ELSEVIER Journal of Financial Economics 49 (1998) 141–160

JOURNAL OF Financial ECONOMICS

Book-to-market ratios as predictors of market returns[1]

Jeffrey Pontiff*, Lawrence D. Schall

School of Business Administration, University of Washington, Seattle, WA 98195, USA

Received 11 August 1995; received in revised form 14 July 1997

Abstract

The book-to-market ratio of the Dow Jones Industrial Average predicts market returns and small firm excess returns over the period 1926–1994. The DJIA book-to-market ratio contains information about future returns that is not captured by other variables such as interest yield spreads and dividend yields. The DJIA book-to-market ratio's predictive ability is specific to the pre-1960 sample. In contrast, the S&P book-to-market ratio provides some predictive ability in the post-1960 period, although this relation is dramatically weaker than the Dow Jones pre-1960 findings. The predictive ability of book-to-market ratios appears to stem from the relation between book value and future earnings. © 1998 Elsevier Science S.A. All rights reserved.

JEL classification: G10; G11

Keywords: Book-to-market; Book value; Cash flow; Forecast; Market return; Prediction; Predictor; Dow Jones Industrial Average; Standard and Poors

1. Introduction

Economists and financial practitioners have sought to identify variables that predict stock returns. This study examines the ability of an aggregate

* Corresponding author. Tel.: 206/543-3021; fax: 206/685-9392; e-mail: pontiff@u.washington.edu.

[1] This paper has benefited from comments from workshop participants at Arizona State University, University of Washington, and the Western Finance Association meetings, as well as Hendrik Bessembinder, Gary Biddle, Michael Brennan, J.B. Chay, Roger Edelen, Eugene Fama (the referee), Wayne Ferson, Mark Huson, S.P. Kothari, Paul Malatesta, Eric Noreen, Charles Nelson, Vance Roley, Andrew Siegel, Jerold Warner, and Simon Wheatley. We would also like to thank David Myers and Denny Stephenus for research assistance, and Kenneth French for providing some of the data used in this study.

PII S0304-405X(98)00020-8

book-to-market ratio to forecast market returns. The use of the book-to-market ratio is motivated by the findings of Fama and French (1992), who show that the book-to-market ratio of individual stocks has the ability to explain cross-sectional variation in stock returns. An independent paper by Kothari and Shanken (1997) uses a Bayesian framework to document that the book-to-market ratio of the Dow Jones Industrial Index (DJIA) predicts market returns over the period 1926–1991. They provide evidence that the book-to-market ratio sometimes predicts negative expected returns, although this finding is not robust to the later half of their sample.

Similar to Kothari and Shanken, we show that during the period 1926–1994 the DJIA book-to-market ratio predicts market returns. In addition, our estimation includes other variables that have been shown to predict market returns, such as default spreads, interest rates, term structure slopes, and dividend yields. Fama (1991) summarizes the studies that provide evidence of return predictability. We also show that the DJIA book-to-market ratio predicts small firm excess returns during this period. We document a structural difference between the post- and pre-1960 subsamples. The ability of the book-to-market ratio to predict returns dissipates in the second half of the sample. One explanation for this result is that the DJIA is less representative of the equities market as a whole in the post-1960 period. We show that the book-to-market ratio of the Standard and Poor's Industrial Index (S&P) is slightly better at forecasting market returns, but we still cannot reject the null hypothesis of no predictive ability.

We provide initial evidence regarding the source of the book-to-market ratio's predictive ability. Similar to the conjecture of Ball (1978) and, more recently, Berk (1995) and Sharathchandra and Thompson (1994), we argue that the book-to-market ratio captures information about expected future returns because book value proxies for expected cash flows. Thus, the book-to-market ratio is the ratio of a cash flow proxy to the current price level. As the discount rate changes, the price level changes, and therefore the book-to-market ratio changes. Holding expected cash flow constant, an increase in the discount rate produces a decrease in market value and thus an increase in the book-to-market ratio. This could explain a positive relation between current book-to-market ratios and future returns (see the Appendix for a brief discussion of this point). Consistent with our hypothesis, the S&P book value is a better predictor of a cash flow surrogate (market earnings) post-1960 than is the DJIA book value. Our results neither support nor refute market efficiency, since our analysis holds regardless of whether discount rates are affected by irrational traders.

In Section 2 we describe the construction of the data used in the paper. Section 3 describes the estimation technique we use to address a small sample bias in slope coefficients. Section 4 presents results regarding the ability of the DJIA book-to-market ratio to predict returns and compares the predictive ability of the DJIA book-to-market ratio with that of the S&P book-to-market ratio. Section 5 examines whether our findings are related to the ability of book value to predict aggregate earnings. Section 6 presents conclusions.

2. Data

2.1. DJIA book-to-market ratio

We calculate our book-to-market ratio from the Dow Jones Industrial Average, which was introduced in 1896. The DJIA is an index of the prices of 30 large U.S. corporations selected to represent a cross section of U.S. industry (the number of companies represented by the DJIA was 20 until the change to 30 in 1928). Changes in the composition of the DJIA are made periodically to reflect changes in the companies and in the economy. The DJIA is the sum of the prices of the 30 stocks divided by a single divisor. Whenever a DJIA stock splits, or a company in the DJIA is replaced by another company, the divisor is adjusted so that the value of the index is unaffected by the split or the change in company. Being a price-weighted average, the relative weight of each company in the index is proportional to the price of one share of that company's stock, and is not proportional to the total value of the company's outstanding shares (as for 'value-weighted' indexes, such as the Standard and Poor's and NYSE indexes).

We use the December year-end book value of the DJIA from the Value Line publication, 'A Long Term Perspective'. This source provides the annual book value of the DJIA from 1920 to 1993. The book value of the DJIA is the sum of the fiscal year-end per-share book values of all stocks in the DJIA divided by the same divisor used in computing the DJIA price index. If a firm in the DJIA does not have a December fiscal year-end, then the computation uses the firm's book value prevailing at the company's fiscal year-end earlier in the calendar year. Currently, 26 of the 30 stocks in the DJIA have December fiscal year-ends. For the beginning of our sample, all stocks in the DJIA had December fiscal year-ends. In order to ensure that our measure reflects information that was available to market participants, we do not incorporate the previous year's book value until March. We construct an aggregate book-to-market ratio by dividing the DJIA book value by the DJIA index's level. A monthly book-to-market ratio is constructed by dividing the most recent year's book value by the contemporaneous level of the DJIA. Thus, our book-to-market ratio in March 1988 is the book value from December 1987 divided by the level of the DJIA in March 1988; our measure in February 1988 is the book value from December 1986 divided by the level of the DJIA in February 1988. We expect this construction to produce a book-to-market variable that incorporates as much contemporaneous information as possible.

2.2. Dividend yields

Other studies have documented that dividend yields predict stock returns (e.g., Fama and French, 1988; Nelson and Kim, 1993). We include a dividend yield measure as an independent variable in our estimation. Dividend yields are constructed using the same method as in Fama and French (1988). Our dividend yield measure, DIV, equals total dividend payments accruing to the Center for Research in Security Prices (CRSP) value-weighted index over the previous 12 months divided by the current level of the index.

2.3. Bond variables

Various studies have demonstrated that interest rates are related to future equity market returns. Short-term rates are negatively related to future stock returns (Fama and Schwert, 1977). Also, the slope of the term structure of interest rates is positively related to future stock returns (Keim and Stambaugh, 1986). Fama (1990) shows that the yield of default-prone bonds minus the yield of less risky bonds is positively related to future returns.

We include three interest rate variables in our estimation to capture information about the short-term rate, the difference between long- and short-term rates, and the difference between corporate rates and risk-free rates. Yield data are from Ibbotson Associates and Citibase. The Ibbotson data start in 1926 and end in 1987, while the Citibase data start in 1947 and end in 1994. In order to use the CRSP data for the entire 1926–1994 period, the Ibbotson and Citibase files must be spliced. We create three variables from the Citibase files: YLD3 is the yield of a T-bill that matures in three months, DEF is the average yield of bonds rated Baa by Moody's minus the average yield of bonds with a Moody's rating of Aaa, and TERM is the average yield of Treasury bonds with more than ten years to maturity minus the yield of T-bills that mature in three months. All three of the Citibase monthly variables are averages of weekly data within the month. Using each of these three Citibase series as dependent variables, we estimate regression equations using the following Ibbotson data as independent variables: the average yield of bonds with more than 15 years to maturity, the average yield of bonds with less than 15 years to maturity, the three month Treasury bill yield, the average yield of Baa-rated bonds, the average yield of Aaa-rated bonds, and the lagged values of each of these variables. Lagged values are used in the regression because the Ibbotson data are from the end of the month, whereas the Citibase data are averages of weekly data.

The results of the regression suggest that the data can be reliably spliced. The adjusted R^2s from the DEF, TERM, and YLD3 regressions are, respectively, 0.82, 0.87, and 0.99. The actual values of these variables are used after 1947 when the Citibase data are available, and the predicted values from the regression analysis are used before 1947. Our decision to use actual Citibase variables post-1947 and predicted values from regressions of Citibase variables on current and lagged Ibbotson data pre-1947 was influenced by the fact that the Citibase data are within-month averages, whereas the Ibbotson data are month-end. Our choice avoids implicitly using future information about the independent variables in the regressions, since the Citibase variables and the predicted variables are not constructed from data from the following month.

2.4. Descriptive statistics

Panel A of Table 1 presents descriptive statistics of the variables used in this study. Five return variables are presented: VWRET, the return of the CRSP value-weighted index; EWRET, the return of the CRSP equally weighted index; SMALL, the difference between the CRSP equally weighted index return and the CRSP value-weighted index return (EWRET–VWRET); SMB, the return of firms with small capitalization minus the return of firms with big capitalization (Fama and French, 1993); and HML, the return of firms with high book-to-market ratios minus the return of firms with low book-to-market ratios (Fama and French, 1993). The autocorrelations of these return variables at lags of one month range from 0.11 to 0.21. For lags of over one month, the autocorrelations are relatively small. Also presented are summary statistics for the independent variables (YLD3, DEF, TERM, DIV, and the book-to-market ratio, BM). All of these variables are highly persistent. One-month autocorrelations range from 0.97 to 0.99. At lag 12, the autocorrelations decrease to levels between 0.60 and 0.90.

Panel B presents contemporaneous cross-correlations. SMB and SMALL are highly correlated with a coefficient of 0.91. This correlation implies that subtracting the returns of VWRET from EWRET creates a good proxy for the small minus big portfolio used by Fama and French (1992). This is important because SMB begins in 1963, whereas SMALL can be constructed back to 1926. Thus, this paper will use SMALL as a proxy for SMB.

The cross-correlations suggest that the book-to-market ratio is strongly contemporaneously correlated with default spread (DEF) and dividend yield (DIV). The correlation with dividend yield is particularly high (0.67) which is reasonable given that prices are used in the denominator of both variables and that both book value and dividends are persistent. The book-to-market ratio and dividend yield are both negatively correlated with the CRSP value-weighted index and the CRSP equally weighted index. This finding is expected since the book-to-market ratio and dividend yield are month-end variables that are denominated with price levels. Positive (negative) market return shocks will produce negative (positive) shocks to price-denominated variables. This relation, combined with the high autocorrelation of the book-to-market ratio and dividend yield series, produces the spurious estimation bias that is discussed next in Section 3.

Table 1
Summary Statistics

Summary statistics computed from data from January 1926 to August 1994 except for statistics involving the variables SMB and HML, which use data from July 1963 to December 1991. The contemporaneous correlations involve correlation of variables for the same month, thus dividend yield and book-to-market are both month-end. VWRET and EWRET are the returns of the CRSP value-weighted index and CRSP equally weighted index, respectively. SMALL is the difference between EWRET and VWRET. SMB is the return of firms with small capitalization minus the return of firms with big capitalization. HML is the return of firms with high book-to-market ratios minus the return of firms with low book-to-market ratios. YLD3 is the yield of a T-bill that matures in 3 months. DEF is the average yield of bonds rated by Moody's Baa minus the average yield of bonds rated by Moody's Aaa. TERM is the average yield of Treasury bonds with greater than 10 years to maturity minus the yield of T-bills that mature in 3 months. DIV is the total dividend payments accruing to the CRSP value-weighted index over the previous twelve months divided by the current level of the index. BM is the book-to-market ratio of the Dow Jones index.

Panel A: Autocorrelations

Name	Mean	Std. Dev.	1	2	3	4	5	6	7	8	9	10	11	12	24	36	48	60
VWRET	0.939	5.58	0.11	− 0.02	− 0.12	0.01	0.07	− 0.03	0.01	0.03	0.08	0.01	− 0.02	− 0.00	0.03	0.02	− 0.02	0.02
EWRET	1.305	7.69	0.21	0.02	− 0.11	− 0.06	− 0.00	− 0.04	0.00	0.02	0.14	0.07	− 0.00	0.02	0.02	0.04	0.00	0.03
SMALL	0.366	3.44	0.18	0.07	− 0.05	− 0.05	− 0.07	− 0.01	0.03	0.07	0.13	0.12	0.10	0.19	0.11	0.10	0.10	0.08
SMB	0.271	2.89	0.19	0.07	− 0.01	0.03	0.00	0.11	− 0.01	0.01	− 0.04	0.06	0.10	0.23	0.06	0.15	0.12	0.05
HML	0.400	2.54	0.18	0.06	− 0.01	− 0.04	− 0.02	0.10	0.03	− 0.03	− 0.02	0.00	0.02	0.07	0.12	0.06	0.02	0.01
YLD3	3.642	3.12	0.99	0.98	0.97	0.96	0.95	0.94	0.94	0.93	0.93	0.92	0.91	0.90	0.81	0.75	0.72	0.70
DEF	0.922	0.38	0.97	0.93	0.90	0.88	0.87	0.84	0.82	0.80	0.78	0.75	0.73	0.71	0.50	0.35	0.35	0.34
TERM	1.520	1.41	0.97	0.93	0.90	0.87	0.85	0.82	0.81	0.79	0.78	0.76	0.74	0.72	0.48	0.30	0.25	0.29
DIV	4.267	1.37	0.97	0.93	0.89	0.87	0.84	0.81	0.78	0.74	0.71	0.67	0.63	0.60	0.36	0.26	0.23	0.25
BM	0.668	0.23	0.97	0.94	0.91	0.89	0.88	0.85	0.83	0.81	0.78	0.74	0.70	0.68	0.43	0.23	0.08	− 0.00

Panel B: Contemporaneous correlations

	EWRET	SMALL	SMB	HML	YLD3	DEF	TERM	DIV	BM
VWRET	0.91	0.42	0.35	− 0.38	− 0.04	0.11	0.05	− 0.15	− 0.07
EWRET		0.75	0.73	− 0.22	− 0.06	0.15	0.09	− 0.10	− 0.01
SMALL			0.91	0.14	− 0.08	0.15	0.12	0.01	0.09
SMB				− 0.08	− 0.07	0.06	0.02	0.03	0.06
HML					0.06	− 0.01	− 0.06	0.06	0.08
YLD3						0.25	− 0.56	− 0.30	0.09
DEF							0.37	0.26	0.50
TERM								0.23	0.08
DIV									0.67

3. Estimation

Our objective is to estimate regressions of various holding period returns on variables such as the book-to-market ratio that are expected to forecast returns. Empirical problems suggest that customary t-statistics do not provide accurate inferences. Both dividend yields and book-to-market ratios are computed by dividing a fairly persistent variable by price levels. Since returns are largely changes in price levels, changes in dividend yields and book-to-market ratios are expected to be negatively related to contemporaneous returns. This induces a spurious bias in parameter estimates from a regression of future returns on these variables (see, for example, Kendall, 1954; Stambaugh, 1986). The intuition behind this bias is that the negative contemporaneous correlation between returns and the persistent independent variables induces a positive estimated relation between future returns and the independent variables, since, in-sample, future returns are usually lower and the contemporaneous independent variables are usually higher. This bias is more pronounced in smaller samples. Slope coefficients in long-run return regressions will have a larger bias than slope coefficients in short-run return regressions, since, for a given sample time period, long-run return regressions contain a smaller number of independent observations than do short-run return regressions.

In order to generate p-values that account for this spurious bias, we use the randomization method of Nelson and Kim (1993). For each of our independent variables, we estimate the following first-order autoregressive process:

$$x_t^i = a + bx_{t-1}^i + e_t^i, \tag{1}$$

where t denotes time period t and i denotes the variable. This equation is estimated for the book-to-market ratio (BM), DEF, TERM, YLD3, and DIV. Since the spurious bias is caused by the contemporaneous correlation between the return and the e_t^i, both the estimates of the e_t^i and the contemporaneous returns are retained. Each grouping of return and e_t^i is randomized. From this randomized series, pseudo-independent variables are created by substituting the randomized e_t^i in the estimate of Eq. (1). Initial values are the x^i picked from a random time period. This process creates a series of pseudo-independent variables and returns that have similar time-series properties as the actual series used to test return predictability, but are generated under the null of no predictability. This randomization is conducted for 1000 iterations, and the parameter estimates form an empirical distribution that is used to calculate p-values. For estimation that involves multiple independent variables, this technique produces p-values that are slightly different from conventional p-values in that the randomized p-values are computed under the null hypothesis that all the independent variables are unrelated to future returns.

4. Regression results

Table 2 shows the regression results for the entire sample period. The left and middle sections of this table present estimations that use the returns of the CRSP value-weighted index (VWRET) and the CRSP equally weighted index (EWRET) as dependent variables. Both monthly and annual (April to March) return holding periods are used. Similar to the findings of cross-sectional studies (Fama and French, 1992), the results indicate that there is a general positive relation between book-to-market ratios and future market returns. This finding is typically stronger for the CRSP equally weighted returns than for the CRSP value-weighted returns. When the other variables are included as independent variables, the predictive ability of the book-to-market ratio is no longer statistically significant for value-weighted returns (in general the book-to-market ratio has lower *p*-values than the other variables), although statistical significance remains when equally weighted returns are used. Other than the book-to-market ratio, the only variable that has a significant ability to predict market returns is the default spread variable (DEF), which has a positive relation to one-month market returns.

The ability of the book-to-market ratio to predict small firm returns in excess of large firm returns (SMALL) is presented in the right-hand columns of Table 2. The book-to-market ratio is positively and significantly related to SMALL for both holding periods. These results hold even with the inclusion of other variables. Default spreads also have an ability to forecast small firm excess returns. In the specification that does not include the book-to-market ratio, dividend yield has a positive relation with SMALL. When the book-to-market ratio is used in the estimation, dividend yield has a negative relation with SMALL, which suggests that this result is influenced by collinearity between dividend yield and the book-to-market ratio.

The above results indicate that the DJIA book-to-market ratio is useful in predicting both market returns and small firm excess returns. In general, the predictive ability of the DJIA book-to-market ratio appears stronger than other variables that are typically used by financial economists to predict returns. Table 3 addresses whether these findings are robust to different subperiods. The results are reestimated for the periods of January 1926 to June 1959 and July 1959 to August 1994. Panel A of Table 3 presents results from the first subperiod. Overall, the ability of the book-to-market ratio to predict market returns (VWRET and EWRET) is stronger in the first subperiod than over the entire sample period. The term structure slope variable (TERM) and the default spread variable (DEF) are, respectively, negatively and positively related to one-month and 12-month returns.

The book-to-market ratio's ability to predict small firm excess returns, SMALL, in the first subperiod is more consistent with the entire sample period than are the VWRET and EWRET findings. In the SMALL regressions

Table 2
Regression results for the entire sample (January 1926–August 1994)

Regressions of CRSP value-weighted market returns (VWRET), equally weighted market returns (EWRET), and the difference between these returns (SMALL), on the Dow Jones book-to-market ratio (BM), default yield premium (DEF), long-term yield premium (TERM), the three-month T-bill rate (YLD3), and the dividend yield of the value-weighted index over the past 12 months (DIV). Returns are calculated for various holding periods subsequent to the independent variables. Nelson and Kim (1993) randomized p-values are in brackets. Parameter estimates and Nelson and Kim p-values are in bold if the p-value is greater than 95% or less than 5%.

VWRET							EWRET							SMALL						
Intercept	BM	DEF	TERM	YLD3	DIV	Adj.- R^2	Intercept	BM	DEF	TERM	YLD3	DIV	Adj.- R^2	Intercept	BM	DEF	TERM	YLD3	DIV	Adj.- R^2
One-month returns																				
−1.08	**3.02**					0.01	**−2.40**	**5.55**					0.03	**−1.31**	**2.53**					0.03
[0.99]	**[0.01]**						**[1.00]**	**[0.00]**						**[1.00]**	**[0.00]**					
− 0.84	3.12	**1.10**	0.01	− 0.12	− 0.22	0.02	− 1.81	**5.56**	**2.53**	− 0.03	− 0.29	**−0.43**	0.04	− 0.97	**2.44**	**1.43**	− 0.04	− 0.18	− 0.20	0.05
[0.78]	[0.07]	**[0.05]**	[0.55]	[0.82]	[0.89]		[0.85]	**[0.02]**	**[0.01]**	[0.69]	[0.93]	**[0.95]**		[0.93]	**[0.02]**	**[0.00]**	[0.64]	[0.64]	[0.93]	
− 0.80		**1.74**	− 0.05	− 0.09	0.12	0.01	− 1.74		**3.66**	− 0.14	− 0.24	0.18	0.03	− 0.94		**1.92**	− 0.09	− 0.15	0.06	0.04
[0.80]		**[0.01]**	[0.62]	[0.76]	[0.53]		[0.89]		**[0.00]**	[0.71]	[0.91]	[0.47]		[0.94]		**[0.00]**	[0.62]	[0.84]	[0.43]	
Annual returns																				
−16.20	**42.18**					0.16	**−43.85**	**91.93**					0.28	**−27.65**	**49.75**					0.31
[0.99]	**[0.01]**						**[0.99]**	**[0.01]**						**[0.99]**	**[0.01]**					
−18.80	45.86	− 4.12	2.30	0.52	− 0.36	0.11	**−43.69**	**103.36**	5.16	3.90	− 1.02	− 3.45	0.28	**−24.89**	**57.50**	9.28	1.60	− 1.54	− 3.09	0.36
[0.92]	[0.11]	[0.60]	[0.35]	[0.62]	[0.55]		**[0.97]**	**[0.04]**	[0.37]	[0.29]	[0.62]	[0.74]		**[0.98]**	**[0.02]**	[0.18]	[0.29]	[0.81]	[0.88]	
−16.30		6.16	0.89	0.72	4.49	0.05	**−38.06**		28.33	0.72	− 0.55	7.49	0.14	**−21.75**		**22.17**	− 0.17	− 1.28	3.00	0.20
[0.90]		[0.30]	[0.47]	[0.41]	[0.10]		**[0.96]**		[0.07]	[0.49]	[0.61]	[0.09]		**[0.96]**		**[0.03]**	[0.52]	[0.79]	[0.14]	

Table 3

Regression results for two subsamples, January 1926 to June 1959 and July 1959 to August 1994

Regressions of CRSP value-weighted market returns (VWRET), equally weighted market returns (EWRET), and the difference between these returns (SMALL) on the Dow Jones book-to-market ratio (BM), default yield premium (DEF), long-term yield premium (TERM), the three-month T-bill rate (YLD3), and the dividend yield of the value-weighted index over the past 12 months (DIV). Returns are calculated for various holding periods subsequent to the independent variables. Nelson and Kim (1993) randomized *p*-values are in brackets. Parameter estimates and Nelson and Kim *p*-values are in bold if the *p*-value is greater than 95% or less than 5%.

Panel A: *January* 1926 *to June* 1959

VWRET							EWRET							SMALL						
Intercept	BM	DEF	TERM	YLD3	DIV	Adj.-R^2	Intercept	BM	DEF	TERM	YLD3	DIV	Adj.-R^2	Intercept	BM	DEF	TERM	YLD3	DIV	Adj.-R^2
One-month returns																				
−2.01	**4.32**					0.02	**−4.39**	**8.46**					0.04	**−2.38**	**4.14**					0.06
[0.98]	**[0.03]**						**[0.99]**	**[0.01]**						**[1.00]**	**[0.00]**					
−1.23	**9.51**	**5.03**	**−2.14**	−0.49	**−0.79**	0.04	−4.71	**13.80**	**8.56**	**−2.68**	−0.48	**−1.01**	0.06	−3.48	**4.29**	**3.54**	−0.54	0.00	−0.21	0.09
[0.53]	**[0.01]**	**[0.01]**	**[0.99]**	[0.80]	**[0.97]**		[0.75]	**[0.02]**	**[0.00]**	**[0.98]**	[0.80]	**[0.95]**		[0.92]	**[0.02]**	**[0.00]**	[0.92]	[0.73]	[0.77]	
−0.24		**4.46**	**−1.43**	−0.65	0.18	0.00	−3.27		**7.75**	**−1.66**	−0.72	0.40	0.03	−3.03		**3.28**	−0.22	−0.07	0.22	0.07
[0.40]		**[0.01]**	**[0.95]**	[0.86]	[0.51]		[0.68]		**[0.00]**	**[0.94]**	[0.86]	[0.35]		[0.91]		**[0.00]**	[0.83]	[0.76]	[0.19]	
Annual returns																				
−33.41	**66.13**					0.23	**−80.63**	**144.84**					0.40	**−47.23**	**78.71**					0.48
[1.00]	**[0.02]**						**[1.00]**	**[0.00]**						**[1.00]**	**[0.00]**					
−30.44	**113.93**	**47.15**	**−20.80**	−3.55	**−6.36**	0.27	−97.77	**211.94**	**83.36**	−20.73	2.20	**−12.70**	0.48	**−67.32**	**98.00**	**36.21**	0.07	5.75	**−6.34**	0.67
[0.72]	**[0.00]**	**[0.03]**	**[0.98]**	[0.66]	**[0.98]**		[0.91]	**[0.00]**	**[0.01]**	[0.94]	[0.54]	**[0.99]**		**[0.98]**	**[0.00]**	**[0.01]**	[0.63]	[0.37]	**[0.99]**	
−11.56		**69.10**	**−23.01**	−10.69	4.67	0.02	−62.63		**124.19**	**−24.84**	−11.07	7.82	0.17	**−51.07**		**55.09**	−1.84	−0.38	3.15	0.40
[0.55]		**[0.00]**	**[0.99]**	[0.86]	[0.24]		[0.81]		**[0.00]**	**[0.97]**	[0.81]	[0.20]		**[0.96]**		**[0.00]**	[0.71]	[0.64]	[0.19]	

Panel B: *July* 1959 *to August* 1994

VWRET							EWRET							SMALL						
Intercept	BM	DEF	TERM	YLD3	DIV	Adj.-R^2	Intercept	BM	DEF	TERM	YLD3	DIV	Adj.-R^2	Intercept	BM	DEF	TERM	YLD3	DIV	Adj.-R^2
One-month returns																				
−0.09	1.55					0.00	−0.32	2.33					0.01	−0.23	0.79					0.00
[0.67]	[0.37]						[0.76]	[0.29]						[0.92]	[0.23]					
−4.82	**−6.73**	**2.86**	−0.31	**−0.69**	**3.35**	0.09	−5.11	**−6.51**	**4.30**	−0.53	**−1.03**	**3.71**	0.10	−0.28	−0.22	**[1.43]**	−0.22	**−0.34**	0.36	0.02
[0.90]	**[0.99]**	**[0.01]**	[0.66]	**[0.99]**	**[0.01]**		[0.89]	**[0.99]**	**[0.00]**	[0.92]	**[0.99]**	**[0.01]**		[0.59]	[0.59]	**[0.01]**	[0.85]	**[0.98]**	[0.23]	
−3.70		**2.01**	0.02	**−0.56**	**1.70**	0.07	−4.02		**3.47**	−0.21	**−0.90**	**2.12**	0.09	−0.32		**1.46**	−0.23	**−0.34**	0.42	0.03
[0.88]		**[0.01]**	[0.50]	**[0.98]**	**[0.03]**		[0.86]		**[0.01]**	[0.71]	**[0.99]**	**[0.02]**		[0.61]		**[0.01]**	[0.87]	**[0.99]**	[0.89]	
Annual returns																				
−0.06	18.48					0.06	−9.78	38.21					0.15	**−9.12**	19.72					0.10
[0.73]	[0.28]						[0.89]	[0.13]						**[0.96]**	[0.07]					
−40.15	**−37.32**	1.13	0.32	−1.48	**23.82**	0.30	**−58.50**	−40.50	−2.97	−2.97	−4.78	**36.45**	0.34	**−18.54**	−13.18	3.59	−3.28	−3.29	12.62	0.15
[0.96]	**[0.95]**	[0.36]	[0.47]	[0.65]	**[0.02]**		**[0.98]**	[0.89]	[0.53]	[0.65]	[0.84]	**[0.02]**		**[0.02]**	[0.74]	[0.67]	[0.79]	[0.90]	[0.06]	
−34.13		−3.47	2.33	−0.45	14.00	0.24	**−52.16**		−0.27	−0.78	−3.65	**25.80**	0.32	**−18.03**		3.21	−3.11	−3.21	**11.79**	0.18
[0.96]		[0.58]	[0.61]	[0.50]	[0.09]		**[0.98]**		[0.47]	[0.50]	[0.78]	**[0.02]**		**[0.03]**		[0.33]	[0.82]	[0.92]	**[0.01]**	

described in Panel A of Table 3, the strongest independent variables are the book-to-market ratio and DEF, both of which have a positive, significant relation to one-month small firm excess returns. In the annual SMALL regressions, the slope coefficient on DEF is insignificant, whereas the coefficient on the book-to-market ratio remains positive and significant.

Panel B of Table 3 presents regression estimates for the second half of the sample (July 1959 to August 1994). These results differ dramatically from the results over the entire period and from the first subperiod. The book-to-market ratio does not have a significant positive relation with any of the return measures (VWRET, EWRET, or SMALL). In fact, when the book-to-market ratio is considered with dividend yields and interest rate spreads as independent variables, there exists a negative and frequently significant relation between the book-to-market ratio and future market returns. Another difference between the second half of the sample and the first half of the sample is that the slope coefficient on dividend yield is positive and typically significant in the second half.

Similar to our post-1959 subsample, Kothari and Shanken (1997) examine the period 1963–1991. Their findings are consistent with ours. Using annual returns on the CRSP value- and equal-weighted indexes, they are unable to reject the null of no return predictive ability for the DJIA book-to-market ratio. They are able to reject the null hypothesis at the 2% level when dividend yields are used to forecast value-weighted returns and at the 12% level when dividend yields are used to forecast equally weighted returns. All of the Kothari–Shanken tests are univariate in that none of the specifications includes both dividend yield and the book-to-market ratio.

Although Table 3 implies that our ability to reject the null of no return predictability differs between the 1926–59 and 1959–94 subsamples, this does not imply that there is a significant difference in the book-to-market ratio's forecasting ability in the two subsamples. One possibility is that the power of the test in the second subsample is too low to reject the null. Table 4 addresses this possibility by testing the hypothesis that the slope coefficients in the first subsample are the same as the slope coefficients in the second subsample. As already noted, the slope coefficients in the first subsample are always larger than the slope coefficients in the second subsample, so the difference variable in Table 4 is always positive. We can reject the hypothesis that the second subsample's slope coefficient is larger than the first subsample's slope coefficient at the 5% level for five of the six slope comparisons. The remaining comparison between one-month value-weighted returns produces a *p*-value of 6%. Overall, these results suggest a structural difference in the ability of the DJIA book-to-market ratio to forecast market returns in the two subsamples.

Using the Fama–French data, we investigate whether returns of the Fama–French high-minus-low book-to-market return factor (HML) can be predicted with an aggregate book-to-market ratio. Unfortunately, this data

Table 4
Test of the difference between the first and second sample book-to-market slope coefficients

Difference between first and second slope coefficients from regression of CRSP value-weighted market returns (VWRET), equally weighted market returns (EWRET), and the difference of these returns (SMALL) on the Dow Jones book-to-market ratio. Nelson and Kim (1993) randomized p-values are in brackets for the hypothesis that the difference in the slope coefficients is zero. Parameter estimates and Nelson and Kim p-values are in bold if the p-value is greater than 95% or less than 5%

VWRET			EWRET			SMALL		
Period 1 BM	Period 2 BM	Difference	Period 1 BM	Period 2 BM	Difference	Period 1 BM	Period 2 BM	Difference
One-month returns								
4.32	1.55	2.77	**8.46**	2.33	**6.13**	**4.14**	0.79	**3.35**
[0.03]	[0.32]	[0.06]	**[0.01]**	[0.29]	**[0.03]**	**[0.00]**	[0.23]	**[0.01]**
Annual returns								
66.13	18.48	**47.65**	**144.84**	38.21	**106.63**	**78.71**	19.72	**58.99**
[0.02]	[0.28]	**[0.04]**	**[0.00]**	[0.13]	**[0.02]**	**[0.00]**	[0.07]	**[0.01]**

series begins in 1963, so we are unable to estimate regressions for the entire 1926–1992 period. Table 5 contains monthly and annual HML return regressions. When the book-to-market ratio is the only conditioning variable, HML cannot be predicted. For the specifications that include the DJIA book-to-market ratio, we cannot reject the null hypothesis that there is no relation between the book-to-market ratio and HML returns. In contrast, dividend yields produce negative and statistically significant slope coefficients in all specifications. For both monthly and annual returns, the coefficient on the term structure slope is insignificant. For the one-month HML return regression that includes the DJIA book-to-market ratio, three-month T-bill yields have a positive and statistically significant slope coefficient, and default spreads have an insignificant coefficient. In contrast, for the annual HML return regression, default spreads have a positive and statistically significant slope coefficient and T-bill yields have an insignificant coefficient.

4.1. Standard and Poor's results

We have found that the Dow Jones Industrial Average book-to-market ratio predicts market returns and small firm excess returns before 1960, but not after 1960. A possible explanation for this is that the number of publicly traded firms has increased and consequently the DJIA is less representative of the market since it consists of a small number of securities. We now investigate the

Table 5
Regression estimation using the returns of high market-to-book firms minus the returns of low market-to book firms, over the time period July 1963 to December 1991

Regressions of Fama–French (1993) high minus low market-to-book portfolio return (HML) on the Dow Jones book-to-market ratio (BM), default yield premium (DEF), long-term yield premium (TERM), the three-month T-bill rate (YLD3), and the dividend yield of the value-weighted index over the past 12 months (DIV). Returns are calculated for various holding periods subsequent to the independent variables. Nelson and Kim (1993) randomized *p*-values are in brackets. Parameter estimates and Nelson and Kim *p*-values are in bold if the *p*-value is greater than 95% or less than 5%

Intercept	BM	DEF	TERM	Yld3	DIV	Adj R^2
One-month HML						
0.16	0.36					– 0.00
[0.76]	[0.25]					
1.48	3.16	– 0.66	0.24	**0.38**	**−1.38**	0.03
[0.27]	[0.06]	[0.93]	[0.18]	**[0.03]**	**[1.00]**	
1.25		– 0.18	0.01	0.26	**−0.61**	0.02
[0.33]		[0.68]	[0.53]	[0.07]	**[0.97]**	
Annual HML						
3.39	2.37					– 0.04
[0.57]	[0.33]					
15.41	19.25	**12.80**	0.17	0.36	**−10.21**	0.11
[0.22]	[0.13]	**[0.03]**	[0.43]	[0.47]	**[0.98]**	
14.72		**15.77**	– 1.48	– 0.59	**−5.30**	0.09
[0.22]		**[0.01]**	[0.73]	[0.77]	**[1.00]**	

performance of the book-to-market ratio of the Standard and Poor's industrial index in the second subsample. The S&P is potentially a better measure of the overall market than the DJIA because it contains 350 stocks compared to 30 stocks for the DJIA. We construct a monthly S&P book-to-market series in an identical manner to the DJIA book-to-market series. In the months of March through December, the book value of the S&P industrials in the previous December is divided by the price level of the S&P. In the months of January and February we use the book value from the December one year prior to the previous December. A comparison of the S&P and DJIA book-to-market ratios for the entire sample is impossible because our S&P book value data start in 1940.

The S&P regression results are reported in Table 6. For every regression, the *p*-values for the slope coefficients on the S&P book-to-market ratio are lower than the *p*-values reported in Panel B of Table 3 for the slope coefficients on the DJIA book-to-market ratio. Although the S&P book-to-market ratio does a better job than the DJIA book-to-market ratio, these findings are still weak. Only one out of four regressions, the annual SMALL return regression,

Table 6
Regression results using the Standard and Poor's book-to-market ratio for the second half of the sample (July 1959–August 1994)

Regressions of CRSP value-weighted market returns (VWRET), equally weighted market returns (EWRET), and the difference of these returns (SMALL) on the Standard and Poor's book-to-market ratio (SPBM). Nelson and Kim (1993) randomized p-values are in brackets. Parameter estimates, Nelson and Kim p-values, and t-statistics are in bold if the p-value is greater than 95% or less than 5%

VWRET			EWRET			SMALL		
Intercept	SPBM	Adj R^2	Intercept	SPBM	Adj R^2	Intercept	SPBM	Adj R^2
One-month returns								
− 0.76	3.18	0.01	− 0.95	4.07	0.01	− 0.19	0.89	0.00
[0.88]	[0.16]		[0.89]	[0.16]		[0.82]	[0.23]	
Annual returns								
− 7.13	34.53	0.13	− 17.94	61.63	0.20	**−10.81**	**27.09**	0.09
[0.81]	[0.20]		[0.93]	[0.09]		**[0.95]**	**[0.04]**	

produces a statistically significant slope coefficient on the S&P book-to-market ratio.

5. The source of the book-to-market ratio's predictive ability

One explanation for the book-to-market ratio's ability to predict returns is that book value proxies for future cash flows. The Appendix illustrates that dividing a cash flow proxy by a current market price produces a variable that is correlated with future returns. This correlation is greater when better expected cash flow proxies are used. Our reasoning builds on Ball (1978) and Berk (1995). Berk argues that market capitalization does not cause a firm's discount rate to be high but that high discount rates cause a firm to have a lower market capitalization.

This explanation is centered on the ability of book value to proxy for cash flow, since dividing a cash flow proxy by market value produces a discount rate proxy. This approach is very general and does not rely on a specific model to generate discount rates. The approach holds whether or not discount rates are related to economy wide risk or are influenced by irrational factors. Thus, our method cannot be used to determine whether the market rationally prices assets.

Our goal in this section is to test this cash flow explanation by comparing the DJIA and S&P results. From the previous section, the S&P book-to-market ratio is better than the DJIA book-to-market ratio in forecasting post-1960 returns. If the ability of the book-to-market ratio to predict returns is related to the ability of book value to predict cash flows, then S&P book values should outperform DJIA book values as cash flow predictors. An independent paper by Hunt (1996) uses a similar approach to explain the book-to-market ratio's cross-sectional predictive ability, although Hunt is unable to relate the book-to-market ratio's ability to proxy for cross-sectional returns to the ability of book value to proxy for cross-sectional cash flows.

In order to examine this explanation, we need to construct a cash flow proxy. We do so from the Compustat active and research files, including back data files for both. We include all publicly traded Compustat firms that have December fiscal year-ends. From this sample, we use as our cash flow proxy annual average earnings weighted by the size of the firm (measured by assets).

The next step is to estimate the relation between book values and future cash flows. Since book values and earnings are highly persistent series, we are concerned about specification problems. Our first step in addressing this issue is to conduct augmented Dickey–Fuller tests of all post-1960 series (DJIA book value, S&P book value, and our earnings measure). This test suggests that all three of these series have a unit root. Following Johansen's (1991) vector autoregression approach, we test for cointegration. The test suggests that the earnings measure and the DJIA book value, as well as the earnings measure and the S&P book value, have one cointegrating vector. From these findings, we pick an estimation specification that relates levels of earnings and levels of book values.

Table 7 reports the time series relation between future annual earnings and book value, and also between three-year earnings and book value. We expect that three-year earnings will provide a better proxy for cash flow than annual earnings, since over longer periods earnings are less influenced by accounting conventions such as depreciation (Dechow, 1994). In the one-year earnings regression, residuals are modeled as a second-order autoregressive process. Since the three-year earnings regression uses annual overlapping data, these regressions model the residuals as a fourth-order autoregressive process.

Table 7 shows that, although both S&P book value and DJIA book value predict future earnings, the S&P is a better predictor than the DJIA. When both book values are included in the estimation, the sign on DJIA book value becomes negative, and we can reject the hypothesis that both book values have the same slope coefficient at the 1% level. Since the S&P book-to-market ratio performs better than the DJIA book-to-market ratio in predicting returns in this subsample, these results are consistent with the explanation that the ability of book value to predict market return is related to the ability of book value to predict cash flow.

Table 7
Regression estimation of the relation between book values and future market earnings (1959–1994)

Regressions of average earnings (from Compustat) on the previous book value of the Standard and Poor's industrials index and the previous book value of the Dow Jones index. One-year earnings regressions model residuals as a second-order autoregressive process, whereas three-year earnings regressions incorporate overlapping data and model residuals as a fourth-order autoregressive process (t-statistics in parentheses).

Intercept	S&P book	Dow book	p-value equal slopes
Annual earnings			
− 69.98		0.47	
(− 0.56)		(3.07)	
− 25.26	3.65		
(− 0.32)	(4.00)		
188.81	10.67	− 1.06	0.01
(1.30)	(2.70)	(− 1.83)	
Three-year earnings			
− 321.20		1.65	
(− 1.13)		(4.67)	
− 134.38	12.52		
(− 0.78)	(6.79)		
245.47	24.30	− 1.82	0.01
(0.80)	(3.00)	(− 1.52)	

6. Conclusion

Previous studies document a strong cross-sectional relation between the book-to-market ratio and future returns. We show that an aggregate measure of the book-to-market ratio forecasts future market returns and the excess returns of small stocks over big stocks. In general, the DJIA book-to-market ratio is a stronger predictor of market returns than are previously examined variables such as interest rate spreads and dividend yields. An aggregate measure of the book-to-market ratio does not appear to forecast the excess returns of high book-to-market stocks over low book-to-market stocks. Our findings are sample specific in that the book-to-market ratio's predictive power only occurs before 1960. After 1960, there is no significant relation.

We find evidence that the book-to-market ratio's predictive ability is related to the ability of book value to forecast future cash flows, because dividing an expected cash flow proxy (book value) by a price level (market value) yields a discount rate proxy. This explanation relies on the time variation of discount rates, and does not depend on whether this variation is attributable to rational or irrational factors. Consistent with this explanation, we find that, after 1960, the S&P book-to-market ratio is a better predictor of market returns than is the DJIA book-to-market ratio, and the S&P book value is better than the DJIA book value in predicting market cash flow. This finding suggests that future research on the cross-sectional relation between book-to-market ratios and returns might benefit from examining the relation between book values and cash flows.

Appendix A.

The relation between the ability of book value to predict cash flows and the ability of the book-to-market ratio to predict returns can be illustrated with the following example. Assume that market prices are determined in the current period, and that shareholders receive a liquidating cash flow next period. Let B be the current book value, where book value proxies for next period's expected cash flow, as follows:

$$B = \mathrm{ECF} + \eta, \tag{A.1}$$

where η is a zero-mean white noise process that causes book value to imperfectly predict future cash flows, and ECF is the expected cash flow that will be distributed next period.

The current price level of the market, M, is determined by a discount rate process, κ, and by the expected cash flow, ECF, as follows:

$$M = \frac{\mathrm{ECF}}{\kappa}. \tag{A.2}$$

In the next period the cash flow, CF, is distributed to shareholders. The realized cash flow is the expected cash flow, ECF, plus ε, where ε is a white noise error term with a mean of zero:

$$\mathrm{CF} = \mathrm{ECF} + \varepsilon. \tag{A.3}$$

Thus, investors cannot perfectly predict next period's cash flow.

We consider the relation between the book-to-market ratio and expected return. The book-to-market measure is computed by dividing book value in Eq. (A.1) by market value in Eq. (A.2), yielding

$$\text{book-to-market} = \kappa + \frac{\eta\kappa}{\mathrm{ECF}}. \tag{A.4}$$

The return that accrues to the security is the security's payoff, CF, divided by the market price level, M, minus one. From Eqs. (A.2) and (A.3),

$$\text{return} = \kappa + \frac{\varepsilon\kappa}{\mathrm{ECF}} - 1. \tag{A.5}$$

All the primary variables (ECF, ε, η, and κ) are assumed to be jointly independent, except we permit dependence between ECF and κ. Dependence between ECF and κ may be attributable to economy wide factors. For example, when expected cash flows are high, agents prefer to save less and borrow more, causing an increase in discount rates (Friedman, 1957). Using Eqs. (A.4) and (A.5), the covariance between the book-to-market ratio and subsequent return is the variance of κ. The correlation between the book-to-market ratio and return is

$$\mathrm{Corr}(B/M, \text{return}) = \frac{\mathrm{Var}(\kappa)}{\sqrt{\left(\mathrm{Var}(\kappa) + \mathrm{Var}(\eta)\mathrm{Var}\left(\frac{\kappa}{\mathrm{ECF}}\right)\right)\left(\mathrm{Var}(\kappa) + \mathrm{Var}(\varepsilon)\mathrm{Var}\left(\frac{\kappa}{\mathrm{ECF}}\right)\right)}}. \tag{A.6}$$

Since we assume no restrictions on the correlation between cash flow and the discount rate,

$$\mathrm{Var}\left(\frac{\kappa}{\mathrm{ECF}}\right) \gtreqless \mathrm{Var}(\kappa)\mathrm{Var}\left(\frac{1}{\mathrm{ECF}}\right).$$

Similarly, we can express the correlation between book value and cash flow as

$$\mathrm{Corr}(\text{book}, \mathrm{CF}) = \frac{\mathrm{Var}(\mathrm{ECF})}{\sqrt{(\mathrm{Var}(\mathrm{ECF}) + \mathrm{Var}(\eta))(\mathrm{Var}(\mathrm{ECF}) + \mathrm{Var}(\varepsilon))}}. \tag{A.7}$$

The variance of η measures the usefulness of book value as a proxy for expected cash flow, and the variance of ε measures the unpredictable component of cash flow. Using Eqs. (A.1) and (A.3), the forecast error between CF and B is $\mathrm{CF} - B = \varepsilon - \eta$, and $\mathrm{Var}(\mathrm{CF} - B) = \mathrm{Var}(\varepsilon) + \mathrm{Var}(\eta)$. From Eq. (A.7), as either the variance of η or the variance of ε increases, the ability of book value to forecast cash flow decreases because book value becomes a noisier proxy for cash flow. For the same reason, from Eq. (A.6) we see that an increase in either the variance of η or the variance of ε will lead to a decrease in the book-to-market ratio's ability to predict return. Thus, the ability of book value to forecast return is positively related to the ability of the book-to-market ratio to forecast cash flow.

References

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

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

Dechow, P.M., 1994. Accounting earnings and cash flows as measures of firm performance: the role of accounting accruals. Journal of Accounting and Economics 18, 3–42.

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

Fama, E.F., 1991. Efficient capital markets: II. Journal of Finance 46, 1575–1617.

Fama, E.F., French, K., 1988. Dividend yields and expected stock returns. Journal of Financial Economics 33, 3–56.

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

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

Fama, E.F., Schwert, G.W., 1977. Asset returns and inflation. Journal of Financial Economics 5, 115–146.

Friedman, M., 1957. A Theory of the Consumption Function. Princeton University Press, Princeton, NJ.

Hunt, A., 1996. An equity cash flow perspective on the book to market equity ratio's ability to predict cross-sectional stock returns. Unpublished doctoral dissertation. University of Washington.

Johansen, S., 1991. Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models. Econometrica 59, 473–480.

Keim, D.B., Stambaugh, R.F., 1986. Predicting returns in bond and stock markets. Journal of Financial Economics 12, 357–390.

Kendall, M.G., 1954. Note on bias in the estimation of autocorrelation. Biometrics 41, 403–404.

Kothari, S.P., Shanken, J., 1997. Book-to-market, dividend yield, and expected market returns: a time-series analysis. Journal of Financial Economics 44, 169–203.

Nelson, C.R., Kim, M.J., 1993. Predictable stock returns: the role of small sample bias. Journal of Finance 48, 641–661.

Sharathchandra, G., Thompson, R., 1994. Book-to-market as a surrogate for priced risk when risk is time varying. Working paper. Southern Methodist University, Dallas, TX.

Stambaugh, R., 1986. Bias in regression with lagged stochastic regressors. Center for Research in Security Prices. Working paper No. 156. University of Chicago.