THE JOURNAL OF FINANCE • VOL. LVIII, NO. 2 • APRIL 2003

The Value Spread

RANDOLPH B. COHEN, CHRISTOPHER POLK, and TUOMO VUOLTEENAHO*

ABSTRACT

We decompose the cross-sectional variance of firms' book-to-market ratios using both a long U.S. panel and a shorter international panel. In contrast to typical aggregate time-series results, transitory cross-sectional variation in expected 15-year stock returns causes only a relatively small fraction (20 to 25 percent) of the total cross-sectional variance. The remaining dispersion can be explained by expected 15-year profitability and persistence of valuation levels. Furthermore, this fraction appears stable across time and across types of stocks. We also show that the expected return on value-minus-growth strategies is atypically high at times when their spread in book-to-market ratios is wide.

INTUITIVELY, BOTH EXPECTED STOCK RETURNS and expected cash-flow growth play a role in determining the market price of a firm's stock and, thus, its book-to-market ratio. Firms with high book-to-market ratios have on average higher subsequent stock returns than firms with low book-to-market ratios (Rosenberg, Reid, and Lanstein (1985), Fama and French (1992), and others). Simultaneously, differences in firms' book-to-market ratios are also related to differences in future expected cash-flow and earnings growth as well as future profitability. Low-book-to-market firms grow faster and are persistently more profitable than high-book-to-market firms. The relative importance of these two contributing factors to cross-sectional variation in firms' book-to-market ratios remains an open empirical question.

We decompose the cross-sectional variance of firms' valuation multiples using a long (1938 to 1997) U.S. panel and a shorter (1982 to 1998) international panel. Our variance decompositions show what fraction of the cross-sectional dispersion in book-to-market ratios is caused by variation in expected stock returns

*Cohen is from Harvard Business School; Polk is from the Kellogg School of Management, Northwestern University; and Vuolteenaho is from the Department of Economics, Harvard University and the NBER. This paper subsumes the March 14, 1996, version of the paper titled "Will the Scaled-Price Effect in Stock Returns Continue?" We would like to thank Ken French and Grantham, Mayo, Van Otterloo & Co. for providing us with some of the data used in this study. We are grateful to Cliff Asness, Ron Bird, John Campbell, John Cochrane, Mike Cooper, Kent Daniel, Gene Fama, Ken French (referee), Bob Hodrick, Steve Kaplan, Matti Keloharju, Owen Lamont, Rafael Laporta, Andy Lutz, Andrei Shleifer, Jeremy Stein, and Maria Vassalou (discussant) for their suggestions. We received useful comments from the seminar participants at the NBER 2000 Summer Institute, GMO Research Conference, Harvard Business School brown bag lunch, Purdue University, and Chicago Quantitative Alliance 2001 Spring Meeting. We thank Qianqiu Liu for excellent research assistance.

(induced either by mispricing or risk) and what fraction is caused by variation in expected cash-flow growth. We find that transitory variation in 15-year expected returns is responsible for only a relatively small fraction of the cross-sectional book-to-market variance. In the long U.S. panel, approximately 75 to 80 percent of the unconditional cross-sectional variance of log book-to-market ratios can be explained by expected future 15-year log profitability and persistence of log book-to-market 15 years into the future, while just 20 to 25 percent can be explained by transitory variation in expected returns. These fractions appear stable across time and across types of stocks.

The observation that cross-sectional variation in book-to-market ratios is related to cross-sectional variation in future profitability is not new. For example, Fama and French (1995) show that a high book-to-market ratio signals persistent poor earnings and profitability and a low book-to-market ratio signals strong earnings and profitability. The novel part of our analysis is that by using a present-value model, we are able to measure exactly how much the cross-sectional variation in expected profitability contributes to the cross-sectional variation in firms' book-to-market ratios. Provided that book-to-market does not behave explosively, an approximate identity equates the current book-to-market ratio with an infinite discounted sum of future stock returns less future profitability. This identity implies that a low current book-to-market ratio has to be justified by either high expected future profitability or low expected future stock returns. Therefore, to a very accurate approximation, all cross-sectional variation in the book-to-market ratios must necessarily be accounted for by cross-sectional variation in expected long-horizon stock returns and/or profitability. By exploiting this identity, our analysis is able to quantitatively, not just qualitatively, comment on the source of the heterogeneity in valuation ratios across firms.

It is interesting to relate our basic cross-sectional result to aggregate time-series results in the previous literature. Earlier studies (Campbell and Shiller (1988), Cochrane (1992), Vuolteenaho (2000), and others) examine the time series of aggregate scaled-price measures to quantify the nature of the information they contain. These papers find that the substantial majority of the information in the time series of scaled prices is about expected returns, and relatively little is about cash flows. For example, Vuolteenaho finds that almost all of the time-series variation in the aggregate book-to-market ratio is due to expected stock returns and approximately none due to expected profitability. Our observation that the cross section of individual stocks looks very different from aggregate indexes in this regard is interesting in its own right, but may also be relevant for the interpretation of the previous aggregate studies. If the cross section of valuation ratios is largely driven by rational cash-flow expectations, the conclusion that aggregate valuation ratios are exclusively driven by irrational investor sentiment is perhaps premature.

Note that our basic decomposition is silent as to whether the information we measure in a firm's book-to-market ratio concerning subsequent returns is due to risk or mispricing. However, even if one takes the extreme stand and characterizes the expected-return component of book-to-market as a market-inefficiency phenomenon, our evidence suggests that at most about 20 to 25 percent of the variation in valuation ratios across firms can be linked to capital-market inefficiencies. Most of that variation represents differences in expected future profitability, heterogeneity that is considerably less controversial.

We refine our basic result by extending our analysis of the cross-sectional heterogeneity of valuation ratios. First, we replicate our measurements using a shorter international panel (1982 to 1998, 23 countries excluding the United States). While the short time dimension of our international sample reduces the statistical precision of our estimates, the international point estimates support the conclusions drawn from the U.S. sample. At the five-year horizon, approximately 40 percent of the variation in country-adjusted book-to-market ratios can be allocated to expected country-adjusted profitability, 15 percent to expected country-adjusted stock returns, and the remaining 45 percent to persistence of country-adjusted book-to-market ratios. The fact that we obtain similar point estimates from a largely independent sample lends credibility to our basic result.

Second, in order to continue to study the relation between the results at different levels of aggregation, we further split the information in book-to-market concerning future returns and future profitability into intra- and interindustry components. We find that the book-to-market effect in returns is mostly an intraindustry effect. This confirms evidence at the monthly horizon presented by Cohen and Polk (1998), Lewellen (1999), and Asness, Porter, and Stevens (2000). More importantly, these results show that controlling for industry effects does not alter our variance decomposition results.

Third, we document statistically strong results concerning the predictability of returns on strategies that go long value stocks and short growth stocks. The above variance-decomposition results suggest that the difference in average returns between value and growth stocks may vary through time with the value spread (i.e., the book-to-market ratio of a typical value stock minus that of a typical growth stock). In other words, if the variance decomposition is constant over time with a nontrivial portion of the variance allocated to expected returns, the level of the value spread should predict returns on value-minus-growth strategies. Stocks whose book equity is cheap should have especially high expected returns at times when their book equity is especially cheap.

We test this hypothesis on the Fama–French (1993) HML portfolio and find that the HML value spread indeed has strong predictive power for HML returns, evidence that the expected return on the HML portfolio varies over time. Our expected-return point estimates suggest that the expected return on HML was zero or negative in the periods 1948 to 1953 and 1977 to 1979. Point estimates obtained from an international analysis are consistent with the U.S. HML-predictability results. Our findings are also consistent with those of Asness, Friedman, Krail, and Liew (2000), who examine the cross section of scaled-price measures and returns on scaled-price portfolios. Using data from a relatively recent period (1982 to 1999) and a sample of U.S. stocks listed on I/B/E/S, they find that value spreads and spreads in projected earnings growth have predictive power for the time series of monthly returns on value-versus-growth strategies.

The remainder of the paper is organized as follows. Section I describes the data. Section II presents the variance-decomposition framework and results.

Section III concentrates on predicting value-minus-growth returns. Section IV concludes.

I. Data

The basic U.S. data come from the merger of three databases. The first of these, 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 second database, the COMPUSTAT annual research file, contains the relevant accounting information for most publicly traded U.S. stocks. The COMPUSTAT accounting information is supplemented by the third database, Moody's book equity information collected by Davis, Fama, and French (2000).[1]

The basic merged data cover the period 1928 to 1997, but we only include the period 1938 to 1997 in tests that interpret book-to-market or price-to-book ratios on a ratio scale (e.g., regressions on the book-to-market ratio). It is comforting, however, that our main regression results are robust to the choice between the 1928–1997 and 1938–1997 periods. Note that in our timing convention, 1938 data is computed by using book values from the end of 1937 and returns through May 1939. In the merged data set, the 1928 to 1997 panel contains 192,295 and the 1938 to 1997 panel 185,441 firm-years.

The logic behind this choice of period is based on disclosure regulation. Before the Securities Exchange Act of 1934, there was essentially no regulation to ensure the flow of accurate and systematic accounting information. Among other things, the act prescribes specific annual and periodic reporting and record-keeping requirements for these companies. The companies required to file reports with the SEC must also "make and keep books, records, and accounts, which, in reasonable detail, accurately and fairly reflect the transactions and disposition of the assets of the issuer." In addition, the legislation introduces the concept of "an independent public or certified accountant" to certify financial statements and imposes statutory liabilities on accountants. After studying the implementation of the act, we decided to exclude the book-equity data up to 1936 from the ratio-scale tests: The 1934 to 1936 period could reasonably be characterized as an initial enforcement period, after which reporting conventions have converged to their steady states.[2]

[1] We thank Kenneth French for providing us with the data.

[2] Merino and Mayper (1999) provide statistics on the enforcement of the 1934 Securities Exchange Act. In the first 10 years of the enforcement of the 1934 act, the SEC began 279 proceedings. Of these proceedings, 272 were begun in the 1933 to 1937 period and only 7 began in the subsequent five-year period. In these proceedings, the SEC identified numerous types of accounting and nonaccounting violations by publicly traded firms. The above indicates that enforcement was broad and active in the first five years, whereas in the 1938 to 1942 period, few enforcement proceedings were initiated. While the total number of identified problems in firms' statements was similar during both the 1933 to 1937 (101 accounting violations) and 1938 to 1942 (83 accounting violations) period, it appears that in the latter period, only the worst violators were examined, in which many errors per registration statement were uncovered.

Detailed data definitions are the following. Book equity is defined as the stockholders' equity, plus balance sheet deferred taxes (data item 74) and investment tax credit (data item 208; if available), plus postretirement benefit liabilities (data item 330; if available), minus the book value of preferred stock. Depending on availability, we use redemption (data item 56), liquidation (data item 10), or par value (data item 130) in that order for the book value of preferred stock. Stockholders' equity used in the above formula is calculated as follows. We prefer the stockholders' equity number reported by Moody's or COMPUSTAT (data item 216). If neither one is available, we measure stockholders' equity as the book value of common equity (data item 60), plus the par value of preferred stock. (Note that the preferred stock is added at this stage, because it is later subtracted in the book equity formula.) If common equity is not available, we compute stockholders' equity as the book value of assets (data item 6) minus total liabilities (data item 181), all from COMPUSTAT.

Each year we create 40 value-weight portfolios by sorting stocks on book-to-market (BE/ME) and track the subsequent stock returns, profitability, and book-to-market ratios of these portfolios. We form portfolios on the BE/ME ratio in May of year t and use this value as the dependent variable in our regressions. We calculate BE/ME as book common equity for the fiscal year ending in calendar year $t-1$, divided by market equity at the end of May of year t.[3] We require the firm to have a valid past BE/ME. Moreover, to eliminate likely data errors, we discard those firms with BE/MEs less than 0.01 and greater than 100. When using COMPUSTAT as our source of accounting information, we require that the firm must be on COMPUSTAT for two years before using the data. This requirement alleviates the potential survivor bias due to COMPUSTAT backfilling data. After imposing these data requirements, the cumulative number of firms sorted into portfolios is 139,310 for the 1928 to 1997 period and 134,467 for the 1938 to 1997 period.

The variables of our regressions are formed as follows. Stock returns correspond to value-weight portfolio returns. When computing firms' stock returns, we include delisting data when available on the CRSP tapes. In some cases, CRSP records delisting prices several months after the security ceases trading and, thus, after a period of missing returns. In these cases, we calculate the total return from the last available price to the delisting price and pro-rate this return over the intervening months. Delisting return is included when computing the market value at which the firm exits. We aggregate firm-level book equities by summing the book-equity data for each portfolio. If book equity is not available for a firm at the end of year t, we assume that the firm's book-to-market ratio has

This decline in proceedings may signal increasing compliance by registrants or declining interest by the SEC in regulatory enforcement. We believe the former cause was the driving force behind the reduced number of new proceedings. It is thus reasonable to characterize the 1933 to 1937 period as the initial and strict enforcement period, and the 1938 to 1942 as the beginning of a steady-state level of enforcement.

[3] Some earlier research (e.g., Fama and French (1992)) uses market equity at the end of year $t-1$ to compute BE/ME. For our decomposition of BE/ME to hold, however, we use the May of year-t market equity as the denominator in BE/ME.

not changed from $t-1$ 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. A portfolio's log profitability is computed from portfolio-level book- and market-equity data and CRSP dividends using the clean-surplus-based formula $\log\{[(1+R_t)ME_{t-1}-D_t]BE_t/(ME_tBE_{t-1})-[1-D_t/BE_{t-1}]\}$, where the simple value-weight stock return is denoted by R, book value by BE, market value by ME, and dividends by D. This formula assumes that earnings this year equal the change in book equity plus dividends and adds an appropriate adjustment for equity offerings.

In addition to the long U.S. panel, we also construct a shorter international panel from a monthly data set created by Grantham, Mayo, Van Otterloo & Co. (GMO) using the same data-manipulation practices as for the U.S. data. GMO's monthly data set is in turn constructed from the Morgan Stanley Capital International data and various real-time data sources. The international panel spans the period 1982 to 1998. The number of countries in the data set begins at 19 and ends at 22, and 23 different countries are included in the panel at various points of time. This international panel contains 27,913 firm-years and does not include U.S. data.

II. Decomposing the Cross-Sectional Variance of Firms' Book-to-Market Ratios

It is well known that firms' BE/ME ratios forecast cross-sectional variation in returns and profitability (see, e.g., Fama and French (1995)). However, calculating how each of these two effects contributes to the cross-sectional spread in firms' BE/MEs requires a quantitative framework. We employ Vuolteenaho's (2000) return-profitability model to decompose the cross-sectional variance of BE/ME ratios into three components: (a) covariance of future market-adjusted stock returns with past market-adjusted BE/ME ratios, (b) covariance of future market-adjusted profitability (i.e., accounting return on equity or ROE) with past market-adjusted BE/ME ratios, and (c) persistence of market-adjusted BE/ME ratios.

The following intuition underlies the decomposition. Suppose two firms (neither one of which pays dividends or issues equity) have different BE/ME ratios. Over time, this value spread can close by either the high BE/ME firm experiencing higher stock returns than the low BE/ME firm or by the high BE/ME firm being less profitable and growing its book equity slower than the low BE/ME firm. On the one hand, if the spread closes via stock returns, the covariance of past relative BE/ME ratios and future relative returns is positive. On the other hand, if the spread closes via profitability, the covariance of past relative BE/ME ratios and future relative profitability is negative. There is also a third possibility – the spread may not close completely. In this case, the difference in BE/ME ratios persists, and the covariance of past and future relative BE/ME ratios is high.

A. Log-Linear Present-Value Model and Variance Decomposition

We make three assumptions. First, because the model is stated in terms of logarithms, one must assume book value, BE, market value, ME, and dividends, D, are strictly positive. Although this assumption is not necessarily satisfied for all individual firms, it is almost certainly satisfied for the BE/ME-sorted portfolios we use in our tests, because we exclude firms with presort book-to-market ratios under 1/100 and over 100. Second, one must assume firms' log BE/ME ratios to be stationary, even though both log book and market equity series have an integrated component. Third, we assume that earnings, dividends, and book-equity series satisfy the clean-surplus relation.[4] The clean-surplus relation ties the income-statement and balance-sheet dynamics together. In that relation, earnings, dividends, and book equity satisfy

$$BE_t - BE_{t-1} = X_t - D_t \tag{1}$$

book value today equals book value last year plus clean-surplus earnings (X_t) less (net) dividends.

The reported earnings, dividends, and book values do not always strictly adhere to the above clean-surplus relation. To exactly satisfy this assumption in our sample, we define our earnings series as the sum of dividends and the change in book equity. This approach is partly dictated by necessity (the early data consist of book-equity series but do not contain earnings).[5]

Accounting and stock returns remain to be defined. Let r_t denote the log stock return and e_t the log clean-surplus accounting return on equity, defined as

$$r_t \equiv \log\left(1 + \frac{\Delta ME_t + D_t}{ME_{t-1}}\right) \tag{2}$$

$$e_t \equiv \log\left(1 + \frac{\Delta BE_t + D_t}{BE_{t-1}}\right). \tag{3}$$

In our data set, we construct the clean-surplus ROE with an appropriate adjustment for equity offerings as

$$e_t = \log\left(\left[\frac{(1+R_t)ME_{t-1} - D_t}{ME_t}\right] \times \left[\frac{BE_t}{BE_{t-1}}\right] - \left[1 - \frac{D_t}{BE_{t-1}}\right]\right). \tag{4}$$

Substituting the log dividend-growth rate, Δd_t, the log dividend-price ratio, δ_t, and the log dividend-to-book-equity ratio, $\psi_t \equiv d_t - b_t$, into the return definitions (2) and (3) yields

$$r_t = \log(\exp(-\delta_t) + 1) + \Delta d_t + \delta_{t-1} \tag{5}$$

$$e_t = \log(\exp(-\psi_t) + 1) + \Delta d_t + \psi_{t-1} \tag{6}$$

Finally, we denote the log BE/ME ratio by θ_t:

[4] The clean-surplus relation has also been used in equity valuation by Ohlson (1995), Feltham and Ohlson (1999), and others.

[5] In their Appendix, Frankel and Lee (1999) list the accounting standards that violate clean-surplus accounting. They argue that these violations of the clean-surplus relation are largely unpredictable. Thus, using reported earnings instead of imputed clean-surplus earning would probably not materially affect our variance-decomposition results.

$$\theta_t \equiv \log(BE_t/ME_t) \tag{7}$$

An inconvenient nonlinear law describes the evolution of a firm's BE/ME if the firm pays dividends. However, a linear model can do a good job of approximating this nonlinear evolution:

$$e_t - r_t = \rho\theta_t - \theta_{t-1} + \kappa_t \tag{8}$$

where ρ is a parameter and κ_t an approximation error. If the firm pays any dividends then $\rho < 1$, and $\rho = 1$ if the firm does not. This approximation can be justified by the following logic. We approximate both stock and accounting returns by a first-order Taylor-series approximation, choosing the same expansion point:

$$r_t = \log(\exp(-\delta_t) + 1) + \Delta d_t + \delta_{t-1} \approx \alpha - \rho\delta_t + \Delta d_t + \delta_{t-1} \tag{9}$$

$$e_t = \log(\exp(-\psi_t) + 1) + \Delta d_t + \psi_{t-1} \approx \alpha - \rho\psi_t + \Delta d_t + \psi_{t-1} \tag{10}$$

Subtracting (10) from (9) yields the linearized accounting identity (8).

Using the convenient linear form of equation (8), one can iterate forward and express the BE/ME ratio as an infinite discounted sum of future stock return less future profitability:

$$\theta_{t-1} = \sum_{j=0}^{N} \rho^j r_{t+j} + \sum_{j=0}^{N} \rho^j(-e_{t+j}) + \sum_{j=0}^{N} \rho^j \kappa_{t+j} + \rho^{N+1}\theta_{t+N} \tag{11}$$

If the BE/ME ratio is well behaved and $\rho < 1$, the last term of (11) converges to zero as $N \to \infty$:

$$\theta_{t-1} = \sum_{j=0}^{\infty} \rho^j r_{t+i} + \sum_{j=0}^{\infty} \rho^j(-e_{t+j}) + \sum_{j=0}^{\infty} \rho^j \kappa_{t+j} \tag{12}$$

This approximate model for a firm's BE/ME provides the foundation for our main variance decomposition.

To derive the variance decomposition, we multiply both sides of (12) by $\tilde{\theta}_{t-1}$, drop the approximation error, and take unconditional expectations:

$$\operatorname{var}(\tilde{\theta}) \approx \sum_{j=0}^{\infty} \rho^j \operatorname{cov}(\tilde{r}_{t+j}, \tilde{\theta}_{t-1}) + \sum_{j=0}^{\infty} \rho^j \operatorname{cov}(-\tilde{e}_{t+j}, \tilde{\theta}_{t-1}) \tag{13}$$

Above, cross-sectionally demeaned quantities are denoted by tildes. Equation (13) equates the unconditional cross-sectional variance of firms' BE/ME ratios with the cross-sectional covariance of future stock and/or accounting returns with past BE/ME ratios. Note that $\operatorname{var}(\tilde{\theta})$ in equation (13) corresponds to the average squared cross-sectionally demeaned BE/ME ratio and that this variance metric is, thus, best interpreted as the typical cross-sectional dispersion in BE/MEs.[6]

[6] In the sections that follow, we will typically refer to these variables without the relative, cross-sectionally demeaned, or log modifier for ease of exposition. Similarly, we will use "unconditional variance" of BE/MEs to refer to the unconditional variance of cross-sectionally

Due to the infinite sums, implementing the above variance decomposition (13) requires an assumed auxiliary statistical model and either the assumption of equal long-run BE/ME ratios for all firms or $\rho < 1$.[7] If one is unwilling to make such assumptions, the infinite sums in (13) can be replaced with finite sums by including an additional term. Working with equation (11) without taking the limit $N \to \infty$ and repeating the above steps yields

$$\operatorname{var}(\tilde{\theta}) \approx \sum_{j=0}^{N} \rho^j \operatorname{cov}(\tilde{r}_{t+j},\, \tilde{\theta}_{t-1}) + \sum_{j=0}^{N} \rho^j \operatorname{cov}(-\tilde{e}_{t+j},\, \tilde{\theta}_{t-1}) + \rho^{N+1} \operatorname{cov}(\tilde{\theta}_{t+N},\, \tilde{\theta}_{t-1}) \quad (14)$$

Compared to (13), equation (14) has a catchall predictability term that captures the profitability and stock-return predictability beyond horizon N, as well as cross-sectional heterogeneity in the long-run BE/ME ratios.

Assuming that the cross-sectional variance of the BE/ME ratios is not zero, one can divide both sides of equation (14) by the unconditional BE/ME variance:

$$1 \approx \frac{\sum_{j=0}^{N} \rho^j \operatorname{cov}(\tilde{r}_{t+j}, \tilde{\theta}_{t-1})}{\operatorname{var}(\tilde{\theta})} + \frac{\sum_{j=0}^{N} \rho^j \operatorname{cov}(-\tilde{e}_{t+j}, \tilde{\theta}_{t-1})}{\operatorname{var}(\tilde{\theta})} + \rho^{N+1} \frac{\operatorname{cov}(\tilde{\theta}_{t+N}, \tilde{\theta}_{t-1})}{\operatorname{var}(\tilde{\theta})} \quad (15)$$

The three terms in (15) represent the fraction of variance attributable to the three sources. This relative variance decomposition is particularly easy to interpret—each component in (15) corresponds to a simple regression coefficient. The predictive regression coefficient of long-horizon returns plus the predictive regression coefficient of the negative of long-horizon profitability plus a measure of the persistence of BE/ME spread must be equal to one.

B. Cross-sectional Variance-Decomposition Results Using the U.S. Panel

Each year, we create 40 value-weight portfolios by sorting stocks on BE/ME and track the subsequent stock returns, profitability, and book-to-market ratios of these portfolios. In 1997 (the last year of the U.S. sample), the low BE/ME portfolio has a value-weight average BE/ME of 0.04, while the highest portfolio is at 2.82. The mean portfolio BE/ME is 0.62, while the standard deviation across portfolios is 0.54. Dispersion in BE/MEs has risen and fallen many times over the years. For example, for 1987, the standard deviation (0.55) is almost identical to the 1997 value, but for 1991 it is more than twice as high (1.18).

Three simple forecasting regressions measure the percentage of information in a firm's BE/ME ratio about the firm's future:

demeaned BE/MEs; the true unconditional variance would include time variation in the cross-sectional mean of BE/MEs.

[7] Past research that decomposes the time-series variance of aggregate portfolios often uses vector autoregressions (VARs) instead of long-horizon regressions. In our cross-sectional application, the long-horizon regression methodology is preferable to the more common VAR methodology. The difficulty with the VAR methodology arises from rebalancing effects. Appendix A discusses this problem in more detail.

$$\begin{aligned}\sum_{j=0}^{N-1}\rho^j\tilde{r}_{k,t+j} &= b(\tilde{r},N)\,\tilde{\theta}_{k,t-1}+\varepsilon(\tilde{r},N,k,t+N-1)\\ \sum_{j=0}^{N-1}\rho^j(-\tilde{e}_{k,t+j}) &= b(-\tilde{e},N)\,\tilde{\theta}_{k,t-1}+\varepsilon(-\tilde{e},N,k,t+N-1)\\ \rho^N\tilde{\theta}_{k,t+N-1} &= b(\tilde{r},N)\,\tilde{\theta}_{k,t-1}+\varepsilon(\tilde{\theta},N,k,t+N-1).\end{aligned}\tag{16}$$

We estimate these regression coefficients using the 40 BE/ME-sorted portfolios and report the combined results in Table I, Panel A. All data are cross-sectionally demeaned, and subscripts refer to portfolio and date: For example, $\tilde{\theta}_{k,t}$ is the aggregate book equity of the stocks in portfolio k at the end of year t divided by the portfolio's aggregate market value, less the average of these ratios across the 40 portfolios at the end of year t. In the system of regressions described in equation (16), $b(\tilde{r},N)+b(-\tilde{e},N)+b(\tilde{\theta},N)\approx 1$.

The $N=1$ row of the table breaks BE/ME into information concerning future one-year returns, concerning future one-year ROE, and concerning one-year-ahead BE/ME. The split is three percent for future returns, 15 percent for future profitability, and 83 percent for persistence of BE/ME. At the one-year horizon, the BE/ME ratio predicts all three variables with the expected signs. Because BE/ME ratios are quite persistent, the largest variance component of BE/ME is due to covariance with next year's BE/ME.

Of course, next year's BE/ME ratio also has information about future returns and ROEs beyond the one-year horizon. Subsequent rows of the table look further ahead to 2, 3, 5, 10, and 15 years. Figure 1 graphs these coefficients as a function of the forecast horizon. At the 15-year horizon, 20 percent of BE/ME information is about returns, 58 percent about profitabilities, and the remaining 26 percent about future BE/ME ratios. Assuming equal long-run BE/MEs for all firms would guarantee that the last component due to the persistence of BE/MEs is informative about future returns and profitability beyond the 15-year horizon. Assuming that most of the remaining information in the 15-year-ahead BE/ME concerns cash flows and not returns and assuming that allocating the persistent component in BE/MEs to profitability would lead to an aggressive interpretation, that 80 percent of the BE/ME variance is due to cash-flow effects and only 20 percent is due to expected-return effects. A more conservative practice is to allocate the persistent component to cash flows and returns in the same proportions as they explain the 15-year resolution of BE/MEs. This assumption would lead us to conclude that 75 percent of the BE/ME variance is caused by variation in expected cash flows and 25 percent by variation in expected returns.

Our procedure imposes that the discount coefficient (ρ) be constant across firms. Because long-run dividend yields and payout ratios may vary systematically with BE/ME ratios, the assumption of a constant discount coefficient may lead to a poor approximation in equation (8). Our regressions provide a natural way to evaluate the effect of the approximation error on the variance-decomposition results. Repeating the derivations in equations (11)–(15) and carrying the approximation error along shows that the variance component due to the approx-

Table I

Decomposition of the Unconditional Variance of Firms' Log BE/ME Ratios

This table decomposes the unconditional variance of firms' BE/ME ratios. The decomposition is estimated from the 1938 to 1997 sample of U.S. stocks. Each data point consists of the log value-weight BE/ME (θ), return (r), and profitability (ROE, e) for 1 of 40 portfolios formed each year by sorting firms on BE/ME. Panel A uses cross-sectionally demeaned data and Panel B industry-adjusted data (using Fama and French's (1997) industry classifications). The first column shows the horizon N. The second column shows the simple predictive regression coefficient of N-period discounted future stock return on BE/ME. The third column shows the predictive coefficient of the negative of N-period discounted future ROE on BE/ME. The fourth column shows the predictive coefficient of N-period-in-the-future discounted BE/ME on BE/ME:

$$\begin{aligned}\sum_{j=0}^{N-1} \rho^j \tilde{r}_{k,t+j} &= b(\tilde{r}, N)\, \tilde{\theta}_{k,t-1} + \varepsilon(\tilde{r}, N, k, t+N-1);\\ \sum_{j=0}^{N-1} \rho^j (-\tilde{e}_{k,t+j}) &= b(-\tilde{e}, N)\, \tilde{\theta}_{k,t-1} + \varepsilon(-\tilde{e}, N, k, t+N-1);\\ \rho^N \tilde{\theta}_{k,t+N-1} &= b(\tilde{\theta}, N)\, \tilde{\theta}_{k,t-1} + \varepsilon(\tilde{\theta}, N, k, t+N-1).\end{aligned}$$

Above, the discount coefficient (ρ) equals 0.96. The predictive coefficients, and thus each row of the table, sum up to approximately one. The point estimates are produced with pooled OLS and the standard errors (in parentheses) with GMM. The applicable GMM standard-error formulas account for cross-sectional and serial correlation of the errors in large samples.

N	Expected Returns	(–) Expected Profitability	Future BE/ME
Panel A: Cross-Sectionally Demeaned U.S. Data, 1938–1997			
1	0.0272 (0.0083)	0.1469 (0.0097)	0.8258 (0.0127)
2	0.0559 (0.0242)	0.2386 (0.0360)	0.7055 (0.0378)
3	0.0809 (0.0441)	0.3006 (0.0613)	0.6194 (0.0607)
5	0.1198 (0.0834)	0.3826 (0.0895)	0.5012 (0.0954)
10	0.1734 (0.0951)	0.4938 (0.1185)	0.3399 (0.1441)
15	0.1996 (0.1097)	0.5832 (0.1754)	0.2646 (0.0967)
Panel B: Industry-Adjusted U.S. data, 1938–1997			
1	0.0236 (0.0078)	0.1209 (0.0080)	0.8590 (0.0086)
2	0.0510 (0.0262)	0.2103 (0.0300)	0.7459 (0.0267)
3	0.0763 (0.0464)	0.2765 (0.0547)	0.6584 (0.0451)
5	0.1163 (0.0840)	0.3679 (0.0857)	0.5359 (0.0786)
10	0.1677 (0.0999)	0.4853 (0.1445)	0.3596 (0.1153)
15	0.1859 (0.1092)	0.5784 (0.1549)	0.2743 (0.0909)

imation error equals $1 - b(\tilde{r}, N) - b(-\tilde{e}, N) - b(\tilde{\theta}, N)$. Thus, the approximation error's share of the unconditional variance in Table I, Panel A, is 0.01 percent at the one-year horizon and -4.74 percent at the 15-year horizon. These computations indicate that the error in the constant-ρ linear approximation does not materially affect our results.

Table I, Panel A, also reports the standard errors for the variance decomposition. Although we use regression-coefficient point estimates obtained using

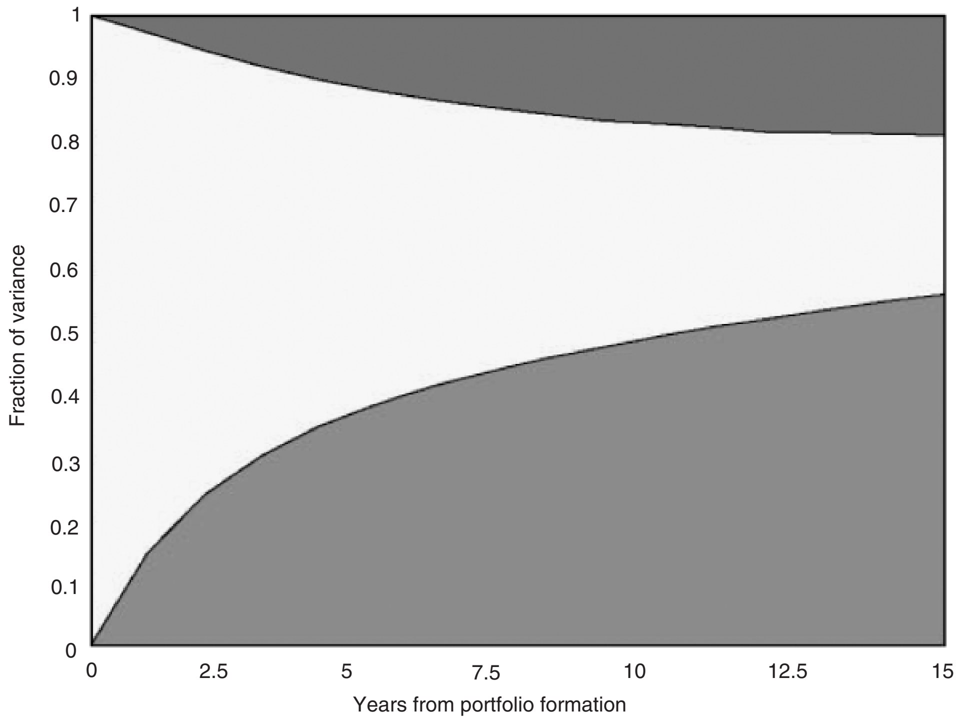

Figure 1. Decomposition of the cross-sectional BE/ME variance. This figure shows the unconditional decomposition of the cross-sectional BE/ME variance estimated from the U.S. data sample (40 portfolios sorted on BE/ME, 1938–1997 period). The horizontal axis shows the horizon N. The height of the top area shows the simple predictive regression coefficient of cross-sectionally demeaned, N-period discounted future stock return on cross-sectionally demeaned BE/ME. The height of the bottom area shows the simple predictive regression coefficient of the negative of cross-sectionally demeaned, N-period discounted future profitability (ROE) on cross-sectionally demeaned BE/ME. The height of the middle area shows the simple predictive regression coefficient of cross-sectionally demeaned, N-period-in-the-future discounted BE/ME on cross-sectionally demeaned BE/ME.

pooled OLS, the usual OLS standard errors are likely to be significantly understated. The problem with OLS standard errors arises from two well-known sources. First, the regression residuals may be correlated in the cross section. If this correlation is related to the explanatory variables, the standard errors are not valid. Second, because we use overlapping dependent variables, the regression residuals are likely to be autocorrelated. A more general and precise statement of these problems is that the covariance matrix of the pooled-regression errors is not proportional to an identity matrix. To calculate appropriate standard errors that account for correlation of the residuals both over time and in the cross section, we adapt Rogers's (1983, 1993) standard-error formulas to our regressions. Appendix B contains the details of these calculations.

Not surprisingly, the statistical evidence that the past BE/ME ratios predict future BE/MEs is overwhelming, as t-statistics in the last column are uniformly high. Predictions of future BE/MEs start out at t-statistics above 70 and remain above 2.0 even 15 years out. The ability of BE/ME to predict profitability at longer horizons is also very convincing. The t-statistics in the early years range from 3 to 11, and even toward the 15-year mark stay above 2.0.

As expected, the standard errors grow with the prediction horizon, because longer horizons lead to a smaller number of independent observations in our sample. The problem of statistical power is particularly severe in the prediction of returns, especially for horizons longer than three years. Although the coefficients on future expected returns grow with the horizon, the standard errors grow even faster and, as a consequence, the t-statistics fall below 2.0 in longer horizons. We cannot reject the hypothesis that none of the BE/ME variance is due to expected returns at conventional levels for the 15-year horizon regression (the associated t-statistic is 1.82). This is consistent with our conclusion that, in the cross section, firms' BE/ME ratios are mainly driven by expected profitability, not by expected returns.[8]

Our main interpretation of these variance-decomposition results is that they help to put the average return difference between value and growth stocks into a price-level perspective. Our results suggest that most of a growth stock's atypical valuation is simply due to high expected profitability rather than due to that stock having a low expected return. For example, Cisco Systems and General Motors have approximately the same book value (Cisco \$29.7 vs. GM \$31.5 billion), but Cisco has a far greater market value (Cisco \$175.6 vs. GM \$31.3 billion).[9] According to our results, most of the difference is because equity on Cisco's books is expected to be more productive than that on GM's. However, some reversion toward normal book-to-market ratios in the form of higher relative stock returns for GM is to be expected.

Our results are consistent with previous research by Vuolteenaho (2002) that finds that firm-level stock returns (approximately changes in prices) are mainly driven by changes in cash-flow expectations, not by changes in expected returns. Our finding that the level of a firm's stock price (normalized by book equity) is also mainly driven by cash-flow expectations adds to the evidence that for individual firms, the expected cash-flow fundamentals are the most important valuation concern.

These cross-sectional results may also be helpful in interpreting variance-decomposition results for valuation multiples of the aggregate stock market. Campbell and Shiller (1988), Cochrane (1992), Vuolteenaho (2000), and others find that

[8] Lamont and Polk (2001) decompose the cross-sectional variance of valuation levels as well. They find that 20 percent of the cross-sectional variance of diversified firms' excess values (over a portfolio of matched single-segment firms) is due to differences in future returns. Note that Lamont and Polk decompose the cross-sectional variance of the difference between diversified and single-segment firms' book-to-market asset ratios, while our analysis focuses on firms' book-to-market equity ratios. Our work also differs from theirs in that our methodology uses long horizon regressions, while Lamont and Polk rely on a firm-level vector autoregression. The fact that both techniques produce similar results should corroborate Lamont and Polk's findings as well as ours.

[9] The Cisco Systems' and General Motors' data are from finance.yahoo.com as of March 7, 2001.

the substantial majority of the information in the time series of aggregate valuation multiples is about expected returns, and little or none is about cash flows. These aggregate return-predictability phenomena are not statistically overwhelming, however: Rejections of the null of no predictability are significant at around the five percent level. Five percent p-values are obtained by exploiting the usual long sample (from the early 1800s to the present date) and by carefully adjusting the size of the test to reflect the small sample.

The cross-sectional dimension of our sample may offer additional evidence on the sources of time-series variation in valuation multiples. First, if one has a prior that the cross-sectional and time-series variation are caused by similar mechanisms, our cross-sectional evidence may tilt posteriors about the source of variation in aggregate valuation multiples more toward cash flows. Second, it may be that the predictable component of aggregate cash flows is simply not variable and persistent enough to have a significant impact on prices, even if the market rationally discounts it into the prices. Contrary to the aggregate time series, the expected cash-flow growth and profitability are known to be highly variable in the cross section. Our tests provide evidence that in situations where economically significant differences in cash-flow fundamentals exist, they are clearly reflected in prices.

C. International Results

Table II estimates the system of regression equations (16) using the international panel. While the international results are noisier, we interpret them as supporting our main results obtained from the U.S. data.

As above, the $N=1$ row of Table II, Panel A, breaks the BE/ME ratio into information concerning the future one-year returns, one-year ROE, and one-year-ahead BE/ME ratios. The split again assigns close to 3 percent to future returns. In the international panel, the remaining portion contains less information about future profitability (11 percent) and more about persistence in BE/MEs (85 percent) than the U.S. panel.

The data availability limits our long-horizon estimates to two, three, and five years. (After five years, the precision of the variance-decomposition estimates and the accuracy of the asymptotic standard errors decrease dramatically.) At the five-year horizon, 18 percent of BE/ME information is about returns, 33 percent about profitability, and the remaining 49 percent about future BE/ME ratios.

One might argue that some portion of the cross-sectional variance of the BE/ME ratios is due to permanent differences in accounting standards across countries. This argument suggests that country-adjusted BE/ME ratios should contain a higher percentage of information concerning transitory returns and/or profitability once this country-specific noise due to accounting standards is removed. Therefore, we refine our variance decomposition of Panel A by adjusting each variable (BE/ME, return, and profitability) by subtracting the appropriate value-weight country measure.

The results (reported in Panel B) are quite similar to the unadjusted decomposition. For the country-adjusted data, 14 percent of the cross-sectional variance of

Table II

Variance Decomposition Estimated from the International Data

This table shows the decomposition of the unconditional variance of firms' BE/ME ratios estimated from the international data sample (1982 to 1998, 23 countries). Panel A uses cross-sectionally demeaned data and Panel B country-adjusted data. Each data point consists of the log value-weight BE/ME (θ), return (r), and profitability (ROE, e) for 1 of 40 portfolios formed each year by sorting firms on BE/ME. The first column shows the horizon N. The second column shows the simple predictive regression coefficient of N-period discounted future stock return on BE/ME. The third column shows the predictive coefficient of the negative of N-period discounted future ROE on BE/ME. The fourth column shows the predictive coefficient of N-period-in-the-future discounted BE/ME on BE/ME:

$$\sum_{j=0}^{N-1} \rho^j \tilde{r}_{k,t+j} = b(\tilde{r}, N)\, \tilde{\theta}_{k,t-1} + \varepsilon(\tilde{r}, N, k, t+N-1);$$

$$\sum_{j=0}^{N-1} \rho^j (-\tilde{e}_{k,t+j}) = b(-\tilde{e}, N)\, \tilde{\theta}_{k,t-1} + \varepsilon(-\tilde{e}, N, k, t+N-1);$$

$$\rho^N \tilde{\theta}_{k,t+N-1} = b(\tilde{\theta}, N)\, \tilde{\theta}_{k,t-1} + \varepsilon(\tilde{\theta}, N, k, t+N-1).$$

Above, the discount coefficient (ρ) equals 0.97. The predictive coefficients, and thus each row of the table, sum up to approximately one. The point estimates are produced with pooled OLS and the standard errors (in parentheses) with GMM. The applicable GMM standard-error formulas account for cross-sectional and serial correlation of the errors in large samples.

Panel A: Cross-Sectionally Demeaned International Data, 1982–1998			
N	Expected Returns	(–) Expected profitability	Future BE/ME
1	0.0348 (0.0155)	0.1114 (0.0102)	0.8546 (0.0156)
2	0.0849 (0.0471)	0.1978 (0.0326)	0.7199 (0.0390)
3	0.1192 (0.0812)	0.2516 (0.0422)	0.6312 (0.0629)
5	0.1775 (0.1181)	0.3288 (0.0466)	0.4899 (0.0834)
Panel B: Country-Adjusted International Data, 1982–1998			
N	Expected Returns	(–) Expected profitability	Future BE/ME
1	0.0317 (0.0131)	0.1188 (0.0077)	0.8546 (0.0139)
2	0.0664 (0.0311)	0.2225 (0.0268)	0.7228 (0.0305)
3	0.0921 (0.0310)	0.2967 (0.0344)	0.6277 (0.0346)
5	0.1401 (0.0410)	0.4110 (0.0274)	0.4726 (0.0339)

country-adjusted BE/ME is due to cross-sectional differences in expected five-year returns and 41 percent due to five-year profitability. As with the unadjusted results, virtually half of the cross-sectional variance is due to five-year persistence in country-adjusted BE/ME.

D. Additional Descriptive Results

D.1. U.S. Intraindustry Results

We also examine the role of industry in the information contained in firms' BE/ME ratios. Persistent differences in accounting standards across U.S. industries could cause persistent BE/ME differences, and such differences in industry

BE/ME would be unrelated to near-term predictability of either profitability or returns.

Table I, Panel B, decomposes the variance of intraindustry BE/MEs using the Fama–French (1996) industry classifications. First, we sort stocks into portfolios based on intraindustry BE/ME and then estimate regressions (16) from industry-adjusted data. Intraindustry BE/ME is defined as the difference between a firm's BE/ME and the value-weight average BE/ME of the industry the firm is in. Thus, the test assets for Table I, Panel B, are portfolios that are more industry-balanced than those in Table I, Panel A, as every industry is likely to have firms that are high, medium, and low in their industry-relative BE/ME ratio.

We find that the importance of future return, ROE, and BE/ME in explaining current BE/ME relative to a firm's industry is quite comparable to their importance in explaining overall BE/ME. Over 15 years, returns explain 19 percent of cross-sectional variation, ROEs 58 percent, and future BE/ME 27 percent, similar to the corresponding numbers in Table I, Panel A.

In Table III, we decompose the variance of raw (i.e., not industry-adjusted) BE/ME into six components by further separating the profitability, return, and BE/ME-persistence components of Table I, Panel A, into inter- and intraindustry parts. This decomposition shows whether interindustry or intraindustry phenomena are the main drivers of the cross section of raw BE/MEs.

Components labeled "intraindustry" in Table III are regression coefficients of industry-adjusted returns, ROE, and future BE/ME on the raw BE/ME ratio. The second set of components, labeled "industry" in Table III, are simply the regression coefficients of the value-weight return, ROE, or future BE/ME for the industry to which the firm belongs on the firm's raw BE/ME ratio. Since a firm's return in any given year is the sum of the return on the firm's industry for that year and the excess return of the firm over that industry return (and since the same can be said for ROE or future BE/ME), we can replace our three-component decomposition with a new six-component decomposition while preserving the identities.

The test portfolios of Table III are created by sorting on the raw BE/ME ratios, and the only difference between the portfolios used in Table I, Panel A, and Table III is that portfolios used in Table III exclude stocks that cannot be assigned to an industry of at least 10 firms. Summing the intraindustry coefficients shows that, regardless of horizon, roughly 80 percent of the information in the raw BE/ME ratios concerns the firm's behavior relative to its industry, while only the remaining small percentage is informative about the industry as a whole.

Looking at returns a single year ahead, the coefficient of intraindustry return is 2.1 percent, while that of industry return is 0.55 percent. Fifteen years out, the intraindustry return explains 16 percent of the cross-sectional variance of firms' BE/ME, while industry return explains only 4.5 percent. The dominance of intraindustry information is even greater in the profitability data. The coefficient of intraindustry ROE is nine times larger than that of industry ROE at the one-year horizon (13 percent of variance for intraindustry vs. 1.4 percent for interindustry) and 19 times larger (50 vs. 2.6 percent) at the 15-year horizon.

Table III

Decomposition of BE/ME Variance into Industry and Intraindustry Components

This table decomposes the variance of firms' BE/ME ratios into six components by splitting each of the three variance components of Table I into industry and intraindustry parts. The decomposition is estimated from the 1938 to 1997 sample of U.S. stocks. We use Fama and French's (1997) industry classifications. Each data point consists of the log value-weight industry and intraindustry BE/ME (θ, θ_{ii}), industry and intraindustry return (r_I, r_{ii}), and industry and intraindustry ROE (e_I, e_{ii}) for 1 of 40 portfolios formed each year by sorting firms on raw (not industry-adjusted) BE/ME. All data are cross-sectionally demeaned. The first column shows the horizon N. The second and third columns show the simple predictive regression coefficients of N-period discounted future industry stock return on BE/ME and the predictive coefficient of N-period discounted industry-adjusted future stock return on BE/ME. The fourth and fifth columns show the simple predictive regression coefficient of the negative of N-period discounted future industry profitability (ROE) on BE/ME and the predictive coefficient of the negative of N-period discounted industry-adjusted future ROE on BE/ME. The sixth and seventh columns show the simple predictive regression coefficient of N-period-in-the-future discounted industry BE/ME on BE/ME and the predictive coefficient of N-period-in-the-future discounted industry-adjusted BE/ME on BE/ME. These six regressions are shown below:

Intra-industry:

$$\sum_{j=0}^{N-1} \rho^j \tilde{r}_{k,t+j,ii} = b(\tilde{r}_{ii}, N)\,\tilde{\theta}_{k,t-1,i} + \varepsilon(\tilde{r}_{ii}, N, k, t+N-1);$$

$$\sum_{j=0}^{N-1} \rho^j (-\tilde{e}_{k,t+j,ii}) = b(-\tilde{e}_{ii}, N)\,\tilde{\theta}_{k,t-1,i} + \varepsilon(-\tilde{e}_{ii}, N, k, t+N-1);$$

$$\rho^N \tilde{\theta}_{k,t+N-1,ii} = b(\tilde{\theta}_{ii}, N)\,\tilde{\theta}_{k,t-1,i} + \varepsilon(\tilde{\theta}_{ii}, N, k, t+N-1);$$

Industry:

$$\sum_{j=0}^{N-1} \rho^j \tilde{r}_{k,t+j,I} = b(\tilde{r}_I, N)\,\tilde{\theta}_{k,t-1,i} + \varepsilon(\tilde{r}_I, N, k, t+N-1);$$

$$\sum_{j=0}^{N-1} \rho^j (-\tilde{e}_{k,t+j,I}) = b(-\tilde{e}_I, N)\,\tilde{\theta}_{k,t-1,i} + \varepsilon(-\tilde{e}_I, N, k, t+N-1);$$

$$\rho^N \tilde{\theta}_{k,t+N-1,I} = b(\tilde{\theta}_I, N)\,\tilde{\theta}_{k,t-1,i} + \varepsilon(\tilde{\theta}_I, N, k, t+N-1).$$

Above, the discount coefficient (ρ) equals 0.96. The predictive coefficients, and thus each row of the table, sum up to approximately one. The point estimates are produced with pooled OLS and the standard errors (in parentheses) with GMM. The applicable GMM standard-error formulas account for cross-sectional and serial correlation of the errors in large samples.

N	Industry's expected returns	Expected industry-adjusted returns	(−) Industry's expected profitability	(−) Expected industry-adjusted profitability	Industry's expected future BE/ME	Expected industry-adjusted BE/ME
1	0.0055 (0.0031)	0.0214 (0.0058)	0.0141 (0.0021)	0.1318 (0.0097)	0.1935 (0.0094)	0.6336 (0.0156)
2	0.0102 (0.0092)	0.0449 (0.0162)	0.0225 (0.0068)	0.2142 (0.0351)	0.1765 (0.0235)	0.5318 (0.0472)
3	0.0141 (0.0158)	0.0654 (0.0301)	0.0280 (0.0105)	0.2713 (0.0598)	0.1639 (0.0347)	0.4587 (0.0758)
5	0.0230 (0.0269)	0.0971 (0.0588)	0.0336 (0.0186)	0.3471 (0.0870)	0.1463 (0.0447)	0.3576 (0.118)
10	0.0442 (0.0444)	0.1382 (0.0683)	0.0376 (0.0168)	0.4416 (0.1015)	0.1064 (0.0406)	0.2279 (0.1609)
15	0.0450 (0.0315)	0.1639 (0.1067)	0.0258 (0.0797)	0.5027 (0.1009)	0.0855 (0.0444)	0.1733 (0.1174)

While it may not be a surprise that most of the information in the BE/ME ratios is about intraindustry performance rather than industry performance, it is interesting to understand the large magnitude of the difference, especially for profitability. Cohen and Polk (1998) and Asness, Porter, and Stevens (2000) decompose book-to-market ratios into inter- and intraindustry components and similarly find that the value effect is primarily intraindustry. These industry results may also be related to Lewellen's (1999) result that controlling for industry does not significantly reduce the degree of cross-sectional spread in sensitivity to the HML factor of Fama and French (1993).

D.2. Does the Variance-Decomposition Vary as a Function of Firm Characteristics?

The relative importance of the three elements of the decomposition—transitory variation in expected returns, transitory variation in profitability, and persistent differences in BE/ME ratios—may be systematically different for different types of firms. We analyze this possibility by initially sorting firms each year into three groups based on a particular firm characteristic. We then sort firms within each group into five portfolios based on BE/ME and estimate the variance decomposition separately within each of these three groups. This conditional variance decomposition is simply the best estimate of the relative variance decomposition (15) for a particular type of firm. Natural firm characteristics to examine are size and the BE/ME ratio itself, which also have the advantage that conditioning on them does not reduce the size of the sample.

Table IV, Panel A, sorts firms into three size groups based on NYSE breakpoints. The percentage variance decomposition appears to be the same for large and small stocks. At the 15-year horizon, for large firms, approximately 16 percent of the information in the BE/ME ratios is due to differences in expected returns, while for small firms this number is 18 percent; however, these differences do not seem economically large and are not statistically significant.

In Table IV, Panel B, we sort firms into three groups based on firm BE/ME to examine how the decomposition depends on BE/ME. Largely by construction, the medium-BE/ME third has an average cross-sectional variance of log BE/ME of just 0.03, while both the high-BE/ME (0.14) and low-BE/ME (0.22) stocks exhibit far more spread. At the one-year horizon, high-BE/ME firms' BE/MEs have more information concerning expected returns. As the horizon increases, the percentage of information in BE/ME ratios about future returns remains relatively low for low-BE/ME firms. It is only at the 15-year horizon that the information content becomes roughly equal.

One might initially be surprised by the result that expected cash flows and returns drive equal fractions of the BE/ME variance for small and large stocks, because previous work (e.g., Fama and French (1993)) has generally indicated that the value effect is stronger among small stocks. Two facts reconcile these potentially conflicting pieces of evidence. First, the split is less similar at short horizons (4.5 percent due to expected one-year returns for small firms and only 2.3 percent for large). Also, at short horizons, large stocks have more persistent BE/ME: The regression coefficient of future one-year BE/ME on current BE/ME is 0.8757 for large stocks, but only 0.8141 for small stocks. It appears that the returns and profitability predicted by BE/ME realize earlier for small stocks than for large stocks. In the long run, these timing effects wash out, however, leaving the total long-run split between return and profitability information similar for small and large stocks.

Second and more importantly, note that since N-year returns in the first equation in the system of regressions in (16) are given by the product, $b(\tilde{r}, N)\tilde{\theta}_{t-1}$, large cross-sectional variation in expected returns can be generated by *either* a large $b(\tilde{r}, N)$ or a large spread in $\tilde{\theta}_{t-1}$. This observation is central to our ability to forecast the returns on value-minus-growth strategies in Section IV and applies here as well. Small stocks do have a stronger value effect in returns, even at long horizons, but not because a unit difference in their BE/ME ratios has more return-prediction ability. Rather, it is simply because their BE/MEs are more disperse. This difference in dispersion is substantial: For the smallest third, the average cross-sectional variance of log BE/ME for the five BE/ME quintiles is 0.53—considerably higher than the corresponding value of 0.39 for the largest third of the sample.

The fact that expected-return and profitability components cause larger swings to the book-to-market ratios of small stocks than to those of large stocks is not surprising. Suppose that all firms draw equal-size projects from a common pool. The projects in the pool have a common fixed (finite) horizon but vary along the dimensions of their profitability and risk. Each firm draws one new project every year, and the only difference between small and large firms is that the latter have drawn more projects in the past.

This simplified model has a number of predictions that match well with the main features of the data. First, the cross-sectional spread in book-to-market ratios is larger for small firms than for large firms, since large firms are better diversified portfolios of investment projects than small firms, pulling their book-to-market ratios toward the mean of the project distribution. Second, although the variance of fitted values in regressions (16) varies across size groups, the relative variance decomposition is roughly equal across size groups, because the projects are drawn from the same distributions. Third, because an additional project has a larger weight for small firms (few existing projects) than for large firms (many existing projects), the book-to-market ratios of small firms are expected to revert to the mean faster than the book-to-market ratios of large firms.

The same simplified model can also account for some of the return-variance-decomposition results in the previous literature. For example, Vuolteenaho (2002) finds that good news about cash flows is typically accompanied by higher expected returns and that this correlation appears to be larger for smaller stocks and about zero for the largest stocks. The simple model can explain this feature of the data, if we assume that the product market is competitive and every new investment project's net present value (NPV) is near zero. If the projects have near-zero NPVs, the correlation between profitability and risk (and thus expected returns) of projects in the pool must be high. Consider a firm that announces it has started a new investment project. Because all projects have near-zero NPVs,

Table IV

Decomposition of BE/ME Variance as Function of Size and BE/ME

This table reports the variance decomposition of firms' BE/ME ratios estimated within size (i.e., market capitalization) and BE/ME groups. First, we sort stocks into three groupings every year based on either size (Panel A) or BE/ME (Panel B). We calculate yearly size breakpoints for the size groupings in Panel A using only NYSE stocks. Second, within each group, we sort stocks into five portfolios on BE/ME ratios. Third, for each of the three size and three BE/ME groups, we estimate variance decompositions as in Table I, but using the five portfolios formed from the stocks in the particular group. The table also reports the cross-sectional variance of the five portfolios' BE/MEs, $\mathrm{var}(\tilde{\theta})$, for each group. The notes of Table I apply.

Panel A: Decomposition of the BE/ME Variance Conditional on Firm Size

	Small ($\rho = 0.99$) $\mathrm{var}(\tilde{\theta}) = 0.53$			Medium ($\rho = 0.98$) $\mathrm{var}(\tilde{\theta}) = 0.46$			Big ($\rho = 0.95$) $\mathrm{var}(\tilde{\theta}) = 0.39$		
N	Expected Returns	$(-)$Expected Profitability	Future BE/ME	Expected Returns	$(-)$Expected Profitability	Future BE/ME	Expected Returns	$(-)$Expected Profitability	Future BE/ME
1	0.0449	0.1404	0.8141	0.0466	0.1636	0.7916	0.0231	0.1036	0.8757
	(0.0107)	(0.0085)	(0.0147	(0.0124)	(0.0090)	(0.0186)	(0.0094)	(0.0039)	(0.0100)
2	0.0946	0.2614	0.6441	0.0940	0.2653	0.6443	0.0442	0.1832	0.7770
	(0.0351)	(0.0391)	(0.0580)	(0.0264)	(0.0246)	(0.0497)	(0.0274)	(0.0152)	(0.0251)
3	0.1350	0.3204	0.5454	0.1259	0.3313	0.5483	0.0640	0.2458	0.6966
	(0.0583)	(0.0455)	(0.0619)	(0.0409)	(0.0379)	(0.0777)	(0.0484)	(0.0304)	(0.0410)
5	0.1899	0.4239	0.3912	0.1830	0.4013	0.4273	0.0920	0.3425	0.5763
	(0.0959)	(0.0748)	(0.0791)	(0.0754)	(0.0628)	(0.1132)	(0.0842)	(0.0691)	(0.0702)
10	0.2030	0.5587	0.3139	0.2310	0.5027	0.2918	0.1326	0.5007	0.3772
	(0.1924)	(0.2140)	(0.0596)	(0.1633)	(0.2271)	(0.1216)	(0.1458)	(0.2090)	(0.1116)
15	0.1787	0.6923	0.2788	0.2568	0.5845	0.2440	0.1574	0.6149	0.2689
	(0.1918)	(0.2919)	(0.1476)	(0.2322)	(0.3638)	(0.1186)	(0.1611)	(0.3045)	(0.0689)

Panel B: Decomposition of the BE/ME Variance Conditional on Firm Book-to-Market

N	Low ($\rho = 0.95$) $\text{var}(\tilde{\theta}) = 0.22$ Expected Returns	Low (−)Expected profitability	Low Future BE/ME	Medium ($\rho = 0.96$) $\text{var}(\tilde{\theta}) = 0.03$ Expected returns	Medium (−)Expected profitability	Medium Future BE/ME	High ($\rho = 0.97$) $\text{var}(\tilde{\theta}) = 0.14$ Expected returns	High (−)Expected profitability	High Future BE/ME
1	0.0102	0.1638	0.8283	0.0375	0.0873	0.8792	0.0240	0.1479	0.8176
	(0.0155)	(0.0117)	(0.0225)	(0.0184)	(0.0069)	(0.0177)	(0.0180)	(0.0179)	(0.0244)
2	0.0234	0.2730	0.7072	0.0568	0.1568	0.7930	0.0541	0.2507	0.6761
	(0.0419)	(0.0473)	(0.0686)	(0.0466)	(0.0291)	(0.0476)	(0.0435)	(0.0613)	(0.0627)
3	0.0362	0.3388	0.6321	0.0805	0.1959	0.7318	0.1022	0.3223	0.5502
	(0.0760)	(0.0822)	(0.1184)	(0.0881)	(0.0558)	(0.0819)	(0.0611)	(0.1033)	(0.0916)
5	0.0573	0.4344	0.5208	0.0890	0.2966	0.6261	0.1614	0.3841	0.4242
	(0.1634)	(0.1268)	(0.2033)	(0.1760)	(0.1105)	(0.1250)	(0.0986)	(0.1275)	(0.1200)
10	0.0985	0.5790	0.3477	0.1222	0.4834	0.4527	0.1427	0.4789	0.3764
	(0.2853)	(0.2744)	(0.3745)	(0.2989)	(0.2958)	(0.1773)	(0.1147)	(0.1715)	(0.2274)
15	0.1349	0.6763	0.2712	0.1476	0.6193	0.3652	0.0943	0.6949	0.2948
	(0.4153)	(0.5269)	(0.3090)	(0.3565)	(0.4163)	(0.1384)	(0.1315)	(0.1767)	(0.2667)

the announcement does not affect the stock price or cause an unexpected instantaneous stock return. If the firm unexpectedly announces a high-risk project, expected future returns on the firm's stock increase. For this high-risk project to have zero NPV, it also must have a high level of expected profitability. Hence, small firms taking zero-NPV projects with varying levels of risk can generate the positive news-correlation pattern observed in the data. This story can also offer an explanation for the cross-sectional size pattern in the correlation of firm-level expected-return and cash-flow news, if (as assumed in the simple model) a single project is a larger fraction of total assets for small firms than for large firms.

In general, the differences we document in the decomposition as a function of size and BE/ME are not economically large. Although the resolution of the information in book-to-market ratios as well as the dispersion in the ratios varies across the subsets of stocks, our general conclusion that most of the cross-sectional dispersion in the BE/ME ratios is due to differences in expected cash flows is consistently true across the subsets of firms we study.

D.3. Does the Variance Decomposition Vary as a Function of Marketwide Instruments?

The relative importance of the three drivers of the value spread may also vary over time. While there are many ways to estimate a conditional version of (16), we take a simple approach and write the three regression coefficients in (16) as linear functions of variables with intuitive predictive content. We limit ourselves to simple marketwide instruments including the median firm's BE/ME ratio, the cross-sectional variance of firms' BE/ME ratios, the cross-sectional variance of firms' profitability, the cross-sectional covariance of firms' BE/ME ratios and profitability, and bond-yield variables. Our choice of marketwide instruments attempts to use information in a particular cross section of BE/ME ratios to draw inferences about the time-series properties of the information in a typical firm's BE/ME ratio. For example, if the correlation between the BE/ME ratios and firm profitability is higher than normal, one might expect the typical firm's BE/ME ratio to contain more information concerning future expected cash flows than future expected returns. The percentage of information in the BE/ME ratio is relatively constant, perhaps surprisingly so. Our tests cannot identify any periods in time where BE/ME ratios contain relatively more information about expected returns and thus do not report these results.

III. Predicting Value Versus Growth Returns

Consider a portfolio that is long high-BE/ME stocks and short low-BE/ME stocks. Fama and French (1993, 1996) popularize such a zero-investment portfolio (HML) in numerous applications. First, Fama and French construct six value-weight portfolios from the intersections of the two size and the three BE/ME groups. The HML portfolio is then formed by buying both the small and the large high-BE/ME portfolios (combined position denoted by H) and selling short both the small and the large low-BE/ME portfolios (combined position denoted by L). The two components of HML are thus high- and low-BE/ME portfolios with about the same weighted-average size.

Our variance decomposition results can be used to motivate a forecasting model for the return on the HML portfolio. Apply (12) to both the H and L portfolios, take conditional expectations, difference, and reorganize:

$$\sum_{j=0}^{\infty} \rho^j E_{t-1} r_{t+j}^{HML} = (\theta_{t-1}^H - \theta_{t-1}^L) + \left[\sum_{j=0}^{\infty} \rho^j E_{t-1} e_{t+j}^H - \sum_{j=0}^{\infty} \rho^j E_{t-1} e_{t+j}^L\right]. \tag{17}$$

This motivates a predictive regression of HML return on HML value spread (the difference in book-to-market ratios of value stocks in portfolio H and growth stocks in portfolio L):

$$R_t^{HML} = a + b\,(\theta_{t-1}^H - \theta_{t-1}^L) + c\,(e_{t-1}^H - e_{t-1}^L) + \varepsilon_t, \tag{18}$$

where the past profitability spread (i.e., the difference between the log ROEs of the two portfolios) is used as a proxy for the spread in discounted future expected profitability. If we assume that the percentage variance decomposition for the cross section of book-to-market ratios is approximately constant over time, as suggested by our results in Section II, the regression coefficient b in (18) is constant, and variation in the value spread is translated proportionally to variation in expected HML returns.

We present OLS estimates of the coefficients in HML forecasting regressions similar to equation (18). Since most of the empirical work that relies on the time series of returns on the HML portfolio uses simple (not log) returns, we use annual simple returns as the dependent variable in our regression. As one might expect an increase in the expected return on HML to be associated with an increase in the volatility of the HML return, we produce, along with an estimate of the conditional mean HML return, maximum-likelihood GLS estimates based on an accompanying model (using the same instruments) of the log variance of the HML portfolio:

$$R_t^{HML} = Z_{t-1}\beta + \varepsilon_t,\ \ \varepsilon_t \sim N[0, \exp(Z_{t-1}\gamma)] \tag{19}$$

where Z_{t-1} are the lagged predictor variables (including the value spread) and β, γ are parameters. Since the above specification produces conditional variance estimates, it enables us to compute a time series of estimated Sharpe ratios as well as estimated expected returns. For a detailed exposition of the iterative estimation procedure and the standard-error formulas, see Greene (1997, pp. 557–569).

Figure 2 displays the log BE/MEs of the high-, medium-, and low-BE/ME portfolios similar to those created by Fama and French (1993). We deviate from Fama and French's methodology in one respect: We use the market-equity figure at the end of May as the denominator in the book-to-market ratio, and record the post-sort returns from the beginning of June to the end of May (of next year). For the identities to hold, it is necessary that the book-to-market denominator and the beginning of the return period correspond to the same point in time. One can

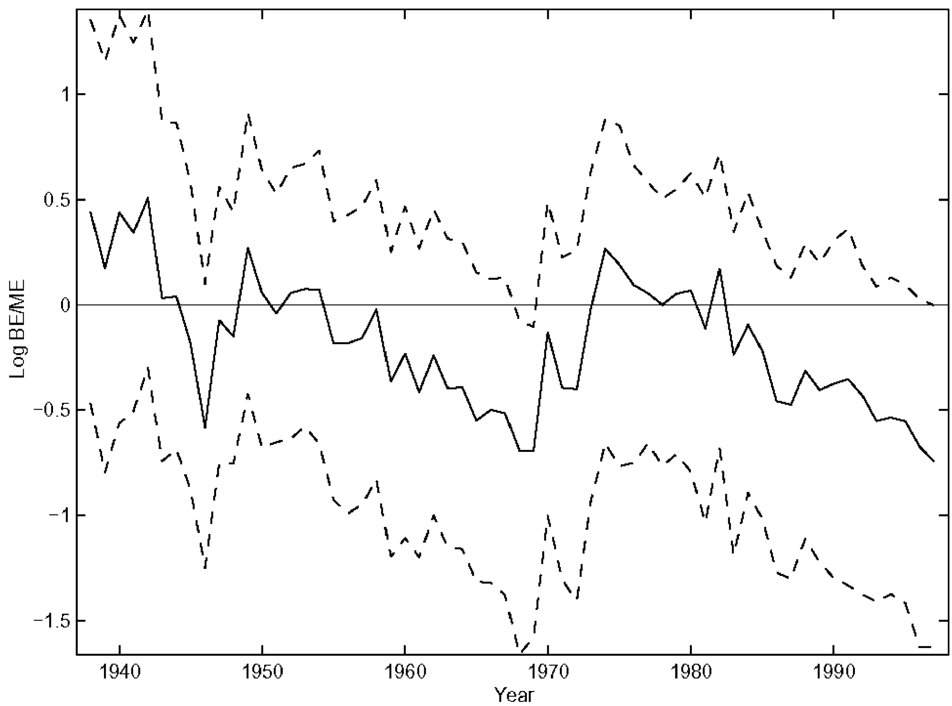

Figure 2. Time-series evolution of BE/MEs (U.S. sample). This figure plots the time-series evolution (1938–1997) of the log book-to-market ratios of H, M, and L portfolios. The portfolios are constructed using Fama and French's (1993) methodology: H is a size-balanced portfolio of high-book-to-market stocks, M of medium-book-to-market stocks, and L of low-book-to-market stocks.

see from the figure that the log BE/MEs of the three size-balanced portfolios move around quite a bit over the 61-year period. While the level of BE/MEs appears volatile and persistent, the HML value spread does appear to be following a mean reverting process.

Table V, Panel A shows the regression results. In that table, we report the coefficients in an OLS regression, the GLS counterpart, and the coefficients (γ) in the exponential-linear conditional-variance model. In the discussion, we focus on the GLS estimates of the conditional mean. As expected, we find that the difference between BE/MEs of the low- and high-BE/ME portfolios, the value spread, is a significant predictor of the return on the HML portfolio. The simple regression coefficient of the value spread is 0.287; the t-statistic is 3.1. This result indicates that the annual expected return on HML is time varying. As the annual standard deviation of the value spread is 8.75 percentage points, this predictive regression implies substantial time variation in the HML premium: The standard deviation of the fitted values of HML is 1.3 times the unconditional mean HML return.

Figure 3 graphs the expected return on the HML portfolio using this specification. As a measure of the economic significance of our finding, Figure 3 also graphs the associated conditional Sharpe ratio. As specified in equation (19), we

Table V

HML return predictability

This table reports regressions predicting the simple return on the HML portfolio. Each regression contains a constant and the HML value spread. In addition, some of the regressions contain combinations of the following variables: the lagged ROE of the HML portfolio, market BE/ME, default yield spread (Moody's BAA less AAA corporate-bond yield), and an interaction term between the value spread and the median BE/ME of the market. The HML value spread is defined as the difference between the log BE/ME of the H portfolio and the log BE/ME of the L portfolio. The HML ROE is defined as the difference between the log ROE of the H portfolio and the log ROE of the L portfolio. These portfolios are constructed following Fama and French's (1993) methodology. The international HML portfolio is country balanced; that is, for each country, the stocks are first sorted into the six elementary portfolios based on country-specific breakpoints and all stocks in each elementary portfolio are value weighted. For each specification, we report three rows. The first row reports the OLS estimates of linear-regression coefficients of HML return on predictor variables. The OLS R^2 in the first row is adjusted for degrees of freedom. The second row reports the maximum-likelihood GLS estimates of the linear-regression coefficients of HML return on predictor variables. The GLS R^2 in the second row is the variance of fitted values (computed using GLS parameter estimates) divided by the sample variance of HML return. The third row shows the maximum-likelihood GLS coefficients in an exponential-linear conditional-variance model, shown in equation (19) of the text. T-statistics are in parentheses.

Panel A: U.S. HML return (1938–1997)

Parameter Estimate (t-Statistic)	Constant	Value Spread	Lagged ROE	Market BE/ME	Default Yield Spread	Value Spread* market BE/ME	R^2	St. dev. $E_{t-1}(R_t)$
OLS return coeff.	− 0.323 (− 2.4)	0.246 (2.8)					0.108	0.040
GLS return coeff.	− 0.376 (− 2.8)	0.287 (3.1)					0.160	0.047
GLS variance coeffi.	− 7.239 (− 4.2)	1.830 (1.6)						
OLS return coeff.	− 0.322 (− 2.4)	0.256 (2.8)	0.001 (0.2)				0.093	0.040
GLS return coeff.	− 0.360 (− 2.7)	0.292 (3.1)	0.003 (0.6)				0.156	0.045
GLS variance coeffi.	− 7.491 (− 4.4)	1.907 (1.6)	− 0.017 (− 0.3)					
OLS return coeff.	− 0.317 (− 2.4)	0.263 (3.0)				0.035 (1.1)	0.111	0.043
GLS return coeff.	− 0.414 (− 2.9)	0.334 (3.2)				0.044 (1.3)	0.226	0.054
GLS variance coeffi.	− 6.971 (− 4.0)	1.957 (1.7)				0.758 (1.8)		
OLS return coeff.	− 0.311 (− 2.4)	0.285 (3.0)	0.004 (0.7)			0.045 (1.3)	0.103	0.044
GLS return coeff.	− 0.425 (− 3.0)	0.376 (3.5)	0.006 (1.0)			0.056 (1.6)	0.254	0.058
GLS variance coeffi.	− 7.537 (− 4.4)	2.452 (2.0)	0.026 (0.4)			0.732 (1.6)		

Table V—*continued*

OLS return coeff.	− 0.339 (− 2.4)	0.269 (2.6)		− 0.011 (− 0.4)	0.094	0.040
GLS return coeff.	− 0.342 (− 2.5)	0.246 (2.4)		0.031 (0.8)	0.186	0.049
GLS variance coeffi.	− 5.508 (− 3.0)	0.092 (0.1)		0.815 (2.0)		
OLS return coeff.	− 0.301 (− 2.3)	0.253 (2.9)	0.055 (1.1)		0.111	0.043
GLS return coeff.	− 0.386 (− 2.8)	0.317 (3.2)	0.075 (1.4)		0.222	0.054
GLS variance coeffi.	− 6.590 (− 3.8)	1.713 (1.5)	1.167 (1.8)			
OLS return coeff.	− 0.380 (− 2.8)	0.388 (3.4)	0.148 (2.1)	− 0.079 (− 1.8)	0.143	0.049
GLS return coeff.	− 0.432 (− 3.0)	0.394 (3.1)	0.128 (1.8)	− 0.041 (− 0.8)	0.249	0.057
GLS variance coeffi.	− 5.899 (− 3.2)	1.563 (0.4)	1.287 (0.3)	− .0578 (− 1.0)		

Panel B: Country-Balanced International HML Return (1983–1998)

Parameter estimate (t-statistic)	Constant	Value Spread	R^2	St. dev. $E_{t-1}(R_t)$
OLS return coeff.	− 0.117 (− 0.5)	0.131 (0.7)	− 0.030	0.013
GLS return coeff.	− 0.151 (− 0.8)	0.159 (1.0)	0.056	0.016
GLS variance coeffi.	− 8.438 (− 1.8)	2.309 (0.6)		

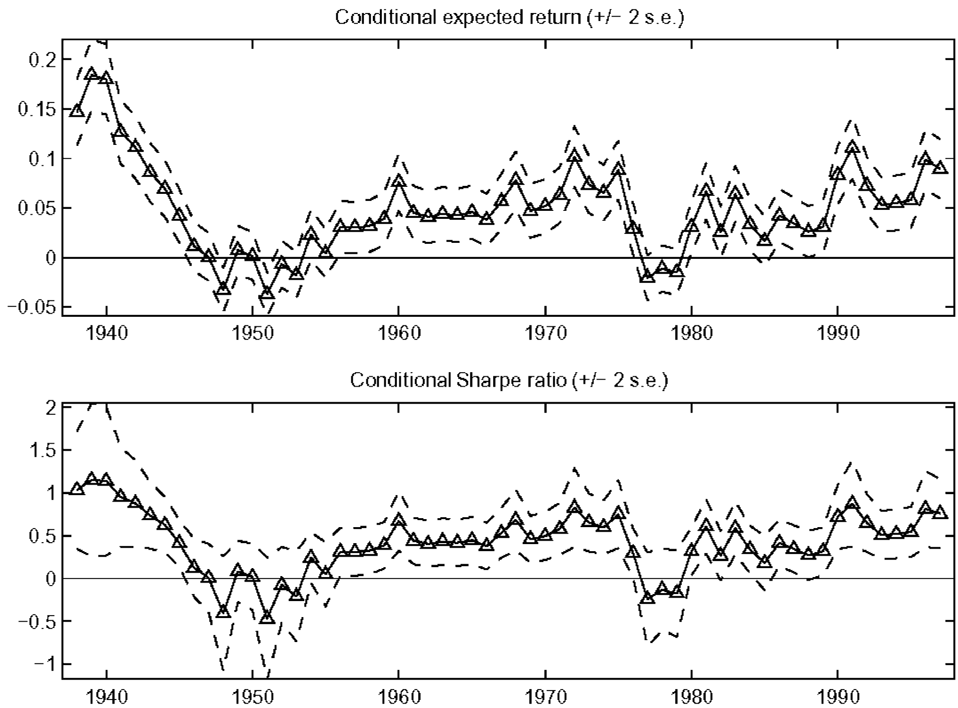

Figure 3. Conditional expected return and Sharpe ratio on HML (U.S. sample). This figure shows a time series of the estimated conditional mean and Sharpe ratio of the HML portfolio. The estimates are produced from maximum-likelihood GLS regressions that relate the expected return and log of the return variance to a constant and the HML value spread using linear regressions.

generate these estimates each year by using the value spread to predict the return and log variance of HML. In general, we find that the point estimates of the HML Sharpe ratio vary considerably over time.

Figure 3 also casts light on an interesting question concerning continuation of the value-minus-growth effect in stock returns. Consider the possibility that the average return on HML was due to a correction of mispricing. Perhaps investors began recognizing these mispricings sometime during the sample period and have been adjusting prices on high- and low-book-to-market stocks, pushing them closer to one another. While the realized returns during this correction may have been high, the expected future return on HML at the end of the sample may be low, and the value-minus-growth anomaly may now be a missed opportunity. This missed-opportunity hypothesis makes a prediction about the level of the value spread: If value stocks are no longer underpriced relative to growth stocks, the cross-sectional spread in book-to-market ratios should have shrunk. Contrary to the prediction of the missed-opportunity hypothesis, the value spread and the fitted values plotted in Figure 3 are slightly above their sample means at the end of the sample (year 1997). Thus, the evidence in the level of the value spread points toward continuation of the value effect and against the missed-opportunity hypothesis.

To examine the robustness of our results to data-snooping concerns, we reestimate this simple forecasting model using the international panel. We look to see whether the international sample provides an out-of-sample test of our finding of time variation in the expected return on value-versus-growth strategies in the U.S. sample. As with the U.S. value spread, our forecasting variable seems well behaved. As the international sample is only 17 years long, we estimate the predictability using country-adjusted variables in the hope of increasing the precision of our regression coefficients. The results, reported in Table V, Panel B, are imprecise, but the coefficients are consistent with the domestic estimates.

Table V, Panel A, also contains multiple regressions with additional predictor variables. In the multiple regression specified by equation (18), the value spread's coefficient and t-statistic are 0.2915 and 3.1, respectively. However, the recent difference in ROE between value and growth stocks does not provide additional predictive power (coefficient of 0.0030, t-statistic 0.6). We also report a specification with the value spread interacted with the median BE/ME ratio of the market. As with the conditional variance decompositions discussed in Section II.D.3, this variable is only marginally significant.

We complete our analysis of time variation in the return on value-versus-growth strategies by investigating what macroeconomic variables are correlated with the value spread. In particular, we regress the value spread on several variables that intuitively might explain movements in the value spread. These regressions are shown in Table VI. We initially regress the value spread on the median BE/ME ratio of stocks in the market. This variable has no explanatory power by itself (t-statistic of -0.12). We turn to the default spread (the difference in yields between BAA and AAA long-term corporate bonds) that, like the median BE/ME ratio, is an indicator of low-frequency movements in business conditions. We find that the value spread has a positive and statistically significant coefficient on the default yield spread. The coefficient is 0.153 with an associated t-statistic of 2.0. Almost a quarter of the variation in the value spread can be linked to movements in this variable alone. We also include the median BE/ME ratio and the lagged profitability spread: The median BE/ME ratio is now statistically and economically significant and the lagged profitability spread is marginally so. The full model explains over 42 percent of the value-spread variance.

When we add both the median BE/ME ratio and the default spread to the return-predictive regressions in Table V, we find that the value spread remains significant (coefficient of 0.394, t-statistic of 3.1), while the median BE/ME ratio is now statistically significant (coefficient of 0.128, t-statistic of 1.8) at the 10 percent level. Unreported orthogonalized regressions show that the component of the value spread that is correlated with the market's BE/ME ratio or default yield spread does not predict HML returns. The predictive ability of the value spread is entirely due to the component that is orthogonal to these marketwide instruments.

Table VI

Explaining the Value Spread

This table reports regressions explaining the spread in the book-to-market ratios of the HML portfolio's components. The regressions contain combinations of the following as explanatory variables: the lagged ROE of the HML portfolio, market BE/ME, and the default yield spread (Moody's BAA less AAA corporate-bond yield). The HML value spread is defined as the difference between the log BE/ME of the H portfolio and the log BE/ME of the L portfolio. The HML ROE is defined as the difference between the log ROE of the H portfolio and the log ROE of the L portfolio. These portfolios are constructed following Fama and French's (1993) methodology. For each specification, we report the OLS estimates of linear-regression coefficients of the HML value spread on predictor variables. The OLS R^2 is adjusted for degrees of freedom. T-statistics are in parentheses.

Constant	Market BE/ME	Default Yield Spread	Lagged ROE	R^2
1.475 (14.6)	− 0.017 (− 0.1)			− 0.016
1.333 (20.4)		0.153 (2.0)		0.239
1.084 (16.4)	− 0.312 (− 3.6)	0.256 (4.6)		0.421
1.043 (15.5)	− 0.322 (− 3.6)	− 0.622 (− 1.1)	0.244(3.9)	0.421

IV. Conclusions

The present-value formula allows us to decompose the cross-sectional variance of firms' book-to-market ratios into three components: (a) covariance of future stock returns with the past book-to-market ratios, (b) covariance of future profitability with the past book-to-market ratios, and (c) persistence of the book-to-market ratios. We estimate this decomposition from a long (1938 to 1997) panel with three simple long-horizon regressions.

Our results suggest that approximately 20 percent of the cross-sectional dispersion in book-to-market ratios can be explained with expected 15-year stock returns, 55 percent with expected 15-year profitability, and 25 percent with 15-year persistence of book-to-market ratios. Intuition and the time-series behavior of stock returns and profitability suggest that the persistence of book-to-market ratios is mostly due to cross-sectional variation in expected profitability beyond the 15-year horizon. Hence, we interpret our regressions as suggesting that approximately 20 to 25 percent of the dispersion in the book-to-market ratios is due to dispersion in expected stock returns and 75 to 80 percent is due to dispersion in expected profitability.

We document similar results for an international panel covering 23 countries (excluding the United States) over the 1982 to 1998 period. As with the domestic panel, most of the variation in the book-to-market ratios, even after adjusting for differences in accounting practices across countries, is due to information concerning expected future profitability.

We also present some time-series evidence on predictability of value-minus-growth returns. Our BE/ME variance decomposition results are constant through time. Therefore, it seems reasonable to conjecture that the expected annual premium on Fama and French's (1993) HML portfolio varies through time as the value spread changes. Our empirical evidence confirms that supposition: The expected return on a value-minus-growth strategy is atypically high at times when the value spread is wide and the market is cheap. The fitted values and the level of the value spread at the end of the sample (1997) suggest that the value-minus-growth effect in stock returns will continue out of the sample.

Appendix A

At first, it may seem that a VAR model would be a simpler and more elegant alternative to our long-horizon regressions. It is tempting to group, for example, HML return, HML value spread, and HML profitability spread as elements of a VAR model's state vector and compute the variance decomposition along the lines of Campbell and Shiller (1988) and Vuolteenaho (2000). It turns out, however, that the economic interpretation of the VAR-based variance decomposition may be materially different from the long-horizon regression variance decomposition we advocate.

The difference between the two methods originates from the fact that the HML portfolio's weights are managed. The key assumption of VARs is that this year's dependent variables are next year's independent variables. Thus, as the definitions of the portfolios change, the VAR incorrectly links the dependent variables for one set of firms to the independent variables for another set. In our long-horizon regressions, all variables are recorded for a fixed set of firms.

To illustrate this point, consider a general managed portfolio series and a VAR model. Let r_t^{t-1}, θ_t^{t-1}, and e_t^{t-1} denote the log return, log book-to-market, and log profitability of the HML portfolio. The superscript in the above variables denotes the time when the buy-and-hold portfolio was formed. Adapting the basic, linearized book-to-market law to this notation yields

$$e_t^{t-1} - r_t^{t-1} \approx \rho\theta_t^{t-1} - \theta_{t-1}^{t-1} \tag{A1}$$

As one can see, the basic identity describes the evolution of a buy-and-hold portfolio's book-to-market.

The objective in our variance decomposition is to measure how θ_{t-1}^{t-1} forecasts r_t^{t-1}, θ_t^{t-1}, and e_t^{t-1}. However, running a VAR with r_t^{t-1}, θ_t^t, and e_t^{t-1} in the state vector will relate θ_{t-1}^{t-1} to r_t^{t-1}, θ_t^t, and e_t^{t-1}, not to r_t^{t-1}, θ_t^{t-1}, and e_t^{t-1}, because, in a VAR, this year's dependent variables are by assumption next year's forecasting variables. Such a VAR will try to forecast both what firms will be included in the managed portfolio and what their book-to-markets will be, not just the latter. Our long-horizon regression, however, will always track the future evolution of returns, book-to-markets, and profitabilities for a single set of firms.

To illustrate the link between a VAR and our long-horizon regressions, we include an additional term due to rebalancing effects to the equation describing the book-to-market evolution of a managed portfolio. Adding $\rho(\theta_t^t - \theta_t^{t-1})$ to both sides of (A1) results in

$$e_t^{t-1} - r_t^{t-1} + \rho(\theta_t^t - \theta_t^{t-1}) \approx \rho\theta_t^t - \theta_{t-1}^{t-1} \tag{A2}$$

Equation (A2) describes the evolution of a managed portfolio series. These terms explain the change in a managed portfolio's book-to-market: managed portfolio profitability, managed portfolio return, and a rebalancing term.

Comparison of equations (A1) and (A2) highlights the difference between a VAR variance decomposition and a long-horizon regression variance decomposition. A long-horizon regression is based on iteration of (A1). A VAR-model maps a vector $[r_{t-1}^{t-2}, \theta_{t-1}^{t-1}, e_{t-1}^{t-2}]$ to $[r_t^{t-1}, \theta_t^t, e_t^{t-1}]$ and, thus, can only implement a variance decomposition based on iterating equation (A2). The more aggressively the portfolio is rebalanced, the larger the difference between the two.

Experiments with a first-order VAR specification indicate that the covariance term,

$$\frac{\operatorname{cov}\left[\sum_{i=0}^{\infty} \rho^{i+1}\left(\theta_{t+i}^{t+i} - \theta_{t+i}^{t+i-1}\right),\ \theta_{t-1}^{t-1}\right]}{\operatorname{var}\theta_{t-1}^{t-1}} \tag{A3}$$

is economically quite large (contributing approximately 30 percent of the total cross-sectional variance). Thus, interpretation of a VAR-based variance decomposition estimated from data corresponding to a managed portfolio is difficult.

Appendix B

In many finance applications, the available data set contains perhaps 70 overlapping cross sections, each with hundreds or even thousands of data points. In such cases, incorrectly assuming that the errors covariance matrix is proportional to an identity matrix can yield standard errors that are severely biased downwards. This bias is due to the fact that error correlations are often systematically related to the explanatory variables. Fortunately, the statistics literature has proposed a solution for a similar problem frequently arising with complex surveys: Rogers's (1983, 1993) robust standard errors. Compared to the popular Fama–MacBeth (1973) procedure, this method has the practical advantage of giving the standard errors for pooled-OLS/WLS coefficients—allowing for, among other things, the use of common time-series variables in the regressions.

A simple exposition of Rogers's (1983, 1993) standard errors starts from the familiar formula for OLS standard errors. Let X denote the panel of explanatory variables, Ω the covariance matrix of the panel of errors, and $X_{t-N+1,t+N-1}$ and $\Omega_{t-N+1,t+N-1}$ a single cluster of explanatory variables and the corresponding error covariance matrix. A cluster is defined as the set of cross sections whose errors are correlated with the errors of the year-t cross section. Assuming that the year-t errors may be dependent within a cluster but are independent across clusters allows writing

$$\begin{aligned}&(X'X)^{-1}X'\Omega X(X'X)^{-1} = \\ &(X'X)^{-1}\sum_{t=1}^{T}\left[X'_{t-N+1,t+N-1}\Omega_{t-N+1,t+N-1}X_{t-N+1,t+N-1}\right](X'X)^{-1}\end{aligned} \tag{B1}$$

We denote regression errors by ε, and notation for fitted values is modified with a hat. As $X'_{t-N+1,t+N-1}\Omega_{t-N+1,t+N11}X_{t-N+1,t+N-1} = E(X'_{t-N+1,t+N-1}\varepsilon_{t-N+1,t+N-1}\varepsilon'_{t-N+1,t+N-1}X_{t-N+1,t+N-1})$, Rogers's standard errors are computed by substituting in-sample estimates of the errors for true errors to get an in-sample variance estimator of regression coefficients:

$$(X'X)^{-1}\sum_{t=1}^{T}\left[X'_{t-N+1,t+N-1}\hat{\varepsilon}_{t-N+1,t+N-1}\hat{\varepsilon}'_{t-N+1,t+N-1}X_{t-N+1,t+N-1}\right](X'X)^{-1} \quad \text{(B2)}$$

Under plausible assumptions, these standard errors are consistent in T; that is, they converge as the time dimension of the panel grows. To ensure that the effect of a single cluster on the coefficient estimates vanishes as more and more clusters are included, Rogers's assumptions include that the time series of $X'_{t-N+1,t+N-1}\varepsilon_{t-N+1,t+N-1}\varepsilon'_{t-N+1,t+N-1}X_{t-N+1,t+N-1}$ is well behaved.

The above standard-error formula can be interpreted as generalized White's standard errors; in the special case of only one observation per cluster (e.g., a univariate time series with serially uncorrelated errors), the standard errors are equivalent to White (1980) heteroskedasticity consistent standard errors. The method can also be interpreted as an application of Hansen's (1982) generalized method of moments or as a multivariate generalization of Hansen–Hodrick (1980) standard errors.

References

Asness, Clifford, Jacques Friedman, Robert Krail, and John Liew, 2000, Style timing: Value versus growth, *Journal of Portfolio Management* 26, 50–60.

Asness, Clifford, R. Burt Porter, and Ross Stevens, 2000, Predicting stock returns using industry relative firm characteristics, Working paper, University of Florida.

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

Cochrane, John H., 1992, Explaining the variance of price-dividend ratios, *Review of Financial Studies* 5, 243–280.

Cohen, Randolph B., and Christopher Polk, 1998, The impact of industry factors in stock returns, Working paper, Kellogg School of Management.

Davis, James L., Eugene F. Fama, and Kenneth R. French, 2000, Characteristics, covariances, and average returns: 1929 to 1997, *Journal of Finance* 55, 389–406.

Fama, Eugene F., and Kenneth R. French, 1992, The cross section of expected stock returns, *Journal of Finance* 47, 427–465.

Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3–56.

Fama, Eugene F., and Kenneth R. French, 1995, Size and book-to-market factors in earnings and returns, *Journal of Finance* 50, 131–155.

Fama, Eugene F., and Kenneth R. French, 1996, Multifactor explanations of asset pricing anomalies, *Journal of Finance* 51, 55–84.

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

Feltham, Gerald A., and James A. Ohlson, 1999, Residual earnings valuation with risk and stochastic interest rates, *Accounting Review* 74, 2–23.

Frankel, Richard, and Charles M. C. Lee, 1999, Accounting diversity and international valuation, Working paper, Johnson Graduate School of Management.

Greene, William H., 1997, *Econometric Analysis* (Prentice Hall, Upper Saddle River, NY).

Hansen, Lars P., 1982, Large sample properties of generalized methods of moments estimators, *Econometrica* 50, 1029–1054.

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

Lamont, Owens, and Christopher Polk, 2001, The diversification discount: Cash flows versus returns, *Journal of Finance* 56, 1693–1721.

Lewellen, Jonathan, 1999, The time-series relations among expected return, risk, and book-to-market, *Journal of Financial Economics* 54, 5–44.

Merino, Barbara, and Alan Mayper, 1999, The rhetoric and reality of American dream: Securities legislation and the accounting profession in the 1930s, Working paper, University of North Texas.

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

Rogers, W. H., 1983, Analyzing complex survey data, Rand Corporation memorandum. Santa Monica, CA.

Rogers, W. H., 1993, Regression standard errors in clustered samples, Stata Technical Bulletin Reprints STB-13–STB-18, 88–94.

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

Vuolteenaho, Tuomo, 2000, 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 returns? *Journal of Finance* 57, 233–264.

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