical discounts outside this interval are rare.

I interpret these price-level results showing that, while cash-flow news is more important than expected-return news, expected-return variation still has an economically significant impact on firm-level stock prices.

B. Variance decomposition of market-adjusted returns as a function of firm size

The covariance matrix of news terms may vary as a function of the firm’s characteristics. One way to allow for such variation is to let the VAR-error covariance matrix Σ vary across firms while assuming that VAR coefficient matrix Γ is common for all stocks. In Table IV, I sort stocks based on firm size (i.e., market value of equity). Then, I estimate separate variance decompositions for stocks in each cell by assuming that the stocks in a particular cell share the same VAR-error covariance matrix Σ and all stocks share the same VAR transition matrix Γ . (Allowing the transition matrices to differ across groups would induce additional complications, because the infinite-sum formulas used in the variance decomposition would have to be modified to account for probability of migration from one size-group to another.)⁵

Contrasting the large-stock results to the small-stock results provides an interesting comparison. As seen from Figure 3, there are clear, (near) monotonic patterns in the news variances and correlation. The news variances, the ratio of expected-return-news to cash-flow-news variance, and the correlation between the news terms appear to be decreasing functions of size.

For the largest stocks (in decile ten), the variance of cash-flow news is eight times that of expected-return news and the estimated correlation between the news terms is statistically insignificant and economically small. Almost all of the large-stock market-adjusted-return variance is due to cash-flow news.

For the smallest stocks (in decile one), cash-flow fundamentals are highly variable with standard deviation of 41%. Expected-return-news standard deviation of 20% is also much higher than that of large stocks'. For small firms, the expected-return and cash-flow news are strongly correlated (correlation .58 with .09 standard error), while this correlation is approximately zero for large stocks. Since the covariance term is so large and positive, the estimated small-stock unexpected-return variance is, in fact, slightly lower than the estimated cash-flow-news variance.

C. Return response to cash-flow news

In this section, I further interpret my variance-decomposition results via the regression coefficient of stock return on cash-flow news. This regression coefficient shows how much the price moves on average if there is \$1 cash-flow news.

Regressing small-stock returns on cash-flow news may give one some insights into why the expected-return-news variance is relatively more important for small stocks than for large stocks. Table IV and Figure 3 examine this regression coefficient as a function of the firm size. For a typical large stock in decile ten, the estimated regression coefficient of return on cash-flow news is close to one (0.99 with 0.07 standard error). For a typical small stock in decile one, this regression coefficient is significantly below one (0.72 with 0.07 standard error). Furthermore, the point estimates are (nearly) monotonically related to the size-decile assignment.

A behavioral-finance interpretation of this result is that the market underreacts to small-stock cash-flow news. Following DeBondt and Thaler (1985), I assume that the fair “subjective” expected return is the same for all stocks (i.e., the market participants try to price the stocks to yield equal expected returns.) I take true cash-flow news as exogenous and assume that everything else that drives prices (w_t) is independent noise. Then, the estimated expected-return-news series reflects biases in the market’s pricing process. Under these (admittedly restrictive) assumptions, define overreaction as $b > 1$ in the following regression:

$$\tilde{r}_t = a + b \tilde{N}_{cf,t} + w_t \quad (11)$$

Conversely, underreaction case corresponds to $b < 1$ and the “correct” reaction case to $b = 1$. According to this interpretation, when there is news about the cash flows of a small stock, the price does not move “enough.” Therefore, expected returns must increase. This overreaction hypothesis ($b > 1$) can be restated in terms of the correlation between the expected-return and cash-flow news. Noting that the one-period expected return is by definition independent of both the cash-flow and expected-return news, the stock market overreacts if expected-return news and cash-flow news are negatively correlated. Conversely, the market underreacts to cash-flow news ($b < 1$) if the two news series are positively correlated.⁶

Another interpretation of my empirical estimates is that, for small stocks, news about higher-than-expected profitability coincides with a temporary increase in risk. In an efficient market, uncertainty about the risks of projects can create an underreaction-like pattern in cash-flow and expected-return news.⁷ Assume for simplicity that the product market is competitive and every new investment project’s net present value (NPV) is zero. Consider a firm that announces it has started a new investment project. Because all projects are zero NPV, the announcement does not affect the stock price or cause an unexpected stock return. If the firm unexpectedly announces a high-risk project, expected future returns on the firm’s stock increase. For this high-risk project to have zero NPV, it also must have high level of cash-flows. Hence, small firms taking zero-NPV projects with varying levels of risk can generate the positive news-correlation pattern observed in the data. This story can also offer an explanation for the size pattern in the firm-level news correlations, if a single project is a larger fraction of total assets for small firms than for large firms. Furthermore, allowing for some positive NPV projects and assuming that recently started projects are riskier than mature projects may allow one to match the price-momentum patterns in the data. (For a related formal model, see Berk, Green, and Naik [1999].)

D. Implied variance decomposition of the equal-weight index portfolio

Past research (e.g., Campbell and Ammer [1993]) suggests that aggregate excess stock returns are predominantly driven by expected-return news. To reconcile my results with these, I calculate the implied excess-return-variance decomposition of an equal-weight index.

First, I estimate a VAR for firm-level excess returns (a firm’s stock return less the return on Treasury bills). I include both firm-level and market-wide variables in the state vector. My objective in including the aggregate variables (cross-sectional medians of the firm-level excess log return, log book-to-market, and

excess log profitability, contained in vector x_t) is to allow common, market-wide variables to affect expected returns and cash flows on all stocks. In this case, the model is written as

$$\begin{bmatrix} z_{i,t} \\ x_t \end{bmatrix} = A + \Gamma \begin{bmatrix} z_{i,t-1} \\ x_{t-1} \end{bmatrix} + u_{i,t}. \quad (12)$$

$z_{i,t}$ is the vector of firm specific variables, and excess log stock return is the first element of this vector. The lower left corner of Γ is constrained to zero (i.e., there is no feedback from firm-level state variables to market-wide state variables.) Because the variables do not necessarily have zero means, an intercept vector A is included in the model.

Table V shows the parameter estimates of the short VAR for excess returns. The firm-specific variables in the state vector of this specification are excess log stock return, the log book-to-market, and excess log profitability. Panel A of Table VI shows the variance decomposition implied by the short VAR, and Panel B shows essentially the same results obtained from a richer specification. Similar to the market-adjusted results, the firm-level cash-flow-news variance is larger than the firm-level expected-return-news variance and the news series are positively correlated. The expected-return-news standard deviation is 22% (variance 0.0465 with 0.0311 standard error) and the cash-flow-news standard deviation is 32% (variance 0.1002 with 0.0247 standard error).

To reconcile the firm-level results with the previous aggregate results, I form the aggregate news series using fitted values of news series from the firm-level excess-return specification shown in Table V. I approximate an equal-weight portfolio's expected-return and cash-flow news with the following formulas:

$$\begin{aligned} r_t(EW) &= N_{cf,t}(EW) - N_{r,t}(EW) \\ N_{cf,t}(EW) &\approx \frac{1}{n(t)} \sum_{i=1}^{n(t)} N_{cf,i,t} = \frac{1}{n(t)} \sum_{i=1}^{n(t)} (e1' + \lambda') u_{i,t} \\ N_{r,t}(EW) &\approx \frac{1}{n(t)} \sum_{i=1}^{n(t)} N_{r,i,t} = \frac{1}{n(t)} \sum_{i=1}^{n(t)} \lambda' u_{i,t} \end{aligned} \quad (13)$$

The above formulas in equation (13) are approximations, because I use log returns. While the simple return on a portfolio is a weighed average of simple returns on individual securities, this relationship does not hold (exactly) for log returns.

Table VII shows the variance decomposition of equal-weight portfolio return calculated from these two aggregated series. Expected-return news is the main force moving the aggregate returns. The expected-return-news standard deviation is 17% (or variance 0.0296 with 0.0297 standard error), while the cash-flow-

news standard deviation 15% (variance 0.0232 with 0.0201 standard error). It appears that while cash-flow information is largely firm specific (such as success or failure of a specific product), expected-return information is predominantly driven by systematic, macroeconomic components. Hence, my firm-level results are consistent with the earlier aggregate results.

In order to better understand the market-wide component in expected-return and cash-flow news, I define a “diversification factor” as the ratio of the equal-weight-index news variance to firm-level news variance. This quantity is related to the average correlation between the news terms. The variance of equal-weight index news can be expressed as a function of the elements of the cross-sectional covariance matrix:

$$\text{var } N_r(EW) = \frac{1}{n^2} \sum_{i=1}^n \text{var}(N_{r,i}) + \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1, j \neq i}^n \text{cov}(N_{r,i}, N_{r,j}) = \frac{1}{n} \overline{\text{var}} + \frac{n-1}{n} \overline{\text{cov}} \quad (14)$$

Denoting average variance and covariance by $\overline{\text{var}}$ and $\overline{\text{cov}}$. Defining the average correlation as $\overline{\text{corr}} \equiv \overline{\text{cov}} / \overline{\text{var}}$ yields an expression for the diversification factor:

$$\frac{\text{var } N_r(EW)}{\overline{\text{var}}} = \frac{1}{n} + \left(\frac{n-1}{n} \right) \overline{\text{corr}} \quad (15)$$

From (15) one can see that for relatively large n , the diversification factor equals the average correlation. The diversification factor is thus a good measure of commonality between individual firms’ news terms.

Table VII shows the diversification factors for expected-return and cash-flow news. The estimated diversification factor for expected-return news is approximately 0.64 (with 0.26 standard error). That is, the index’s expected-return news variance is 0.64 times that of a typical stock’s. Cash-flow news is more idiosyncratic, as indicated by a diversification factor of only 0.23 (with 0.15 standard error).

While there are a number of reasons to expect the cash-flow news to diversify to a great extent, note that there is also an argument that expected-return news should diversify, although to a more limited extent. Following Ferson and Harvey (1991), one can assume that the firm-level expected returns are described by a multi-factor asset-pricing model, i.e., expected return on an asset equals the asset’s factor loadings times the factor premia. Factor loadings of firms may move around idiosyncratically, contributing to the idiosyncratic firm-level expected-return-news variance. In contrast, the factor loadings of portfolios are relatively stable, resulting in low expected-return-news variance.⁸

E. Additional robustness checks

E.1 Approximation error

I investigate the size of the cumulative approximation error κ_t in equation (3) with an additional VAR specification. I add a fourth variable, market-adjusted clean-surplus ROE (\tilde{e}^{CS}) to the state vector. Addition of this fourth variable enables me to calculate the cumulative approximation error. Table VIII reports the covariance matrix of expected-return news, cash-flow news computed using the indirect method (i.e., computing cash-flow news as a residual and thus including the approximation error in the cash-flow-news term), cash-flow news computed using a direct method (i.e., directly calculating the change in the discounted sum of clean-surplus ROEs), and the approximation error (κ_t).

Three observations are apparent from Table VIII. First, the approximation error is positively correlated with expected-return news and negatively correlated with both indirect and direct cash-flow news. Second, direct computation of the cash-flow news results in a larger cash-flow-news variance than indirect computation. The indirect method is thus the more conservative choice relative to my finding that the cash-flow-news variance dominates the firm-level returns. Third, the magnitude of the approximation error is so low that the choice between indirect and direct methods is inconsequential to the results. The standard deviation of κ is 3% (variance 0.0009 with 0.0003 standard error).

E.2 Potential selection biases

Data must satisfy a number of requirements to be included in the sample. Therefore, it is important to clarify whether the VAR estimates are subject to a selection bias. On the one hand, one can assume that the population data are generated by the VAR in equation (6) and that the probability of the data $[z_{i,t}, z_{i,t-1}]$ being included in the sample depends on $z_{i,t-1}$. Under these assumptions, the VAR parameters are estimated consistently, as long as I do not impose any data requirements on $z_{i,t}$, the dependent variables of my regressions. On the other hand, one can assume that the observed data are generated by the VAR in equation (6) and the unobserved data by a different model. In this case, the variance-decomposition results apply to the return on a managed portfolio that is tilted towards mature listed firms. This managed portfolio invests into a single firm until the firm disappears from the sample, after which the money is invested into another firm that both satisfies the data requirements and has the same values of the state variables as the

disappeared firm. This interpretation is closely related to the “follow-the-money” approach of Elton, Gruber, and Blake (1996) in the context of mutual-fund studies.

E.3 Horizon issues

The measurement horizon may be an important factor to the regression coefficient of return on cash-flow news. The existing literature⁹ suggests that securities with long strings of good cash-flow news (typically three to five years) receive atypically high valuations and, hence, have lower expected returns in the future. Because the VAR completely specifies the state-variable dynamics, the VAR parameters enable me to compute the variance decomposition for lower frequencies.

I decompose the q -period discounted return into q -period cash-flow and expected return news. I take the change in expectations of equation (2) from $t-1$ to $t+q-1$ and reorganize:

$$\sum_{j=0}^{q-1} \rho^j r_{t+j} - E_{t-1} \sum_{j=0}^{q-1} \rho^j r_{t+j} = \sum_{j=0}^{q-1} \rho^j N_{cf,t+j} - \rho^q \Delta E_{t+q-1,t-1} P_{r,t+q-1}, \quad (16)$$

where $\Delta E_{t+q-1,t-1}$ denotes the change in expectations from $t-1$ to $t+q-1$. Defining the two components of q -period return as q -period cash-flow news (N_{cf}^q) and expected-return news (N_r^q):

$$N_{cf,t+q-1}^q \equiv \sum_{j=0}^{q-1} \rho^j N_{cf,t+j}, \quad N_{r,t+q-1}^q \equiv \rho^q \Delta E_{t+q-1,t-1} P_{r,t+q-1}. \quad (17)$$

From equation (17) one can see that q -period cash-flow news is a discounted sum of one-period cash-flow news. The expected-return news term is related to the change in expectations about the log price of returns q -periods in the future. Appendix 3 contains a derivation, as well as formulae for computing the news terms and variance decomposition from the VAR parameters.

Figure 4 shows the variance-decomposition results as a function of the return measurement horizon, calculated from the short VAR in Table II. The expected-return-news variance increases up to a five-year horizon, and then begins a slow decay. The cash-flow-news variance grows nearly linearly, because the news terms are serially independent and ρ is close to one. The figure also plots the regression coefficient of q -year returns on q -year cash-flow news. The regression coefficient begins from 0.81 at one-year horizon, rising to 0.95 at the four-year horizon and to the 1.00 at ten-year horizon. Hence, the point estimates suggest positive correlation between 3-to-5-year cash-flow and expected-return news.

E.4 1975-1996 subsample

Kothari, Shanken, and Sloan (1995) argue that book-to-market related return predictability is spuriously induced to the COMPUSTAT database by the process of back-filling the data for successful firms. In order to show that this back-filling bias is not driving my results, I re-estimate the model using 1975-1996 subsample (22 years, 29123 firm-years). Because back-filling is not a serious problem in COMPUSTAT data in the later period and because I require an extensive history of data, this subsample should be free of look-ahead biases. Using 1975-1996 subperiod and the short VAR specification to decompose market-adjusted returns confirms my earlier results. (Full results are available on request.)

E.5 Log returns vs. simple returns

The variation in expected log returns does not necessarily imply variation in expected simple returns. For example, if log returns are conditionally normal, the conditional expected simple return equals

$$E_{t-1}(1 + R_t) = \exp\left[E_{t-1}(\log(1 + R_t)) + \frac{1}{2} \text{var}_{t-1}(\log(1 + R_t))\right]. \quad (18)$$

Hence, expected log return may change simply because the conditional log-return variance changes.

To examine this issue, I estimate the VAR with simple instead log returns and set the discount coefficient ρ to 1.¹⁰ The results suggest that the variance decomposition is not driven by the choice between log and simple returns. The ratio of expected-return-news variance to total unexpected-return variance (0.2412 for log returns and 0.2355 for simple returns) and the news correlation (0.40 for log returns to 0.26 for simple returns) are practically unaltered.

IV. Conclusions

The present-value formula enables one to divide unexpected stock return into two components: changes in cash-flow expectations (i.e., cash-flow news) and changes in discount rates (i.e., expected-return news). As shown by Campbell and Shiller (1988a), stock return volatility must originate from volatile cash-flow and/or expected-return news. While previous research investigates aggregate portfolios, this paper measures the importance of cash-flow and expected-return news for firm-level stock returns.

The analysis yields essentially three results. First, firm-level stock returns are predominantly driven by cash-flow news. For excess log returns, the variance of expected-return news is approximately one half of the variance of cash-flow news. For market-adjusted log returns, the variance of expected-return news is one fifth of the cash-flow-news variance.

Second, I find that cash-flow news is positively correlated with shocks to expected returns for a typical stock. Good news about cash flows is typically accompanied by higher expected returns. This correlation appears to be larger for smaller stocks and about zero for the largest stocks.

Third, cash-flow news is more easily diversified away in portfolios than expected-return news. For an equal-weight portfolio the cash-flow-news variance is only three quarters of the expected-return-news variance. This finding suggests that cash-flow information is largely firm-specific and that expected-return information is predominantly driven by systematic, market-wide components.

Appendix 1: Approximate expression for the book-to-market ratio

Market and accounting returns (i.e., ROE) can be expressed as

$$r_t + f_t \equiv \log\left(\frac{M_t + D_t}{M_{t-1}}\right) = \log\left(1 + \frac{\Delta M_t + D_t}{M_{t-1}}\right) = \log(1 + R_t + F_t) \quad (\text{A1.2})$$

$$e_t \equiv \log\left(\frac{B_t + D_t}{B_{t-1}}\right) = \log\left(1 + \frac{\Delta B_t + D_t}{B_{t-1}}\right) = \log(1 + E_t) \quad (\text{A1.3})$$

Substituting the log dividend-growth rate, Δd_t , the log dividend-price ratio, δ_t , and the log dividend-to-book-equity ratio, $\gamma_t \equiv d_t - b_t$, to the return definitions (A1.2) and (A1.3):

$$r_t + f_t = \log(\exp(-\delta_t) + 1) + \Delta d_t + \delta_{t-1}. \quad (\text{A1.4})$$

$$e_t = \log(\exp(-\gamma_t) + 1) + \Delta d_t + \gamma_{t-1}. \quad (\text{A1.5})$$

The nonlinear functions (A1.4) and (A1.5) can be approximated around $\hat{\delta}$ and $\hat{\gamma}$. Specifically, use some convex combination of the unconditional means of the variables as an expansion point for both functions. Subtracting $r_t + f_t$ from e_t , the approximate expression is

$$e_t - r_t - f_t = \log(\exp(-\gamma_t) + 1) - \log(\exp(-\delta_t) + 1) + (\gamma_{t-1} - \delta_{t-1}) \approx \rho\theta_t - \theta_{t-1}. \quad (\text{A1.6})$$

Above, the log book-to-market ratio is denoted by θ_t . Note that as dividend yields drop, the approximation becomes more accurate, while ρ approaches unity.

Finally, the one period approximation is iterated forward to yield:

$$\theta_{t-1} = \sum_{j=0}^N \rho^j r_{t+j} + \sum_{j=0}^N \rho^j f_{t+j} - \sum_{j=0}^N \rho^j e_{t+j} + \sum_{j=0}^N \rho^j \xi_{t+j} + \rho^{N+1} \theta_{t+N} \quad (\text{A1.7})$$

In equation (A1.7), ξ_t denotes the approximation error in equation (A1.6). Because $\rho < 1$, the limit $N \rightarrow \infty$ of equation (A1.7) converges to equation (2) in the text. (The approximation error of equation (2) in the text is defined as $k_{t-1} \equiv \sum_{j=0}^{\infty} \xi_{t+j}$.)

Which value to pick for ρ is an empirical question. In my data, the three model variables (profitability, stock return, and the lagged book-to-market ratio) and the approximation explain 99.82% of the variation in the book-to-market ratios. The 99.82% R^2 is achieved with $\rho = .967$. Because ρ is estimated accurately and my main results are not sensitive to the ρ -choice, I use this ρ -value in the analysis and treat it as a constant. Table A1.I also shows similar regressions using US GAAP ROE instead of clean-surplus ROE (row 2). Using US GAAP ROE will reduce the R^2 , but does not materially affect ρ .

Appendix 2: Cross-correlation consistent standard errors

In many finance applications, the available data set contains perhaps twenty cross-sections, each with thousands of data points. In such cases, incorrectly assuming that the errors are cross-sectionally uncorrelated can yield standard errors that are biased downwards by a factor of five (Fama and French [2000b]). The statistics literature has proposed two solutions for a similar problem frequently arising with complex surveys: Rogers' (1983, 1993) robust standard errors and the jackknife standard errors of Shao and Rao (1993). Both methods yield asymptotically correct standard errors for the OLS and WLS estimators under a general cross-correlation structure. Compared to the popular Fama-MacBeth (1973) procedure, these two methods have the practical advantage of giving the standard errors for pooled-OLS/WLS coefficients – allowing for, among other things, the use of common time-series variables in the regressions.

Let X denote the panel of explanatory variables, Ω the covariance matrix of the panel of errors, and X_t and Ω_t a single cross-section of explanatory variables and the corresponding error covariance matrix. Assuming that the errors are independent across cross-sections allows writing

$$(X'X)^{-1} X' \Omega X (X'X)^{-1} = (X'X)^{-1} \sum_{t=1}^T [X_t' \Omega_t X_t] (X'X)^{-1}. \quad (\text{A2.1})$$

Denote regression errors by ε , and notation for fitted values is modified with a hat. Since $X_t' \Omega_t X_t = E(X_t' \varepsilon_t \varepsilon_t' X_t)$, Rogers substitutes estimated errors for true errors to get a variance estimator of regression coefficients:

$$(X'X)^{-1} \sum_{t=1}^T [X_t' \hat{\varepsilon}_t \hat{\varepsilon}_t' X_t] (X'X)^{-1} \quad (\text{A2.2})$$

Under plausible assumptions, the standard errors are consistent in T , i.e., converge to the population values as the time dimension of the panel grows. The above standard error formula can be interpreted as generalized White's standard errors; in the special case of only one observation per cross-section, the standard errors are equivalent to White (1980) heteroscedasticity consistent standard errors.

Rogers (1993) performs a small-scale Monte Carlo study of the finite sample properties. According to his findings, “[a]s long as the largest cluster is 5 percent or less of the total sample, [the] bias [in variance formula (12)] should be negligible.” Since I weigh cross-sections equally and have over forty cross-sections, small sample biases are unlikely to affect my regression results.

An alternative to Rogers' robust standard errors is the jackknife method. Shao and Rao (1993) prove that a delete-cross-section jackknife method produces consistent standard errors for a linear regression even if the errors are cross-sectionally dependent. The jackknife begins by computing the normal OLS or WLS regression coefficient estimate from the entire sample (denote this by "normal estimate"). Then, T new samples of size $T-1$ are recorded omitting one cross-section at a time. Next, the regression coefficient estimates are computed from each one of the new samples (denote these estimates "delete-cross-section estimates"). Then, $T-1$ pseudo-values are calculated as T times the normal estimate less $T-1$ times the delete-cross-section estimate. Finally, the variance of pseudo-values is the estimated variance of the regression coefficient estimator. For details, see Shao and Rao (1993).

Appendix 3: q-period return decomposition

I decompose the q -period discounted return into q -period cash-flow and expected-return news. I take the change in expectations of equation (2) from $t-1$ to $t+q-1$ and ignore the approximation error:

$$0 = (E_{t+q-1} - E_{t-1}) \sum_{j=0}^{\infty} \rho^j (e_{t+j} - f_{t+j}) - (E_{t+q-1} - E_{t-1}) \sum_{j=0}^{\infty} \rho^j r_{t+j} \Leftrightarrow \quad (\text{A3.1})$$

$$\begin{aligned} \sum_{j=0}^{q-1} \rho^j r_{t+j} - E_{t-1} \sum_{j=0}^{q-1} \rho^j r_{t+j} &= (E_{t+q-1} - E_{t+q-2} + E_{t+q-2} \dots - E_{t-1}) \sum_{j=0}^{\infty} \rho^j (e_{t+j} - f_{t+j}) \\ &\quad - (E_{t+q-1} - E_{t-1}) \sum_{j=q}^{\infty} \rho^j r_{t+j} \end{aligned} \quad (\text{A3.2})$$

Substituting the definitions of one-period cash-flow news (4) and atypical discount (9):

$$\sum_{j=0}^{q-1} \rho^j r_{t+j} - E_{t-1} \sum_{j=0}^{q-1} \rho^j r_{t+j} = \sum_{j=0}^{q-1} \rho^j N_{cf,t+j} - \rho^q \Delta E_{t+q-1,t-1} p_{r,t+q-1}, \quad (\text{A3.3})$$

where $\Delta E_{t+q-1,t-1}$ denotes the change in expectations from $t-1$ to $t+q-1$.

The VAR can be used to calculate the news terms and variance decompositions. Equations (A3.4) and (A3.5) state the relevant expressions:

$$N_{cf,t+q}^q \equiv \sum_{j=1}^q \rho^{j-1} (e1' + \lambda') u_{t+j-1}, \quad N_{r,t+q}^q \equiv \rho^{q-1} \sum_{j=1}^q \lambda' \Gamma^{q-j} u_{t+j-1} \quad (\text{A3.4})$$

$$\begin{aligned}
\text{var } N_{cf}^q &\equiv \sum_{j=1}^q \rho^{j-1} (e1' + \lambda') \Sigma [\rho^{j-1} (e1' + \lambda')]', \\
\text{var } N_r^q &\equiv \rho^{2(q-1)} \sum_{j=1}^q \lambda' \Gamma^{q-j} \Sigma [\lambda' \Gamma^{q-j}]' \\
\text{cov} [N_{cf}^q, N_r^q] &= \rho^{q-1} \sum_{j=1}^q \rho^{j-1} (e1' + \lambda') \Sigma [\lambda' \Gamma^{q-j}]'
\end{aligned} \tag{A3.5}$$

References

- Banz, R. W., 1981, The relationship between return and market value of common stocks, *Journal of Financial Economics* 9, 3-18.
- Berk, J. B., R. C. Green, and V. Naik, 1999, Optimal investment, growth options, and security returns, *Journal of Finance* 54, 1553-1607.
- Bernard, V. L., and J. K. Thomas, 1989, Post-earnings-announcement drift: Delayed price response or risk premium?, *Journal of Accounting Research* 27, 1-36.
- Bernard, V. L., and J. K. Thomas, 1990, Evidence that stock prices do not fully reflect the implications of current earnings for future earnings, *Journal of Accounting and Economics* 13, 305-340.
- Bhandari, L. C., 1988, Debt/equity ratio and expected common stock returns: empirical evidence, *Journal of Finance* 43, 507-528.
- Campbell, J. Y., 1991, A variance decomposition for stock returns, *Economic Journal* 101, 57-179.
- Campbell, J. Y., and J. Ammer, 1993, What moves the stock and bond markets? A variance decomposition for long-term asset returns, *Journal of Finance* 48, 3-37.
- Campbell, J. Y., and R. J. Shiller, 1988a, The dividend-price ratio and expectations of future dividends and discount factors, *Review of Financial Studies* 1, 195-228.
- Campbell, J. Y., and R. J. Shiller, 1988b, Stock prices, earnings, and expected dividends, *Journal of Finance* 43, 661-676.
- Cochrane, J. H., 1991, Volatility tests and efficient markets: review essay, *Journal of Monetary Economics* 27, 463-85.
- Cochrane, J. H., 1992, Explaining the variance of price-dividend ratios, *Review of Financial Studies* 5, 243-280.
- Daniel, K., D. Hirshleifer, and A. Subrahmanyam, 1998, Investor psychology and securities market under- and overreactions, *Journal of Finance* 53, 1839-1885.
- DeBondt, W. F. M. and R. H. Thaler, 1985. Does the stock market overreact?, *Journal of Finance* 40, 557-581.
- Elton, E. J., M. J. Gruber, and C. R. Blake, 1996. Survivorship bias and mutual fund performance, *Review of Financial Studies* 9, 1097-1120.

- Fama, E. F. and K. R. French, 2000a, Disappearing dividends: Changing firm characteristics or increased reluctance to pay?, Working paper, University of Chicago Graduate School of Business.
- Fama, E. F. and K. R. French, 2000b, Forecasting profitability and earnings, *Journal of Business* 73, 161-75.
- Fama, E. F. and J. MacBeth, 1973, Risk, return and equilibrium: Empirical tests, *Journal of Political Economy* 81, 607-636.
- Ferson, W. E., and C. R. Harvey, 1991, The variation of economic risk premiums, *Journal of Political Economy* 99, 385-415.
- Haugen, R. A., and N. L. Baker, 1996, Commonality in the determinants of expected stock returns, *Journal of Financial Economics* 41, 401-439.
- Jegadeesh, N. and S. Titman, 1993, Returns to buying winners and selling losers: implications for stock market efficiency, *Journal of Finance* 48, 65-91.
- Jegadeesh, N. and S. Titman, 1999, Profitability of momentum strategies: An evaluation of alternative explanations, *Journal of Finance*, forthcoming.
- Jolls, C., 1998, Stock repurchases and incentive compensation, NBER Working Paper 6467.
- Kothari, S. P., J. Shanken, J., and R. G. Sloan, 1995, Another look at the cross-section of expected stock returns, *Journal of Finance* 50, 185-224.
- Lakonishok, J., A. Shleifer, and R. W. Vishny, 1994, Contrarian investment, extrapolation, and risk, *Journal of Finance* 49, 1541-1578.
- Rogers, W. H., 1983, Analyzing complex survey data, Rand Corporation memorandum, Santa Monica, CA.
- Rogers, W. H., 1993, Regression standard errors in clustered samples, *Stata Technical Bulletin Reprints* STB-13 – STB-18, 88-94.
- Rosenberg, B., and K. Reid, and R. Lanstein, 1985, Persuasive evidence of market inefficiency, *Journal of Portfolio Management* 11, 9-17.
- Shao, J. and J. N. K. Rao, 1993, Jackknife inference for heteroscedastic linear regression models, *Canadian Journal of Statistics* 21, 377-385.
- Shumway, T. G., 1997, The delisting bias in CRSP data, *Journal of Finance* 52, 327-340.
- Vuolteenaho, T., 2000, Understanding the aggregate book-to-market ratio and its implications to current equity-premium expectations. Working paper, Harvard University Department of Economics.

White, H., 1980, A heteroscedasticity-consistent covariance matrix estimator and a direct test of heteroscedasticity, *Econometrica* 48, 817-838.

Table I: Descriptive statistics

Panel A reports means, standard deviations, and percentiles (minimum, 25%, 50%, 75%, and maximum) of excess log return, r , log US GAAP return on equity in excess of risk-free rate, $e^{GAAP} - f$, log leverage, lev , and log book-to-market, θ .

Panel B shows these descriptive statistics for market-adjusted log return, \tilde{r} , market-adjusted log US GAAP return on equity, \tilde{e}^{GAAP} , market-adjusted log leverage, \tilde{lev} , and market-adjusted log book-to-market, $\tilde{\theta}$. The variables are market-adjusted by subtracting the cross-sectional average each year.

The descriptive statistics are estimated from the pooled 1954-1996 CRSP-COMPUSTAT intersection, consisting of 36791 firm-years.

Panel A: Descriptive statistics, basic data

| Variable | Mean | St. Dev. | Min | 25%-pct | Median | 75%-pct | Max |
|----------------|---------|----------|---------|---------|---------|---------|--------|
| r | 0.0327 | 0.3223 | -2.1497 | -0.1283 | 0.0419 | 0.2157 | 2.0014 |
| $e^{GAAP} - f$ | 0.0075 | 0.2664 | -2.5755 | -0.0027 | 0.0428 | 0.0083 | 3.8693 |
| lev | -0.5794 | 0.4987 | -5.5215 | -0.7831 | -0.4684 | -0.2395 | 0.0000 |
| θ | -0.2069 | 0.6131 | -3.7796 | -0.5783 | -0.1673 | 0.2016 | 3.9827 |

Panel B: Descriptive statistics, market-adjusted data

| Variable | Mean | St. Dev. | Min | 25%-pct | Median | 75%-pct | Max |
|--------------------|------|----------|---------|---------|--------|---------|--------|
| \tilde{r} | 0 | 0.2867 | -2.1718 | -0.1466 | 0.0063 | 0.1602 | 1.9011 |
| \tilde{e}^{GAAP} | 0 | 0.2640 | -2.5210 | -0.0186 | 0.0280 | 0.0793 | 3.8861 |
| \tilde{lev} | 0 | 0.4944 | -4.8967 | -0.1998 | 0.1034 | 0.3296 | 0.6453 |
| $\tilde{\theta}$ | 0 | 0.5606 | -3.3295 | -0.1034 | 0.0448 | 0.3491 | 3.8282 |

Table II: Short VAR for market-adjusted returns

The table reports the parameter estimates for the short VAR. The model variables include market-adjusted log stock return, \tilde{r}_t , (the first element of the state vector z), the market-adjusted log book-to-market ratio, $\tilde{\theta}_t$, (the second element), and market-adjusted log profitability, $\tilde{\epsilon}_t$, (the third element). The parameters in the table correspond to the following system:

$$z_{i,t} = \Gamma z_{i,t-1} + u_{i,t}, \quad \Sigma = E(u_{i,t}u'_{i,t})$$

For each parameter, I report three numbers. The first number (bold) is a weighted least squares estimate of the parameter, where observations are weighted such that each cross-section receives an equal weight. The second number (in parentheses) is a robust standard error computed using the Rogers' (1983, 1993) method. The third number (in brackets) is a robust jackknife standard error computed using a jackknife method outlined by Shao and Rao (1993).

I use the CRSP-COMPUSTAT intersection 1954-1996 as the sample, in total 36791 firm-years.

| Coefficient estimates for the first order market-adjusted VAR | | | | (estimate), (s.e.), [j.s.e.] | | |
|---|---------------------------------------|--|---------------------------------------|--|--|---------------------------------------|
| | Γ | | | Σ | | |
| \tilde{r}_t | 0.1182 (0.0224) [0.0229] | 0.0477 (0.0131) [0.0136] | 0.1464 (0.0308) [0.0315] | 0.0668 (0.0040) [0.0040] | -0.0544 (0.0027) [0.0028] | 0.0130 (0.0019) [0.0020] |
| $\tilde{\theta}_t$ | 0.0554 (0.0327) [0.0336] | 0.8953 (0.0174) [0.0180] | 0.0570 (0.0494) [0.0512] | -0.0544 (0.0027) [0.0028] | 0.0967 (0.0151) [0.0152] | 0.0114 (0.0021) [0.0021] |
| $\tilde{\epsilon}_t$ | 0.1042 (0.0110) [0.0112] | -0.0264 (0.0061) [0.0063] | 0.4939 (0.0595) [0.0620] | 0.0130 (0.0019) [0.0020] | 0.0114 (0.0021) [0.0021] | 0.0344 (0.0053) [0.0052] |

Table III: Variance decomposition of market-adjusted returns

The table reports a variance decomposition of market-adjusted returns and other derived statistics calculated from the two VAR specifications. Both VAR specifications have the structure

$$z_{i,t} = \Gamma z_{i,t-1} + u_{i,t}, \quad \Sigma = E(u_{i,t}u'_{i,t})$$

The short VAR (Panel A) state vector includes market-adjusted log stock return, \tilde{r} , the market-adjusted log book-to-market ratio, $\tilde{\theta}$, and market-adjusted log profitability, \tilde{e} .

The long VAR in Panel B includes four lags of market-adjusted log stock return, \tilde{r} , the market-adjusted log book-to-market ratio, $\tilde{\theta}$, two lags of market-adjusted log profitability, \tilde{e} , two lags of market-adjusted leverage, $\tilde{l}e v$, and the size variable, $s\tilde{z}e$.

The first number (bold) is a point estimate computed using the weighted least-squares estimates of the parameters. The second number (in brackets) is a robust jackknife standard error computed using a jackknife method outlined by Shao and Rao (1993).

I use the CRSP-COMPUSTAT intersection 1954-1996 as the sample, in total 36791 firm-years.

Panel A: Distribution of the expected-return and cash-flow news from the short VAR (estimate), [j.s.e.]

| <i>Cov. matrix</i> | Nr | Ncf | | |
|---------------------------|---------------------------|---------------------------|---|---------------------------|
| Expected-return news (Nr) | 0.0161 [0.0069] | 0.0147 [0.0074] | Correlation between expected-return and cash flow news: | 0.4092 [0.1200] |
| Cash-flow news (Ncf) | 0.0147 [0.0074] | 0.0801 [0.0130] | Ratio of expected-return-news variance to total unexpected-return variance: | 0.2412 [0.0984] |

Panel B: Distribution of the expected-return and cash-flow news from the long VAR (estimate), [j.s.e.]

| <i>Cov. matrix</i> | Nr | Ncf | | |
|---------------------------|---------------------------|---------------------------|---|---------------------------|
| Expected-return news (Nr) | 0.0203 [0.0107] | 0.0195 [0.0176] | Correlation between expected-return and cash flow news: | 0.4679 [0.2392] |
| Cash-flow news (Ncf) | 0.0195 [0.0176] | 0.0856 [0.0285] | Ratio of expected-return-news variance to total unexpected-return variance: | 0.3029 [0.1501] |

Table IV: Variance decomposition as a function of firm size

The table reports a local variance decomposition of market-adjusted returns for small and large stocks. The VAR specification has the structure

$$z_{i,t} = I\tilde{z}_{i,t-1} + u_{i,t}, \quad \Sigma(\text{characteristics}) = E(u_{i,t}u'_{i,t}|\text{characteristics})$$

The state vector includes market-adjusted log stock return, \tilde{r} , the market-adjusted log book-to-market ratio, $\tilde{\theta}$, and market-adjusted log profitability, $\tilde{\epsilon}$. The characteristic affecting the error covariance matrix is firm size (i.e., market value of equity)

The table estimates a separate variance decomposition for each of the size deciles. For every cell, I report five statistics: the variance of expected-return news (var(Nr)), the variance of cash-flow news (var(Ncf)), the covariance between the news terms (N-Cov), the correlation between the news terms (N-Corr), and the regression coefficient of return on cash-flow news (b).

For each statistic, I report two numbers. The first number (bold) is a point estimate computed using the weighted least-squares estimates of the parameters. The second number (in brackets) is a robust jackknife standard error computed using a jackknife method outlined by Shao and Rao (1993).

I use the CRSP-COMPUSTAT intersection 1954-1996 as the sample, in total 36791 firm-years.

| Variance decomposition as a function of the firm size | | | | | |
|---|---------------------------|---------------------------|---------------------------|---------------------------|---------------------------|
| | Var(Nr) | Var(Ncf) | N-Cov | N-Corr | b |
| Small | 0.0410 [0.0166] | 0.1673 [0.0303] | 0.0476 [0.0191] | 0.5751 [0.0876] | 0.7154 [0.0683] |
| 2 | 0.0326 [0.0133] | 0.1394 [0.0246] | 0.0359 [0.0153] | 0.5320 [0.1039] | 0.7426 [0.0729] |
| 3 | 0.0210 [0.0092] | 0.1013 [0.0179] | 0.0197 [0.0101] | 0.4264 [0.1251] | 0.8059 [0.0724] |
| 4 | 0.0160 [0.0069] | 0.0834 [0.0140] | 0.0127 [0.0072] | 0.3477 [0.1322] | 0.8479 [0.0660] |
| 5 | 0.0149 [0.0067] | 0.0831 [0.0147] | 0.0144 [0.0082] | 0.4105 [0.1420] | 0.8263 [0.0738] |
| 6 | 0.0105 [0.0044] | 0.0636 [0.0099] | 0.0074 [0.0045] | 0.2866 [0.1389] | 0.8834 [0.0585] |
| 7 | 0.0088 [0.0044] | 0.0516 [0.0077] | 0.0048 [0.0041] | 0.2262 [0.1562] | 0.9063 [0.0712] |
| 8 | 0.0071 [0.0036] | 0.0423 [0.0064] | 0.0031 [0.0031] | 0.1777 [0.1598] | 0.9273 [0.0672] |
| 9 | 0.0053 [0.0029] | 0.0373 [0.0054] | 0.0012 [0.0026] | 0.0844 [0.1865] | 0.9682 [0.0688] |
| Big | 0.0040 [0.0022] | 0.0319 [0.0046] | 0.0003 [0.0021] | 0.0257 [0.1961] | 0.9909 [0.0681] |

Table V: Short VAR for excess returns with aggregate variables

This table reports the parameter estimates for the short VAR for excess returns. The model includes both firm specific and aggregate variables. The firm specific variables included are excess log stock return, r , (the first element of the state vector z), the log book-to-market ratio, θ , (the second element), and excess log profitability, $e^{GAAP} - f$, (the third element). The aggregate variables (vector x), are cross-sectional median excess log return, cross-sectional median log book-to-market, and cross-sectional median excess log profitability.

The parameters in the table correspond to the following system, where the left-lower block of Γ is constrained to a zero matrix:

$$\begin{bmatrix} z_{i,t} \\ x_t \end{bmatrix} = A + \Gamma \begin{bmatrix} z_{i,t-1} \\ x_{t-1} \end{bmatrix} + u_{i,t}, \quad \Sigma = E(u_{i,t}u'_{i,t})$$

For each parameter, I report three numbers. The first number (bold) is a weighed least squares estimate of the parameter. The second number (in parentheses) is a robust standard error computed using the Rogers' (1983, 1993) method. The third number (in brackets) is a robust jackknife standard error computed using a jackknife method outlined by Shao and Rao (1993).

I use the CRSP-COMPUSTAT intersection 1954-1996 as the sample, in total 36791 firm-years.

| Coefficient estimates for the first order market-adjusted VAR | | | | (estimate), (s.e.), [j.s.e.] | | | |
|---|--|---------------------------------------|--|---------------------------------------|--|--|--|
| | A | Γ | | | | | |
| r_t | -0.0527 (0.0611) [0.0724] | 0.1238 (0.0231) [0.0240] | 0.0531 (0.0140) [0.0150] | 0.1333 (0.0299) [0.0309] | -0.4311 (0.1187) [0.1381] | 0.2563 (0.1111) [0.1358] | 3.0362 (1.1974) [1.5273] |
| θ_t | -0.0068 (0.0841) [0.0945] | 0.0585 (0.0287) [0.0298] | 0.8891 (0.0185) [0.0194] | 0.0501 (0.0399) [0.0414] | 0.2240 (0.1478) [0.1694] | -0.1664 (0.1161) [0.1398] | -1.6814 (1.3478) [1.6535] |
| $e_t - f_t$ | -0.0261 (0.0057) [0.0065] | 0.1051 (0.0111) [0.0113] | -0.0252 (0.0061) [0.0063] | 0.5007 (0.0590) [0.0615] | -0.1017 (0.0214) [0.0245] | 0.0255 (0.0123) [0.0140] | 0.7328 (0.1028) [0.1183] |
| <i>median</i> r_t | 0.0093 (0.0532) [0.0636] | 0 | 0 | 0 | -0.2428 (0.1547) [0.1593] | 0.1718 (0.1030) [0.0992] | 1.3922 (1.0209) [1.2156] |
| <i>median</i> θ_t | -0.0229 (0.0572) [0.0623] | 0 | 0 | 0 | 0.2019 (0.1664) [0.1722] | 0.8554 (0.1108) [0.1014] | -0.3454 (1.0981) [1.1201] |
| <i>median</i> $e_t - f_t$ | 0.0060 (0.0038) [0.0047] | 0 | 0 | 0 | 0.0120 (0.0110) [0.0113] | -0.0070 (0.0073) [0.0093] | 0.8207 (0.0727) [0.0750] |

Coefficient estimates for the first order market-adjusted VAR (continued)

| Σ | r_t | θ_t | $e_t - f_t$ | <i>median</i> r_t | <i>median</i> θ_t | <i>median</i> $e_t - f_t$ |
|------------------------------|--|--|--|--|--|--|
| | r_t | 0.0883 (0.0074) [0.0082] | -0.0772 (0.0074) [0.0082] | 0.0130 (0.0019) [0.0020] | 0.0191 (0.0050) [0.0056] | -0.0199 (0.0056) [0.0062] |
| θ_t | -0.0772 (0.0074) [0.0082] | 0.1256 (0.0213) [0.0216] | 0.0114 (0.0022) [0.0021] | -0.0205 (0.0055) [0.0061] | 0.0234 (0.0067) [0.0073] | -0.0005 (0.0003) [0.0003] |
| $e_t - f_t$ | 0.0130 (0.0019) [0.0020] | 0.0114 (0.0022) [0.0021] | 0.0346 (0.0053) [0.0052] | 0.0000 (0.0003) [0.0002] | 0.0000 (0.0003) [0.0002] | 0.0001 (0.0000) [0.0000] |
| <i>median</i> r_t | 0.0191 (0.0050) [0.0056] | -0.0205 (0.0055) [0.0061] | 0.0000 (0.0003) [0.0002] | 0.0183 (0.0046) [0.0051] | -0.0190 (0.0051) [0.0057] | 0.0006 (0.0002) [0.0002] |
| <i>median</i> θ_t | -0.0199 (0.0056) [0.0062] | 0.0234 (0.0067) [0.0073] | 0.0000 (0.0003) [0.0002] | -0.0190 (0.0051) [0.0057] | 0.0212 (0.0059) [0.0065] | -0.0006 (0.0002) [0.0003] |
| <i>median</i> $e_t - f_t$ | 0.0006 (0.0002) [0.0003] | -0.0005 (0.0003) [0.0003] | 0.0001 (0.0000) [0.0000] | 0.0006 (0.0002) [0.0002] | -0.0006 (0.0002) [0.0003] | 0.0001 (0.0000) [0.0000] |

Table VI: Variance decomposition of excess returns

This table reports the variance decomposition of excess returns and other derived statistics calculated from two VAR specifications. Both models include both firm specific and aggregate variables. The VAR specifications have the structure shown below, with the left-lower block of Γ is constrained to a zero matrix:

$$\begin{bmatrix} z_{i,t} \\ x_t \end{bmatrix} = A + \Gamma \begin{bmatrix} z_{i,t-1} \\ x_{t-1} \end{bmatrix} + u_{i,t}, \quad \Sigma = E(u_{i,t}u'_{i,t})$$

The short VAR (used to compute the variance decomposition in Panel A) contains the following variables. Firm specific variables included are excess log stock return, r , the log book-to-market ratio, θ , and excess log profitability, $e^{GAAP} - f$. The aggregate variables (vector x), are cross-sectional median excess log return, cross-sectional median log book-to-market, and cross-sectional median excess log profitability.

The long VAR (used to compute the variance decomposition in Panel B) contains the following variables. Firm specific variables include three lags of excess log stock return, the log book-to-market ratio, two lags of excess log profitability, two lags of leverage, lev , and my size variable, $size$. The aggregate variables are cross-sectional median excess log return, cross-sectional median log book-to-market, and cross-sectional median excess log profitability.

The first number (bold) is a point estimate computed using the weighted least-squares estimates of the parameters. The second number (in brackets) is a robust jackknife standard error computed using a jackknife method outlined by Shao and Rao (1993).

I use the CRSP-COMPUSTAT intersection 1954-1996 as the sample, in total 36791 firm-years.

| Panel A: Distribution of the expected-return and cash-flow news from the short VAR | | | | (estimate), [j.s.e.] |
|---|---------------------------|---------------------------|---|-----------------------------|
| <i>Cov. matrix</i> | Nr | Ncf | | |
| Expected-return news (Nr) | 0.0465 [0.0311] | 0.0292 [0.0213] | Correlation between expected-return and cash flow news: | 0.4279 [0.1643] |
| Cash-flow news (Ncf) | 0.0292 [0.0213] | 0.1002 [0.0247] | Ratio of expected-return-news variance to total-unexpected-return variance: | 0.5264 [0.3416] |

| Panel B: Distribution of the expected-return and cash-flow news from the long VAR | | | | (estimate), [j.s.e.] |
|--|---------------------------|---------------------------|---|-----------------------------|
| <i>Cov. matrix</i> | Nr | Ncf | | |
| Expected-return news (Nr) | 0.0454 [0.0297] | 0.0292 [0.0271] | Correlation between expected-return and cash flow news: | 0.4292 [0.2239] |
| Cash-flow news (Ncf) | 0.0292 [0.0271] | 0.1021 [0.0378] | Ratio of expected-return-news variance to total-unexpected-return variance: | 0.5097 [0.3206] |

Table VII: Implied variance decomposition of the equal-weight index returns

This table reports variance decompositions for the equal-weight index returns implied by two VAR specifications used in Table VI. I first compute the news series for each firm using the particular VAR. I then form two new series by taking the cross-sectional average of expected-return and cash-flow news. The reported variance decomposition is based on the sample covariance matrix of these two series. I also report a diversification factor, i.e., the ratio of firm level news variance to aggregate news variance. The first number (bold) is the point estimate. The second number (in brackets) is a robust jackknife standard error computed using a jackknife method outlined by Shao and Rao (1993).

| Panel A : Variance decomposition for equal-weight portfolio, the short VAR | | | (estimate), [j.s.e.] | |
|---|---------------------------|---------------------------|--|---------------------------|
| <i>Cov. matrix</i> | Nr | Ncf | | |
| Expected-return news (Nr) | 0.0296 [0.0297] | 0.0154 [0.0197] | Diversification factor for expected-return news: | 0.6373 [0.2582] |
| Cash-flow news (Ncf) | 0.0154 [0.0197] | 0.0232 [0.0201] | Diversification factor for cash-flow news: | 0.2309 [0.1475] |

| Panel B : Variance decomposition for equal-weight portfolio, the long VAR | | | (estimate), [j.s.e.] | |
|--|---------------------------|---------------------------|--|---------------------------|
| <i>Cov. matrix</i> | Nr | Ncf | | |
| Expected-return news (Nr) | 0.0278 [0.0273] | 0.0144 [0.0178] | Diversification factor for expected-return news: | 0.6132 [0.2869] |
| Cash-flow news (Ncf) | 0.0144 [0.0178] | 0.0222 [0.0192] | Diversification factor for cash-flow news: | 0.2172 [0.1445] |

Table VIII: The size of the cumulative approximation error κ

This table reports the covariance matrix of expected-return news, indirectly computed cash-flow news, directly computed cash-flow news, and the approximation error. The news covariance matrix is computed from the VAR parameters:

$$z_{i,t} = \Gamma z_{i,t-1} + u_{i,t}, \quad \Sigma = E(u_{i,t}u'_{i,t})$$

The VAR state vector includes (in order) market-adjusted log stock return, \tilde{r} , the market-adjusted log book-to-market ratio, $\tilde{\theta}$, market-adjusted log GAAP profitability, \tilde{e}^{GAAP} , and market-adjusted log clean-surplus profitability, \tilde{e}^{CS} . Define $e1' \equiv [1 \ 0 \ \dots \ 0]$, $e4' \equiv [0 \ \dots \ 0 \ 1]$, and $\lambda' \equiv e1' \rho \Gamma (I - \rho \Gamma)^{-1}$. Expected-return news can then be conveniently expressed as $\lambda' u_{i,t}$, indirect cash-flow news as $(e1' + \lambda') u_{i,t}$, direct cash-flow news as $e4' (I - \rho \Gamma)^{-1} u_{i,t}$, and the approximation error as the difference between indirect and direct cash-flow news. The covariance matrix of these four terms is computed using the VAR-error covariance matrix Σ .

For each element of the covariance matrix, I report two numbers. The first number (bold) is the point estimate. The second number (in brackets) is a robust jackknife standard error computed using a jackknife method outlined by Shao and Rao (1993).

I use the CRSP-COMPUSTAT intersection 1954-1996 as the sample, in total 36791 firm-years.

| Covariance matrix of news terms and the approximation error. | | | | (estimate), [j.s.e.] |
|--|---------------------------|----------------------------|----------------------------|----------------------------|
| <i>Cov. matrix</i> | Nr | Ncf (indirect) | Ncf (direct) | κ |
| Expected-return news (Nr) | 0.0161 [0.0069] | 0.0144 [0.0075] | 0.0132 [0.0074] | 0.0012 [0.0004] |
| Indirect cash-flow news (Ncf) | 0.0144 [0.0075] | 0.0795 [0.0132] | 0.0809 [0.0133] | -0.0014 [0.0004] |
| Direct cash-flow news (Ncf) | 0.0132 [0.0074] | 0.0809 [0.0133] | 0.0832 [0.0134] | -0.0023 [0.0007] |
| Cumulative approximation error (κ) | 0.0012 [0.0004] | -0.0014 [0.0004] | -0.0023 [0.0007] | 0.0009 [0.0003] |

Table A1.I: Return on equity (ROE) and the approximate identity

The table estimates an expansion point for the approximate identity relating the book-to-market ratio, stock return, and return on equity. The table shows regression estimates of profitability less stock return plus lagged book-to-market on future book-to-market.

The table contains two rows. The first row regresses the excess log clean-surplus ROE, $e_t^{CS} - f_t$, less excess log stock return, r_t , plus lagged log book-to-market, θ_{t-1} , on log book-to-market, θ_t . The second row regresses the excess log US GAAP ROE, $e_t^{GAAP} - f_t$, less excess log stock return, r_t , plus lagged log book-to-market, θ_{t-1} , on log book-to-market, θ_t . f_t is log one plus the interest rate. Clean-surplus return on equity is calculated using the formula

$$e_t^{CS} = \log \left\{ \left[\frac{(1 + R_t + F_t)M_{t-1} - D_t}{M_t} \right] \left[\frac{B_t}{B_{t-1}} \right] + \left[\frac{D_t}{B_{t-1}} \right] \right\}$$

where M denotes market and B book equity, D dividends, and F the interest rate.

I use the CRSP-COMPUSTAT intersection 1954-1996 as the sample, in total 36791 firm-years.

Accuracy of the approximation

| Y-variable | Intercept | Discount coef. ρ | X-variable | R ² |
|---|-----------|-----------------------|------------|----------------|
| $(e_t^{CS} - f_t) - r_t + \theta_{t-1}$ | -0.000 | 0.967 | θ_t | 99.82% |
| $(e_t^{GAAP} - f_t) - r_t + \theta_{t-1}$ | -0.017 | 0.987 | θ_t | 88.25% |

Figure 1: Cumulative response of returns to shocks

The figure contains two graphs. The first graph shows the cumulative response of returns to a 25% unexpected return. The impulse response function is calculated from the short VAR in Table II. The 25% return shock is induced by setting the first element of the VAR-error vector to 0.25. The other elements of the VAR-error vector are set to their conditional expectations, conditional on the first element being equal to 0.25.

The second graph shows the cumulative response of returns to a 25% cash-flow shock. The typical 25% cash-flow shock is induced by setting the VAR-error vector to a constrained maximum likelihood value, imposing a constraint cash-flow news equals 0.25.

The solid horizontal line is set to 25% on both graphs. Dashed lines denote +/- standard-error bounds, calculated with the jackknife.

Figure 2: Histogram of atypical discounts

The figure graphs a histogram of fitted values of market-adjusted atypical discounts, 36791 observations. The market-adjusted atypical discount is defined as:

$$\tilde{p}_{r,t-1} \equiv \sum_{j=0}^{\infty} \rho^j E_{t-1} \tilde{r}_{t+j}$$

The fitted values of atypical discounts are computed from the short VAR in Table II.

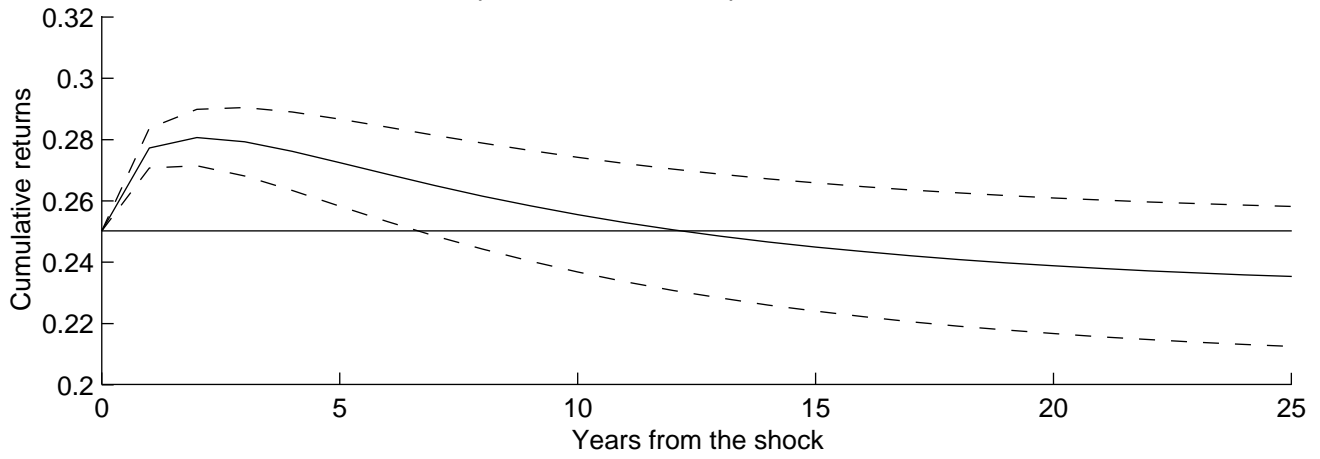
Figure 3: Local variance decompositions for size deciles

The figure graphs the local variance decomposition for the size deciles. The upper-left plot shows the variance of expected-return news, the upper-right plot the variance of cash-flow news, the lower-left plot the regression coefficient of return on cash-flow news, and the lower-right plot the news correlation as a function of the firm size. Dashed lines denote +/- standard-error bounds, calculated with the jackknife. For details, see Table IV.

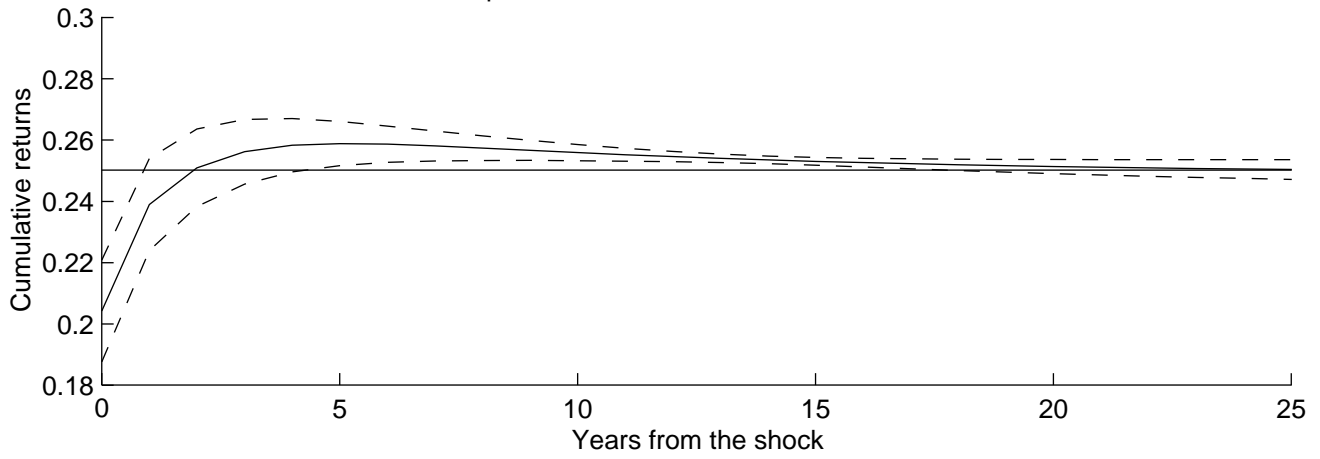
Figure 4: Variance decompositions for different return measurement horizons

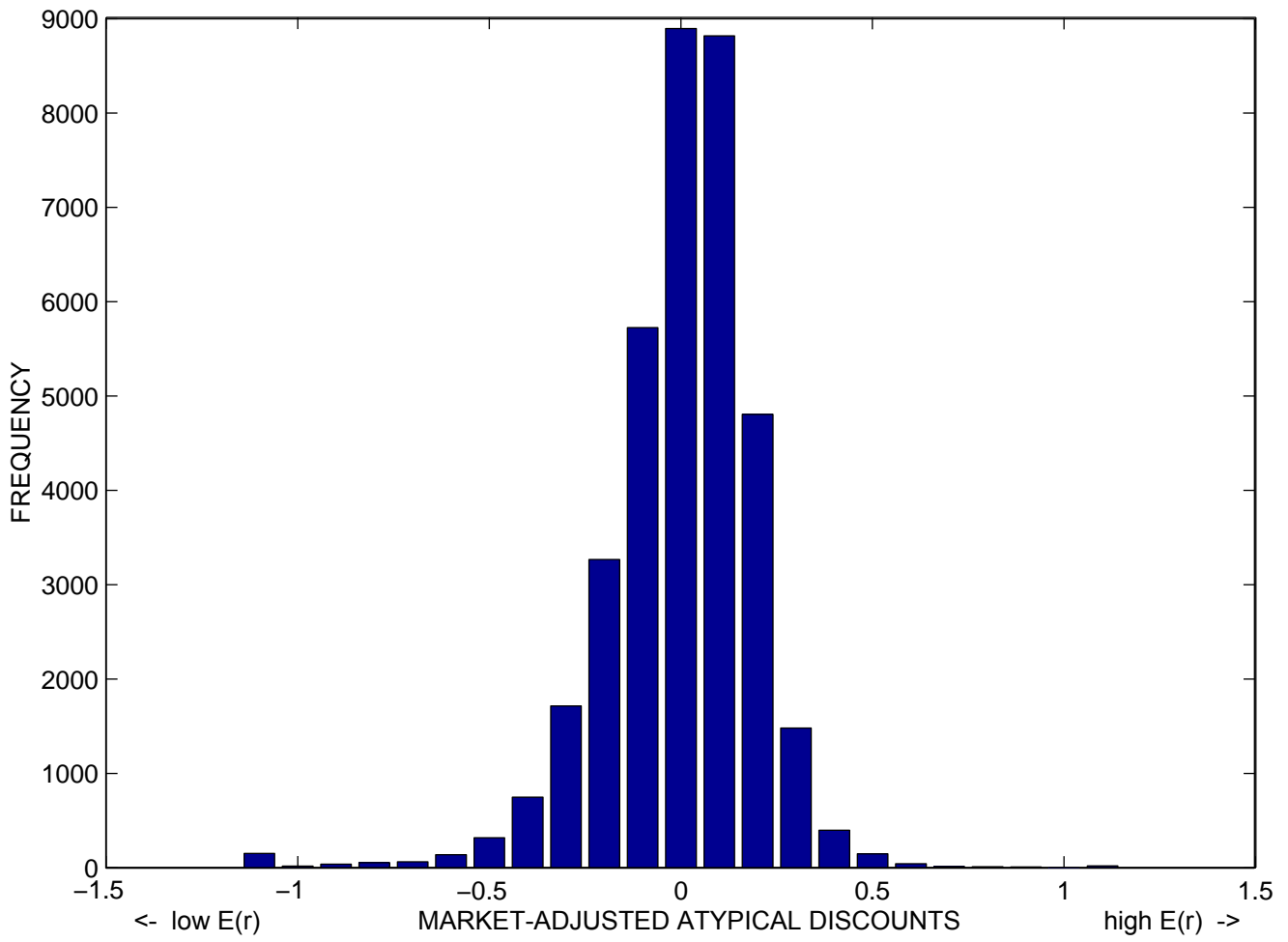
The figure graphs variance decompositions for different return measurement horizons. Figures are computed from the short VAR shown in Table II, using the formulas in Appendix 3. For all graphs, the x-axis signifies the return measurement interval in years. The upper-left plot shows the variance of expected-return news, the upper-right plot the variance of cash-flow news, the lower-left plot the regression coefficient of return on cash-flow news, and the lower-right plot the news correlation. Dashed lines denote +/- standard-error bounds, calculated with the jackknife.

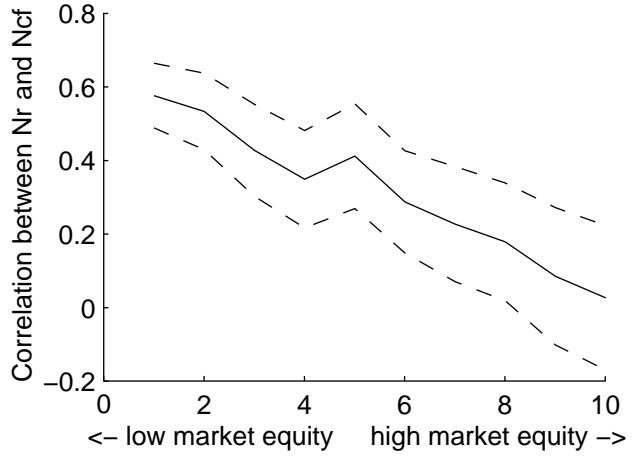
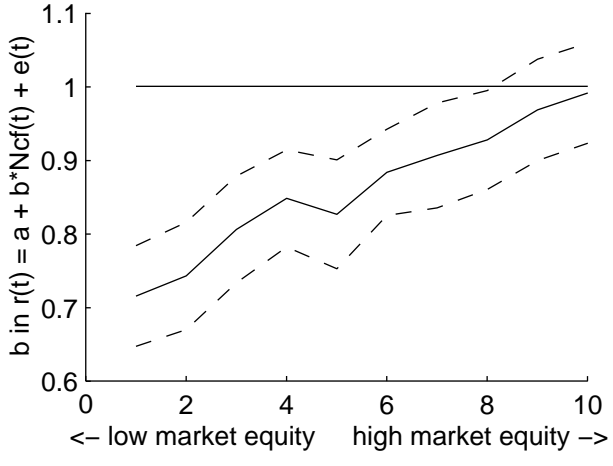
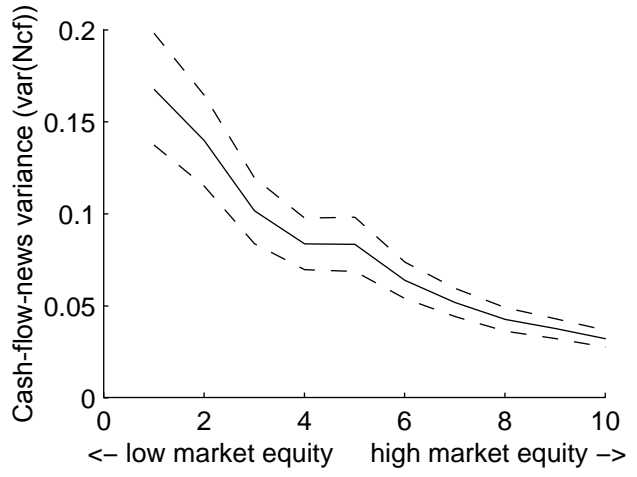
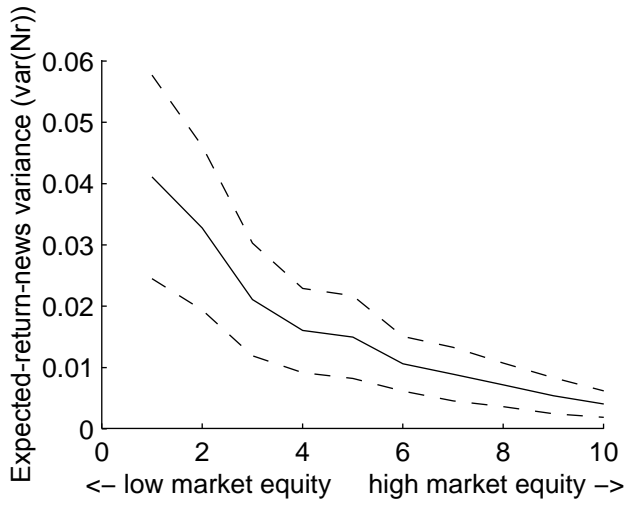
Response to a 25% unexpected return

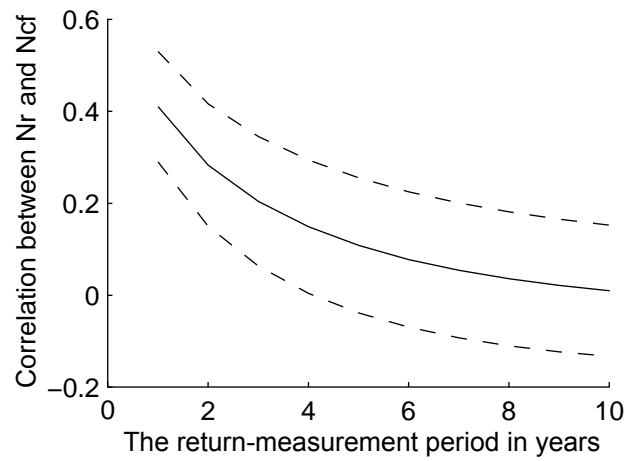
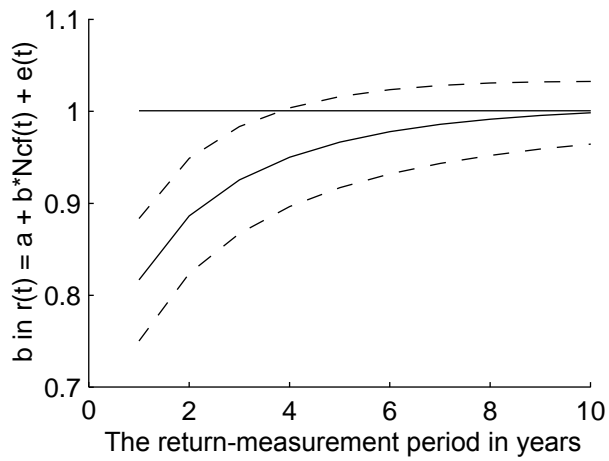
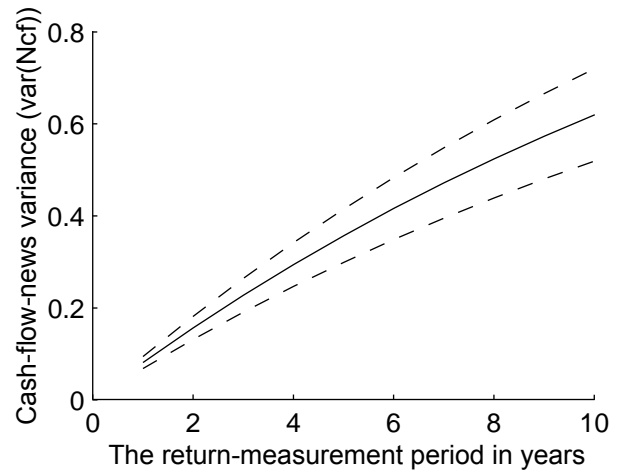
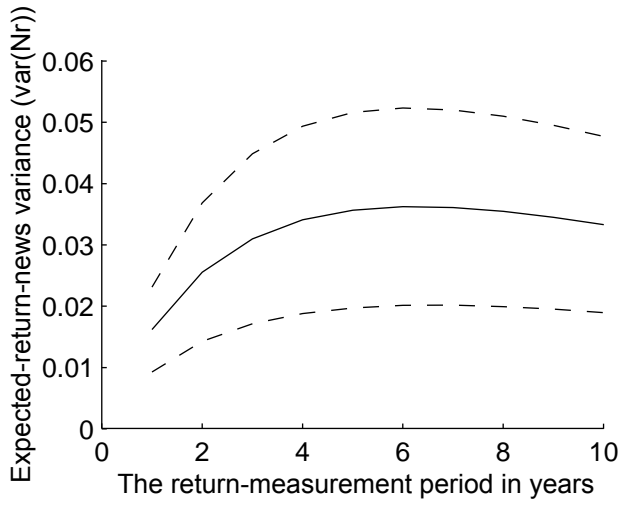


Response to a 25% cash-flow news









¹ Modeling dividend policy is difficult for two reasons. First, the time-series stability of a firm's dividend policy is suspect (Fama and French [2000a]). For example, changing tax laws and increasing use of (non-dividend-protected) executive stock options may have caused near-permanent shifts in a typical firm's dividend policy (Jolls [1998]). Second, because one can observe a wide range of dividend policies (from a high, stable pay-out ratio to not paying dividends at all), a homogeneous VAR would not adequately capture this persistent cross-sectional heterogeneity.

² Market-adjusted or benchmark-adjusted returns can be decomposed, as well. Apply equation (3) to individual firm-level stock returns and market returns separately, and subtract the latter from the former. As a result, the (unexpected) market-adjusted stock return can be decomposed into components due to above-market expected stock returns and ROEs. When the discussion applies only to market-adjusted quantities, I modify the notation by a tilde. For example, \tilde{r}_i denotes the market-adjusted log stock return.

³ The delisting-return assumptions follow Shumway's (1997) results. Shumway tracks a sample of firms whose delisting returns are missing from CRSP and finds that performance-related delistings are associated with a significant negative return, on average approximately -30 %. This assumption is unimportant to my final results, however.

⁴ In the case of market-adjusted returns and homogenous VAR, this typical risk premium is identically zero.

⁵ Because I constrain the transition matrix Γ to be equal across size groups, the size-related heterogeneity in the variance-decomposition results might be an artifact of the restriction of equal transition matrices across size groups. To investigate this possibility, I assume that each firm's assignment to a size quartile is permanent and estimate a separate Γ matrix for each size quartile. The results obtained under these assumptions have larger standard errors but the point estimates support my conclusions: The cash-flow-news variance (Q1:0.1662, Q2:0.0843, Q3:0.0488, Q4:0.0379), expected-return-news variance (Q1:0.0416, Q2:0.0147, Q3:0.0064, Q4:0.0061), news correlation (Q1:0.63, Q2:0.37, Q3:0.07, Q4:0.16), and ratio of expected-return-news to total-return variance (Q1:0.41, Q2:0.20, Q3:0.12, Q4:0.16) generally decrease with firm size.

⁶ The above discussion raises a more general methodological point. A typical overreaction study examines the univariate autocovariance function of returns and infers overreaction from long-horizon negative autocorrelation of returns. Unfortunately, as noted by Campbell (1991) in a different context, a univariate time-series approach cannot unambiguously estimate both the variance of expected-return news and the regression coefficient of return and cash-flow news. It is possible that positive cash-flow news implies

higher expected future returns, yet returns are negatively autocorrelated. Concluding overreaction to relevant cash-flow news (as I define it) from negative long-horizon autocorrelations is erroneous, because the long-horizon negative autocorrelation may be induced by expected-return variation that is orthogonal to cash-flow news, i.e., w_t .

⁷ I thank Jonathan Berk and Robert Merton for clarifying discussions on this topic.

⁸ I thank the anonymous referee for suggesting this interpretation.

⁹ See the summary of empirical evidence in Daniel, Hirshleifer, and Subrahmanyam (1998).

¹⁰ Unreported regressions indicate that the approximation is most accurate for simple returns when discount coefficient is set close to one.