Earnings, retained earnings, and book-to-market in the cross section of expected returns*

Ray Ball[†1], Joseph Gerakos[2], Juhani T. Linnainmaa[3,4] and Valeri Nikolaev[1]

[1]University of Chicago Booth School of Business, United States

[2]Tuck School of Business at Dartmouth College, United States

[3]University of Southern California Marshall School of Business, United States

[4]National Bureau of Economic Research, United States

September 25, 2017

Abstract

We delve into what causes the relation between book-to-market and the cross section of stock returns. Book value of equity consists of two main components that we expect contain different information about expected returns: retained earnings and contributed capital. Retained earnings-to-market subsumes book-to-market's power to predict the cross section of stock returns in pre- and post-Compustat U.S. data as well as in international data. Contributed capital-to-market has no predictive power. Retained earnings represent the accumulated difference between earnings and dividends. We show that retained earnings-to-market's predictive power stems entirely from accumulated earnings. Our thesis is that retained earnings-to-market—and, by extension, book-to-market—predicts returns because it is a good proxy for earnings yield (Ball, 1978; Berk, 1995).

JEL classification: G11, G12, M41.

Keywords: Book-to-market; Contributed capital; Earnings yield; Mispricing; Retained earnings; Value premium.

*Ray Ball is a trustee of the Harbor Funds. Joseph Gerakos is affiliated with Citadel LLC. Juhani Linnainmaa is affiliated with Citadel LLC and Research Affiliates. Harbor Funds, Citadel LLC, and Research Affiliates did not provide funding or data for this study and none of the authors has a financial interest in the outcomes of this research. We thank Dirk Black, Vedran Capkan (discussant), Mikhail Chernov (discussant), Eugene Fama, Jon Lewellen, Clemens Sialm, the referee, and workshop and conference participants at Cubist Systematic Strategies, the 2017 Midwest Finance Association Annual Meetings, the 14th Annual Conference in Financial Economics Research by Eagle Labs, the 2017 Macquarie Quantitative Conference, and the University of Iowa for their comments. We thank Michael Roberts for providing the historical accounting data.

[†]Corresponding author. Mailing address: University of Chicago Booth School of Business, 5807 South Woodlawn Avenue, Chicago, IL 60637, United States. E-mail address: Ray.Ball@ChicagoBooth.edu. Telephone number: +1 (773) 834-5941.

Earnings, retained earnings, and book-to-market in the cross section of expected returns

Abstract

We delve into what causes the relation between book-to-market and the cross section of stock returns. Book value of equity consists of two main components that we expect contain different information about expected returns: retained earnings and contributed capital. Retained earnings-to-market subsumes book-to-market's power to predict the cross section of stock returns in pre- and post-Compustat U.S. data as well as in international data. Contributed capital-to-market has no predictive power. Retained earnings represent the accumulated difference between earnings and dividends. We show that retained earnings-to-market's predictive power stems entirely from accumulated earnings. Our thesis is that retained earnings-to-market—and, by extension, book-to-market—predicts returns because it is a good proxy for earnings yield (Ball, 1978; Berk, 1995).

JEL classification: G11, G12, M41.

Keywords: Book-to-market; Contributed capital; Earnings yield; Mispricing; Retained earnings; Value premium.

1 Introduction

We delve into what causes the relation between book-to-market and the cross section of stock returns. Book value of equity consists of two main parts that we expect contain different information about expected returns: retained earnings and contributed capital. We show that retained earnings-to-market subsumes book-to-market's predictive power in the cross section of stocks returns. Contributed capital-to-market, by contrast, has no predictive power. This result holds not only in the 1964–2016 U.S. data and its two halves, but also in international and pre-Compustat U.S. data. We show that book-to-market lost predictive power after 1990 because its correlation with retained earnings-to-market dropped substantially. Moreover, although earnings-to-price is informative about returns even when controlling for book-to-market, retained earnings-to-market subsumes this predictor as well. Our results suggest that book-to-market predicts returns because it is correlated with retained earnings-to-market, which is a good proxy for the firm's underlying earnings yield (Ball, 1978; Berk, 1995).

Retained earnings comprise the accumulation of the total earnings a firm generated over its history to date less accumulated dividend distributions.[1] We expect retained earnings-to-market to be informative about expected returns for two reasons. First, when deflated by market values, earnings likely share common economic determinants with expected stock returns (Ball, 1978; Berk, 1995). Second, accounting effects, which reduce the informativeness of bottom-line net income (Novy-Marx, 2013; Ball, Gerakos, Linnainmaa, and Nikolaev, 2015, 2016), effectively "average out" through accumulation because noise in accounting accrual estimates reverses in subsequent periods. For example, over-estimates of uncollectible receivables reduce current-year earnings, but the effect reverses when the error is corrected. Similarly, one-time items such as asset impairment charges impact current-period earnings, but they substitute for depreciation or other charges that would have been made against earnings over an extended period. Hence, individual-year accounting effects have an attenuated effect on the accumulated past earnings component of retained earnings.

[1] "Earnings" here refer to bottom-line net income, which is what accountants transfer into retained earnings each period. As described below, some stock repurchase transactions reduce retained earnings.

We therefore expect retained earnings-to-market to be a better proxy than current-period earnings-to-price for the economic determinants of expected stock returns.

Accumulated past equity issuances less past repurchases compose the contributed capital component of the book value of equity. As a first-order effect, we see no reason to expect net capital contributed over the life of the firm to be substantially related to the cross section of returns. The fact that investors contributed capital to a firm does not reveal information about its risk class (or other determinants of expected return); it merely indicates that investors are prepared to accept them. Recent net issuances could, however, lead to a negative relation between contributed capital and stock returns (i.e., the "net issuance anomaly").[2] We therefore expect either no relation or a weak negative relation between contributed capital and the cross section of expected returns.

The effect of deflators on the correlation between the two main book value components provides an early indication that contributed capital and retained earnings contain different information about equity valuation. When they are deflated by the book value of equity, contributed capital and retained earnings are almost perfectly negatively correlated, because, on average, they jointly represent 95% of the book value deflator.[3] However, when deflated by market value of equity, the correlation between the two measures is -0.19. Introducing price in the denominator substantially alters the relation between the variables, suggesting that the two components contain different information about equity valuation.

We start by estimating Fama and MacBeth (1973) regressions to compare the information in retained earnings and contributed capital. We find that retained earnings-to-market subsumes book-to-market in explaining the cross section of average returns. By contrast, contributed capital-to-market has no explanatory power over the cross section of stock returns either on its own or when controlling for retained earnings-to-market. Moreover, when we include both book-to-market and contributed capital-to-market, they are both significant but with opposing signs, and the difference

[2]See Ikenberry, Lakonishok, and Vermaelen (1995), Loughran and Ritter (1995), Daniel and Titman (2006), and Pontiff and Woodgate (2008).

[3]A third component is "Accumulated Other Comprehensive Income," which is an accumulation of largely transitory past gains and losses. We describe this component in the following section. On average, it represents 5% of the book value of equity. We find that—as is the case for contributed capital—the ratio of this component to the market value of equity has no explanatory power for the cross section of expected returns.

between the estimated coefficients is similar to our estimate for retained earnings-to-market on its own. Thus, book-to-market only predicts stock returns because it contains retained earnings.

In spanning regressions, we show that a retained earnings-based factor contains all of the valuable information in current-period's earnings-to-price. This result holds in both halves of our 1964–2016 sample period. It is consistent with our thesis that the accumulation of past earnings in retained earnings reduces timing issues in accounting that affect individual-year earnings, making retained earnings-to-market a better proxy for underlying earnings yield.

Retained earnings-to-market reliably explains the cross section of average returns further into the future than book-to-market itself. Its explanatory power persists at least five years ahead. If book-to-market predicts returns by assisting in identifying over- and under-valued securities, our results imply the market must be mispricing earnings, and that this mispricing takes years to correct.

We next dig deeper into the composition of the book value of equity and investigate the accumulated earnings and accumulated dividends components of retained earnings. Here our evidence indicates that retained earnings only predict stock returns because they contain past earnings. The accumulated dividends component of retained earnings is uninformative. This result is consistent with book-to-market's explanatory power arising because it is correlated with accumulated earnings and therefore provides a good proxy for underlying earnings yield.

We confirm this interpretation by conducting our own accumulation of past earnings over increasing windows, ending with the most recent year and extending back up to five years. In Fama and MacBeth (1973) regressions with earnings accumulated backward over increasing windows, the coefficient on this backward-accumulated earnings measure increases monotonically with the earnings accumulation window and its t-value increases almost monotonically. The coefficient and t-value for book-to-market decrease monotonically, losing significance after earnings is accumulated back two years.

Several additional results are consistent with our thesis and provide comfort that our results do not represent statistical artifacts (Lo and MacKinlay, 1990; Lewellen, Nagel, and Shanken, 2010;

Harvey, Liu, and Zhu, 2016). First, we find similar results for non-U.S. firms. Second, when we split the U.S. sample into sub-periods, we find that retained earnings-to-market predicts returns both pre- and post-1990 even though book-to-market fails to do so in the latter sub-period. We show that book-to-market loses its predictive power in the latter sub-period because its correlation with retained earnings-to-market drops. Third, we obtain similar results for U.S. firms over a pre-Compustat period from 1938 through 1964.

We draw the following conclusions. First, book-to-market explains the cross section of average returns only because of its retained earnings-to-market component. Second, retained earnings-to-market has explanatory power because retained earnings attenuate accounting effects on individual-year earnings yields, which in turn share common economic determinants with expected stock returns (Ball, 1978; Berk, 1995).[4] Although we find that retained earnings-to-market is a good proxy for a firm's underlying earnings yield, we do not claim that retained earnings is better predictor of future earnings and cash flows than the myriad of variables examined in prior research.[5] Instead, our goal is to demonstrate the source of book-to-market's predictive power for the cross section of stock returns.

Third, we cast doubt on whether book value of equity functions in value strategies as an indicator of fundamental value. Why would a minority component of book value be a useful benchmark for the entire market value, especially when (scaled by market value) it is almost uncorrelated with the other components?

Our evidence highlights the value of digging into accounting numbers, whose components generally contain different information about the cross section of stock returns. Novy-Marx (2013) and Ball, Gerakos, Linnainmaa, and Nikolaev (2015, 2016) show that decomposing "bottom line" earnings into operating versus non-operating components and into accruals versus cash flow com-

[4]Our results can be thought of as a cross-sectional version of Shiller's cyclically adjusted price-to-earnings (CAPE) ratio, which uses a moving average of the S&P 500 aggregate earnings to predict the market return (Campbell and Shiller, 1988). Just as the CAPE smooths out transitory shocks to aggregate income, retained earnings smooths out transitory shocks at the firm-level.

[5]For example, the implied cost of capital literature combines forecasts of earnings or cash flows with current prices to estimate expected returns (Easton, 2009; Etcherling, Eierle, and Ketterer, 2015).

ponents increases the predictive power over the cross section of average returns. This study obtains similar insights from decomposing "bottom line" book value of equity.

2 Book value of equity, retained earnings, and contributed capital

Our basic thesis is that not all of the components of the book-to-market ratio's numerator—the book value of equity—contain the same information about the cross section of stock returns, and that differences in their informativeness shed light on the source and interpretation of book-to-market's predictive ability for the cross section of returns.

The book value of common equity can be decomposed as follows, with the Compustat data items in brackets:

$$\text{Common/Ordinary Equity [CEQ]} = \text{Contributed capital} + \text{Retained earnings} + \text{Other},$$

where:

$$\begin{aligned}
\text{Contributed capital} &= \text{Common/Ordinary Stock [CSTK]} \\
&\quad + \text{Capital Surplus/Share Premium Reserve [CAPS]} \\
&\quad - \text{Treasury Stock [TSTK]} \\
\text{Retained earnings} &= \text{Retained Earnings [RE]} \\
&\quad - \text{Accumulated Other Comprehensive Income [ACOMINC]} \\
\text{Other} &= \text{Accumulated Other Comprehensive Income [ACOMINC].}
\end{aligned}$$

Contributed capital represents the net contribution of capital from shareholders that a firm receives from issuances and repurchases of its common stock. It consists of capital contributions that are recorded as the par value of common stock [CSTK] plus capital paid in excess of par value (i.e., "additional paid in capital") [CAPS], net of returns of capital to shareholders that are recorded as the book value of treasury stock [TSTK]. The par value of common stock and the capital paid in excess of par value do not differ in economically important ways and can therefore

be combined.[6] Treasury stock is the cost of stock repurchased from shareholders (but not retired). If the firm does not retire the repurchased stock, the repurchase cost is reported on the balance sheet as a negative offset to contributed capital. Firms typically do not retire the repurchased stock because cancellation forgoes options to reissue the stock on the market or under executive compensation schemes.

Retained earnings are the earnings (i.e., net income) accumulated since the firm's inception less accumulated distributed dividends. Retained earnings increase as the firm generates and books earnings, and decrease as the firm books losses or declares dividends. A corporation cannot create earnings through trading in its own capital stock, so treasury stock transactions never increase retained earnings. These transactions can, however, reduce retained earnings. As discussed in the appendix, companies can account for treasury stock using either the cost or par method. Under the cost method, retained earnings are reduced when previously repurchased treasury stock is reissued at a price lower than that paid to repurchase it.[7] Under the par method, retained earnings are reduced by the difference between the repurchase price and the amount originally received when the stock was issued.[8] Retained earnings can therefore become negative if a firm generates a series of book losses either during a growth phase or due to poor economic performance, or it engages in certain stock repurchase transactions.

A third and typically much smaller component of the book value of equity is accumulated other comprehensive income [ACOMINC]. Accumulated other comprehensive income is a technical account that accumulates the amount of various paper (i.e., not realized in cash) gains and losses that primarily originate in transitory shocks to prices of financial assets in which companies have either long or short positions. Accounting rules exclude these shocks from earnings until they

[6]For example, the par value of common stock is commonly set to an arbitrarily small amount such as one dollar or one cent. This practice circumvents restrictions in some jurisdictions against issuing stock at a price below par.

[7]Treasury stock transactions can reduce retained earnings by economically significant amounts. Microsoft, for example, historically granted generous amounts of stock options to employees and also purchased substantial amounts of treasury shares. The options generally had exercise prices substantially below the cost to Microsoft of buying back the shares that were reissued to employees. This difference between the repurchase and strike prices largely explains the $29.46 billion retained deficit it reported at June 30, 2007.

[8]The cost method appears to be more prevalent method. In 2010, the American Institute of Certified Public Accountants surveyed 500 firms about their accounting policies (AICPA, 2010). Of these 500 firms, 340 engaged in stock repurchases with 321 using the cost method and 19 using the par method.

subsequently are realized, and "park them" in the meantime in a separate book value-of-equity account. These items include unrealized gains and losses on those marketable securities that are designated as "securities available for sale," unrealized gains and losses on cash flow hedging instruments, unrealized gains and losses on pension plan assets net of liabilities, and foreign currency translation adjustments. If these paper gains and losses are later realized, they are removed from accumulated other comprehensive income, recognized as earnings, and then moved into retained earnings. If they are not realized, they remain on the balance in accumulated other comprehensive income.

Accumulated other comprehensive income measures price changes and liabilities that are largely unrelated to firms' operations (we remove financial firms from our sample). We therefore do not expect this component to be informative about firms' expected returns. Even though U.S. GAAP does not include accumulated other comprehensive income in retained earnings, and firms therefore report retained earnings without it, Compustat adds ACOMINC to their retained earnings variable [RE]. Because we expect this component to differ from retained earnings in terms of informativeness about the cross section of stock returns, we back it out from retained earnings and study its contribution separately.

Book value of equity therefore evolves over time as a function of net capital transactions with shareholders (new issuances less treasury stock purchases), net earnings retention (earnings less dividends), and some transitory gains and losses due to shocks to asset prices. Consequently, book-to-market ratios consist of several components with potentially different implications for asset pricing. In particular, retained earnings (when scaled by current market equity) likely is a proxy for underlying earnings yield and hence for expected returns. We expect no such effect for contributed capital and accumulated other comprehensive income.

3 Data

Our primary sample is U.S.-listed securities over 1964–2016. (We also study non-U.S. markets and different periods).We take monthly stock returns and dividend histories from the Center for

Research in Security Prices (CRSP) and annual accounting data from Compustat. We start our sample with all firms traded on NYSE, Amex, and NASDAQ, and exclude securities other than ordinary common shares. We exclude financial firms, which are defined as firms with one-digit standard industrial classification codes of six. Delisting returns are taken from CRSP; if a delisting return is missing and the delisting is performance-related, we impute a return of -30% for NYSE and Amex firms and -55% for Nasdaq firms (Shumway, 1997; Shumway and Warther, 1999; Beaver, McNichols, and Price, 2007). We match the firms on CRSP against Compustat, and lag annual accounting information by six months. For example, if a firm's fiscal year ends in December, we assume that this information is public by the end of the following June. We start our sample in July 1964 and end it in December 2016. We start the sample in 1964 as opposed to the usual 1963 start year because Compustat did not collect the retained earnings variable for most firms until the 1963 fiscal year.[9] The sample consists of firms with non-missing market value of equity, book-to-market, current month returns, and returns for the prior one-year period.

In Fama and MacBeth (1973) regressions, we exclude microcaps to avoid having them exert undue influence (Novy-Marx, 2013). Following Fama and French (2008), we define microcaps as stocks with market values of equity below the 20^{th} percentile of the NYSE market capitalization distribution. These comprise only 3.2% of aggregate market capitalization. In Fama and MacBeth (1973) regressions, we re-compute the explanatory variables every month. In portfolio sorts, and when constructing return factors, we include all stocks and rebalance the portfolios annually at the end of June.

We generate two measures of book-to-market that differ in their numerators. First, we follow Fama and French and calculate the book value of equity as shareholders' equity, plus balance sheet deferred taxes, plus balance sheet investment tax credits, plus postretirement benefit liabilities, and minus preferred stock. We set missing values of balance sheet deferred taxes and investment tax credits equal to zero. To calculate the value of preferred stock, we set it equal to the redemption

[9]Compustat reports retained earnings (RE) for just 27.0% of the firms for fiscal years ending during 1962. For fiscal years ending in 1963, this fraction is 86.5%. From 1964 through the end of the sample in 2016, this fraction is almost always above 99%.

value if available, or else the liquidation value or the carrying value, in that order. If shareholders' equity is missing, we set it equal to the value of common equity if available, or total assets minus total liabilities.[10] We then use the Davis, Fama, and French (2000) book values of equity from Ken French's website to fill in missing values.

For the second measure of book-to-market, we use Compustat's book value of common shareholders' equity measure [CEQ]. The benefits of the second measure are that it reflects the book value of common equity reported on firms' balance sheets, and that it can be exactly broken down into the components that we expect to be differently priced, including contributed capital and retained earnings.

Table 1 presents descriptive statistics for book-to-market and its components. Panel A presents time-series averages of the distributions. The first two rows compare the Fama and French book-to-market measure and book-to-market based on the book value of equity reported on balance sheets. The distributions of the two measures are almost identical with means of 0.83 and 0.79 and medians of 0.68 and 0.64. The third row describes the distribution of retained earnings-to-market. This measure is more skewed, with a mean of -0.14 and a median of 0.24.

The next two rows present the distributions of contributed capital-to-market and accumulated other comprehensive income-to-market. The mean of contributed capital-to-market is similar to the means of the book-to-market measures (0.94 vs. 0.83 and 0.79), but the median is smaller (0.38 vs. 0.68 and 0.64). Accumulated other comprehensive income-to-market is the smallest component with a mean of 0.03 and a median of 0.01.

Panel B presents the distributions of the components as a percentage of the book value of equity.[11] On average, contributed capital represents a larger percentage of the book value of equity (54%). Fewer than four percent of the sample firms have negative contributed capital. On average, retained earnings are 41% of the book value of equity, and 15% of the sample firms have negative

[10] See `http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data_Library/variable_definitions.html` and Cohen, Polk, and Vuolteenaho (2003, p. 613) for a detailed discussion of how this book value of equity measure is calculated.

[11] In Panel B, we restrict the sample to firms with positive book values of equity and drop firms for which a single component accounts for more than 200% or less than -100% of the book value of equity. These restrictions are why the fractions of negative values (reported in column "Pct < 0") in Panel A differ from those reported in Panel B.

retained earnings. Accumulated other comprehensive income represents the smallest share of the book value of equity, with a mean of 5%.

Panel C presents the distributions of the ratios of the components to the total book value of equity for the six Fama and French portfolios. Several patterns emerge from these distributions. First, larger firms have higher percentages of retained earnings, lower frequencies of negative retained earnings, and lower percentages of contributed capital. Second, of the six portfolios, large growth firms have the highest percentage of retained earnings (mean = 62%) and the lowest percentage of contributed capital (mean = 35%) while small growth firms have the lowest percentage of retained earnings (mean = 28%) and the highest percentage of contributed capital (mean = 71%).

Panel D presents Pearson and Spearman correlations for book-to-market, retained earnings-to-market, contributed capital-to-market, and accumulated other comprehensive income-to-market. All three components positively correlate with book-to-market. Because retained earnings and contributed capital comprise 95% of book value of equity, they would be almost perfectly negatively correlated when deflated by book value of equity. In the data (not reported in Panel D), the Pearson and Spearman correlations are -0.96 and -0.95. Retained earnings-to-market and contributed capital-to-market, however, are only slightly negatively correlated; the Pearson and Spearman correlations -0.19 and -0.17. These correlations indicate that each explains less than four percent of the other's variance. Moreover, the change in correlations when switching the denominator from the book value of equity to the market value of equity implies that they contain different information about market values. That is, the market appears to *price* these two components differently. This difference, in turn, suggests that they contain different information about expected stock returns.

4 The cross section of returns

4.1 Retained earnings versus contributed capital

Panel A of Table 2 presents our Fama and MacBeth (1973) regressions that use the book-to-market definition as per Fama and French. We require that the book value of equity is positive and we take the natural logarithm of all ratios. Column 1 reports the baseline regressions that include as control variables size, prior one-month return, and prior one-year return skipping a month. As expected, book-to-market is statistically significant (coefficient of 0.23 with a t-value of 3.31) and in line with prior studies.

In column 2, we replace book-to-market with retained earnings-to-market and an indicator variable for negative retained earnings.[12] The coefficient on retained earnings-to-market is positive and its t-value is greater than the t-value for book-to-market (4.47 vs. 3.31). In column 3, we include book-to-market along with retained earnings-to-market. Book-to-market is no longer statistically significant—its t-value is 0.79. By contrast, the t-value on retained earnings-to-market is 4.66, showing that the information in retained earnings-to-market subsumes the information in book-to-market.[13]

In columns 4, and 5, we run similar regressions that replace retained earnings-to-market with contributed capital-to-market. When contributed capital-to-market is included on its own, it is statistically insignificant (coefficient of 0.00 with a t-value of 0.08). When we include it along with book-to-market in column 5, book-to-market's coefficient and t-value are larger than when book-to-market is included on its own in column 1 (coefficients, 0.34 versus 0.23; t-values, 4.26 versus 3.31). Moreover, contributed capital-to-market is now negative and significant (coefficient of -0.12 with a t-value of -4.18). These estimates imply that it is the difference—that is, retained earnings-to-market—that predicts returns.

In columns 6 and 7, we repeat the regressions with accumulated other comprehensive income-to-

[12] When retained earnings are negative, we replace the log of retained earnings-to-market with zero and include an indicator variable for negative values.

[13] This result is for all but microcaps. Controlling for retained earnings-to-market, book-to-market has predictive power among microcaps.

market. When we include both measures in column 7, accumulated other comprehensive income-to-market is statistically insignificant and the t-value on book-to-market is similar in magnitude to that in column 2. In column 8, we include retained earnings-to-market, contributed capital-to-market, and accumulated other comprehensive income-to-market along with indicators for negative values for the three ratios. In this specification, only retained earnings-to-market is significant (t-value = 4.91).

Panel B of Table 2 repeats the analysis in Panel A with the book-to-market ratio based on the book value of equity reported on the firm's balance sheet (i.e., without the adjustments implemented by Fama and French). Reported book-to-market also is statistically significant, with a coefficient of 0.24 (vs. 0.23) and a t-value of 3.49 (vs. 3.31). Similar to Panel A, book-to-market loses its statistical significance when we include retained earnings-to-market in the regression (coefficient of 0.05 with t-value of 0.70). By contrast, retained earnings-to-market is highly significant with coefficient of 0.18 and t-value of 4.67. The estimates in the remaining columns closely mimic the findings in Panel A. This analysis suggests that the findings above are not sensitive to the definition of the book value of equity.

Overall, the regressions in Panels A and B show that book-to-market predicts the cross section of returns only because it contains retained earnings. The other components of the book value of equity provide no significant information about the cross section of average returns, and removing these other components (i.e., using retained earnings alone) increases the t-value. These results are consistent with our thesis and imply that in asset pricing tests it is preferable to use the retained earnings-to-market ratio rather than the book-to-market ratio.

4.2 Addressing negative observations

In Panel C of Table 2, we examine whether the previous results are sensitive to how we handle negative observations of the book value of equity and retained earnings. In columns 1 and 2, we expand Panel A's sample to include firms with negative book values of equity. When the book value of equity is negative, we replace the log of book-to-market with zero and include an indicator

variable for negative values. In column 1, the coefficient and t-value for book-to-market are same as those presented in Panel A. When we add retained earnings-to-market in column 2, book-to-market is statistically insignificant (t-value = 0.84) and retained earnings-to-market is again significant (t-value = 4.58).

Our main results in Panel A include firms with negative retained earnings, and we control for such firms with an indicator variable. In columns 3 and 4 we present specifications that examine whether this treatment of negative values influences our estimates. It could be, for example, that the non-linearity arising from the indicator variable affects the coefficients on retained earnings-to-market. Columns 3 and 4 restrict the sample to firm-years with positive retained earnings. Again, our results for retained earnings-to-market remain essentially unchanged (t-value = 4.85).

4.3 Earnings-to-price

Retained earnings-to-market, by definition, relates to another predictor of stock returns, earnings-to-price. The difference between the two predictors is that while the numerator in earnings-to-price is the most recent earnings, the numerator in retained earnings-to-market includes the entire history of a firm's earnings.

In Panel D of Table 2, we examine the relation between book-to-market, retained earnings-to-market, and earnings-to-price. In column 1, we present baseline results that include earnings-to-price along with control variables. The coefficient on earnings-to-price is positive and significant (t-value = 3.30). In column 2, we include both book-to-market and earnings-to-price. Both coefficients are significant but attenuate from the baseline regression presented in column 1 of this Panel and that of Panel A. In column 3, we include retained earnings-to-market and earnings-to-price. In this specification, retained earnings-to-market is statistically significant (t-value = 4.43), while earnings-to-price is not (t-value = 1.46). Earnings-to-price therefore has little, if any, information about the cross section of average returns when controlling for retained earnings-to-market.

5 Portfolio sorts

We next perform portfolio tests, which provide a potentially more robust method to evaluate predictive ability. These portfolio sorts do not impose the parametric assumptions embedded in the Fama and MacBeth (1973) regressions. Table 3 reports excess returns, CAPM alphas, and three-factor model alphas, together with their t-values, for quintiles sorted on book-to-market, retained earnings-to-market, and contributed capital-to-market.

With respect to excess returns, the high-minus-low portfolios generate positive returns for book-to-market (39 basis points per month with a t-value of 2.66) and retained earnings-to-market (47 basis points per month with a t-value of 2.88). The high-minus-low portfolio for contributed capital-to-market, by contrast, generates only 14 basis points a month with a t-value of 0.97. The pattern in CAPM alphas is the same: book-to-market and retained earnings-to-market spread returns, while contributed capital-to-market does not.

The three-factor model alphas display a different pattern. The high-minus-low portfolio for book-to-market has a negative but insignificant alpha (-14 basis points per month with a t-value of -1.35). The high-minus-low portfolio for contributed capital-to-market has a negative and significant alpha (-25 basis points per month with a t-value of -2.32). This result is the portfolio-sort equivalent of the Fama and MacBeth (1973) regression reported in column 5 in Panel A of Table 2. In that regression, the coefficient on contributed capital-to-market is significantly negative because it is the difference, retained earnings-to-market, that predicts returns.[14] The three-factor model alpha on the retained earnings-to-market high-minus-low portfolio, by contrast, is close to zero (11 basis points per month with a t-value of 0.99).

In Table A.1 of the appendix, we use two-way sorts to further demonstrate that the predictive power in book-to-market arises from its retained earnings component. In Panel A, we first sort stocks into quintiles based on retained earnings-to-market and then, within each retained earnings-to-market quintile, sort stocks into quintiles based on book-to-market. The CAPM alphas for

[14] Gerakos and Linnainmaa (2017) find a similar result when they decompose book-to-market into prior changes in the market value of equity and a residual component. They find that a high-minus-low strategy that trades the residual component earns a significantly negative three-factor model alpha.

the book-to-market high-minus-low portfolios are not significantly different from zero within any of the retained earnings-to-market quintiles. In Panel B, we reverse the order and first sort on book-to-market and then, within each book-to-market quintile, on retained earnings-to-market. In this specification, the CAPM alphas for the retained earnings-to-capital high-minus-low portfolios are statistically significant in four of the five quintiles, showing that retained earnings-to-market has information content holding book-to-market constant. The average of the five high-minus-low portfolios—which represents a book-to-market-neutral strategy—has a CAPM alpha of 46 basis points per month; this alpha is associated with a t-value of 4.31.

In sum, the evidence based on portfolio sorts reconciles with our findings from Fama and MacBeth (1973) regressions. Both specifications show that book-to-market predicts returns only because of the retained-earnings component; neither book-to-market nor contributed capital-to-market carries additional information about the cross section of stock returns.

6 Retained earnings and contributed capital factors

We next construct factors that capture the relation between average returns and the major components of book value of equity. To construct the factors, we follow the six-portfolio methodology used in Fama and French (2015); this is also the methodology that Fama and French (1993) use to construct the HML factor. We first sort stocks by size into small and big sub-groups depending on whether a company is below or above the median NYSE market capitalization breakpoint. We then perform an independent sort of stocks into high (i.e., above the 70^{th} NYSE percentile breakpoint), low (i.e., below the 30^{th} NYSE percentile breakpoint), and intermediate portfolios based on the ratio of the particular component of the book value of equity (i.e., retained earnings or contributed capital) to the market value of equity. We construct each factor by taking the average of the two high component portfolios minus the average of the two low component portfolios.

Panel A of Table 4 presents the average annualized returns, standard deviations, and t-values for HML and the factors based on the three components. We present these statistics for two versions of HML. The first is standard HML and the second is based on our sample (HML*). The difference

between the two versions of HML is that we exclude financials and firms with missing retained earnings from our sample. The two versions of HML have similar annualized average returns (4.32 and 4.07) and similar t-values (3.18 and 2.92). When we calculate a version of HML based on retained earnings-to-market value (HML_{RE}), the average return increases to 5.13 and the t-value to 3.33.[15] By contrast, when we calculate a version of HML based on contributed capital-to-market (HML_{CC}), the average return is not significantly different from zero (0.33 with a t-value of 0.29).

Given the relation between retained earnings-to-market and earnings-to-price, we also present average annualized returns, standard deviations, and t-values for a factor based on earnings-to-price (E/P). The E/P and the HML factors have similar average annualized returns (4.14 vs. 4.32 and 4.07) and similar t-values (2.69 vs. 3.18 and 2.92). The E/P factor, however, has a lower average annualized return and t-value than HML_{RE}.

Panel B presents correlations between the factors. There are several important takeaways from the correlations. First, the correlation between HML and HML_{RE} is 0.80, while the correlation between HML and HML_{CC} is only 0.37. Second, HML and HML_{RE} have similar correlations with MKT (-0.26 and -0.37), while the correlation between MKT and HML_{CC} is approximately zero. Third, HML and HML_{RE} have similar negative correlations with SMB (-0.20 and -0.31), while HML_{CC} is positively correlated with SMB (0.27). That is, although retained earnings and contributed capital represent similar proportions of the book value of equity, HML and HML_{RE} have similar correlations with the other factors, while HML_{CC} does not. E/P resembles HML_{RE} in its correlations with MKT and SMB.

Panel C measures the information contents of HML, HML_{RE}, HML_{CC}, and E/P. We estimate spanning regressions in which the dependent variable is the monthly return on the factor of interest

[15] Prior research finds that the book-to-market premium is to a large extent driven by small-growth stocks. Because the average returns on these stocks are low, they would need to be sold to capture the full value premium. The concern is that it might be prohibitively expensive to short these stocks. We can measure their influence on factor premiums by constructing the factors from a universe of stocks that excludes them. We define these stocks as the corner portfolio of the 25 Fama and French portfolios. When we remove them and then form HML and HML_{RE} from the remaining stocks, the premiums are 31 basis points with a t-value of 2.67 and 39 basis points with a t-value of 3.16. That is, even if we remove small-growth stocks, both HML and HML_{RE} are statistically significant. Because t-values are proportional to Sharpe ratios, the Sharpe ratios decrease by 8.6% for HML and 5.2% for HML_{RE}. Moreover, if anything, HML_{RE} appears more "robust" to removing small-growth stocks that might be expensive to short.

and the independent variables are the other factors. If the intercept is significant in these regressions, then the left-hand side factor is valuable to an investor who already trades all the right-hand side factors.

In the first two columns, the dependent variable is the monthly return on HML. In column 1, we include as explanatory variables MKT, SMB, and HML_{RE}. The intercept for this regression is approximately zero with a t-value of 0.47. The insignificant intercept implies that an investor already trades MKT, SMB, and HML_{RE} would not benefit from HML. In column 2, we replace HML_{RE} with HML_{CC}. The intercept of this regression is 0.46 with a t-value of 4.73, implying that there is useful information in HML if an investor already trades a factor based on the contributed capital-to-market along with MKT and SMB.

In the next two columns, we replace the dependent variable with the monthly return on the retained earnings-based factor (HML_{RE}). The intercepts are positive and significant with t-values of 2.72 and 5.02 when we include HML in column 3 and HML_{CC} in column 4. These regressions show that the HML_{RE} is valuable to an investor who already trades MKT, SMB, and either the standard or contributed capital-based HML factors.

In columns 5 and 6, the dependent variable is the monthly return on the contributed capital-based factor (HML_{CC}). When we include HML as an explanatory variable, the intercept is negative and significant (t-value $= -2.13$), suggesting that investor who already trades MKT, SMB, and HML would be better off shorting HML_{CC}. This negative alpha, once more, represents the same effect picked up by both the Fama and MacBeth (1973) regressions and portfolio sorts. The alpha is negative because it is the difference between book-to-market and contributed capital-to-market (i.e., retained earnings-to-market) that predicts returns. Indeed, when we replace HML with HML_{RE} in column 6, the intercept is not significantly different from zero (t-value $= -0.17$).

In the final three columns, the dependent variable is the monthly return on the earnings-to-price-based factor (E/P). When we include HML as an explanatory variable, the intercept is positive and significant (t-value $= 2.59$), suggesting that an investor who already trades MKT, SMB, and HML would be better off by also trading E/P. When we replace HML with HML_{RE} in column 8,

the intercept is not significantly different than zero (t-value = 1.18), showing that E/P has no additional information beyond that contained in HML_{RE}. These estimates are consistent with the Fama and MacBeth (1973) regressions reported in Panel D of Table 2, which show that retained earnings-to-market, but not book-to-market, subsumes earnings-to-price's predictive power. By contrast, when we replace HML_{RE} with HML_{CC} in column 9, the intercept is significant (t-value = 4.90), suggesting that E/P is valuable to an investor who trades HML_{CC}.

Asness, Frazzini, Israel, and Moskowitz (2015) find that E/P better predicts the cross section of stock returns than book-to-market since 1990, while Fama and French (1992) find that the combination of book-to-market and size subsumes E/P in the earlier period. When we estimate the spanning tests post- and pre-1990, we find that retained earnings-to-market subsumes the information content in current-period E/P during both sub-periods.

Overall, the spanning regressions show that when combined with MKT and SMB, the retained earnings-based factor captures the valuable information in HML. By contrast, HML and HML_{CC} do not capture all of the information embedded in the retained earnings-based factor. Moreover, the retained earnings-based factor captures the valuable information in E/P, which is consistent with individual-year accounting issues have an attenuated effect on the accumulated past earnings component of retained earnings.

7 Analysis of what explains the retained earnings premium

In this section, we dig deeper into the composition of retained earnings and explore what underlies the difference between the information in contributed capital and retained earnings for expected returns.

7.1 Accumulated dividends versus accumulated earnings

Retained earnings represent the difference between accumulated earnings and accumulated dividends over the firm's history. Hence, the information in retained earnings-to-market for the cross section of returns could be due to differences in firms' payout policies rather than (as we hypoth-

esize) underlying earnings yield. For example, a firm could have low retained earnings scaled by market value because it has a low earnings yield, or because historically it paid out a large portion of its earnings as dividends. To rule out the possibility that dividend policy rather than the firm's underlying earnings yield drives our results, we use CRSP data to estimate the cumulative amount of dividends distributed by the firm since it went public.[16]

As a baseline, column 1 of Table 5 repeats column 3 of Panel A of Table 2. This specification includes both book-to-market and retained earnings-to-market. In column 2, we include the natural logarithm of the ratio of cumulative dividends to the market value equity. In this specification, retained earnings-to-market remains positive and significant (coefficient of 0.16 with a t-value of 4.37), book-to-market remains insignificant (coefficient of 0.06 with a t-value of 0.85), and cumulative dividends-to-market is negative and insignificant (t-value = -0.57). The insignificance of the accumulated dividends component of retained earnings is consistent with retained earnings-to-market's explanatory power over the cross section arising because it is a good proxy for underlying earnings yield.

7.2 Fama and MacBeth (1973) regressions with earnings cumulated over different back windows

To examine more closely whether the accumulation of earnings in retained earnings is responsible for retained earnings' predictive ability, we construct our own accumulation of earnings backward in time. We accumulate over windows ending with the most recent year and extending back up to five years. We deflate the sum by the most recent market value of equity and then take the natural logarithm.[17]

We start the sample in July of 1967, which is five years after Compustat's 1962 start date. Because Compustat was created for industry professionals, the initial database was backfilled for

[16] We cannot directly observe the accumulated earnings component of retained earnings for older firms, because Compustat data do not go back to their origin. However, dividends data are obtainable from CRSP back to 1926, so with the exception of firms created before that date we can observe the accumulated dividends component of retained earnings.

[17] When the backward-accumulated earnings are negative, we set the log of accumulated earnings-to-market equal to zero and include an indicator variable for negative values.

surviving firms listed as of the start date (Ball and Watts, 1977).[18] Commencing our sample in 1967 frees our Fama and MacBeth (1973) regressions from this survivorship bias. All of the regressions include as unreported controls size, prior one-month return, and prior one-year returns skipping a month. Our thesis implies that this accumulation measure should increasingly subsume book-to-market's predictive power as we expand the accumulation period.

Table 6 presents the Fama and MacBeth (1973) regressions that include this backward-accumulated earnings measure. For the top half of the table, we accumulate net income. For all of the windows, the coefficient on accumulated earnings-to-market remains statistically positive and increases monotonically with the window length. The t-values increase almost monotonically, with the largest being for accumulating back five years (t-value = 3.63). Importantly, the coefficients and their associated t-values for book-to-market monotonically decrease in the length of the back accumulation, starting with a coefficient of 0.167 and t-value of 2.40 for the current period's earnings alone and declining to a coefficient of 0.08 and t-value of 1.39 for earnings accumulated five years back.

In the bottom half of the table, we use income before extraordinary and special items as the earnings measure, again deflated by the most recent market value of equity. For these regressions, the coefficients and t-values on book-to-market continue to monotonically decrease as we extend the earnings accumulation horizon, and the coefficients and t-values on retained earnings-to-market now monotonically increase. The coefficients and t-values for this accumulated earnings-to-price measure are larger than when we use net income. This result occurs because income before extraordinary and special items is affected less by individual-year accounting effects that reduce the informativeness of bottom-line net income for expected returns (Novy-Marx, 2013; Ball et al., 2015, 2016)

8 Predicting average returns over increasing horizons

We next compare how far ahead book-to-market and its retained earning-to-market predict returns. We modify the Fama and MacBeth (1973) regressions in Table 2 by replacing the current

[18]Consequently, returns for firms with Compustat data are truncated prior to the start date, thereby leading to biased coefficient estimates. For a discussion of this point, see pages 799–802 of Wooldridge (2010).

values of book-to-market and retained earnings-to-market with increasingly stale values, but retaining the current values of the control variables. Thus, we assume the investor knows the current values of the control variables, but does not know the current values of the balance sheet measures. Would an investor still benefit from these stale measures? How far ahead do these measures predict the cross section of returns?

Figure 1 plots average monthly Fama and MacBeth (1973) regression t-values for both variables, increasing the horizon in one-month increments up to five years. The cross-sectional regressions are estimated for each month from July 1969 through December 2016. By starting in 1969, we ensure that we have data for all lags because our data start in 1964. Control variables are updated each month, but the book-to-market and retained earnings-to-market variables are not. Figure 1 shows that both book-to-market and retained earnings-to-market reliably predict returns several years ahead. Book-to-market has significant explanatory power for approximately three years. The effect of retained earnings-to-market is stronger, and persists for about five years.

These results are difficult to reconcile with the notion of book value being a benchmark for identifying market mispricing. Why would only one component of the book value of equity indicate mispricing of the entire equity? Why would the correction of mispricing occur gradually over a horizon that extends at least as far as five years, especially bearing in mind that all of this accounting information is made publicly accessible at essentially zero cost for all firms and all years? In addition, value strategies have been implemented since at least Graham and Dodd (1934).

We interpret the results in Figure 1 as reflecting two factors. First, the accumulation of earnings in retained earnings "washes out" accounting issues that affect earnings in individual years but reverse over time. Then, when deflated by current market value of equity, they proxy for expected returns (Ball, 1978; Berk, 1995). Second, the ability of variables deflated by current market values to predict stock returns inevitably declines over longer horizons, because expected returns are unlikely to be constant over time.

9 Results for different samples

In this section, we present results from different countries and time periods that corroborate our main results and address concerns that they represent statistical artifacts (Lo and MacKinlay, 1990; Lewellen et al., 2010; Harvey et al., 2016).

9.1 Evidence for U.S. sub-periods

Asness et al. (2015) and Fama and French (2016) show that book-to-market is a not a significant predictor of returns after 1990. We therefore split the U.S. sample in 1990 to evaluate whether retained earnings-to-market has predictive power across our sample period. The pre-1990 sample begins in July 1964 and ends in June 1990. The post-1990 sample begins in July 1990 and end in December 2016.

Table 7 presents CAPM alphas and their associated t-values for quintile portfolios and high-minus-low portfolios sorted on book-to-market, retained earnings-to-market, and contributed capital-to-market. When we examine the book-to-market portfolios, the high-minus-low CAPM alpha is statistically significant for the U.S. prior to 1990 (55 basis points with a t-value of 2.67). However, consistent with Asness et al. (2015) and Fama and French (2016), book-to-market is not a significant predictor of the cross section for the U.S. post-1990 (33 basis points with a t-value of 1.61). By contrast, the retained earnings-to-market portfolios spread the CAPM alphas essentially as well post-1990 as in the prior sub-period; the post-1990 CAPM alpha is statistically significant (64 basis points with a t-value of 2.67) and similar in magnitude to the pre-1990 CAPM alpha (60 basis points with a a t-value of 2.96). The CAPM alphas for the contributed capital-to-market portfolios are not significant for either of the U.S. sub-periods.

We next examine why book-to-market loses its predictive power after 1990 while retained earnings-to-market does not. The reason turns out to be that the relation between book-to-market and retained earnings-to-market weakens in the second half of the sample. Our interpretation is that book-to-market became a less effective proxy for underlying earnings-yield. We show this effect in the evolution of cross-sectional means and correlations.

In Panel A of Figure 2, we plot the annual cross-sectional means of book-to-market, retained earnings-to-market, and contributed capital-to-market over our sample period. Up to 1983, the cross-sectional means of book-to-market, retained earnings-to-market, and contributed capital-to-market are highly correlated. After 1983, retained earnings-to-market diverges while book-to-market and contributed capital-to-market remain correlated. These results suggest that book-to-market lost its predictive power in the second half of the sample because it was less correlated with retained earnings-to-market.

After 1978, Fama and French (2001) find a substantial shift in the composition of public firms that was due to a surge in new listings that were unprofitable and did not pay dividends. In Panel B, we therefore plot the annual cross-sectional means of book-to-market by six cohorts based on when a firm makes its first appearance on CRSP (1926–1969, 1970–1979, 1980–1989, 1990–1999, 2000–2009, and 2010–2016). In general, the annual cross-sectional means of book-to-market move similarly over the six cohorts, indicating that the compositional change documented by Fama and French (2001) did not affect book-to-market.

In Panel C, we similarly plot the annual cross-sectional means of retained earnings-to-market for the six cohorts. Here, we find evidence of the compositional change. The means for the first two cohorts (1926–1969 and 1970–1979) are positive for almost every year in our sample. By contrast, the means for the later cohorts are negative for almost every year after 1980. In Panel D, we plot the annual cross-sectional correlations between book-to-market and retained earnings-to-market and between book-to-market and contributed capital-to-market. The trend for the cross-sectional correlations is similar to the trend for the cross-sectional means. Starting after 1980, there was a marked drop in the correlation between book-to-market and retained earnings-to-market. The correlation for every year prior to 1982 was above 0.75. After 1982, the correlation ranges from 0.15 to 0.6. By contrast, the correlation between book-to-market and contributed capital-to-market remains above 0.6 for almost every year in the sample.

These results suggest that book-to-market predicted the cross section in the first half of our sample because book-to-market and retained earnings-to-market were highly correlated. In the

second half of our sample period, book-to-market lost its predictive power for the cross-section of returns because the change in the composition of public firms reduced book-to-market's correlation with retained earnings-to-market.

9.2 Evidence for the world excluding the U.S.

We next use international data to compare the information contents of book-to-market, retained earnings-to-market, and contributed capital-to-market. We explore only developed markets excluding the U.S.: North America (Canada), Europe (Austria, Belgium, Denmark, Finland, France, Germany, Greece, Iceland, Ireland, Italy, Luxembourg, the Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, and the United Kingdom), and Asia-Pacific (Australia, Hong Kong, Japan, New Zealand, and Singapore). The sample period begins in November 1989 and ends in June 2016.

We continue to measure accounting variables at end of December in year $t-1$ and accumulate monthly returns from the beginning of July in year t through the end of June in year $t+1$. We then form portfolios at the end of each June and compute value-weighted returns on the stocks over the next year. The world excluding the U.S. sorts use "large" stocks, with the large-stock cutoff determined as the point at which the largest firms in each region account for 90% of total market capitalization. In the appendix, we describe how we construct the sample and variables.

Table 7 also presents CAPM alphas and their associated t-values for the world excluding U.S. portfolios. The retained earnings-to-market portfolios spread the CAPM alphas essentially as well as for the U.S. sub-periods. The high-minus-low CAPM alpha for retained earnings-to-market is statistically significant (53 basis points with a t-value of 3.45) and larger than for book-to-market (46 basis points with a t-value of 2.86). Here too, the CAPM alpha for the high-minus-low contributed capital-to-market portfolio is insignificant (5 basis points with a t-value of 0.33).

9.3 Evidence for the pre-1964 period

To further corroborate our main results, we next investigate pre-Compustat U.S. data. To do so, we employ the historical accounting data used in Graham, Leary, and Roberts (2015) and

Linnainmaa and Roberts (2017). This database combines Standard and Poor's Compustat database with accounting data from Moody's Industrial and Railroad manuals along with the historical book value of equity data provided by Ken French. Their database excludes financials and utilities. We merge their database with the CRSP database to obtain returns.

The Securities Exchange Act of 1934 was enacted to ensure the flow of accurate and systematic accounting information. Cohen et al. (2003) analyze the historical SEC enforcement records and determine that post-1936 accounting data is of sufficiently high quality to employ in empirical analysis. They characterize 1935 and 1936 as an initial enforcement period and drop these years from their sample. Following their timing conditions, our return data start in July 1938 and the associated book values of equity start in 1937.

This database does not include retained earnings. It does, however, include net income and the book value of equity. To proxy for retained earnings, we accumulate each firm's past net income for up to 20 years and then regress book-to-market on accumulated past net income. We take the fitted value from this regression as our proxy for retained earnings-to-market and the residual as our proxy for contributed capital-to-market.

Table 8 presents CAPM alphas and associated t-values for quintiles sorted on book-to-market and the proxies for retained earnings-to-market and contributed capital-to-market. We form portfolios at the end of each June and compute value-weighted returns on the stocks over the next year. The sorts use NYSE breakpoints. For the high-minus-low portfolios, only retained earnings-to-market has a significant CAPM alpha (t-value = 3.20).[19]

The CAPM alphas on the individual retained earnings-to-market portfolios are similar in magnitude across the different samples presented in Tables 7 and 8: the U.S. 1964–1990, the U.S. post-1990, the U.S. pre-1964, and the World excluding the U.S. post-1990. The high-minus-low CAPM

[19]The fact that the high-minus-low portfolio based on retained earnings-to-market earns a positive CAPM alpha in these pre-Compustat data is noteworthy. Ang and Chen (2007) note that although HML's average return is statistically significant in the pre-1963 data, its CAPM alpha is not. The reason for this finding is that, before 1963, high book-to-market firms' market betas were typically higher than those of low book-to-market firms. When we examine average excess returns, we find that they are similar to the CAPM alphas for retained earnings-to-market. These results reveal, indirectly, that contributed capital in the pre-1963 data is responsible such marked differences in market betas.

alphas for the four samples are 0.60, 0.64, 0.51, and 0.53. Similarly, not one of the CAPM alphas is significant for the high-minus-low contributed capital-to-market portfolios.

10 Conclusion

Generations of investors and scholars have calibrated the market value of a firm's equity against its book value to help identify market under- or over-valuation. Book-to-market then is viewed as a crude indicator of whether prices have strayed from "fundamental" or "intrinsic" values. Consequently, book-to-market has been a common element of "value" investment strategies that have been popular at least since publication of Graham and Dodd's (1934) classic text. For example, Investopedia states: "The book-to-market ratio attempts to identify undervalued or overvalued securities by taking the book value and dividing it by market value. In basic terms, if the ratio is above 1 then the stock is undervalued; if it is less than 1, the stock is overvalued."[20] Among scholars, many interpret the abundant evidence that book-to-market reliably predicts stock returns (Rosenberg, Reid, and Lanstein, 1985; Chan, Hamao, and Lakonishok, 1991) as evidence that market value can be calibrated against book value to indicate systematic mispricing. For example, Lakonishok, Shleifer, and Vishny (1994, p. 1541) summarize their interpretation as follows:

> "For many years, scholars and investment professionals have argued that value strategies outperform the market. These value strategies call for buying stocks that have low prices relative to earnings, dividends, book assets, or other measures of fundamental value. While there is some agreement that value strategies produce higher returns, the interpretation of why they do so is more controversial. This article provides evidence that value strategies yield higher returns because these strategies exploit the suboptimal behavior of the typical investor and not because these strategies are fundamentally riskier."

We provide a different interpretation. Book-to-market consists of two major and economically different components: retained earnings and contributed capital, both deflated by the market value of equity. We expect one to contain substantial information about expected returns, but not

[20] http://www.investopedia.com/terms/b/booktomarketratio.asp, accessed on January 12, 2017.

the other. Consistent with this expectation, we show that retained earnings-to-market entirely subsumes book-to-market in predicting returns even though retained earnings represent averages only 41% of the book value of equity and retained earnings-to-market explains less than 50% of book-to-market's variance. By contrast, contributed capital-to-market has no ability to predict future returns when controlling for retained earnings-to-market. These results are evident in U.S. and non-U.S data over different periods.

We also explore what drives the power of retained earnings to predict future returns. Our conclusion is that retained earnings-to-market is comparatively free of the individual-year accounting issues that affect current-period earnings yield, and consequently is a good proxy for underlying earnings yield, which has a direct conceptual link with expected returns.

REFERENCES

AICPA, 2010. Accounting trends & techniques. Tech. rep., American Institute of Certified Public Accountants (AICPA).

Ang, A., Chen, J., 2007. CAPM over the long run: 1926–2001. Journal of Empirical Finance 14, 1–40.

Asness, C., Frazzini, A., Israel, R., Moskowitz, T., 2015. Fact, fiction, and value investing. Journal of Portfolio Management 42, 75–92.

Ball, R., 1978. Anomalies in relationships between securities' yields and yield surrogates. Journal of Financial Economics 6, 103–126.

Ball, R., Gerakos, J., Linnainmaa, J., Nikolaev, V., 2015. Deflating profitability. Journal of Financial Economics 117, 225–248.

Ball, R., Gerakos, J., Linnainmaa, J., Nikolaev, V., 2016. Accruals, cash flows, and operating profitabilty in the cross section of stock returns. Journal of Financial Economics 121, 28–45.

Ball, R., Watts, R., 1977. Reply to Salamon and Smith. Journal of Finance 32, 1802–1808.

Beaver, W., McNichols, M., Price, R., 2007. Delisting returns and their effect on accounting-based market anomalies. Journal of Accounting and Economics 43, 341–368.

Berk, J., 1995. A critique of size-related anomalies. Review of Financial Studies 8, 275–286.

Campbell, J., Shiller, R., 1988. Stock prices, earnings, and expected dividends. Journal of Finance 43, 661–676.

Chan, L., Hamao, Y., Lakonishok, J., 1991. Fundamenals and stock returns in Japan. Journal of Finance 46, 1739–1764.

Cohen, R., Polk, C., Vuolteenaho, T., 2003. The value spread. Journal of Finance 58, 609–641.

Daniel, K., Titman, S., 2006. Market reactions to tangible and intangible information. Journal of Finance 61, 1605–1643.

Davis, J., Fama, E., French, K., 2000. Characteristics, covariances, and average returns: 1929 to 1997. Journal of Finance 55, 389–406.

Easton, P., 2009. Estimating the cost of capital implied by market prices and accounting data. Fourndations and Trends in accounting 2, 241–364.

Etcherling, F., Eierle, B., Ketterer, S., 2015. A review of the literature on methods of computing the implied cost of capital. International Review of Financial Analysis 42, 235–252.

Fama, E., French, K., 1992. The cross-section of expected stock returns. Journal of Finance 47, 427–465.

Fama, E., French, K., 1993. Common risk factors in the returns of stocks and bonds. Journal of Financial Economics 33, 3–56.

Fama, E., French, K., 2001. Disappearing dividends: changing firm characteristics or lower propensity to pay? Journal of Financial Economics 60, 3–43.

Fama, E., French, K., 2008. Average returns, B/M, and share issues. Journal of Finance 63, 2971–2995.

Fama, E., French, K., 2015. A five-factor asset pricing model. Journal of Financial Economics 116, 1–22.

Fama, E., French, K., 2016. Dissecting anomalies with a five-factor model. Review of Financial Studies 29, 69–103.

Fama, E., MacBeth, J., 1973. Risk, return, and equilibrium: Empirical tests. Journal of Political Economy 81, 607–636.

Gerakos, J., Linnainmaa, J., 2017. Decomposing value. Review of Financial Studies, forthcoming.

Graham, B., Dodd, D., 1934. Security analysis principles and technique. McGraw-Hill.

Graham, J., Leary, M., Roberts, M., 2015. A century of capital structure: The leveraging of corporate America. Journal of Financial Economics 118, 658–683.

Harvey, C., Liu, Y., Zhu, H., 2016. . . . and the cross-section of expected returns. Review of Financial Studies 29, 5–68.

Hou, K., Karolyi, G., Kho, B.-C., 2011. What factors drive global stock returns? Review of Financial Studies 24, 2527–2574.

Ikenberry, D., Lakonishok, J., Vermaelen, T., 1995. Market underreaction to open market share repurchases. Journal of Financial Economics 39, 181–208.

Ince, O., Porter, R., 2006. Individual equity return data from Thomson Datastream: Handle with care! Journal of Financial Research 29, 463–479.

Lakonishok, J., Shleifer, A., Vishny, R., 1994. Contrarian investment, extrapolation, and risk. Journal of Finance 49, 1541–1578.

Lewellen, J., Nagel, S., Shanken, J., 2010. A skeptical appraisal of asset pricing tests. Journal of Financial Economics 96, 175–194.

Linnainmaa, J., Roberts, M., 2017. The history of the cross section of stock returns. Review of Financial Studies, forthcoming.

Lo, A., MacKinlay, A., 1990. Data-snooping biases in tests of financial asset pricing models. Review of Financial Studies 3, 431–467.

Loughran, T., Ritter, J., 1995. The new issues puzzle. Journal of Finance 50, 23–51.

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

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

Rosenberg, B., Reid, K., Lanstein, R., 1985. Persuasive evidence of market inefficiency. Journal of Portfolio Management 11, 9–16.

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

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

Wooldridge, J., 2010. Econometric Analysis of Cross Section and Panel Data. The MIT Press, Cambridge, Massachusetts, second ed.

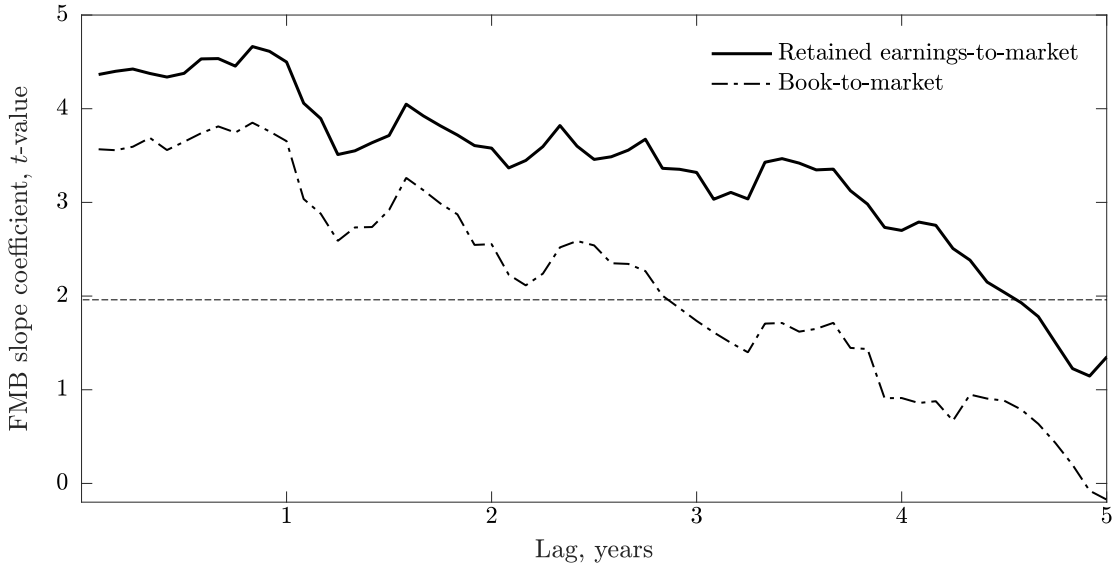
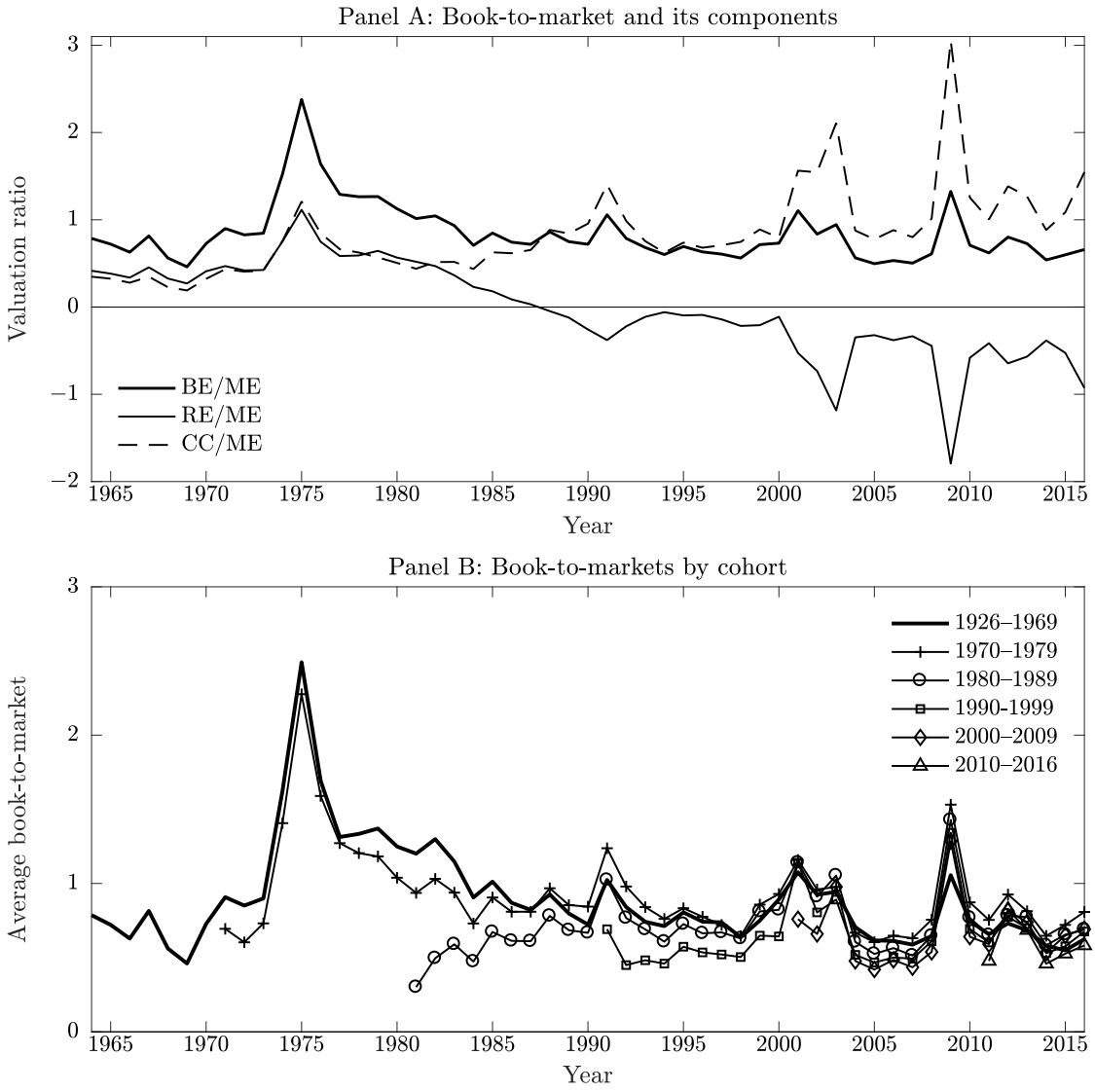


Figure 1: **Comparison of the predictive power of lagged BE/ME and lagged RE/ME.** This figure plots t-values associated with the Fama and MacBeth (1973) regression slopes for log(BE/ME) and log(RE/ME) from cross-sectional regressions that predict monthly returns. The regressions are estimated monthly using data from July 1969 through December 2016 for stocks with a market value of equity above the 20th percentile of the NYSE market capitalization distribution (all but microcaps). The regressions are estimated separately using book-to-market or retained earnings-to-market as the main regressor. The other regressors are: prior one-month return, prior one-year return skipping a month, and log-size. We lag all regressors by the value indicated on the x-axis. The estimates at $x = 2$, for example, explain cross-sectional variation in returns using the values of log(BE/ME) and log(RE/ME) recorded two years earlier. The dashed line indicates the threshold for statistical significance at the 5% level.


Panel A: Book-to-market and its components
Valuation ratio
Year
BE/ME
RE/ME
CC/ME
Panel B: Book-to-markets by cohort
Average book-to-market
Year
1926–1969
1970–1979
1980–1989
1990-1999
2000–2009
2010–2016

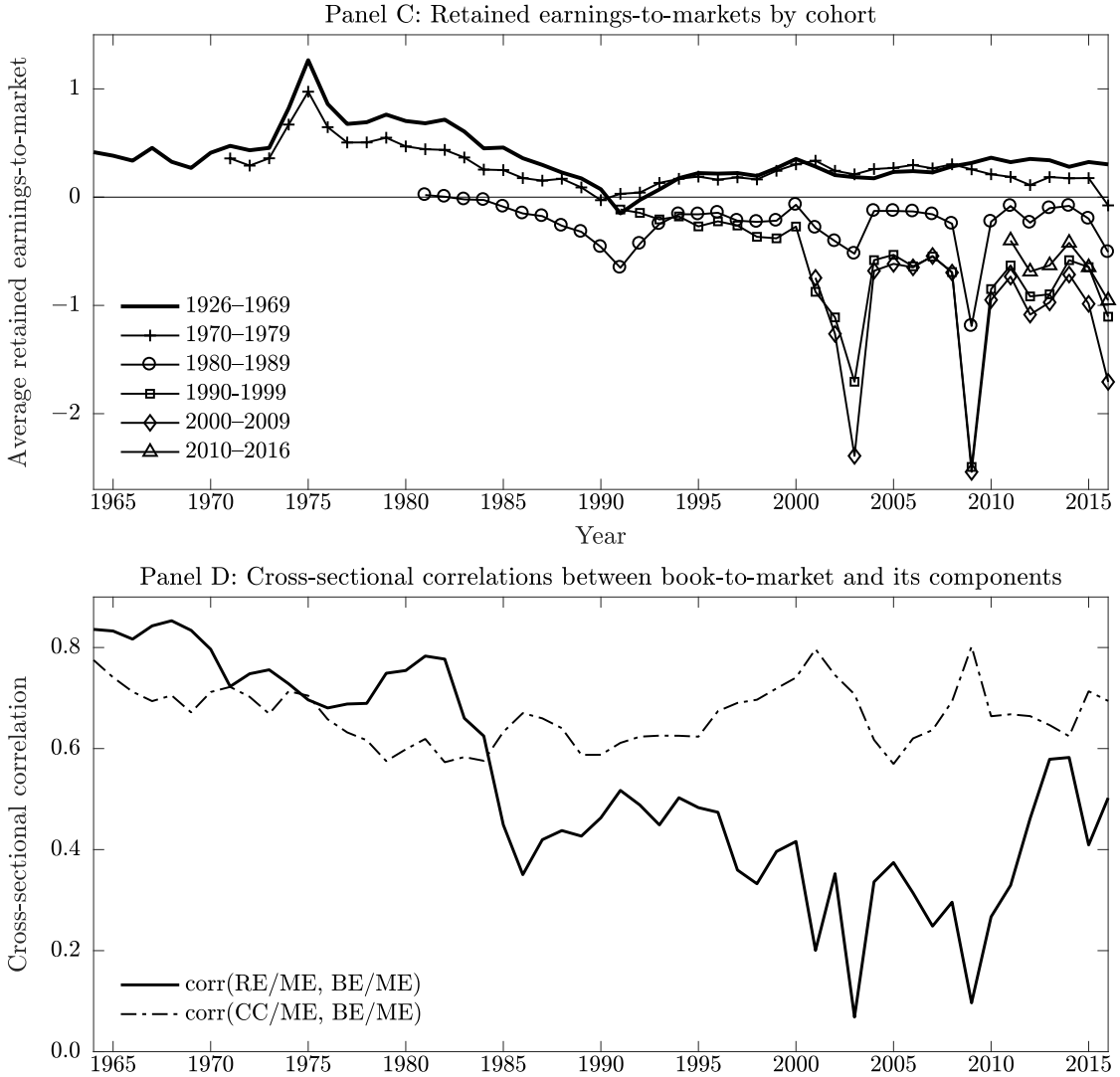


Figure 2: **Annual cross-sectional means and correlations of book-to-market and its components.** Panel A plots the cross-sectional means of book-to-market, retained earnings-to-market, and contributed capital-to-market from 1964 through 2016. Panel B plots the cross-sectional means of book-to-market for firms classified by cohort. A firm's cohort is determined by its first appearance on CRSP. Panel C plots the cross-sectional means of retained earnings-to-market by cohort. Panel D plots cross-sectional correlations between (a) book-to-market and retained earnings-to-market and (b) book-to-market and contributed capital-to-market.

Table 1: Summary statistics and correlations for retained earnings, contributed capital, and book value of equity

Panel A presents distributions of book-to-market and the ratios of its components to the market value of equity (ME). Panel B presents distributions of the ratios of the components of the book value of equity (BE) to the total book value of equity. The components of the book value of equity are contributed capital (CC), retained earnings (RE), and accumulated other comprehensive income (AOCI). Panel C presents the distributions of the ratios of the components to the total book value of equity for the six Fama and French portfolios. Panel D presents Pearson and Spearman correlations for the measures when they are deflated by the December market value of equity. Panel A includes all firms. Panels B through D include firms with positive book values of equity and require that each book-value-of-equity component is between −100% and 200% of the book value of equity. Our sample period starts in July 1964 and ends in December 2016. The sample consists of all but microcap firms, which are stocks with market values of equity at or above the 20$^{\text{th}}$ percentile of the NYSE market capitalization distribution.

Panel A: Distributions of book-to-market and components of the book value of equity scaled by the market value of equity

				Percentiles				
Ratio	Mean	SD	Pct < 0	10th	25th	50th	75th	90th
Book-to-market	0.83	2.20	3%	0.18	0.37	0.68	1.11	1.69
Reported book-to-market	0.79	1.88	3%	0.18	0.36	0.64	1.04	1.62
Retained earnings-to-market	−0.14	3.61	27%	−0.97	−0.12	0.24	0.53	0.89
Contributed capital-to-market	0.94	3.04	4%	0.05	0.16	0.38	0.86	1.94
AOCI-to-market	0.03	0.69	24%	−0.01	−0.00	0.01	0.05	0.16

Panel B: Distributions of the components of the book value of equity scaled by the book value of equity (%)

				Percentiles				
Component	Mean	Pct < 0	SD	10th	25th	50th	75th	90th
Retained earnings (RE)	41%	15%	42%	−15%	17%	45%	70%	88%
Contributed capital (CC)	54%	4%	42%	8%	24%	48%	78%	113%
AOCI	5%	20%	12%	−1%	−0%	2%	8%	18%

Panel C: Components of the book value of equity scaled by the book value of equity by Fama-French portfolios (%)

		RE/BE			CC/BE			AOCI/BE		
Size	Book-to-market	Mean	Median	Pct < 0	Mean	Median	Pct < 0	Mean	Median	Pct < 0
Small	Growth	28%	32%	26%	71%	65%	2%	2%	0%	25%
	Neutral	40%	44%	16%	57%	51%	3%	3%	1%	21%
	Value	40%	44%	15%	55%	49%	3%	5%	2%	19%
Big	Growth	62%	66%	7%	35%	30%	11%	4%	3%	20%
	Neutral	55%	58%	5%	36%	31%	8%	9%	8%	14%
	Value	41%	41%	8%	43%	39%	2%	16%	16%	10%

Panel D: Correlations among book value of equity components

	Pearson correlations				Spearman rank correlations			
	BE/ME	RE/ME	CC/ME	AOCI/ME	BE/ME	RE/ME	CC/ME	AOCI/ME
BE/ME	1				1			
RE/ME	0.54	1			0.57	1		
CC/ME	0.68	−0.19	1		0.60	−0.17	1	
AOCI/ME	0.37	0.17	0.12	1	0.32	0.29	0.00	1

Table 2: Retained earnings and contributed capital in Fama-MacBeth regressions

This table presents average Fama and MacBeth (1973) regression slopes and their t-values from cross sectional regressions that predict monthly returns. The regressions are estimated monthly using data from July 1964 through December 2016. The sample consists of all but microcap firms with positive book value of equity and non-missing values for retained earnings and contributed capital. All but microcap firms are stocks with market values of equity at or above the 20^{th} percentile of the NYSE market capitalization distribution. Variables are trimmed at the 1^{st} and 99^{th} percentiles based on book-to-market, size, prior one-month return, and prior one-year return skipping a month. Regressors log(RE/ME), log(CC/ME), and log(AOCI/ME) are set to zero when RE ≤ 0, CC ≤ 0, or AOCI ≤ 0 is non-positive. Indicator variables at the bottom of the table identify these observations. Panel A presents our main results; Panel B presents results that use reported book value of equity instead of Fama and French (1992) book value of equity; Panel C presents results for alternative samples; and Panel D presents results that include earnings-to-price as an additional regressor. The second to the last row reports pseudo t-values from tests that examine whether all regressors other than book-to-market, size, prior one-month return, and prior one-year return skipping a month are jointly zero. This t-value is computed by converting the p-value from the Hotelling's test into a z-score. In column 2 of Panel A, for example, the t-value of 4.17 is associated with the test that log-retained earnings-to-market and the RE ≤ 0 indicator variable are jointly zero.

Panel A: Main regressions

	Regression							
Regressor	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
log(ME)	−0.07	−0.10	−0.09	−0.09	−0.09	−0.10	−0.07	−0.09
	(−1.88)	(−2.75)	(−2.42)	(−2.51)	(−2.43)	(−2.86)	(−2.05)	(−2.73)
$r_{1,1}$	−3.14	−2.99	−3.23	−2.89	−3.21	−2.87	−3.21	−3.22
	(−7.28)	(−6.91)	(−7.61)	(−6.51)	(−7.54)	(−6.57)	(−7.52)	(−7.60)
$r_{12,2}$	0.84	0.82	0.83	0.76	0.83	0.79	0.83	0.81
	(4.45)	(4.31)	(4.46)	(3.98)	(4.49)	(4.16)	(4.41)	(4.33)
log(BE/ME)	0.23		0.06		0.34		0.22	
	(3.31)		(0.79)		(4.26)		(3.52)	
log(RE/ME)		0.19	0.17					0.19
		(4.47)	(4.66)					(4.91)
log(CC/ME)				0.00	−0.12			0.01
				(0.08)	(−4.18)			(0.32)
log(AOCI/ME)						0.02	0.00	0.01
						(2.11)	(0.49)	(1.05)
Indicator variables								
RE ≤ 0		−0.62	−0.55					−0.62
		(−3.08)	(−2.81)					(−3.27)
CC ≤ 0				0.14	0.26			−0.03
				(2.01)	(2.79)			(−0.43)
AOCI ≤ 0						−0.15	0.00	−0.03
						(−1.37)	(0.06)	(−0.33)
Pseudo t-value for joint sig. of add'l regressors		4.17	4.30	1.71	3.76	1.74	0.43	4.24
Avg. Adj. R^2	5.48%	5.50%	6.12%	4.77%	5.86%	4.90%	5.76%	6.16%

Panel B: Regressions using reported book value of equity

	Regression			
Regressor	(1)	(2)	(3)	(4)
log(ME)	−0.07	−0.09	−0.08	−0.07
	(−1.75)	(−2.40)	(−2.23)	(−2.02)
$r_{1,1}$	−3.12	−3.20	−3.18	−3.21
	(−7.18)	(−7.50)	(−7.44)	(−7.50)
$r_{12,2}$	0.83	0.82	0.83	0.82
	(4.37)	(4.37)	(4.42)	(4.37)
log(Reported BE/ME)	0.24	0.05	0.35	0.22
	(3.49)	(0.70)	(4.53)	(3.61)
log(RE/ME)		0.18		
		(4.67)		
log(CC/ME)			−0.13	
			(−4.32)	
log(AOCI/ME)				0.01
				(0.96)
Indicator variables				
RE ≤ 0		−0.56		
		(−2.76)		
CC ≤ 0			0.26	
			(2.80)	
AOCI ≤ 0				−0.05
				(−0.57)
Pseudo t-value for joint sig. of add'l regressors		4.27	3.91	0.54
Avg. Adj. R^2	5.36%	6.00%	5.73%	5.72%

Panel C: Alternative samples

	Include firms with negative book value of equity		Require firms to have positive retained earnings	
Regressor	(1)	(2)	(3)	(4)
log(ME)	−0.07	−0.09	−0.09	−0.09
	(−1.83)	(−2.35)	(−2.47)	(−2.63)
$r_{1,1}$	−3.11	−3.20	−3.55	−3.61
	(−7.22)	(−7.54)	(−8.10)	(−8.32)
$r_{12,2}$	0.84	0.83	0.73	0.74
	(4.49)	(4.50)	(3.67)	(3.74)
log(BE/ME)	0.23	0.06	0.18	0.00
	(3.31)	(0.84)	(2.58)	(−0.01)
log(RE/ME)		0.17		0.19
		(4.58)		(4.85)
Indicator variables				
BE ≤ 0	0.11	0.40		
	(0.53)	(1.87)		
RE ≤ 0		−0.56		
		(−2.87)		
Pseudo t-value for joint sig. of add'l regressors		4.23		4.85
Avg. Adj. R^2	5.52%	6.16%	5.51%	5.86%

Panel D: Earnings-to-price, book-to-market, and retained earnings-to-market

	Regression		
Regressor	(1)	(2)	(3)
log(ME)	−0.10	−0.09	−0.10
	(−2.71)	(−2.35)	(−2.83)
$r_{1,1}$	−3.04	−3.29	−3.14
	(−7.08)	(−7.81)	(−7.40)
$r_{12,2}$	0.81	0.84	0.82
	(4.35)	(4.56)	(4.44)
log(BE/ME)		0.16	
		(2.53)	
log(RE/ME)			0.16
			(4.43)
log(E/P)	0.19	0.13	0.07
	(3.30)	(3.09)	(1.46)
Indicator variables			
RE ≤ 0			−0.48
			(−2.69)
E ≤ 0	−0.66	−0.52	−0.26
	(−2.70)	(−2.55)	(−1.19)
Pseudo t-value for joint sig. of add'l regressors	2.71	2.35	2.83
Avg. Adj. R^2	5.49%	6.19%	6.09%

Table 3: Returns on portfolios sorted by book-to-market, retained earnings-to-market, and contributed capital-to-market

This table reports value-weighted average excess returns and three-factor model alphas for portfolios sorted by book-to-market, retained earnings-to-market, and contributed capital-to-market. We sort stocks into quintiles based on NYSE breakpoints at the end of each June and hold the portfolios for the following year. The sample starts in July 1964 and ends in December 2016.

	Book-to-market component			Book-to-market component		
Quintile	Total	Retained earnings	Contributed capital	Total	Retained earnings	Contributed capital
	Excess returns			***t*-values**		
1 (low)	0.44	0.33	0.52	2.26	1.42	3.00
2	0.53	0.44	0.46	2.91	2.54	2.41
3	0.57	0.64	0.50	3.21	3.82	2.59
4	0.65	0.74	0.52	3.79	4.21	2.96
5 (high)	0.83	0.80	0.66	4.24	4.08	3.19
H−L	0.39	0.47	0.14	2.66	2.88	0.97
	CAPM alphas			***t*-values**		
1 (low)	−0.08	−0.27	0.06	−1.32	−2.93	1.12
2	0.04	−0.03	−0.06	0.79	−0.56	−1.09
3	0.11	0.20	−0.01	1.64	3.63	−0.22
4	0.22	0.29	0.07	2.86	3.97	0.98
5 (high)	0.36	0.32	0.16	3.54	3.35	1.53
H−L	0.44	0.59	0.10	3.04	3.71	0.74
	Three-factor model alphas			***t*-values**		
1 (low)	0.12	−0.09	0.17	3.11	−1.24	4.09
2	0.03	0.06	−0.01	0.58	1.49	−0.14
3	0.01	0.15	−0.03	0.19	2.83	−0.42
4	0.01	0.15	−0.03	0.22	2.38	−0.46
5 (high)	0.02	0.02	−0.07	0.27	0.28	−0.86
H−L	−0.10	0.11	−0.25	−1.35	0.99	−2.32

Table 4: Information content of HML-factors based book value of equity, retained earnings, contributed capital, and earnings-to-price

Panel A shows the annualized average returns and standard deviations of the monthly factors. Panel B shows the Pearson correlations. The factors are the market return minus the risk free rate, MKT; size, SMB; value, HML; two factors constructed from the retained earnings and contributed capital components of the book value of equity, HML_{RE} and HML_{CC}; and a factor based on earnings-to-price, E/P. These additional factors are formed using the same six-portfolio methodology as standard HML. HML is the factor provided by Ken French; HML^* is constructed using our sample, which excludes financials and firms with missing retained earnings. Panel C measures the information content of HML, HML_{RE}, HML_{CC}, and E/P by reporting estimates from spanning regressions. The left-hand side variable is the monthly return on each of these factors. The explanatory variables are the market and size factors, and one of the HML factors. The sample starts in July 1964 and ends in December 2016.

Panel A: Average returns and standard deviations

	Factor						
	MKT	SMB	HML	HML^*	HML_{RE}	HML_{CC}	E/P
Average annualized return	5.96	2.91	4.32	4.07	5.13	0.32	4.14
Standard deviation	15.45	10.78	9.83	10.10	11.15	8.03	11.16
t-value	2.79	1.95	3.18	2.92	3.33	0.29	2.69

Panel B: Correlations

	Factor					
Factor	MKT	SMB	HML	HML_{RE}	HML_{CC}	E/P
MKT	1					
SMB	0.30	1				
HML	−0.26	−0.20	1			
HML_{RE}	−0.37	−0.31	0.80	1		
HML_{CC}	0.01	0.27	0.37	−0.04	1	
E/P	−0.43	−0.39	0.74	0.87	0.00	1

Panel C: Spanning regressions

	Dependent variable								
	HML		HML_{RE}		HML_{CC}		E/P		
Regressor	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Intercept	0.03	0.46	0.20	0.59	−0.17	−0.02	0.20	0.07	0.54
	(0.47)	(4.73)	(2.72)	(5.02)	(−2.13)	(−0.17)	(2.59)	(1.18)	(4.90)
MKT	0.02	−0.12	−0.10	−0.22	0.01	−0.04	−0.15	−0.08	−0.24
	(1.08)	(−5.13)	(−5.70)	(−7.86)	(0.52)	(−1.70)	(−8.01)	(−5.21)	(−9.44)
SMB	0.03	−0.25	−0.12	−0.23	0.26	0.22	−0.21	−0.12	−0.33
	(1.47)	(−7.19)	(−4.68)	(−5.59)	(9.76)	(7.23)	(−7.95)	(−5.92)	(−8.47)
HML			0.84		0.36		0.74		
			(31.24)		(12.33)		(26.35)		
HML_{RE}	0.72					0.02		0.80	
	(31.24)					(0.50)		(38.99)	
HML_{CC}		0.54		0.03					0.12
		(12.33)		(0.50)					(2.51)
N	630	630	630	630	630	630	630	630	630
Adj. R^2	64.1%	26.1%	67.6%	17.2%	25.5%	7.4%	64.9%	78.4%	26.7%

Table 5: Retained earnings, contributed capital, and payout policy in Fama-MacBeth regressions

This table presents average Fama and MacBeth (1973) regression slopes and their t-values from cross sectional regressions that predict monthly returns. The regressions are estimated monthly using data from July 1964 through December 2016. The sample consists of all but microcap firms with positive book value of equity and non-missing values for retained earnings and contributed capital. All but microcap firms are stocks with market values of equity at or above the 20^{th} percentile of the NYSE market capitalization distribution. log(Cum. dividends/ME) is the ratio of cumulative dividends distributed since the firm's IPO to the current market value of equity. Variables are trimmed at the 1^{st} and 99^{th} percentiles based on the four explanatory variables used in the first column: book-to-market, size, prior one-month return, and prior one-year return skipping a month.

	Regression	
Regressor	(1)	(2)
log(ME)	−0.09	−0.09
	(−2.42)	(−2.63)
$r_{1,1}$	−3.23	−3.32
	(−7.61)	(−7.88)
$r_{12,2}$	0.83	0.82
	(4.46)	(4.44)
log(BE/ME)	0.06	0.06
	(0.79)	(0.85)
log(RE/ME)	0.17	0.16
	(4.66)	(4.37)
log(Cum. dividends/ME)		0.00
		(−0.57)
Indicator variables		
RE ≤ 0	−0.55	−0.51
	(−2.81)	(−2.65)
Cum. dividends $= 0$		−0.04
		(−0.41)
Avg. Adj. R^2	6.12%	6.49%

Table 6: Fama-MacBeth regressions with cumulative earnings

Each line shows estimates from a separate set of Fama-MacBeth regressions in which we accumulate earnings over different windows. Our measure of accumulated earnings is the natural logarithm of the sum of earnings over the window deflated by the most recent market value of equity. The regressions include as (unreported) controls size, prior one-month return, and prior one-year return skipping a month. The sample consists of all but microcap firms with positive book value of equity and non-missing values for retained earnings and contributed capital. All but microcap firms are stocks with market values of equity at or above the 20^{th} percentile of the NYSE market capitalization distribution. The table presents results Fama and MacBeth (1973) regressions that includes accumulated net income and accumulated income before extraordinary items over horizons out to five years in the past. In the regressions that accumulate earnings up to five years, the sample consist of firms with at least five years of earnings data. The sample starts in July 1967 and ends in December 2016.

Window,	log(BE/ME)		log(Earnings/ME)		Earnings < 0		Avg.
years	EST	t-value	EST	t-value	EST	t-value	adj. R^2
Baseline	0.220	3.18					5.17%
			Earnings = Accumulated net income				
$t-1$	0.167	2.68	0.107	2.48	−0.507	−2.55	5.88%
$t-2$ to $t-1$	0.153	2.56	0.109	2.28	−0.432	−2.24	5.93%
$t-3$ to $t-1$	0.133	2.28	0.127	2.54	−0.410	−2.19	5.93%
$t-4$ to $t-1$	0.101	1.75	0.165	3.10	−0.457	−2.37	5.98%
$t-5$ to $t-1$	0.080	1.39	0.191	3.63	−0.497	−2.57	5.98%
			Earnings = Accumulated income before extraordinary items				
$t-1$	0.148	2.41	0.137	2.80	−0.780	−3.20	6.02%
$t-2$ to $t-1$	0.115	2.01	0.178	3.32	−0.729	−3.17	6.04%
$t-3$ to $t-1$	0.085	1.51	0.192	3.33	−0.696	−3.15	6.04%
$t-4$ to $t-1$	0.056	1.01	0.227	3.81	−0.738	−3.39	6.05%
$t-5$ to $t-1$	0.043	0.78	0.241	3.95	−0.743	−3.37	6.08%

Table 7: Returns on portfolios sorted by book-to-market, retained earnings-to-market, and contributed capital-to-market: U.S. sub-period and international evidence

This table reports monthly CAPM alphas for quintiles sorted by book-to-market, retained earnings-to-market, and contributed capital-to-market. We form portfolios at the end of each June and compute value-weighted returns on the stocks over the next year. The U.S. sorts use NYSE breakpoints. The World excluding U.S. sorts use a large-stock cutoff, which is the point at which the largest firms in each region account for 90% of total market capitalization. The pre-1990 U.S. sample begins in July 1964 and ends in June 1990. The post-1990 samples begin in July 1990 and end in December 2016. The World excluding U.S. sample includes all countries except the U.S. and emerging markets.

	CAPM alphas			t-values		
	Pre-1990	Post-1990		Pre-1990	Post-1990	
Quintile	U.S.	U.S.	World ex U.S.	U.S.	U.S.	World ex U.S.
			Book-to-market			
1 (low)	−0.12	−0.03	−0.26	−1.38	−0.38	−3.04
2	−0.03	0.13	0.06	−0.47	1.75	0.93
3	0.10	0.12	0.07	1.18	1.19	1.35
4	0.32	0.13	0.16	3.08	1.13	2.14
5 (high)	0.43	0.29	0.23	3.05	2.02	2.32
H−L	0.55	0.33	0.49	2.67	1.61	3.06
			Retained earnings-to-market			
1 (low)	−0.26	−0.33	−0.32	−2.62	−2.19	−3.42
2	−0.04	−0.02	−0.06	−0.69	−0.28	−0.95
3	0.13	0.29	0.14	1.72	3.82	2.41
4	0.29	0.33	0.18	3.28	2.93	2.28
5 (high)	0.35	0.31	0.23	2.64	2.23	2.60
H−L	0.60	0.64	0.55	2.96	2.67	3.64
			Contributed capital-to-market			
1 (low)	0.02	0.11	0.03	0.30	1.63	0.34
2	−0.03	−0.08	0.01	−0.46	−1.06	0.11
3	0.02	−0.03	0.03	0.23	−0.38	0.58
4	0.10	0.03	0.11	1.11	0.29	1.70
5 (high)	0.18	0.09	0.09	1.34	0.56	0.95
H−L	0.15	−0.03	0.06	0.86	−0.14	0.45

Table 8: Returns on portfolios sorted by book-to-market, retained earnings-to-market, and contributed capital-to-market: Pre-1964 evidence

This table reports monthly CAPM alphas for portfolios covering the period 1938 through 1964. We sort the quintiles based on book-to-market and proxies for retained earnings-to-market and contributed capital-to-market. We form portfolios at the end of each June and compute value-weighted returns on the stocks over the next year. The sorts use NYSE breakpoints. To create the proxies for retained earnings-to-market and contributed capital-to-market, we regress book-to-market on income accumulated up to 20 years in the past. We use the fitted value from this regression as the proxy for retained earnings-to-market and residual as the proxy for contributed capital-to-market. For this analysis, we obtain accounting data from the Graham et al. (2015) database and stock returns from CRSP.

	Book-to-market component			Book-to-market component		
Quintile	Total	Retained earnings	Contributed capital	Total	Retained earnings	Contributed capital
	CAPM alphas			***t*-values**		
1 (low)	0.03	−0.13	0.08	0.47	1.61	1.32
2	0.02	0.12	0.01	0.33	2.02	0.10
3	0.29	0.19	0.15	3.54	2.94	1.87
4	0.15	0.20	0.15	1.33	2.69	1.29
5 (high)	0.36	0.38	0.29	2.40	3.26	2.06
H−L	0.33	0.51	0.20	1.68	3.20	1.12

APPENDIX

Share issuances

Par value shares have a face value assigned to them. Such shares may be issued at par, above par or below par. When par value shares are issued exactly at par, cash is debited and common stock or preferred stock account is credited. In case of issuance above par, cash account is debited for the total cash received by the company, common stock or preferred stock is credited for the par value multiplied by number shares issued and additional paid-in capital account is credited for the excess of cash received over the par value multiplied by number of shares issued. When par value shares are issued below par, cash is debited for the actual amount received, common stock or preferred stock is credited for the total par value, and discount on capital is debited for the excess of total par value over cash received. The discount on capital is part of shareholders' equity and it appears as a deduction from other equity accounts on balance sheet.

Share repurchases

How do share repurchases affect the book value of common equity? There are two governing legal principles involved:

- A corporation cannot be its own shareholder, so treasury stock cannot be recorded as assets. The debit therefore must be to reduce stockholders' equity.
- A corporation cannot create earnings through trading in its own capital stock, so treasury stock transactions generally increase or decrease contributed capital. Some treasury stock transactions decrease retained earnings, but never increase.

There are two methods of accounting for treasury stock: the cost method and the par method. Their use depends on what the company intends to do with the repurchased stock. The cost method is used when the company might want to reissue the shares in future. It records the amount paid to repurchase stock as increasing treasury stock, which is a contra account to stockholders' equity and therefore has a debit balance. No distinction is made between par value and the premium paid in the purchase transaction. The corresponding credit reduces cash. If treasury stock subsequently is sold at a price greater than its repurchase cost, the gain is recorded as additional paid-in-capital (treasury stock). If a subsequent sale is at less than the repurchase cost, the loss is recorded as a reduction in additional paid-in-capital (treasury stock) and, if that account is fully depleted, the balance is a reduction in retained earnings.

The par method is used if the board retires the stock when it is repurchased. The stock is legally cancelled and common stock and additional paid-in-capital are reduced by the amounts recorded when the stock was originally issued to stockholders. If the repurchase price is greater than the amount originally received when the stock was issued, the "loss" reduces retained earnings. If it is less, the "gain" increases additional paid-in-capital.

The par value method permanently reduces the Stockholders Equity accounts. The cost method reduces them temporarily, using a contra account that is shown on the balance sheet as a deduction from Stockholders Equity. The cost method appears to be more prevalent method. Of the 500 firms surveyed by the American Institute of Certified Public Accountants about their accounting policies in 2010, 340 firms engaged in stock repurchases (AICPA, 2010). Of those firms, 321 used the cost method and 19 used the par method.

World excluding U.S. data

We obtain international stock returns from Datastream and international accounting data from Worldscope. We use the Thomson Reuters QI Direct Table to link the two databases. We include inactive companies and exclude financial firms, depositary receipts, real estate investment trusts, preferred shares, and warrants. We retain the major equity security issued by a firm and use the returns from its primary exchange listing. We further restrict the sample to stocks traded on major exchanges. In most countries, these are the exchanges on which the majority of a country's stocks are traded. We compute returns based on the end of month return index (RI) provided by Datastream ($RI_t/RI_{t-1} - 1$). We set equal to zero monthly returns greater than 300% that reverse within one month (Ince and Porter, 2006; Hou, Karolyi, and Kho, 2011). To further remove outliers, we exclude monthly returns in the bottom and top 0.1% of the returns distribution in each country and require that a firm has a minimum of 12 monthly returns during the sample period.

We compute the market value of equity as the end of year share price (Worldscope item 5085) times the number of shares outstanding (Worldscope item 5301). We then use the book value of common equity (Worldscope item 3501) to construct the book-to-market ratio. Retained earnings are defined as the sum of retained earnings excluding reserves (Worldscope item 3495), unappropriated and other appropriated reserves (Worldscope item 3494 plus item 3490), and equity in untaxed reserves (Worldscope item 3490).[21] When reserves data are missing, we assume there are none and set them to zero. We calculate contributed capital as the sum of common stock (Worldscope item 3480) and capital surplus (Worldscope item 3481) less treasury stock (Worldscope item 3499). We set the last two items equal to zero if missing. We compute accumulated other comprehensive income as the difference between the book value of equity and the sum of retained earnings and contributed capital. We do not use negative values of book-to-market or retained earnings-to-market when calculating the portfolio breakpoints.

[21] Regulations in some countries require companies to set aside and classify a proportion of earnings or certain of its components as reserves, with restrictions on their distribution as dividends.

Table A.1: Two-way portfolio sorts: Book-to-market and retained earnings-to-market

This table reports CAPM alphas and t-values associated with those alphas for value-weighted portfolios sorted by book-to-market and retained earnings-to-market. Panel A sorts stocks first into quintiles by retained earnings-to-market and then, *conditional* on the retained earnings quintile, into quintiles by book-to-market. Panel B reverses the order of the sorts. Each sort uses NYSE breakpoints. The portfolios are rebalanced annually at then end of June. The estimates in the bottom right corners, preceded by the ↳ signs, correspond to the averages of the five strategies above them. In Panel A, this strategy is a retained earnings-neutral strategy that buys firms with high book-to-markets and sells those with low book-to-markets; in Panel B, it is a book-to-market-neutral strategy that buys firms with high retained earnings-to-market and sells those with low retained earnings-to-market. The sample starts in July 1964 and ends in December 2016.

Panel A: Conditional portfolio sorts on retained earnings-to-market and book-to-market

	BE/ME (Conditional sort)					
RE/ME	Low	2	3	4	High	H − L
			Monthly CAPM alphas			
Low	−0.36	−0.21	−0.23	−0.18	−0.02	0.33
2	0.07	−0.10	−0.13	−0.07	0.10	0.03
3	0.21	0.08	0.02	0.23	0.44	0.23
4	0.30	0.21	0.44	0.34	0.32	0.03
High	0.34	0.28	0.32	0.22	0.52	0.18
H − L	0.70	0.49	0.55	0.39	0.54	↳ 0.16
			***t*-values**			
Low	−2.46	−1.66	−2.17	−1.49	−0.15	1.52
2	0.83	−1.13	−1.35	−0.69	0.88	0.17
3	2.39	0.88	0.18	2.21	3.60	1.54
4	3.25	1.99	3.95	3.15	2.61	0.19
High	3.22	2.42	2.59	1.56	3.20	1.21
H − L	3.40	2.56	3.26	2.29	3.06	↳ 1.50

Panel B: Conditional portfolio sorts on book-to-market and retained earnings-to-market

	RE/ME (Conditional)					
BE/ME	Low	2	3	4	High	H − L
			Monthly CAPM alphas			
Low	−0.44	−0.22	0.02	−0.01	0.19	0.63
2	−0.23	−0.08	0.00	0.20	0.28	0.52
3	−0.23	0.01	0.15	0.24	0.31	0.54
4	−0.01	0.35	0.34	0.25	0.36	0.38
High	0.24	0.36	0.37	0.32	0.48	0.24
H − L	0.68	0.58	0.35	0.34	0.29	↳ 0.46
			***t*-values**			
Low	−2.92	−1.80	0.22	−0.16	2.18	3.28
2	−2.12	−0.88	−0.05	2.34	2.80	3.31
3	−2.03	0.11	1.47	2.33	3.00	3.51
4	−0.11	2.99	3.25	2.10	3.01	2.58
High	1.60	2.81	2.85	2.32	2.86	1.30
H − L	3.15	2.80	1.91	1.91	1.48	↳ 4.31