NBER WORKING PAPER SERIES

EXPLAINING RETURNS WITH CASH-FLOW PROXIES

Peter Hecht
Tuomo Vuolteenaho

Working Paper 11169
http://www.nber.org/papers/w11169

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
March 2005

Earlier drafts of this paper were circulated under the titles "Stock return variation and expected dividends: a reinterpretation" and "Stock return variation and expected dividends: a comment." Corresponding author: Peter Hecht, (617) 495 6171, phecht@hbs.edu. We are indebted to John Cochrane, Eugene Fama, S.P. Kothari, Jay Shanken, and many anonymous referees for their suggestions. We would also like to thank Tom Berglund, Randy Cohen, George Constantinides, Campbell Harvey, J. B. Heaton, Jussi Keppo, Owen Lamont, Eva Liljeblom, Christopher Polk, William Schwert, Vesa Puttonen, Richard Thaler, and the seminar participants at the Swedish School of Economics and University of Chicago for helpful comments. Vuolteenaho wishes to thank the Foundation for Economic Education, the Emil Aaltonen Foundation, and the Oscar Mayer Foundation for financial support. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

Explaining Returns with Cash-Flow Proxies
Peter Hecht and Tuomo Vuolteenaho
NBER Working Paper No. 11169
March 2005
JEL No. E440, G100, G120, G140, G350

ABSTRACT

Stock returns are correlated with contemporaneous earnings growth, dividend growth, future real activity, and other cash-flow proxies. The correlation between cash-flow proxies and stock returns may arise from association of cash-flow proxies with one-period expected returns, cash-flow news, and/or expected-return news. We use Campbell's (1991) return decomposition to measure the relative importance of these three effects in regressions of returns on cash-flow proxies. In some of the popular specifications, variables that are motivated as proxies for cash-flow news also track a nontrivial proportion of one-period expected returns and expected-return news. As a result, the R2 from a regression of returns on cash-flow proxies may overstate or understate the importance of cash-flow news as a source of return variance.

Peter Hecht
Harvard Business School
Cambridge, MA 02163
phecht@hbs.edu

Tuomo Vuolteenaho
Harvard University
Department of Economics
Littauer Center
Cambridge, MA 02139
and NBER
t-vuolteenaho@harvard.edu

Most valuation theories identify time-varying one-period expected returns, news about cash flows, and news about future expected returns as the three possible sources of return variation. Barring the existence of infinitely-lived bubbles in asset prices, if an asset's return is high, its current one-period expected return must have been high, expected cash flows must have increased (i.e., cash-flow news must have been positive), and/or future expected returns decreased (i.e., expected-return news must have been negative.)

Fama (1990), Schwert (1990), Kothari and Shanken (1992), Liew (1995) and others regress aggregate stock returns on cash-flow proxies and find that cash-flow proxies explain returns well. The typical interpretation of these regressions is that cash-flow news are an important source of observed return variation. For example, Kothari and Shanken (p. 178) state, "In this paper, we find that proxies for the market's expectations of future dividends explain a substantial fraction of return variation." Some also interpret the high explanatory power of these types of regressions more generally as evidence of market rationality.

Using firm-level data, an extensive accounting literature focuses on the contemporaneous correlation of stock returns and earnings. Despite the statistically reliable positive association between stock returns and earnings, Ball and Brown (1968), Beaver, Clarke, and Wright (1979), Beaver, Lambert, and Morse (1980), Easton and Harris (1991), Collins, Kothari, Shanken, and Sloan (1994), and others find that the explained fraction of stock return variation is significantly less than one (typically under 10 percent). Lev (1989) and others suggest that the relatively low explanatory power stems from earnings' lack of timeliness and/or value-irrelevant noise in earnings.

Irrespective of motivation, however, the explanatory power of cash-flow proxies may arise from the correlation of cash-flow proxies with one-period expected returns, cash-flow news, and/or expected-return news. If expected-return variation is responsible for the high explanatory power of the aggregate regressions, these R^2s should not be interpreted as evidence of cash-flow news driving the returns. Similarly, if expected-return news is highly variable and positively correlated with cash-flow news, the low R^2s in regressions of firm-level returns on earnings do not necessarily imply that earnings are a noisy or delayed measure of the cash-flow-generating ability of the firm. Even if earnings are a clean signal of cash-flow news, expected-return effects (due to variation in risk-adjusted discount rates and/or mispricing) can garble the earnings-returns relation.

The idea that correlation between a cash-flow proxy and stock return may be due to any of the three components is not novel. Fama (1990), Schwert (1990), Kothari and Shanken (1992), Campbell and Ammer (1993), and others recognize that when stock returns are regressed on cash-flow proxies, any of the three effects may be driving the regression coefficients. They do not, however, clearly quantify the relative importance of these three effects. Thus, in the end, it is still unclear why cash-flow proxies are or are not related to stock returns.

We measure the significance of the above three effects on various regressions of stock returns on cash-flow proxies. We empirically reinterpret the regression results by Beaver, Clarke, and Wright (1979), Fama (1990), Schwert (1990), Easton and Harris (1991), Kothari and Shanken (1992), Collins, Kothari, Shanken, and Sloan (1994) who regress aggregate or firm-level stock returns on cash-flow proxies. Using Campbell's (1991) return decomposition, a regression of returns on cash-flow proxies can be decomposed into three separate regressions, each corresponding to one component of return. Our procedure first assumes and estimates a simple return-generating process. This return-generating process then implies realized values for one-period expected returns, cash-flow news, and expected-return news. Once the three return components are extracted, we regress them individually on the cash-flow proxies. These three regressions expose what drives the regression coefficients in the original returns-on-cash-flow-proxies specification.

Our empirical results show that many of the variables in Fama's (1990) and Schwert's (1990) regressions generally proxy the return components in both an intended and unintended way. The default and term yield spread positively track one-period expected returns. However, the default spread is negatively related to cash-flow news, making the positive link between stock returns and the default spread insignificant. While shocks to the default and term yield spread positively track revisions in future expected returns, i.e., expected-return news, shocks to the term spread are also positively correlated with cash-flow news, producing an insignificant positive relationship between stock returns and the term spread shocks. Between the current and future dividend growth rates, only the current dividend growth is significantly positively correlated with cash-flow news. Since the current dividend growth is also strongly negatively correlated with expected-return news, the positive relationship between stock returns and the current dividend growth is much larger in magnitude.

Similar to Fama (1990) and Schwert (1990), many of the variables in Kothari and Shanken (1992) also proxy the return components in an intended and unintended way. For Kothari and Shanken's returns-on-dividend growth-rates specifications, the contemporaneous and future dividend growth rates track more than just cash-flow news. The contemporaneous dividend growth also strongly negatively tracks expected-return news while the one-year lead dividend growth positively tracks expected-return news. This result helps explain the (perhaps surprising) insignificant negative relationship between one-year lead dividend growth and stock returns. The second- and third-year lead dividend growth also positively track one-period expected returns, increasing the correlation with stock returns.

The additional measurement-error proxies in Kothari and Shanken's full specification also play a dual role in the regression. Because future dividend growth rates may be partially anticipated by the market, the realized dividend growth is a noisy proxy of new information about dividends. Kothari and Shanken include lagged dividend yield and future returns to help "clean-up" this measurement error in future dividend growth rates. In addition to reducing measurement error, our analysis shows that a nontrivial amount of the incremental explanatory power of the "clean-up" variables is due to expected-return effects. The lagged dividend yield positively tracks the one-period expected return while the future returns are positively correlated with expected-return news.

The firm-level results differ from the aggregate results in an interesting way. Consistent with the findings of Vuolteenaho (2002) and Cohen, Gompers, and Vuolteenaho (2003), our estimated VARs imply strong positive correlation between firm-level expected-return and cash-flow news. This positive correlation may in some cases have dramatic effects on firm-level regressions of returns on cash-flow proxies. For example, Easton and Harris (1991) regress returns on contemporaneous earnings (normalized by lagged price). Estimating a loosely similar regression equation in our sample yields a regression coefficient of 0.75 and R^2 of 10 percent. However, when we regress the estimated cash-flow news on the same earnings variable, we get a coefficient of 1.30 and R^2 of 27 percent. Although the level of earnings variable does a decent job tracking cash-flow news, it also tracks the negative of expected-return news with a coefficient of -0.70 and R^2 of 19 percent and the level of expected returns with a coefficient of 0.15 and R^2 of 9 percent. In the end, the association of the earnings variable with cash-flow and expected-return news partially cancel each other, leaving the original specification with a low coefficient and R^2.

The rest of the paper is organized as follows. Section one presents the framework for decomposing the stock return and the regression coefficient of returns on cash-flow proxies. Section two decomposes and reinterprets aggregate regressions while section three decomposes and reinterprets firm-level regressions. Section four concludes.

I. Framework

A. Decomposing returns and regressions of returns on cash-flow proxies

Using the log-linear approximation of dividend-price ratio derived by Campbell and Shiller (1988, refer to Appendix A), Campbell (1991) decomposes the definition of return.

$$\begin{aligned} r_t &\approx E_{t-1}r_t + (E_t - E_{t-1})\sum_{i=0}^{\infty}\rho^i \Delta d_{t+i} - (E_t - E_{t-1})\sum_{j=1}^{\infty}\rho^j r_{t+j} \\ &= E_{t-1}r_t + N_{cf,t} - N_{r,t} \end{aligned} \tag{1}$$

Above, $r \equiv$ log return, $\rho \equiv$ a constant determined by the average dividend yield and set to 0.96 for the remainder of the paper, and $\Delta d \equiv$ log dividend growth. The realized stock return can be split into three components: One-period expected return ($E_{t-1}r_t$), cash-flow news ($N_{cf,t}$), and negative of expected-return news ($-N_{r,t}$). Ignoring the (very small) linearization error, this decomposition accounts for the entire return: the stock return equals the sum of these three components.

Using the above return decomposition, a regression of returns on cash-flow proxies can be decomposed into three component regressions, one corresponding to conditional one-period expected return, cash-flow news, and expected-return news each. Consider a typical regression of returns on cash-flow proxies:

$$r_t = X_t(\phi_T)\beta + \varepsilon_t \tag{2}$$

The regression (2) explains returns with cash-flow proxies $X_t(\phi_T)$, where ϕ_T denotes the information set at the end of the world, recognizing the possibility that some of the variables may not be known at the end of the return period. Relative to r_t, X_t can contain past, contemporaneous, and future realizations.

Using Campbell's (1991) return decomposition, $r_t \approx E_{t-1}r_t + N_{cf,t} - N_{r,t}$ (defined in equation (1)), the original regression can be split into three component regressions:

$$\begin{aligned} E_{t-1}r_t &= X_t(\phi_T)\beta_{Er} + \varepsilon_{Er,t} \\ N_{cf,t} &= X_t(\phi_T)\beta_{Ncf} + \varepsilon_{Ncf,t} \\ -N_{r,t} &= X_t(\phi_T)\beta_{-Nr} + \varepsilon_{-Nr,t} \end{aligned} \tag{3}$$

Since the explanatory variables in each of the three component regressions are the same as in the original regression (2), we can think of the original regression (2) as the sum of the three component regressions[1]:

$$r_t = X_t(\phi_T)\left(\beta_{Er} + \beta_{Ncf} + \beta_{-Nr}\right) + \left(\varepsilon_{Er,t} + \varepsilon_{Ncf,t} + \varepsilon_{-Nr.t}\right) \tag{4}$$

From (3) and (4) it is clear that the cash-flow proxies, X_t, can explain stock returns well for various reasons, irrespective of the motivating logic. The X_t can explain the level of one-period expected returns, cash-flow news, or expected-return news – or any combination of the three. Similarly, if the relationship between X_t and cash-flow news is offset by the relationship between X_t and expected-return news and/or one-period expected returns, it is possible for X_t to explain cash-flow news but not returns.

For example, returns on a broad stock index, such as the S&P500, are positively correlated with future industrial production (Fama (1990), Schwert (1990)). This correlation may be driven by any of the three sources. First, the correlation may arise from high future production being associated with high one-period expected returns. Second, positive shocks to production may be associated with positive cash-flow news. Third, positive shocks to production may be associated with negative expected-return news. If the first and/or third source drives the correlation, concluding from this correlation that changing cash-flow expectations drive returns would be erroneous.

B. Estimating one-period expected returns, cash-flow news, and expected-return news

The above regression decomposition requires the estimated series of one-period expected returns, cash-flow news, and expected-return news as inputs. As demonstrated by Campbell (1991), a vector autoregressive (VAR) model provides a convenient way to implement this return decomposition and to calculate estimates of the three series. In particular, the VAR model enables solving the infinite expected sums in equation (1) with simple, closed-form formulas.

For the aggregate decompositions, we first estimate a VAR to calculate the expected-return and expected-return news terms[2]. Let z_t be a vector of state variables describing the economy at time *t*. In particular, let the first element of z_t be the aggregate log stock return. The state vector is assumed to follow a linear law:

$$z_t = A + \Gamma\, z_{t-1} + u_t\,. \tag{5}$$

The error term u_t is assumed to have a covariance matrix Σ and to be independent of everything known at $t-1$.

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

$$\lambda' \equiv e1'\rho\Gamma(I-\rho\Gamma)^{-1}. \tag{6}$$

Taking expectations of (5) yields an expression for the one-period expected return, $e1'(A+\Gamma z_{t-1})$. The definition (6) introduced by Campbell (1991) simplifies the other expressions considerably: Expected-return news can now be conveniently expressed as $\lambda' u_t$. Like Campbell (1991), we then solve for the cash-flow-news term as the residual (realized return minus expected return plus expected-return news), $(e1'+\lambda')u_t$. 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, Campbell's (1991) 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 does directly calculating the change in the discounted sum of expected future dividends. This concern turns out not to be an important one, however: Including dividend growth into the VAR state vector and computing cash-flow news directly yields very similar results.

The firm-level and aggregate methodologies are slightly different. For the firm-level VAR, $z_{i,t}$ represents a vector of firm-specific state variables. The first element of $z_{i,t}$ is the firm's market-adjusted log stock return. An individual firm's state vector is assumed to follow a linear law:

$$z_{i,t} = \Gamma z_{i,t-1} + u_{i,t}. \tag{7}$$

Since the firm-level state variables are cross-sectionally demeaned, we can safely omit the intercept and aggregate variables present in the aggregate regressions. Similar to the aggregate methodology, the error term $u_{i,t}$ is assumed to have a covariance matrix Σ and to be independent of everything known at $t-1$. As before, the firm-level expected-return and cash-flow news can be expressed as $\lambda' u_{i,t}$ and $(e1'+\lambda')u_{i,t}$, respectively. Most of our firm-level specifications restrict the VAR parameters to be constant over time and across firms, but some of our firm-level specifications will relax this restriction by allowing for cross-sectional variation in the VAR parameters (and including a constant term in (7)). All of our standard-error computations allow for general correlation structure of errors across firms.

While the VAR framework is convenient for estimating the series of one-period expected returns, cash-flow news, and expected-return news in an internally consistent manner, it is incapable of distinguishing between rational variation in expected returns and variation due to mispricing. For example, even if the model of market equilibrium requires constant expected returns and investors thus subjectively expect constant returns, the actual rationally expected returns might vary over time if investors forecast cash flows with systematic errors. To the extent that the VAR state vector contains variables that capture systematic errors in cash flow forecasts, the VAR framework will attribute this portion of the return variation due to mispricing to the time-varying expected returns and expected return news.[3]

However, even if expected return variation is due to mispricing, the VAR return decomposition still provides a means for isolating the relationship between returns and cash flows. Even if some investors have irrational expectations, there should be other investors with rational expectations, and it is interesting to understand asset price behavior from the perspective of these investors. For example, if cash flow news and mispricing (which shows up as expected return news) are not uncorrelated, a simple correlation between returns and cash flow news or a regression of returns on cash flow news will over- or underestimate the true relationship between returns and cash flows. The VAR framework provides a means for controlling for the correlation structure between expected return and cash flow news, irrespective of whether the expected return variation is rational.

Lastly, the VAR framework has other potential limitations, too. The linear time series structure of the VAR may not provide a good approximation to the true return data generating process. Even if the linear assumption were reasonable, the variance decomposition results are conditional on the information set included in the state vector. Depending on the correlation between the omitted and included state variables, it is possible to arrive at incorrect conclusions if the VAR is misspecified. While Appendix B addresses some of the specification concerns by considering alternative state vectors, it is important to interpret all of our results with these cautions in mind.

II. Empirical decompositions of aggregate regressions

A. Aggregate data

The main data for the aggregate regressions come from the Center for Research in Security Prices (CRSP), Global Financial Data (GFD), Shiller (2000), Schwert (1990), and Macaulay (1938), 1881 to 2001. All variables are of annual frequency.

Motivated by the vast return predictability literature (see, for example, Fama and French (1989) and Campbell and Shiller (1988, 1998)), the state vector of our main VAR includes annual log return (LRET), log price-earnings ratio (LPE), and the spread of the ten-year constant maturity bond log yield over the three-month commercial paper log yield (LTERM). For 1881-1925, LRET is the log return on the aggregate stock market index used in Schwert (1990). After 1925, LRET is the log return on the CRSP value-weight index. LPE is from Shiller (2000), calculated as the price of the S&P 500 divided by the trailing ten-year moving average of aggregate earnings of companies in the S&P 500. The bond and commercial paper yields come from GFD.

In addition to the main VAR state variables, we also use cash-flow and expected-return proxies motivated by previous research. Fama (1990) and Schwert (1990) explain aggregate returns with the lagged dividend yield, spread between long-term and short-term government bonds, yield spread between low-grade and high-grade grade corporate bonds (LDEF), shocks to both yield spreads, and contemporaneous and future industrial-production growth. Similar to Fama (1990) and Schwert (1990), for our set of expected-return proxies, we use LDEF, LTERM, shocks to the default spread (LDEFS), and shocks to LTERM (LTERMS). For our cash-flow proxies, we use the current, one-year lead, two-year lead, and three-year lead annual log dividend growth rate (LGD, LGDF1, LGDF2, and LGDF3). Default spread is the spread of Moody's Baa corporate bond log yield over the Aaa corporate bong log yield obtained from GFD. Since Moody's Baa corporate bond yield is unavailable before 1919, for 1881-1918, LDEF is calculated as the spread of the average top five Macaulay railroad bond log yields (excluding the maximum) over the average bottom five Macaulay railroad bond log yields (excluding the minimum) obtained from Macaulay (1938). The shocks are the estimated residuals from a first-order autoregressive process. The dividend growth data comes from Schwert (1990) for 1881-1925 and CRSP for 1926-2001.

Kothari and Shanken (1992) explain aggregate returns with the lagged dividend yield, contemporaneous and future dividend growth, and future returns. Similar to Kothari and Shanken (1992), for our other two sets of cash-flow proxies, we use the annual log dividend yield (LDP), LGD, LGDF1, LGDF2, LGDF3, one-year lead LRET (LRETF1), two-year lead LRET (LRETF2), and three-year lead LRET (LRETF3). For 1881-1925, these data are for the aggregate stock market index used in Schwert (1990). For 1926-2001, these data are for the CRSP value-weight index obtained from CRSP.

In Appendix B, we explore alternative aggregate VAR specifications. Additional variables in these specifications include the default spread (LDEF), annual log dividend yield (LDP), annual log dividend-earnings ratio (LDE), log book-to-market ratio (LBM), log return on equity (LROE), and the stochastically detrended three-month commercial paper log yield (LDTY). Our book-to-market and profitability data are constructed by Vuolteenaho (2001). LDTY is the spread between the current three-month commercial paper log yield and the average three-month commercial paper log yield over the previous twelve months.

Throughout the paper, all relevant variables are deflated by the commercial paper/T-bill wealth index obtained from GFD. The inflation series is unreliable for the early part of our sample.

Table I shows the descriptive statistics for the aggregate data. Each panel shows means, standard deviations, minimums, percentiles of the variables, and maximums. Panel A shows the descriptives for the VAR state variables. Panel B shows the descriptives for Fama (1990) and Schwert's (1990) and Kothari and Shanken's (1992) cash-flow and expected-return proxies.

B. Estimated process for aggregate stock returns

The state vector of our main VAR (equation (8)) includes LRET, LPE, and LTERM.

$$z_t = A + \Gamma z_{t-1} + u_t, \quad z_t = \left[LRET_t \quad LPE_t \quad LTERM_t \right]', \quad \Sigma = E\left(u_t u_t'\right) \tag{8}$$

This time series model generates our three components of returns: One-period expected return ($E_{t-1}r_t$), cash-flow news ($N_{cf,t}$), and negative of expected-return news ($-N_{r,t}$). We consider alternative VAR specifications in Appendix B.

Table II contains the VAR parameter estimates and a variance decomposition obtained using a 1881-1998 sample.[4] Consistent with the previous return-predictability findings (e.g., Keim and Stambaugh (1986), Campbell and Shiller (1988), and Fama and French (1989)), LPE and LTERM are statistically significant predictors of return. Both LTERM and, especially, LPE are highly persistent. Consistent with the variance decomposition literature, expected-return news is much more volatile than cash-flow news (e.g., Campbell and Ammer (1993)). The ratio of expected-return-news variance to cash-flow-news variance is 5.34. The correlation between the two estimated news series is -0.29. This negative correlation is consistent with a story where unexpected good economic times are associated with lower than normal investor risk aversion, producing lower equilibrium expected future returns.

C. Reinterpreting regressions of returns on real activity

Fama's (1990) and Schwert's (1990) main cash-flow proxy is current and future industrial-production growth. We regress LRET and its three estimated components (one-period expected return, cash-flow news, and negative of expected-return news) individually on explanatory variables (defined above) similar in spirit to Fama (1990) and Schwert (1990):

$$
\begin{aligned}
LRET_t &= B_0 + B_1 LDEF_{t-1} + B_2 LTERM_{t-1} + B_3 LDEFS_t + B_4 LTERMS_t \\
&\quad + B_5 LGD_t + B_6 LGDF1_{t+1} + B_7 LGDF2_{t+2} + B_8 LGDF3_{t+3} + \varepsilon_t \\
E_{t-1} r_t &= B_{Er,0} + B_{Er,1} LDEF_{t-1} + B_{Er,2} LTERM_{t-1} + B_{Er,3} LDEFS_t + B_{Er,4} LTERMS_t \\
&\quad + B_{Er,5} LGD_t + B_{Er,6} LGDF1_{t+1} + B_{Er,7} LGDF2_{t+2} + B_{Er,8} LGDF3_{t+3} + \varepsilon_{Er,t} \\
N_{cf,t} &= B_{Ncf,0} + B_{Ncf,1} LDEF_{t-1} + B_{Ncf,2} LTERM_{t-1} + B_{Ncf,3} LDEFS_t + B_{Ncf,4} LTERMS_t \\
&\quad + B_{Ncf,5} LGD_t + B_{Ncf,6} LGDF1_{t+1} + B_{Ncf,7} LGDF2_{t+2} + B_{Ncf,8} LGDF3_{t+3} + \varepsilon_{Ncf,t} \\
-N_{r,t} &= B_{-Nr,0} + B_{-Nr,1} LDEF_{t-1} + B_{-Nr,2} LTERM_{t-1} + B_{-Nr,3} LDEFS_t + B_{-Nr,4} LTERMS_t \\
&\quad + B_{-Nr,5} LGD_t + B_{-Nr,6} LGDF1_{t+1} + B_{-Nr,7} LGDF2_{t+2} + B_{-Nr,8} LGDF3_{t+3} + \varepsilon_{-Nr,t}
\end{aligned} \tag{9}
$$

Our regressions deviate from Fama's and Schwert's specifications slightly. Fama and Schwert use quarterly industrial-production growth rates. Due to data availability, we use annual aggregate dividend growth rates.

As shown in the first row of Table III Panel A, the above explanatory variables explain 53 percent (adjusted R^2, period 1882-1998) of the total return variation. Our adjusted R^2 is similar in magnitude to Fama's 59 percent (period 1953-1987) and Schwert's 39 percent (period 1918-1988). Similar to Fama and Schwert, the term spread and term-spread shock are statistically insignificant. Similar to Schwert but unlike Fama, we find that the default-spread shock has a negative and significant coefficient in our regression. Unlike Fama's and Schwert's, the default spread is statistically insignificant. Also, our results do not indicate that stock returns are more related to future than contemporaneous dividend growth. The slope on the contemporaneous dividend growth is large and highly significant (coefficient 0.51, s.e. 0.09).

Rows two through four of Table III Panel A show the regressions of the individual return components. The Fama-Schwert proxies explain 67 percent of the time variation of estimated expected returns. Since the proxies include one of the prediction variables in our VAR, LTERM, this is not too surprising. It is important to note, however, that LDEF plays a much stronger role in tracking expected returns (coefficient 4.52, s.e. 1.79). This link between LDEF and expected returns is impossible to identify from looking only at the total return specification.

The Fama-Schwert proxies explain roughly half of the estimated cash-flow-news series (46 percent adjusted R^2). Most of the explanatory power comes from the contemporaneous dividend growth (coefficient 0.16, s.e. 0.05). The future dividend growth rates are only marginally positive. A nontrivial proportion of the explanatory power is due to LTERMS (coefficient 1.92, s.e. 0.83). This significant positive relationship partially explains why the shock to the term spread plays almost no role in explaining total return variation. Furthermore, cash-flow news is statistically negatively related to the lagged default spread (coefficient –2.23, s.e. 0.87). (No lagged variables would track the news series in sample if all lagged variables in the second-stage regression were included in the VAR.) Similar in spirit to the LTERMS situation, the negative relationship between LDEF and cash-flow news gives the false impression that LDEF does not track expected returns when only looking at the total return specification.

Fama's and Schwert's proxies explain 41 percent of the estimated expected-return news series. As expected, the lagged variables are not related to our news series. Both LDEFS (coefficient –17, s.e. 4.1) and LTERMS (coefficient –0.88, s.e. 1.1) are negatively related to the negative of expected-return news. Even though LTERM is included in our VAR as a predictor variable, the results for LDEFS are stronger. The contemporaneous dividend growth is reliably negatively correlated with future expected returns, i.e. positively correlated with our negative of expected-return news series (coefficient 0.31, s.e. 0.07). This result explains why the coefficient on LGD in the total return regression is so large relative to the future dividend growth coefficients. The fact that LGD is negatively correlated with future expected returns makes the regression coefficient of returns on dividend growth an upward biased estimator of the regression coefficient of cash-flow news on dividend growth.

In sum, many of Fama's and Schwert's proxy variables behave in both an intended and unintended way. Although their expected return proxies, LDEF and LTERM, capture a large proportion of the expected return variation, both variables are negatively related to cash-flow news and the negative of expected return news. This additional relationship breaks their link with expected returns when looking at total return variation. The dividend growth variables track cash-flow news, but most of the contemporaneous dividend growth's explanatory power in the total returns regression originates from its large positive correlation with the negative of expected return news. Shocks to the default and term spread are negatively correlated with the negative of expected return news. However, LTERMS is positively related to cash-flow news, reversing the relationship between LTERMS and total returns. In general, measuring a variables relationship (or lack of relationship) with total returns provides a distorted picture on whether that variable is (or is not) tracking a particular component of total return. Empirically, this distortion is of first order importance.

D. Reinterpreting regressions of returns on dividend growth rates

Instead of general economic activity, Kothari and Shanken (1992, KS hereafter) use dividend growth rates as their main cash-flow proxies. As above, we individually regress the three estimated components of return on the cash-flow proxies used in KS's dividends-only specification:

$$\begin{aligned}
LRET_t &= b_0 + b_1 LGD_t + b_2 LGDF1_{t+1} + b_3 LGDF2_{t+2} + b_4 LGDF3_{t+3} + \varepsilon_t \\
E_{t-1}r_t &= b_{Er,0} + b_{Er,1} LGD_t + b_{Er,2} LGDF1_{t+1} + b_{Er,3} LGDF2_{t+2} + b_{Er,4} LGDF3_{t+3} + \varepsilon_{Er,t} \\
N_{cf,t} &= b_{cf,0} + b_{cf,1} LGD_t + b_{cf,2} LGDF1_{t+1} + b_{cf,3} LGDF2_{t+2} + b_{cf,4} LGDF3_{t+3} + \varepsilon_{cf,t} \\
-N_{r,t} &= b_{-Nr,0} + b_{-Nr,1} LGD_t + b_{-Nr,2} LGDF1_{t+1} + b_{-Nr,3} LGDF2_{t+2} + b_{-Nr,4} LGDF3_{t+3} + \varepsilon_{-Nr,t}
\end{aligned} \tag{10}$$

The results are shown in Table III Panel B1. We find that the contemporaneous and future dividend growth rates explain 39.21 percent of the return variation (first row of Table III Panel B1). Most of the explanatory power comes from the contemporaneous dividend growth rate (coefficient 0.73, s.e. 0.12). Although two out of the three future dividend growth rates are positive, only LGDF3 is mildly significant (coefficient 0.14, s.e. 0.09).

Rows two through four of Table III Panel B1 show the regressions on the individual return components. The KS proxies explain 12 percent of the time variation of estimated expected returns. Except for the LGDF1 slope, each of the coefficients is of the same sign as the coefficients estimated in the total return regression. In particular, both LGDF2 (coefficient 0.09, s.e. 0.04) and LGDF3 (coefficient 0.09, s.e. 0.04) are significant and similar in size to their corresponding slopes in the total return regression.

With respect to the expected-return news series (row four of Table III Panel B1), the KS proxies explain 26 percent of the time variation in expected-return news. The contemporaneous dividend growth is highly significant and positive (coefficient 0.45, s.e. 0.09), explaining over 60% of the slope size estimated in the total return regression. This result is consistent with the aggregate VAR variance decomposition (bottom of Table II): Good news about cash flows is associated with lower future returns (higher negative expected return news). Relative to its slope in the other regressions, the magnitude of the LGDF1 coefficient is large and negative (coefficient -0.13, s.e. 0.09), explaining its negative relationship with total returns in row 1.

The VAR estimated cash-flow news is marginally related to the contemporaneous and future dividend growth rates. Row three of Table III Panel B1 reports an adjusted R^2 of 24 percent. Although all of the dividend growth variables are positively related to cash-flow news, only the LGD coefficient is economically and statistically significant (coefficient 0.20, s.e. 0.07). However, the slope on LGD still only accounts for approximately one quarter of its positive relationship with total returns. The magnitudes of the other slopes are all less than 0.06.

The above results are simultaneously good and bad news for KS's dividends-only specification. On a positive note, KS's cash-flow proxies do partially explain the cash-flow-news series estimated from the VAR. All of KS's cash-flow proxies are positively related to cash-flow news. However, similar to the results using the Fama-Schwert proxies, KS's total returns on dividend growth regressions hides the fact that contemporaneous and future dividend growth are also significantly related to expected-return news. In addition, contemporaneous and future dividend growth rates are also correlated with one-period expected returns for KS's total returns on dividend growth regressions. This interaction distorts the overall and relative "cash-flow" importance of each cash-flow proxy when looking only at the coefficients from the total return on dividend growth regressions. The coefficients from the total return regression tend to be two to five times as large as their counterparts measured in the cash-flow news regression.

E. Reinterpreting regressions of returns on dividend growth rates and clean-up variables

KS point out that using realized dividend growth rates as proxies for changes in expectations induces an errors-in-variables bias. The portion of realized dividend growth that was expected is measurement error. Also, the portion of realized dividend growth representing information that arrived after the return period is measurement error. In light of these errors-in-variables problems, KS introduce "clean-up variables." In order to account for the market's expectation at the beginning of the return period, KS use the lagged, annual log dividend yield (LDP). In order to control for the information that is contained in the future dividend growth rates but arrived after the return period, KS add future returns 3 years out (LRETF1, LRETF2, and LRETF3) to the final time-series regression.

We examine the explanatory power of a regression similar in spirit to KS's full specification:

$$
\begin{aligned}
LRET_t &= b_0 + b_1 LDP_{t-1} + b_2 LGD_t + b_3 LGDF1_{t+1} + b_4 LGDF2_{t+2} + \\
&\quad b_5 LGDF3_{t+3} + b_6 LRETF1_{t+1} + b_7 LRETF2_{t+2} + b_8 LRETF3_{t+3} + \varepsilon_t \\
E_{t-1} r_t &= b_{Er,0} + b_1 LDP_{t-1} + b_{Er,2} LGD_t + b_{Er,3} LGDF1_{t+1} + b_{Er,4} LGDF2_{t+2} + \\
&\quad b_{Er,5} LGDF3_{t+3} + b_{Er,6} LRETF1_{t+1} + b_{Er,7} LRETF2_{t+2} + b_{Er,8} LRETF3_{t+3} + \varepsilon_{Er,t} \\
N_{d,t} &= b_{Nd,0} + b_{Nd,1} LDP_{t-1} + B_{Nd,2} LGD_t + b_{Nd,3} LGDF1_{t+1} + b_{Nd,4} LGDF2_{t+2} + \\
&\quad b_{Nd,5} LGDF3_{t+3} + b_{Nd,6} LRETF1_{t+1} + b_{Nd,7} LRETF2_{t+2} + b_{Nd,8} LRETF3_{t+3} + \varepsilon_{Nd,t} \\
-N_{r,t} &= b_{-Nr,0} + b_{-Nr,1} LDP_{t-1} + b_{-Nr,2} LGD_t + b_{-Nr,3} LGDF1_{t+1} + b_{-Nr,4} LGDF2_{t+2} + \\
&\quad b_{-Nr,5} LGDF3_{t+3} + b_{-Nr,6} LRETF1_{t+1} + b_{-Nr,7} LRETF2_{t+2} + b_{-Nr,8} LRETF3_{t+3} + \varepsilon_{-Nr,t}
\end{aligned}
\tag{11}
$$

Including the "clean-up variables" increases the full regression R^2 from 39 percent to 45 percent (row one of Table III Panel B2). Both LDP (coefficient 0.12, s.e. 0.07) and LRETF2 (coefficient -0.25, s.e. 0.09) play a significant role. Although statistically insignificant, the slopes on LRETF1 (coefficient -0.05, s.e. 0.11) and LRETF3 (coefficient -0.15, s.e. 0.09) are similar in sign to LRETF2. Like the results from KS's "dividends only" specification (Table III Panel B1), contemporaneous dividend growth (coefficient 0.74, s.e. 0.11) plays an important role. Including the "clean-up variables" increases the economic and statistical significance of the future dividend growth variables LGDF2 (coefficient 0.29, s.e. 0.11) and LGDF3 (coefficient 0.22, s.e. 0.11).

Rows two through four of Table III Panel B2 decompose KS's full regression. First, the point estimates indicate that lagged dividend yield tracks one-period expected returns (coefficient 0.11, s.e. 0.04). Its relationship with one-period expected returns almost completely explains the positive correlation between the lagged dividend yield and total returns. Second, all of the future return explanatory variables are negatively correlated with the negative of expected-return news, i.e. future realized returns are positively correlated with future expected returns. Of the three future return explanatory variables, the two-year lead return is the most economically and statistically significant (coefficient –0.26, s.e. 0.08). Third, the coefficients on the dividend growth variables exhibit a pattern similar to those found in KS's "dividends only" specification (Table III Panel B1). Although they are all positively related to the cash-flow news series (LGD: coefficient 0.20, s.e. 0.07, LGDF1: coefficient 0.01, s.e. 0.05, LGDF2: coefficient 0.04, s.e. 0.05, LGDF3: coefficient 0.04, s.e. 0.05), the relationship between dividend growth rates and total returns is still mainly a result of the strong positive correlations between dividend growth and expected returns and/or dividend growth and the negative of expected-return news. Relative to the other dividend growth variables, LGD has the strongest positive relationship with one-period expected returns and the negative of expected-return news (coefficient 0.10, s.e. 0.05 and coefficient 0.44, s.e. 0.09, respectively). Although less significant, LGDF2 and LGDF3 are positively related to one-period expected returns and the negative of expected-return news too (LGDF2: coefficient 0.08, s.e. 0.05 and coefficient 0.17, s.e. 0.10, LGDF3: coefficient 0.05, s.e. 0.04 and coefficient 0.13, s.e. 0.10, respectively). In sum, the picture emerging from the regression decomposition is that a nontrivial amount of the increased explanatory power gained by adding the clean-up variables comes from expected-return effects. The relationship between the dividend growth variables and cash-flow news does not appear to be materially affected by the inclusion of the clean-up variables.

We would also like to make the point that adding clean-up variables can in many cases obscure the interpretation of regressions of returns on cash-flow proxies. In Appendix C, we use an analytical and simulation argument to show why, for example, Kothari and Shanken's (1992) main specification and other specifications that include future returns as explanatory variables are especially prone to track not only cash-flow news but also one-period expected returns and expected-return news. As a part of their empirical evidence, Kothari and Shanken (1992) estimate a regression, which explains returns with lagged dividend yield, contemporaneous and future dividend growth rates, and future returns. Via the Campbell-Shiller (1988) log-linear framework, we illustrate in Appendix C that this specification is one future dividend yield short of an approximate identity. As more and more future dividend growth rates and returns are added to the regression, the future dividend yield becomes less important to the identity, producing a higher and higher population regression R^2. Therefore, in the limit, a regression of returns on the lagged dividend yield, contemporaneous and future dividend growth rates, and future returns is *guaranteed* to track all three components of returns, irrespective of the true drivers of stock return variation.

III. Empirical decompositions of firm-level regressions

A. Firm-level data

The basic data come from the CRSP-COMPUSTAT intersection, 1954 to 1998. 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 U.S. stocks. All variables are of annual frequency.

In order to be included in the VAR sample, a firm-year must satisfy the following CRSP and COMPUSTAT data requirements. We require all firms to have a December fiscal-year end of *t*-1, in order to align accounting variables across firms. A firm must have *t*-1, *t*-2, and *t*-3 book equity available, where *t* denotes time in years. We require *t*-1 and *t*-2 net income and long-term debt data. A valid market-equity figure must be available for *t*-1, *t*-2, and *t*-3. We require that there is a valid trade during the month immediately preceding the period t return, ensuring that any return predictability is not spuriously induced by stale prices or other similar market micro-structure issues. We also require at least one monthly return observation during each of the preceding five years, from *t*-1 to *t*-5. We 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. We carefully avoid imposing any COMPUSTAT or CRSP requirements on year *t* data, because these data are used in the dependent variables in our VAR.

The stock return is defined as the annual value-weight return on a firm's common stock issues (typically one). If no year *t* return data are available, we substitute zeros for both returns and dividends. Annual returns are compounded from monthly returns, recorded from the beginning of May to the end of April. If a firm is delisted but the delisting return is missing, we investigate the reason for disappearance. If the delisting is performance-related, we assume a -30 percent delisting return. Otherwise, we assume a zero delisting return.

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

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

ROE is the earnings (COMPUSTAT data item 172) over the last period's book equity. When earnings are missing, we use the clean-surplus formula to compute a proxy for earnings; i.e., earnings equals the change in book equity plus dividends. In every case, we do not allow the firm to lose more than its book equity. That is, we define the net income as a maximum of the 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 ROE is truncated to –100 percent.

Campbell's (1991) return decomposition uses log returns. The log transformations may cause problems if some stock returns are close to –100 percent. We follow Vuolteenaho (2002) and solve this complication by redefining the firm as a portfolio of 90 percent common stock and 10 percent 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.

In addition to the VAR state variables, we also construct cash-flow proxies motivated by previous research. Beaver, Clarke, and Wright (1980) explain returns with simple earnings growth. We construct a similar variable as the ratio of year t earnings per share (COMPUSTAT data item 58) over year t-1 earnings per share (EPS). We set values outside the [exp(-1), exp(1)] range to these range endpoints. When regressing returns on this variable, we also exclude firm-years with either year t or year t-1 EPS less than zero.

Easton and Harris (1991) regress returns on earnings and earnings change normalized by lagged stock price. We compute these variables as EPS(t)/P(t-1) and (EPS(t)-EPS(t-1))/P(t-1), where P(t-1) is the end-of-April price of the common share class with highest market capitalization. Following Easton and Harris, we exclude firm-years with EPS(t-1)/P(t-1), EPS(t)/P(t-1), or (EPS(t)-EPS(t-1))/P(t-1) outside the range [-1.5, 1.5], when regressing returns on these variables.

Collins, Kothari, Shanken, and Sloan (1994) explain returns with contemporaneous and future log earnings growth, future returns, and lagged earnings-price ratio. When regressing returns on these variables, we truncate the data for firm-years with any of the right-hand-side log return or log earnings-growth variables outside the [-1, 1] range. Also, we delete firm years with any of the t-1, t, t+1, t+2, t+3 EPS data negative.

Table IV shows the time-series means of cross-sectional descriptive statistics.[5] Each panel shows means, standard deviations, minimums, percentiles of the variables, and maximums. Panel A shows the descriptives for the VAR state variables. A notable feature of the descriptive statistics is that firm-level log ROE is almost as variable as firm-level log returns (standard deviation of 0.20 vs. 0.27). Panels B, C, and D show the descriptives for Beaver, Clarke, and Wright's (1980), Easton and Harris's (1991), and Collins, Kothari, Shanken, and Sloan's (1994) cash-flow proxies, respectively.

B. Estimated processes for firm-level stock returns

Our firm-level VAR is designed to capture the following empirical return-predictability results. Historically, 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"). We avoid modeling corporate dividend policy by excluding any dividend-based variables from the VAR.

In estimating the VAR coefficient matrix, we concentrate on robustness and simplicity. We 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, we 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 results are not sensitive to the choice between OLS and WLS. We use Shao and Rao's (1993) jackknife method to calculate cross-correlation consistent standard errors.

First, we consider a parsimonious VAR specification that includes market-adjusted log stock return, log book-to-market, and log return on equity as the state variables. Only one lag of each is used to predict the state-vector evolution. The parameter estimates (presented in Table V) imply that expected market-adjusted 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 is 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 0.3 correlation).

The variance decomposition implied by the VAR is also shown in Table V, and cash-flow news is the main driver of market-adjusted firm-level stock returns. The expected-return-news standard deviation is 16 percent (variance 0.026 with 0.0074 standard error) and the cash-flow-news standard deviation is 36 percent (variance 0.13 with 0.012 standard error). The ratio of expected-return-news variance to total-unexpected-return variance is approximately 1/5. The correlation between the two estimated news series is 0.56 and more than five standard errors from zero.

The firm-level variance-decomposition results differ from the aggregate results in an interesting way. Consistent with the findings of Vuolteenaho (2002) and Gompers, Cohen, and Vuolteenaho (2003), our estimated VARs imply strong positive correlation between firm-level expected-return and cash-flow news. This positive correlation may in some cases have dramatic effects on firm-level regressions of returns on cash-flow proxies, as discussed below.

C. Reinterpreting regressions of returns on cash-flow proxies

We consider linear regressions with six sets of explanatory variables in our analysis of firm-level regressions. These selected specifications are motivated by classic earnings-returns association studies in the accounting literature. More recent accounting literature has explored more elaborate earnings-returns specifications. Hughes and Ricks (1987) and others study earnings announcements and proxy for unexpected earnings with analysts' forecast errors. Easton and Zmijewski (1989), Collins and Kothari (1989), Ball, Kothari, and Watts (1993), and others allow for cross-sectional and time-series variation in the earnings-response coefficients and relate it to economically motivated instruments. Lang (1991) studies the relation of the earnings-response coefficients to investor learning about the earnings process. Cheng, Hopwood, McKeown (1992), Freeman and Tse (1992), Das and Lev (1994), Beneish and Harvey (1998), and others relax the linearity assumption of the earnings-return relation. We leave the task of investigating the effect of expected returns, cash-flow news, and expected-return news on these more refined specifications to future research.

The explanatory variable in Table VI Panel A, GEPS, is the gross earnings per share (EPS) growth rate, EPS(t)/EPS(t-1). This explanatory variable is motivated by a study by Beaver, Clarke, and Wright (1980). Following Beaver, Clarke, and Wright, we exclude firm-years with EPS(t) or EPS(t-1) missing, as well as observations with negative EPS(t-1). To ensure that outliers do not dominate the regression results, we set EPS(t)/EPS(t-1) greater than exp(1) to exp(1) and EPS(t)/EPS(t-1) less than exp(-1) to exp(-1).

Consistent with the previous literature, the explanatory power of earnings growth is tiny with a pooled R^2 of about one percent. Some of the literature obtains slightly higher estimates, e.g., four percent in Collins, Kothari, Shanken, and Sloan's (1994) Table IV Panel C. The discrepancy between our results and theirs can be explained by the sample composition: Their sample includes only large, surviving firms, while our pooled data set is dominated by small firms, most of which disappear during our sample period. The explanatory power of earnings growth is also very limited in the component regressions: One-period expected returns zero percent, cash-flow news 1.8 percent, and minus expected-return news 0.10 percent.

The explanatory variable used in Table VI Panel B1 is the normalized level of earnings, motivated by Easton and Harris (1991). We drop firm years with either EPS(t) or EPS(t-1) missing are exclude outliers. Estimating the total-return regression equation in our sample yields a regression coefficient of 0.75 (s.e. 0.08) and R^2 of 10 percent. However, when we regress the estimated cash-flow news on the same earnings variable, we get a coefficient of 1.3 (s.e. 0.17) and R^2 of 27 percent. Although the level of earnings variable does a decent job tracking cash-flow news (a high R^2 of 27 percent), it also tracks the negative of expected-return news with a coefficient of -0.70 (s.e. 0.17) and R^2 of 19 percent and the level of expected returns with a coefficient of 0.15 (s.e. 0.05) and R^2 of 9 percent. In the end, the association of the earnings variable with cash-flow and expected-return news partially cancel each other, leaving the original specification with a lower coefficient and R^2. Including the normalized change in earnings (Panel B2) results in generally similar patterns. Since both cash-flow proxies negatively track the negative of expected-return news, the cash-flow-news R^2 of 31 percent is much higher than the total-return R^2 of 10 percent.

Panels C1-3 use variables similar to those of Kothari, Collins, Shanken, and Sloan (1994). Kothari, Collins, Shanken, and Sloan regress returns on subsets of the following variables: Contemporaneous and three leads of future log earnings growth rates, three leads of future returns, and the lagged earnings-price ratio. In our sample, specifications that include only earnings growth rates (Panel C1 and C2) yield R^2s from 2.1 to 4.4 percent in returns regressions and R^2s from 6.4 to 8.0 percent in cash-flow-news component regressions. Consistent with these results, the earnings growth coefficients from the cash-flow news regression are almost always larger than their total returns regression coefficient counterparts. This is a result of the negative relationship between earnings growth and minus expected returns.

Following Kothari and Shanken (1992), Kothari, Collins, Shanken, and Sloan (1994) include future returns and lagged earning-price ratio to "clean up" the measurement error due to the expected component of earnings growth. Adding these variables (Panel C3) increases the return-regression R^2s from 4 to 11 percent in our sample. The component regressions show, however, that a significant fraction of the improvement in the R^2s is due to correlation of these clean-up variables with one-period expected returns; the one-period expected-return regression R^2 increases from under one percent to 51 percent. Most of this increased explanatory powers is due to the lagged earning-price ratio (coefficient 0.51, s.e. 0.16). All else equal, firms with higher (lower) earning-price ratios have higher (lower) expected stock returns.

In sum, the firm-level regression results illustrate how the correlation between total returns and cash-flow proxies can understate or overstate the importance of cash-flow news as a source of return variance. In the case of Easton and Harris (1991), the importance of cash-flow news is understated. In contrast, the specification of Kothari, Collins, Shanken, and Sloan (1994) overstates the importance of cash-flow news.

IV. Conclusions

The Campbell-Shiller present-value formula enables one to divide stock returns into three components: one-period expected returns, changes in cash-flow expectations (i.e., cash-flow news), and changes in expected returns (i.e., expected-return news). Stock return volatility must originate from volatile one-period expected returns, cash-flow news, and/or expected-return news. Similarly, the correlation between stock returns and various cash-flow proxies must originate from three sources: association of cash-flow proxies with one-period expected returns, cash-flow news, and/or expected-return news.

Previous research has found that selected aggregate cash-flow proxies, such as contemporaneous and future dividens growth rates, explain a large fraction of aggregate stock return variance. Our empirical analysis shows that cash-flow proxies explain aggregate stock returns well (i.e., with a high R^2), because high realizations of cash-flow proxies are associated with declines in expected stock returns.

In contrast, at the individual stock level, earnings and other cash-flow proxies explain a relatively small fraction of annual stock returns. Our empirical results help explaining this perhaps surprising finding:

High realizations of cash-flow proxies are associated with increases in expected returns, depressing the explanatory power in regressions of stock returns on cash-flow proxies in disaggregate data. Although the level of earnings variable does a decent job tracking cash-flow news at the individual stock level, it also tracks the negative of expected-return news with an opposite sign. In the end, the association of the earnings variable with cash-flow and expected-return news partially cancel each other, leaving the original returns-on-earnings specification with a low R^2.

Appendix A: Campbell-Shiller log-linear approximation of return

This appendix represents the log-linear approximation of dividend-price ratio derived by Campbell and Shiller (1988, CS hereafter). The intuition behind this linearization is the following: If prices rise (decline) today, then there must be a rise (decline) in subsequent dividends and or a decline (rise) in subsequent returns. Otherwise, the price increase today has set off an infinitely-lived bubble. Because we rule such bubbles out by assumption, today's return must be justified by future dividends and/or returns.

Following CS, define log return r_t by

$$\begin{aligned} r_t &\equiv \log(P_t + D_t) - \log(P_{t-1}) \\ &= \log(1 + P_t / D_t) + \log(D_t / D_{t-1}) + \log(D_{t-1} / P_{t-1}) \end{aligned} \qquad \text{(A1)}$$

where P_t and D_t denote the period t ex-dividend price and dividend payment, respectively. Substituting the log dividend-price ratio, $\delta_t = \log(D_t) - \log(P_t)$, and log dividend growth rate, $\Delta d_t = \log(D_t) - \log(D_{t-1})$, into (1) yields

$$r_t = \log(\exp(-\delta_t) + 1) + \Delta d_t + \delta_{t-1} \qquad \text{(A2)}$$

Equation (A2) is the exact definition of a one-period log return. It can be approximated by a first-order Taylor-series expansion around $\hat{\delta}$:

$$r_t \approx k - \rho\delta_t + \Delta d_t + \delta_{t-1} \text{ , where} \qquad \text{(A3)}$$

$$k \equiv \log(\exp(-\hat{\delta}) + 1) + \rho\hat{\delta} \text{ and } \rho \equiv \frac{\exp(-\hat{\delta})}{\exp(-\hat{\delta}) + 1}$$

While any $\hat{\delta}$ can be used, the mean δ provides an accurate approximation if the variability of δ is small. Reasonable values lead to a discount-coefficient (ρ) close to 0.96. In the extreme case of constant dividend yield, the linearization becomes exact.

The one-period approximation can be iterated to yield a multi-period approximation. Solving for δ_t from (A3) and iterating forward $N-1$ times generates

$$\delta_t \approx k\frac{1-\rho^N}{\rho - 1} - \sum_{i=0}^{N-1} \rho^i \Delta d_{t+i+1} + \sum_{j=0}^{N-1} \rho^j r_{t+j+1} + \rho^N \delta_{t+N} . \qquad \text{(A4)}$$

Substituting (A4) back into (A3) gives a linearized return identity

$$r_t \approx k\frac{1-\rho^{N+1}}{1-\rho} + \delta_{t-1} + \sum_{i=0}^{N}\rho^i \Delta d_{t+i} - \sum_{j=1}^{N}\rho^j r_{t+j} - \rho^{N+1}\delta_{t+N} . \quad \text{(A5)}$$

If there are no infinitely-lived bubbles, the discounted future dividend yield converges to zero as N approaches infinity. Then, the log return is a linear function of current log dividend yield and discounted contemporaneous and future dividend growth rates and future returns:

$$r_t \approx \frac{k}{1-\rho} + \delta_{t-1} + \sum_{i=0}^{\infty}\rho^i \Delta d_{t+i} - \sum_{j=1}^{\infty}\rho^j r_{t+j} . \quad \text{(A6)}$$

All of the terms in (A6) are realized values. The above mathematics is just accounting. As stated by CS (p. 200), the above result "…*has been obtained only by the linear approximation… There is no economic content to* [(A6)]." The above return formulas hold for all assets: stocks, bonds, portfolio trading strategies, etc. As long as the dividend-price ratio does not explode, the dynamic accounting identity in (A6) holds.

Appendix B: Additional robustness checks

A. Aggregate stock returns

In order to gauge how sensitive our aggregate results are to the state vector used to generate the one-period expected return, cash-flow news, and expected-return news series, this appendix briefly summarizes our results from repeating the analysis of sections II.C, II.D, and II.E using eight different state vectors motivated by the return predictability literature. The tables we refer to in this appendix are not included in this published version to save space; however, they are available on the home pages of the authors as well as that of the NBER.

An unpublished table (Table BI) summarizes the results for alternative state vectors by reporting the adjusted R^2 from each of the three component regressions using the three sets of "cash-flow proxy" explanatory variables. Using the original state vector (LRET, LPE, and LTERM), the Fama-Schwert variables produced R^{2}'s of 67, 46, and 41 percent for one-period expected returns, cash-flow news, and expected-return news, respectively. Across all of the various state vectors, the mean R^2 is 52, 37, and 43 percent for one-period expected returns, cash-flow news, and expected-return news, respectively. For the KS "dividends only" variables, the original R^{2}'s (12, 24, and 26 percent) are close to the mean values across alternative state vectors (39, 10, and 7 percent). Similarly, for the KS "dividends only" and "clean-up" variables combined, the original R^{2}'s (36, 26, and 34 percent) are close to the mean values (40, 17, and 31 percent). In general, our results appear to be reasonably robust across various sets of return predictor variables suggested by the literature.

B. Firm-level stock returns

We also estimate a cross-section of VAR models allowing for firm-specific variation in the VAR parameters. Since we do not impose any forward-looking survival requirements, many of our firms have very short time series. Estimating separate VAR parameters for each firm using OLS or WLS is, therefore, infeasible.

Instead of computing firm-by-firm OLS or WLS estimates, we follow Swamy's (1970) random-coefficient-model (RCM) approach. RCM is based on the assumption that firms' VAR parameters are random variables drawn from a common distribution. The parameters of this common distribution, dubbed hyperparameters, are estimated with maximum likelihood. Given these hyperparameters, maximum-likelihood predictors of individual firms' VAR parameters can be computed using the prior-likelihood approach of Lee and Griffiths (1979). The prior-likelihood approach uses the estimated population distribution of coefficients as the prior information and uses the Bayes rule to compute posterior-mean predictions of the firm-specific coefficients.

An unpublished table (Table BII) shows the results obtained from these random-coefficient-model specifications. Independent of methodological specifics, our earlier results appear robust to possible firm-specific variation in the parameters.

An unpublished table (Table BIII) investigates the sensitivity of our results to variations in the firm-level VAR's lag length. Across the three alternative specifications, the results are very similar to the results for a first-order VAR. In addition, the firm-level results are robust to adding aggregate state variables to the VAR forecasting firm-level stock returns.

Appendix C: The danger in adding clean-up variables to the regression

In this Appendix, we use Campbell and Shiller's (CS hereafter) log-linear framework (refer to Appendix A) to illustrate one possible explanation of why KS's additional clean-up variables are prone to tracking one-period expected returns and expected-return news. Using lagged dividend yield and future realized returns to "clean-up" measurement error from realized dividend growth rates is one future dividend yield short of mapping out the CS realized return *identity*. With enough future dividend growth rates and returns, the future dividend yield becomes irrelevant, and KS' full regression mechanically produces a population R^2 that tends to 100 percent.

Compare the CS linearized return identity to the KS full regression:

$$r_t = k\frac{1-\rho^4}{1-\rho} + \delta_{t-1} + \sum_{i=0}^{3}\rho^i \Delta d_{t+i} - \sum_{j=1}^{3}\rho^j r_{t+j} - \rho^4\delta_{t+3} + \kappa_{3,t} \tag{C1}$$

$$r_t = B_0 + B_1\delta_{t-1} + \sum_{i=0}^{3} B_{i+2}\Delta d_{t+i} + \sum_{j=1}^{3} B_{j+5} r_{t+j} + \varepsilon_t \tag{C2}$$

r is log return, δ log dividend yield, and Δd log dividend growth. The first of the two equations, (C1), is just a first-order Taylor series approximation of the definition of log return iterated forward three times, where $\kappa_{3,t}$ denotes the linearization error. (C2) is very similar to KS's full regression. The only differences between the CS identity and the KS regression are the future log dividend-price ratio term, δ_{t+3}, linearization error, $\kappa_{3,t}$, and error term ε_t.

Our analysis (in Table CI of an unpublished appendix) shows that the linearization error is small. In regressions that include the future dividend yield, one minus the regression R^2 can be interpreted as the size of the linearization error. With one forward iteration, the CS linearization produces an R^2 of 99.8761 percent. Even after iterating forward 15 times, the compounded approximation error is extremely small – the regression still produces an R^2 of 99.5134 percent. It is, therefore, safe to assume that κ_t is constant and does not affect the regression R^2.

Even if the linearization error is trivially small, KS's full regression is not a perfect identity, because the future dividend-yield term is omitted. The size of this omitted variable error depends on the number of future returns and dividend growth rates included in the regression. As we add leads to the regression but omit the future dividend yield, the adjusted R^2 increases steadily from 39.95 percent with one lead up to 62.66 percent with fifteen leads. Analytical arguments suggest that in a large sample with covariance stationary dividend yield, the omitted variable error converges to zero as more terms are included.

Taking the 15-lead regression with 105 data points as an example, the first seven future dividend growth and the first seven future return terms are significant at the 5 percent level. Furthermore, the dividend coefficients and future return coefficients tend to decay at an exponential rate, as predicted by the CS linearization. Although the empirical work of KS (1992) focuses on the parsimonious three-lead model, KS's stated theoretical motivation for the "clean-up variables" does not constrain the number of leads in their regression. Thus, if KS's stated theoretical motivation is taken literally (and perhaps out of context, since KS never include more than three leads), the return in the 15-lead regression is caused by significant revisions of long-term expected future dividend growth rates. In our opinion, this interpretation is implausible, considering the difficulties in predicting dividend growth rates one year ahead with a time-series model (see, e.g., Cochrane (1994)). We see this result as a manifestation of the linearized identity, which helps to possibly explain why explanatory variables similar to KS's "clean-up" variables are prone to picking up expected-return effects, such as those documented in Table III Panel B2.

To further explore regressions with only a small number of leads, we calibrate various realistic economies and examine the large-sample properties of these regressions with simulated data. Specifically, we estimate a number of return-generating processes with a conditional maximum-likelihood method, while constraining the ratio of expected-return-news variance to cash-flow-news variance to values ranging from 1/5 to 5. As a result, we have a set of calibrated simulation economies in which the significance of cash-flow news as a driver of stock returns varies by construction. Note that these economies are not otherwise perverse, since the constrained maximum-likelihood procedure chooses parameters to best match the data while satisfying the constraint.

These simulation results are summarized in the unpublished Figure C1. The x-axis of the graph varies the $\text{var}(N_r)/\text{var}(N_{cf})$ ratio implied by the calibrated VAR model. For each $\text{var}(N_r)/\text{var}(N_{cf})$ value, we estimate the VAR model parameters from the data with constrained maximum likelihood, constraining $\text{var}(N_r)/\text{var}(N_{cf})$ to the particular pre-specified value. The $\text{var}(N_r)/\text{var}(N_{cf})$ values range from 0.2 to 5. In addition, we constrain the model to be stationary by limiting the maximum absolute eigenvalue of the transition matrix. After this calibration exercise, we have 31 points in the VAR parameter space, each corresponding to a single simulation economy.

In our simulations, we concentrate on KS's dividends-only and full specifications, because the required variables can be conveniently computed from the VAR data. The figure plots five population R^2s as functions of the ratio of expected-return-news variance to cash-flow-news variance. First, we plot KS's full regression R^2. The regression R^2s (marked with plus signs) flat-line at around 40 to 45 percent. Second, we plot KS's dividends-only specification R^2. Again, the R^2 is roughly constant at 23 percent (the solid line). This clearly illustrates that it can be erroneous to interpret either KS R^2 as measuring the relative importance of cash-flow news as a driver of realized stock return variation. Both R^2s are insensitive to the importance of cash-flow news. These R^2s can also be misleading as measures of a regression R^2 of returns on cash-flow news. The dashed line plots the R^2 of returns on cash-flow news in these economies. Unlike the KS R^2s, these R^2s do not flat-line. Instead, the dashed line slopes down as the importance of cash-flow news decreases.

The poor performance of both KS R^2s as measures of the importance of cash-flow news may be due to a nonzero correlation between the cash-flow and expected-return news process. Even if the KS explanatory variables were tracking the cash-flow-news series, a nonzero correlation between the two news series would bias both KS R^2s. KS briefly discussed this point. To examine this possibility, we repeat our simulation procedure by imposing an additional constraint corr(N_r, N_{cf})=0. The results (which correspond to the unpublished Figure C2) suggest that none of the above conclusions change because of this additional constraint, highlighting a deeper problem for KS-type regression specifications.

From the general perspective of any returns on cash-flow proxies regression, some of our calibration results may be somewhat surprising. For example, how can the returns-on-dividend-growth R^2 be so completely disconnect from the returns-on-cash-flow-news R^2, even when corr(N_r, N_{cf})=0? We conjecture that the data can simultaneously be very informative about the correlation of dividend growth and stock return and not very informative about the significance of cash-flow news as a driver of stock returns. We suspect that this phenomenon roots from the fact that the cash-flow news depend on not only the contemporaneous correlation of returns and dividend growth but also on the infinite-horizon predictability of dividend growth. The latter is estimated imprecisely from the data and, therefore, the maximum-likelihood calibration method matches the var(N_r)/var(N_{cf}) constraint by altering the long-horizon predictability of dividend growth. More specifically, when the var(N_r)/var(N_{cf}) constraint is high (low), positive cash-flow shocks lead to downward (upward) revisions in future cash-flow growth expectations. This interaction between the cash-flow shock and the revisions in future cash-flow growth expectations disconnects the returns-on-dividend-growth R^2 from the returns-on-cash-flow-news R^2.

On the positive side, one of KS's diagnostics regression behaves as expected in both figures. As the importance of expected-return news increases, the large-sample R^2 of returns on clean-up variables only increases steadily. However, the large-sample properties of these regressions do not answer the question of whether KS's diagnostics regression is useful. To examine the "power" of KS's diagnostics regression, on top of the large-sample R^2s in the unpublished figures, we indicate the fraction of F-tests that reject at the 5%

level in samples of 71 observations. The rejection probability increases with the population R^2s, but even for large R^2s, KS's pre-test diagnostics is hardly perfect. In the calibrated economies in which the variance of expected-return news is higher than that of cash-flow news, the pre-test rejects only approximately 40 to 25 percent of the samples depending on whether the constraint is imposed on the news correlation. Nevertheless, these small-sample experiments suggest that KS's diagnostic helps in ruling out the samples that would lead to the most misleading inferences.

References

Ball, R. and P. Brown, 1968. An empirical investigation of accounting income numbers, *Journal of Accounting Research* 6, 159-178.

Ball, R, S. P. Kothari and Ross L. Watts, 1993. Economic determinants of the relation between earnings changes and stock returns, *The Accounting Review* 68, 622-638.

Beaver, W., R. Clarke, and W. Wright, 1979. The association between unsystematic security returns and the magnitude of earnings forecast errors, *Journal of Accounting Research* 17, 316-340.

Beaver, W., R. Lambert and D. Morse, 1980. The information content of security prices, *Journal of Accounting and Economics* 2, 3-28.

Beneish, M. D., and C. R. Harvey, 1998. Measurement error and nonlinearity in the earnings-returns relation of large firms, *Review of Quantitative Finance and Accounting* 11, 219-247.

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., Ammer J., 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., Mei J., 1993. Where do betas come from? Asset price dynamics and the sources of systematic risk, *Review of Financial Studies* 6, 567-592.

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

Cheng, C. S., William S. Hopwood, and James C. McKeown, 1992. Non-linearity and specification problems in unexpected earnings response regression model, *The Accounting Review* 67, 579-598.

Cochrane, J. H., 1994. Permanent and transitory components of GNP and stock prices, *Quarterly Journal of Economics* 109, 241-265.

Cohen, R., P. Gompers, and T. Vuolteenaho, 2003. Who underreacts to cash-flow news? Evidence from trading between individuals and institutions, *Journal of Financial Economics* 66, 409-462.

Collins, D. and S.P. Kothari, 1989. An analysis of intertemporal and cross-sectional determinants of earnings response coefficients, *Journal of Accounting and Economics* 11, 143-181.

Collins, D., S. P. Kothari, J. Shanken and G. Sloan, 1994. Lack of timeliness versus noise as explanations for low contemporaneous return-earnings association, *Journal of Accounting and Economics* 18, 289-324.

Das, Somnath and Baruch Lev, 1994. Nonlinearity in the returns-earnings relation: tests of alternative specifications and explanation, *Contemporary Accounting Research*, vol. 11, no. 1-II, 353-379.

DeBondt, Werner F. M. and Richard H. Thaler, 1985. Does the stock market overreact?, *Journal of Finance* 40, 557-581.

Easton, P. and T. Harris, 1991. Earnings as an explanatory variable for returns, *Journal of Accounting Research* 29, 19-36.

Easton, P. and Mark E. Zmijewski, 1989. Cross-sectional variation in the stock market response to accounting earnings announcements, *Journal of Accounting and Economics* 11, 117-141.

Fama, E. F., 1990. Stock returns, expected returns and real activity, *Journal of Finance* 45, 1089-1108.

Fama*, E. F., French, K. R., 1989.* Business conditions and expected returns on stocks and bonds*, Journal of Financial Economics* 25, 23-49.

Fama, Eugene F. and James MacBeth, 1973. Risk, return and equilibrium: empirical tests, *Journal of Political Economy* 81, 607-636.

Freeman, Robert N. and Senyo Y. Tse, 1992. A nonlinear model of security price responses to unexpected earnings, *Journal of Accounting Research* 30, 185-209.

Haugen, Robert A., and Nardin L. Baker, 1996. Commonality in the determinants of expected stock returns, *Journal of Financial Economics* 41, 401-439.

Hughes, J. and William E. Ricks, 1987. Associations between forecast errors and excess returns near to earnings announcements, *The Accounting Review* LXII, 158-175.

Jegadeesh, Narasimhan, and Sheridan Titman, 1993. Returns to buying winners and selling losers: implications for stock market efficiency, *Journal of Finance* 48, 65-91.

Keim, D. B., Stambaugh, R. F., 1986. Predicting returns in the stock and bond markets, *Journal of Financial Economics* 17, 357-390.

Kothari, S. P., Shanken, J., 1992. Stock return variation and expected dividends: a time series and cross-sectional analysis, *Journal of Financial Economics* 31, 177-210.

Lang, M., 1991. Time-varying stock price response to earnings induced by uncertainty about the time-series process of earnings, *Journal of Accounting Research* 29, 229-257.

Lee, L. and W. Griffiths, 1979. The Price likelihood and best linear unbiased prediction in stochastic coefficient linear models, University of New England Working Papers in Econometrics and Applied Statistics No. 1.

Lev, B., 1989. On the usefulness of earnings and earnings research: lessons and directions from two decades of empirical research, *Journal of Accounting Research* 27, 153-202.

Liew, J., 1995. Global stock returns, money, and inflation, Unpublished dissertation, University of Chicago, Graduate School of Business.

Rosenberg, Barr, Kenneth Reid, and Ronald Lanstein, 1985. Persuasive evidence of market inefficiency, *Journal of Portfolio Management* 11, 9-17.

Schwert, G. W., 1990. Stock returns and real activity: a century of evidence, *Journal of Finance* 45, 1237-1257.

Shao, Jun, and J. N. K. Rao, 1993. Jackknife inference for heteroscedastic linear regression models, *Canadian Journal of Statistics* 21, 377-385.

Swamy, P. A. V. B., 1970. Efficient inference in a random coefficient regression model, *Econometrica* 38, 311-323.

Vuolteenaho, Tuomo, 2001. Understanding the aggregate book-to-market ratio and its implications to current equity-premium expectations, working paper, Harvard University Department of Economics.

Vuolteenaho, Tuomo, 2002. What drives firm-level stock returns?, *Journal of Finance* 57, 233-264.

Table I: Descriptive statistics of the aggregate data

Panel A shows descriptives of the VAR state variables. LRET is the annual log stock return on the Schwert (1881-1925) or CRSP value-weight index (1926-present), LPE is the log price-earnings ratio, and LTERM is the spread of the ten-year constant maturity bond log yield over the three-month commercial paper log yield. All relevant variables are deflated by the commercial paper/T-bill wealth index. Sample period is 1881-1998.

Panels B shows descriptives of the explanatory variables in aggregate regressions of returns on cash-flow and expected return proxies. The variables in Panel B are motivated by Fama's (1990), Schwert's (1990), and Kothari and Shanken's (1992) cash-flow and expected return proxies. Expected return proxies are the log dividend-price ratio (LDP, also considered a cash-flow proxy), the spread of Moody's Baa corporate bond log yield over the Aaa corporate bond log yield (LDEF), and the spread of the ten-year constant maturity bond log yield over the three-month commercial paper log yield (LTERM). From 1881-1918, LDEF is the spread of the average top five Macaulay railroad bond log yields (excluding the maximum) over the average bottom five Macaulay railroad bond log yields (excluding the minimum). Expected-return news proxies are shocks to the default (LDEFS) and term (LTERMS) spread. The shocks are the estimated residuals from a first-order autoregressive process. Cash-flow proxies are current and future dividend growth rates and future returns. LGD, LGDF1, LGDF2, and LGDF3 are the current, 1-year, 2-year, and 3-year leads of the annual log dividend growth rate. LRETF1, LRETF2, and LRETF3 are the 1-year, 2-year, and 3-year leads of LRET. All relevant variables are deflated by the commercial paper/T-bill wealth index. Sample period is 1881-2001.

Panel A: VAR state variables

VARIABLE:	mean	std.dev.	min	25%-tile	median	75%-tile	max
LRET	0.0474	0.1888	-0.6151	-0.0638	0.0567	0.1866	0.4388
LPE	2.7581	0.3105	1.8478	2.5875	2.7072	3.0024	3.6235
LTERM	-0.0022	0.0175	-0.0626	-0.0131	-0.0001	0.0115	0.0310

Panel B: Fama's (1990), Schwert's (1990), and Kothari and Shanken's (1992) variables

VARIABLE:	mean	std.dev.	min	25%-tile	median	75%-tile	max
LDP	-3.1196	0.2866	-4.0785	-3.2810	-3.1301	-2.9403	-2.4188
LDEF	0.0103	0.0066	0.0030	0.0063	0.0077	0.0124	0.0441
LTERM	-0.0022	0.0175	-0.0626	-0.0131	-0.0001	0.0115	0.0310
LDEFS	-0.0000	0.0037	-0.0077	-0.0017	-0.0006	0.0004	0.0247
LTERMS	0.0003	0.0152	-0.0700	-0.0061	0.0011	0.0087	0.0537
LGD	-0.0073	0.1631	-0.6646	-0.0811	-0.0022	0.0637	0.6706
LGDF1	-0.0072	0.1632	-0.6646	-0.0811	-0.0022	0.0637	0.6706
LGDF2	-0.0097	0.1661	-0.6646	-0.0849	-0.0022	0.0637	0.6706
LGDF3	-0.0088	0.1660	-0.6646	-0.0849	-0.0011	0.0637	0.6706
LRETF1	0.0491	0.1891	-0.6151	-0.0638	0.0787	0.1866	0.4388
LRETF2	0.0484	0.1898	-0.6151	-0.0638	0.0787	0.1866	0.4388
LRETF3	0.0486	0.1895	-0.6151	-0.0638	0.0787	0.1866	0.4388

Table II: Aggregate VAR

This table shows the parameter estimates and some implied properties of a VAR(1) specification with the following variables: LRET is the annual log stock return on the Schwert (1881-1925) or CRSP value-weight index (1926-present), LPE is the log price-earnings ratio, and LTERM is the spread of the ten-year constant maturity bond log yield over the three-month commercial paper log yield. All relevant variables are deflated by the commercial paper/T-bill wealth index. Sample period is 1881-1998. The VAR specification has the structure

$$z_t = \mathrm{A} + \Gamma z_{t-1} + u_t, \quad \Sigma = E(u_t u_t')$$

In each cell, the first number is an OLS point estimate. The second number (in brackets) is a bootstrapped standard error.

Coefficient estimates for the first order VAR

	A	Γ			Σ		
LRET	0.4842	0.0641	-0.1575	2.4005	0.0323	0.0302	0.0003
	[0.1573]	[0.0909]	[0.0561]	[1.0560]	[0.0050]	[0.0047]	[0.0002]
LPE	0.4253	0.1185	0.8483	1.7012	0.0302	0.0314	0.0003
	[0.1562]	[0.0894]	[0.0557]	[1.0464]	[0.0047]	[0.0048]	[0.0002]
LTERM	0.0105	-0.0139	-0.0039	0.5073	0.0003	0.0003	0.0002
	[0.0132]	[0.0078]	[0.0047]	[0.0815]	[0.0002]	[0.0002]	[0.0001]

Covariances and correlations (shaded)	One-period expected return	Cash-flow news	Expected-return news
One-period expected return	0.0042	0.0000	0.0000
	[0.0022]	[0.0000]	[0.0000]
Cash-flow news	0.0000	0.0041	-0.0027
	[0.0000]	[0.0012]	[0.0019]
Expected-return news	0.0000	-0.2875	0.0219
	[0.0000]	[0.2028]	[0.0051]

Table III: Decomposing aggregate regressions of returns on cash-flow and expected-return proxies

This table decomposes aggregate regressions of returns (r) on cash-flow and expected return proxies into three components: (1) the regression coefficient of one-period expected return (Er) on cash-flow and expected return proxies, (2) the regression coefficient of cash-flow news (Ncf) on cash-flow and expected return proxies, and (3) the regression coefficient of minus expected-return news (-Nr) on cash-flow and expected return proxies. The table shows estimated regression coefficients and R^2s.

One-period expected returns, cash-flow news, and expected-return news are computed from the VAR model in Table II. In each cell, the first number is an OLS point estimate. The second number (in brackets) is a bootstrapped standard error. The standard errors account for the fact the dependent variables are computed from an estimated VAR model.

The explanatory variables in Panel A are motivated by Fama (1990) and Schwert's (1990) cash-flow and expected return proxies. Expected return proxies are the spread of Moody's Baa corporate bond log yield over the Aaa corporate bond log yield (LDEF) and the spread of the ten-year constant maturity bond log yield over the three-month commercial paper log yield (LTERM). From 1881-1918, LDEF is the spread of the average top five Macaulay railroad bond log yields (excluding the maximum) over the average bottom five Macaulay railroad bond log yields (excluding the minimum). Cash-flow proxies are the current, 1-year, 2-year, and 3-year leads of the annual log dividend growth rate (LGD, LGDF1, LGDF2, and LGDF3). Expected-return news proxies are shocks to the term (LTERMS) and default (LDEFS) spread. The shocks are the estimated residuals from a first-order autoregressive process. All relevant variables are deflated by the commercial paper/T-bill wealth index. Sample period for the dependent variable is 1882-1998.

The explanatory variables in Panels B1 and B2 are motivated by Kothari and Shanken's (1992) cash-flow proxies. LDP is the log dividend-price ratio. LGD, LGDF1, LGDF2, and LGDF3 are the current, 1-year, 2-year, and 3-year leads of the annual log dividend growth rate. LRETF1, LRETF2, and LRETF3 are the 1-year, 2-year, and 3-year leads of LRET (return). All relevant variables are deflated by the commercial paper/T-bill wealth index. Sample period for the dependent variable is 1882-1998.

Panel A: Fama (1990) and Schwert's (1990) variables

	const.	LDEF	LTERM	LDEFS	LTERMS	LGD	LGDF1	LGDF2	LGDF3	% R^2
r	0.0462	0.6954	1.1748	-20.7046	0.5613	0.5138	-0.0957	-0.0025	0.0732	53.18
	[0.0285]	[2.4631]	[0.7295]	[4.4417]	[0.7983]	[0.0886]	[0.0833]	[0.0776]	[0.0771]	[7.49]
Er	0.0064	4.5169	1.9816	-0.7997	-0.4822	0.0425	0.0299	0.0358	0.0279	66.64
	[0.0261]	[1.7855]	[1.0443]	[2.2219]	[0.4901]	[0.0417]	[0.0355]	[0.0327]	[0.0301]	[17.01]
Ncf	0.0238	-2.2281	-0.2359	-2.7035	1.9213	0.1608	0.0289	0.0339	0.0286	46.07
	[0.0098]	[0.8680]	[0.2667]	[2.6044]	[0.8347]	[0.0497]	[0.0334]	[0.0297]	[0.0286]	[15.73]
-Nr	0.0160	-1.5934	-0.5709	-17.2014	-0.8777	0.3105	-0.1546	-0.0722	0.0167	41.08
	[0.0205]	[1.9572]	[0.5738]	[4.0903]	[1.0739]	[0.0728]	[0.0831]	[0.0704]	[0.0670]	[8.14]

Panel B1: Kothari and Shanken's (1992) variables

	const.	LGD	LGDF1	LGDF2	LGDF3	% R^2
r	0.0548	0.7278	-0.0119	0.0945	0.1367	39.21
	[0.0083]	[0.1156]	[0.0935]	[0.0911]	[0.0892]	[6.82]
Er	0.0502	0.0828	0.0632	0.0932	0.0948	12.19
	[0.0084]	[0.0551]	[0.0449]	[0.0440]	[0.0408]	[5.27]
Ncf	0.0025	0.1963	0.0574	0.0403	0.0272	23.81
	[0.0018]	[0.0693]	[0.0374]	[0.0346]	[0.0338]	[9.96]
-Nr	0.0021	0.4488	-0.1325	-0.0390	0.0146	25.76
	[0.0026]	[0.0915]	[0.0870]	[0.0768]	[0.0744]	[6.43]

Panel B2: Kothari and Shanken's (1992) variables

	const.	LDP	LGD	LGDF1	LGDF2	LGDF3	LRETF1	LRETF2	LRETF3	% R^2
r	0.4664	0.1238	0.7387	0.0544	0.2901	0.2214	-0.0528	-0.2535	-0.1535	45.25
	[0.2217]	[0.0670]	[0.1116]	[0.1272]	[0.1142]	[0.1120]	[0.1097]	[0.0906]	[0.0923]	[7.04]
Er	0.4049	0.1143	0.0978	0.0882	0.0755	0.0546	-0.0065	0.0089	0.0286	36.25
	[0.1314]	[0.0402]	[0.0493]	[0.0497]	[0.0474]	[0.0442]	[0.0438]	[0.0384]	[0.0384]	[13.93]
Ncf	0.0802	0.0256	0.1963	0.0094	0.0441	0.0381	0.0651	-0.0000	-0.0285	26.34
	[0.0570]	[0.0168]	[0.0709]	[0.0471]	[0.0477]	[0.0462]	[0.0425]	[0.0406]	[0.0398]	[10.25]
-Nr	-0.0188	-0.0161	0.4447	-0.0431	0.1705	0.1287	-0.1114	-0.2624	-0.1537	33.88
	[0.1120]	[0.0326]	[0.0918]	[0.1093]	[0.0966]	[0.0956]	[0.0989]	[0.0811]	[0.0817]	[7.24]

Table IV: Descriptive statistics of the firm-level data

This table presents the time-series averages of selected cross-sectional descriptive statistics. First, the descriptive statistics are computed for each cross-section, and then the time-series of descriptive statistics are averaged. Sample period is 1954-1999.

The variables are defined as follows. Panel A shows descriptives of the VAR state variables. LRET is the log stock return, LBM log book-to-market equity, and LROE log accounting return on equity. Variables in Panel A have been transformed by unlevering the firm 10 percent (see text for details).

Panels B, C, and D show descriptives of the explanatory variables in firm-level regressions of returns on cash-flow proxies. The variable in Panel B is motivated by Beaver, Clarke, and Wright's (1980) cash-flow proxy. GEPS is the gross earnings per share (EPS) growth rate, EPS(*t*)/EPS(*t*-1). Observations with EPS(*t*) or EPS(*t*-1) missing are excluded, as well as observations with EPS(*t*-1)<1. EPS(*t*)/EPS(*t*-1) greater than exp(1) are set to exp(1) and EPS(*t*)/EPS(*t*-1) less than exp(-1) to exp(-1).

Panel C shows descriptives for variables similar to Easton and Harris's (1991) cash-flow proxies. EPL is EPS(*t*) normalized by P(*t*-1). DEPL is the change in earnings, EPS(*t*)-EPS(*t*-1), normalized by price, P(*t*-1). Observations with either EPS(*t*) or EPS(*t*-1) missing are excluded. Observations with EPS(*t*)/P(*t*-1), EPS(*t*-1)/P(*t*-1), or (EPS(*t*)-EPS(*t*-1))/P(*t*-1) outside the range [-1.5,1.5] are excluded.

Panel D uses variables similar to those of Kothari, Collins, Shanken, and Sloan (1994). LGEPS is the log EPS growth rate, ln(EPS(*t*)/EPS(*t*-1)). LGEPSF1, LGEPSF2, and LGEPSF3 are 1-year, 2-year, and 3-year leads of LGEPS. LRET is the log return. LRETF1, LRETF2, and LRETF3 are 1-year, 2-year, and 3-year leads of LRET. ELPL is EPS(*t*-1)/P(*t*-1). Observations of LGEPS, LGEPSF1, LGEPSF2, LGEPSF3, LRET, LRETF1, LRETF2, and LRETF3 greater than 1 are set to 1, and observations less than -1 are set to -1. Observations with *t*-1, *t*, *t*+1, *t*+2, or *t*+3 return missing or EPS negative or missing are excluded.

Panel A: VAR state variables, 42075 firm years

VARIABLE:	mean	std.dev.	min	25%-tile	median	75%-tile	max
LRET	0.0848	0.2651	-1.0775	-0.0590	0.0873	0.2307	1.2356
LBM	-0.2627	0.5553	-2.3873	-0.5757	-0.2184	0.0841	2.1378
LROE	0.0751	0.1952	-1.5908	0.0628	0.1024	0.1391	0.9249

Panel B: Beaver, Clarke, and Wright's (1980) variable, 36237 firm years

VARIABLE:	mean	std.dev.	min	25%-tile	median	75%-tile	max
GEPS:	1.0631	0.4670	0.3678	0.7870	1.0342	1.2147	2.6782

Panel C: Easton and Harris's (1991) variables, 41279 firm years

VARIABLE:	mean	std.dev.	min	25%-tile	median	75%-tile	max
EPL	0.0649	0.0969	-0.8256	0.0469	0.0784	0.1063	0.5156
DEPL	-0.0019	0.0984	-0.8149	-0.0194	0.0032	0.0195	0.8429

Panel D: Collins, Kothari, Shanken, and Sloan's (1994) variables, 23225 firm years

VARIABLE:	mean	std.dev.	min	25%-tile	median	75%-tile	max
LGEPS	0.0134	0.3744	-0.9945	-0.1426	0.0538	0.1905	0.9791
LGEPSF1	0.0078	0.3740	-0.9952	-0.1488	0.0538	0.1870	0.9660
LGEPSF2	-0.0025	0.3769	-1.0000	-0.1631	0.0469	0.1793	0.9676
LGEPSF3	-0.0113	0.3848	-1.0000	-0.1813	0.0420	0.1781	0.9702
LRETF1	0.1193	0.2260	-0.6687	-0.0159	0.1134	0.2489	0.8740
LRETF2	0.1094	0.2267	-0.6738	-0.0248	0.1058	0.2410	0.8625
LRETF3	0.1011	0.2356	-0.7665	-0.0368	0.0996	0.2380	0.8681
ELPL	0.0879	0.0414	0.0064	0.0634	0.0851	0.1068	0.4487

Table V: Firm-level VAR

This table shows a firm-level VAR specification parameter estimates and some implied properties of the model. The state vector includes log stock return (LRET), log book-to-market ratio (LBM), and log ROE (LROE) as state variables. All variables are market-adjusted (i.e., cross-sectionally demeaned.) The VAR has the structure

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

In each cell, the first number is a pooled-OLS point estimate. The second number (in brackets) is a cross-correlation-consistent robust jackknife standard error computed using the jackknife method of Shao and Rao (1993).

Coefficient estimates for the first order VAR, market-adjusted data

	Γ			Σ		
LRET	0.1376 [0.0202]	0.0507 [0.0141]	0.0891 [0.0314]	0.0917 [0.0062]	-0.0711 [0.0049]	0.0212 [0.0021]
LBM	0.1470 [0.0297]	0.8691 [0.0178]	0.1016 [0.0471]	-0.0711 [0.0049]	0.1126 [0.0152]	0.0172 [0.0021]
LROE	0.1858 [0.0114]	-0.0055 [0.0060]	0.6042 [0.0533]	0.0212 [0.0021]	0.0172 [0.0021]	0.0594 [0.0059]

Covariances and correlations (shaded)	One-period expected return	Cash-flow news	Expected-return news
One-period expected return	0.0044 [0.0011]	0 [0]	0 [0]
Cash-flow news	0 [0]	0.1315 [0.0121]	0.0330 [0.0068]
Expected-return news	0 [0]	0.5620 [0.1075]	0.0262 [0.0074]

Table VI: Decomposing firm-level regressions of returns on cash-flow proxies

This table decomposes firm-level regressions of returns (r) on cash-flow proxies into three components: (1) the regression coefficient of one-period expected return (Er) on cash-flow proxies, (2) the regression coefficient of cash-flow news (Ncf) on cash-flow proxies, and (3) the regression coefficient of minus expected-return news (-Nr) on cash-flow proxies. The table shows estimated regression coefficients and R^2s.

One-period expected returns, cash-flow news, and expected-return news are computed from the VAR model in Table V. Regressions are estimated by pooling the sample. In each cell, the first number is a pooled-OLS point estimate. The second number (in brackets) is a cross-correlation-consistent robust jackknife standard error computed using the jackknife method of Shao and Rao (1993). The standard errors account for the fact the dependent variables are computed from an estimated VAR model.

The cash-flow proxies are defined as follows. The variable in Panel A, GEPS, is the gross earnings per share (EPS) growth rate, EPS(t)/EPS(t-1). Observations with EPS(t) or EPS(t-1) missing are excluded, as well as observations with EPS(t-1)<1. EPS(t)/EPS(t-1) greater than exp(1) are set to exp(1) and EPS(t)/EPS(t-1) less than exp(-1) to exp(-1).

The variables in Panels B1 and B2 are the normalized level and change of earnings. EPL is EPS(t) normalized by P(t-1). DEPL is the change in earnings, EPS(t)-EPS(t-1), normalized by price, P(t-1). Observations with either EPS(t) or EPS(t-1) missing are excluded. Observations with EPS(t)/P(t-1), EPS(t-1)/P(t-1), or (EPS(t)-EPS(t-1))/P(t-1) outside the range [-1.5,1.5] are excluded.

Panels C1-3 use variables similar to those of Kothari, Collins, Shanken, and Sloan (1994). LGEPS is the log EPS growth rate, ln(EPS(t)/EPS(t-1)). LGEPSF1, LGEPSF2, and LGEPSF3 are 1-year, 2-year, and 3-year leads of LGEPS. LRET is the log return. LRETF1, LRETF2, and LRETF3 are 1-year, 2-year, and 3-year leads of LRET. ELPL is EPS(t-1)/P(t-1). Observations of LGEPS, LGEPSF1, LGEPSF2, LGEPSF3, LRET, LRETF1, LRETF2, and LRETF3 greater than 1 are set to 1, and observations less than -1 are set to -1. Observations with t-1, t, t+1, t+2, or t+3 return missing or EPS negative or missing are excluded.

Panel A: Beaver, Clarke, and Wright's (1980) variable, 36237 firm years

	const.	GEPS	% R^2
r	0.0773	0.0302	1.03
	[0.0258]	[0.0036]	[0.26]
Er	0.1096	0.0009	0.03
	[0.0333]	[0.0014]	[0.12]
Ncf	-0.0049	0.0341	1.79
	[0.0225]	[0.0041]	[0.37]
-Nr	-0.0275	-0.0048	0.10
	[0.0161]	[0.0029]	[0.11]

Panel B1: Easton and Harris's (1991) variables, 41279 firm years

	const.	EPL	% R^2
r	0.0375	0.7521	9.61
	[0.0261]	[0.0784]	[1.88]
Er	0.0959	0.1466	8.74
	[0.0312]	[0.0475]	[3.61]
Ncf	-0.0927	1.3015	27.28
	[0.0308]	[0.1720]	[4.10]
-Nr	0.0343	-0.6960	19.48
	[0.0118]	[0.1690]	[3.18]

Panel B2: Easton and Harris's (1991) variables, 41279 firm years

	const.	EPL	DEPL	% R^2
r	0.0417	0.6850	0.1233	9.78
	[0.0263]	[0.1101]	[0.0775]	[1.74]
Er	0.0862	0.2987	-0.2797	29.27
	[0.0297]	[0.0852]	[0.0821]	[4.22]
Ncf	-0.0714	0.9659	0.6170	31.24
	[0.0284]	[0.1259]	[0.1357]	[4.61]
-Nr	0.0270	-0.5796	-0.2140	20.67
	[0.0114]	[0.1386]	[0.1041]	[3.81]

Panel C1: Collins, Kothari, Shanken, and Sloan's (1994) variables, 23225 firm years

	const.	LGEPS	% R^2
r	0.1240	0.0742	2.11
	[0.0231]	[0.0115]	[0.79]
Er	0.1168	-0.0050	0.41
	[0.0344]	[0.0048]	[0.72]
Ncf	0.0314	0.1018	6.36
	[0.0222]	[0.0120]	[1.74]
-Nr	-0.0242	-0.0226	0.89
	[0.0150]	[0.0098]	[0.75]

Panel C2: Collins, Kothari, Shanken, and Sloan's (1994) variables, 23225 firm years

	const.	LGEPS	LGEPSF1	LGEPSF2	LGEPSF3	% R^2
r	0.1228	0.0761	0.0379	-0.0482	-0.0408	4.35
	[0.0232]	[0.0138]	[0.0171]	[0.0117]	[0.0123]	[1.34]
Er	0.1168	-0.0051	-0.0003	-0.0003	-0.0008	0.42
	[0.0344]	[0.0040]	[0.0042]	[0.0040]	[0.0029]	[0.65]
Ncf	0.0307	0.1005	0.0178	-0.0393	-0.0303	8.06
	[0.0222]	[0.0125]	[0.0126]	[0.0084]	[0.0080]	[1.94]
-Nr	-0.0247	-0.0193	0.0204	-0.0086	-0.0097	2.04
	[0.0152]	[0.0102]	[0.0079]	[0.0058]	[0.0063]	[0.63]

Panel C3: Collins, Kothari, Shanken, and Sloan's (1994) variables, 23225 firm years

	const.	LGEPS	LGEPSF1	LGEPSF2	LGEPSF3	LRETF1	LRETF2	LRETF3	ELPL	% R^2
r	0.0205	0.1273	0.0719	-0.0243	-0.0288	-0.0898	-0.0029	-0.0442	1.2425	11.16
	[0.0326]	[0.0124]	[0.0191]	[0.0125]	[0.0155]	[0.0622]	[0.0475]	[0.0543]	[0.2341]	[2.61]
Er	0.0666	0.0140	0.0091	0.0063	0.0033	0.0034	0.0023	0.0034	0.5092	51.03
	[0.0270]	[0.0062]	[0.0062]	[0.0053]	[0.0034]	[0.0051]	[0.0034]	[0.0033]	[0.1642]	[9.40]
Ncf	-0.0341	0.1323	0.0389	-0.0252	-0.0236	-0.0597	0.0077	-0.0272	0.7773	12.51
	[0.0209]	[0.0123]	[0.0152]	[0.0098]	[0.0104]	[0.0418]	[0.0353]	[0.0354]	[0.1923]	[2.07]
-Nr	-0.0120	-0.0191	0.0238	-0.0054	-0.0085	-0.0334	-0.0129	-0.0204	-0.0440	3.04
	[0.0127]	[0.0115]	[0.0099]	[0.0054]	[0.0070]	[0.0258]	[0.0161]	[0.0206]	[0.1122]	[1.06]

[1] This approach is similar in spirit to Campbell and Mei (1993).

[2] We do not consider the effects of model uncertainty.

[3] We thank the referee for motivating this example.

[4] Since some of our specifications have three year leads of LGD and LRET as explanatory variables, in order to be consistent with our reinterpretation regressions, we estimate the VAR only through 1998. For this reason, the dependent variable series for each of our future regressions ends in 1998.

[5] The methodology of time-series means of cross-sectional descriptives implicitly adds time dummies. Thus, the statistics should be compared to pooled estimates computed from market-adjusted data.