Characteristics, Covariances, and Average Returns: 1929 to 1997

Author(s): James L. Davis, Eugene F. Fama and Kenneth R. French

Source: *The Journal of Finance*, Feb., 2000, Vol. 55, No. 1 (Feb., 2000), pp. 389-406

Published by: Wiley for the American Finance Association

Stable URL: https://www.jstor.org/stable/222559

THE JOURNAL OF FINANCE • VOL. LV, NO. 1 • FEBRUARY 2000

Characteristics, Covariances, and Average Returns: 1929 to 1997

JAMES L. DAVIS, EUGENE F. FAMA, and KENNETH R. FRENCH*

ABSTRACT

The value premium in U.S. stock returns is robust. The positive relation between average return and book-to-market equity is as strong for 1929 to 1963 as for the subsequent period studied in previous papers. A three-factor risk model explains the value premium better than the hypothesis that the book-to-market characteristic is compensated irrespective of risk loadings.

FIRMS WITH HIGH RATIOS OF BOOK VALUE to the market value of common equity have higher average returns than firms with low book-to-market ratios (Rosenberg, Reid, and Lanstein (1985)). Because the capital asset pricing model (CAPM) of Sharpe (1964) and Lintner (1965) does not explain this pattern in average returns, it is typically called an anomaly.

There are four common explanations for the book-to-market (BE/ME) anomaly. One says that the positive relation between BE/ME and average return (the so-called value premium) is a chance result unlikely to be observed out of sample (Black (1993), MacKinlay (1995)). Out-of-sample evidence is, however, provided by Chan, Hamao, and Lakonishok (1991), Capaul, Rowley, and Sharpe (1993), and Fama and French (1998). They document strong relations between average return and BE/ME in markets outside the United States. Using a rather small sample of firms, Davis (1994) finds that the relation between average return and BE/ME observed in recent U.S. returns extends back to 1941. We extend Davis' data back to 1926, and we expand the coverage to all NYSE industrial firms. We find that the value premium in pre-1963 returns is close to that observed for the subsequent period in earlier work. These results argue against the sample-specific explanation for the value premium.

The second story for the value premium is that it is not an anomaly at all. The higher average returns on high BE/ME stocks are compensation for risk in a multifactor version of Merton's (1973) intertemporal capital asset pricing model (ICAPM) or Ross's (1976) arbitrage pricing theory (APT). Consistent with this view, Fama and French (1993) document covariation in returns

* Kansas State University (Davis), University of Chicago (Fama), and the Massachusetts Institute of Technology (French). Support for data collection from Dimensional Fund Advisors and the comments of Kent Daniel, John Heaton, René Stulz, Sheridan Titman, and two referees are gratefully acknowledged.

related to BE/ME beyond the covariation explained by the market return. Fama and French (1995) show that there is a BE/ME factor in fundamentals (earnings and sales) like the common factor in returns. The acid test of a multifactor model is whether it explains differences in average returns. Fama and French (1993, 1996) propose a three-factor model that uses the market portfolio and mimicking portfolios for factors related to size (market capitalization) and BE/ME to describe returns. They find that the model largely captures the average returns on U.S. portfolios formed on size, BE/ME, and other variables known to cause problems for the CAPM (earnings/price, cashflow/price, past sales growth, and long-term past return). Fama and French (1998) show that an international version of their multifactor model seems to describe average returns on portfolios formed on scaled price variables in 13 major markets.

The third explanation for the value premium says it is due to investor overreaction to firm performance. High BE/ME stocks tend to be firms that are weak on fundamentals like earnings and sales, while low BE/ME stocks tend to have strong fundamentals. Investors overreact to performance and assign irrationally low values to weak firms and irrationally high values to strong firms. When the overreaction is corrected, weak firms have high stock returns and strong firms have low returns. Proponents of this view include DeBondt and Thaler (1987), Lakonishok, Shleifer, and Vishny (1994), and Haugen (1995).

The final story for the value premium, suggested by Daniel and Titman (1997), is that it traces to the value characteristic, not risk. For example, a behavioral story that does not require overreaction is that investors like growth stocks (strong firms) and dislike value stocks (weak firms). The result is a value premium (low prices and high expected returns for value stocks relative to growth stocks) that is not due to risk. The behavioral overreaction story can also be viewed as a variant of the characteristics model. In general, the characteristics model covers anything that produces a premium for the value characteristic relative to the growth characteristic and is not the result of risk.

Daniel and Titman (1997) argue that past research cannot distinguish the risk model from the characteristics model. The problem is that the value and growth characteristics are associated with covariation in returns. For example, industries move through periods of distress and growth. When portfolios are formed to capture a risk factor related to relative distress, they pick up return covariation within industries that is always present but for the moment happens to be associated with growth or distress. In this view, the value premium seems to be related to the covariance of returns with a common distress factor, when in fact it is due to the growth and distress characteristics. As a result, one cannot distinguish the risk story from the characteristics story in the typical tests that focus on common factors.

Daniel and Titman (1997) suggest a clever way to break this logjam. If characteristics (growth and distress) drive expected returns, there should be firms with characteristics that do not match their risk loadings. For example, there should be some strong firms in distressed industries. In the characteristics model, these firms have low returns because they are strong. But they can have high loadings on a distress risk factor if the factor is in part due to covariation of returns within industries. Thus, the returns on these firms will be too low, given their risk loadings. Conversely, there are distressed firms in strong industries. Because they are distressed, they have high returns, but in terms of risk loadings they look like strong firms. If characteristics drive prices, their returns will be too high given their risk loadings.

In short, the characteristics hypothesis says that relative distress drives stock returns, and BE/ME is a proxy for relative distress. Low BE/ME (characteristic of strong firms) produces low stock returns, irrespective of risk loadings. Similarly, high BE/ME stocks (distressed firms) have high returns, regardless of risk loadings. In contrast, the risk story says expected returns compensate risk loadings, irrespective of the BE/ME characteristic. It is clear, then, that the empirical key to distinguishing the risk model from the characteristics model is to find variation in risk loadings unrelated to BE/ME.

To identify independent variation in characteristics and risk loadings, Daniel and Titman (1997) form portfolios by triple-sorting stocks on size, BE/ME, and risk loadings. We use a similar approach, but for a much longer time period. Daniel and Titman study returns from July 1973 to December 1993, 20.5 years. Our tests cover the July 1929 to June 1997 period, 68 years. The extended sample period enhances the power of the tests of the characteristics model against the risk model.

Our results are easily summarized. The evidence of Daniel and Titman (1997) in favor of the characteristics model is special to their rather short sample period. In the more powerful tests for our 68-year period, the risk model provides a better story for the relation between BE/ME and average return.

We begin (Section I) with summary statistics for the market, size, and value premiums of the risk model. Sections II to IV present regression tests of the risk model, including the central tests of the risk model against the characteristics model (Section III). Section V concludes.

I. Summary Statistics for the Premiums

Using *Moody's Industrial Manuals*, we collect book common equity (BE) from 1925 to 1996 for all NYSE industrial firms that do not have BE data on COMPUSTAT. To keep the task manageable, we do not collect BE for financial firms, transportation firms, and utilities. To expand the sample of firms, beginning in 1954 we merge the hand-collected data for NYSE industrials with the COMPUSTAT data for NYSE, AMEX, and Nasdaq industrials and nonindustrials.

The number of firms in our sample grows steadily through time. On the first portfolio formation date, June 1929, we have BE data for 339 NYSE firms. By June 1953, we have 756 NYSE firms. With the addition of COM-

PUSTAT data in June 1954 the sample increases to 834 NYSE firms. On the last portfolio formation date, June 1996, the sample contains 4,562 NYSE, AMEX, and Nasdaq firms.

Like Daniel and Titman (1997), our alternative to the characteristics model is that expected returns conform to the three-factor asset pricing model in Fama and French (1993),

$$\mathrm{E}(R_i) - R_f = b_i[\mathrm{E}(R_M) - R_f] + s_i\mathrm{E}(\mathrm{SMB}) + h_i\mathrm{E}(\mathrm{HML}). \tag{1}$$

R_i is the return on asset i, R_f is the riskfree interest rate, and R_M is the return on the value-weight market portfolio. SMB is the difference between the returns on a portfolio of small stocks and a portfolio of big stocks, constructed to be neutral with respect to BE/ME. Specifically, in June of each year we use independent sorts to allocate the NYSE, AMEX, and Nasdaq stocks in our sample to two size groups and three BE/ME groups. Big stocks (B) are above the median market equity of NYSE firms and small stocks (S) are below. Similarly, low BE/ME stocks (L) are below the 30th percentile of BE/ME for NYSE firms, medium BE/ME stocks (M) are in the middle 40 percent, and high BE/ME stocks (H) are in the top 30 percent. We form six value-weight portfolios, S/L, S/M, S/H, B/L, B/M, and B/H, as the intersections of the size and BE/ME groups. For example, S/L is the value-weight return on the portfolio of stocks that are below the NYSE median in size and in the bottom 30 percent of BE/ME. SMB is the difference between the equal-weight averages of the returns on the three small stock portfolios and the three big stock portfolios:

$$\mathrm{SMB} = (\mathrm{S/L} + \mathrm{S/M} + \mathrm{S/H})/3 - (\mathrm{B/L} + \mathrm{B/M} + \mathrm{B/H})/3. \tag{2}$$

Similarly, HML is the difference between the return on a portfolio of high BE/ME stocks and the return on a portfolio of low BE/ME stocks, constructed to be neutral with respect to size:

$$\mathrm{HML} = (\mathrm{S/H} + \mathrm{B/H})/2 - (\mathrm{S/L} + \mathrm{B/L})/2. \tag{3}$$

The correlation between SMB and HML for the July 1929 to June 1997 period is only 0.13. Thus, SMB indeed seems to provide a measure of the size premium that is relatively free of BE/ME effects, and HML is a measure of the BE/ME premium relatively free of size effects.

Table I shows summary statistics for $R_M - R_f$, SMB, and HML for July 1929 to June 1997 and for two subperiods that break in July 1963, the start date in Fama and French (1992, 1993, 1995, 1996). We split the sample at this date to test whether the later period is unusual. The two subperiods are also equal in length, 34 years. Although we have $R_M - R_f$, SMB, and HML back to July 1926, the results in Table I start in July 1929. This is for consistency with the regressions in later tables that use the first three years to form portfolios on preformation estimates of the HML slopes in equation (1).

Table I

Summary Statistics for Monthly Percent Three-Factor Explanatory Returns

R_f is the one-month Treasury bill rate (from Ibbotson Associates). R_M is the value-weight return on all NYSE, AMEX, and Nasdaq stocks with book equity data for the previous calendar year. At the end of June of each year t (1926 to 1996), stocks are allocated to two groups (S or B) based on whether their June market capitalization, ME (stock price times shares outstanding), is below or above the median for NYSE stocks on CRSP. Stocks are allocated in an independent sort to three book-to-market equity (BE/ME) groups (L, M, or H) based on breakpoints for the bottom 30 percent, middle 40 percent, and top 30 percent of the values of BE/ME for the NYSE stocks in our sample. BE is the stockholders' book equity, plus balance sheet deferred taxes and investment tax credit (if available), minus the book value of preferred stock. Depending on availability, we use redemption, liquidation, or par value (in that order) for the book value of preferred stock. Stockholders' equity is the value reported by Moody's or COMPUSTAT, if it is available. If not, we measure stockholders' equity as the book value of common equity plus the par value of preferred stock, or the book value of assets minus total liabilities (in that order). The BE/ME used to form portfolios in June of year t is book common equity for the fiscal year ending in calendar year $t - 1$, divided by market equity at the end of December of $t - 1$. Six portfolios (S/L, S/M, S/H, B/L, B/M, and B/H) are formed as the intersections of the two size and the three BE/ME groups. Value-weight monthly returns on the portfolios are calculated from July of year t to June of $t + 1$. SMB is (S/L + S/M + S/H)/3 − (B/L + B/M +B /H)/3. HML is (S/H + B/H)/2 − (S/L + B/L)/2. The sample includes all NYSE industrials that have BE data either in *Moody's Industrial Manuals* or on COMPUSTAT for fiscal years ending in the 1925 to 1996 period. After June 1954, the sample for year t also includes NYSE, AMEX, and Nasdaq firms with BE data on COMPUSTAT for the fiscal year ending in the preceding calendar year. To be included in the portfolios formed in June of year t (here and in all following tables), firms must also have COMPUSTAT or CRSP data on ME for December of year $t - 1$ and June of year t. We do not use negative BE firms when calculating the breakpoints for BE/ME or when forming the size-BE/ME portfolios. Only firms with ordinary common equity (as classified by CRSP) are included in the tests.

	$R_M - R_f$	SMB	HML	S/L	S/M	S/H	B/L	B/M	B/H
7/29–6/97: 816 months									
Ave	0.67	0.20	0.46	1.05	1.30	1.53	0.89	1.04	1.34
Std	5.75	3.26	3.11	7.89	7.49	8.38	5.65	6.19	7.41
t(Ave)	3.34	1.78	4.24	3.80	4.96	5.21	4.52	4.78	5.16
7/29–6/63: 408 months									
Ave	0.82	0.19	0.50	1.09	1.22	1.49	0.81	1.01	1.40
Std	6.89	3.65	3.59	9.01	9.13	10.57	6.50	7.73	9.52
t(Ave)	2.41	1.07	2.80	2.44	2.71	2.85	2.52	2.64	2.98
7/63–6/97: 408 months									
Ave	0.52	0.21	0.43	1.01	1.38	1.57	0.98	1.06	1.27
Std	4.32	2.83	2.54	6.60	5.38	5.37	4.65	4.12	4.38
t(Ave)	2.44	1.53	3.38	3.10	5.17	5.88	4.24	5.20	5.87
7/73–12/93: 246 months									
Ave	0.51	0.33	0.50	1.23	1.60	1.76	0.96	1.20	1.44
Std	4.79	2.75	2.74	6.88	5.64	5.68	5.22	4.53	4.67
t(Ave)	1.68	1.88	2.87	2.81	4.46	4.87	2.90	4.17	4.83

The average value of the market premium, $R_M - R_f$, for the full 68-year sample period is 0.67 percent per month (t-statistic = 3.34). The market premium for the first half of the sample period is 0.82 percent per month, versus 0.52 percent for the second half. Both are about 2.4 standard errors from zero. Thus, not surprisingly, there is a strong market premium in returns.

There is also a reliable value premium in returns. The average HML return is large for the full July 1929 to June 1997 (0.46 percent per month, t-statistic = 4.24) and for the two 34-year subperiods (0.50 percent per month, t-statistic = 2.80 for July 1929 to June 1963, and 0.43 percent per month, t-statistic = 3.38 for July 1963 to June 1997). Previous studies of the relation between BE/ME and average return typically use sample periods that start after June 1963. (Davis (1994), who uses a limited sample that extends back to 1941, is an exception.) The returns for July 1929 to June 1963 confirm that the premium of value (high BE/ME) stock returns over growth (low BE/ME) stock returns observed in earlier work is not special to the post-1963 period.

Loughran (1997) argues that there is not much of a value premium in the average returns on large stocks for 1963 to 1995. Though nontrivial in magnitude, Table I confirms that the value premium for large stocks (the average B/H − B/L return) for July 1963 to June 1997, 0.29 percent per month, is lower than the value premium for small stocks (the average S/H − S/L return), 0.56 percent per month. For the earlier July 1929 to June 1963 period, however, large stocks produce a bigger value premium than small stocks, 0.59 versus 0.40 percent per month. In the returns for the overall July 1929 to June 1997 period, the value premium for large stocks, 0.45 percent, is quite similar to that for small stocks, 0.48 percent.

In contrast to the market premium and the value premium, the size effect in Table I is puny. The average SMB return for July 1929 to June 1997 is 0.20 percent per month (t-statistic = 1.78). The weak size effect is in part due to the fact that the six components of SMB are value-weight portfolios. More important is the fact that SMB is neutral with respect to BE/ME. By way of contrast, without the control for BE/ME, the average difference between the returns for July 1929 to June 1997 on value-weight portfolios of stocks below and above the NYSE median is 0.33 percent per month (t-statistic = 2.44). This simple size premium is 65 percent larger than the average SMB return because small stocks tend to have higher BE/ME than big stocks, and a size premium that is not neutral with respect to BE/ME in part reflects the value premium in returns.

In line with their hypothesis about the value premium in average returns, Daniel and Titman (1997) postulate that the size premium traces to the size characteristic rather than the SMB risk loading in equation (1). The small average SMB return is, however, fatal for tests of this hypothesis. The characteristics model predicts that SMB loadings s_i in equation (1) that are unrelated to the size characteristic do not affect average return. But a test of this hypothesis has little power when the average SMB return is low since the risk model also predicts that differences in SMB loadings produce small differences in average returns. Like the tests in Daniel and Titman (1997), our tests (in an earlier version of this paper) confirm that the size effect is not large enough to produce a conclusive contest between the risk model and the characteristics model. In contrast, the large value premium (average HML return) allows more powerful tests of the predictions of the two models about the relation between average return and BE/ME.

II. Sorts on Size and BE/ME

Our main interest is testing the characteristics model against the risk model given in equation (1). But the exercise is pointless if equation (1) is not a reasonably good approximation for expected returns. Moreover, we are (not surprisingly) interested in examining how well the three-factor model works outside the post-1963 period for which it was developed. To address these issues, we begin by testing equation (1) on portfolios formed on size and BE/ME, an approach also used in earlier work (e.g., Fama and French (1993, 1996)). Examining size-BE/ME portfolios also allows us to judge whether our pre-1963 samples (which until 1954 are restricted to NYSE industrials) have systematically different size and BE/ME characteristics than the samples for later years (which cover all COMPUSTAT firms).

A. *Three-Factor Regressions*

Table II summarizes estimates of the three-factor regression,

$$R_i - R_f = a_i + b_i(R_M - R_f) + s_i \text{SMB} + h_i \text{HML} + \epsilon_i, \tag{4}$$

for the post-formation returns on nine portfolios formed in June each year as the intersections of independent sorts of stocks into three size groups and three BE/ME groups. The table shows that sorting on size and BE/ME produces strong orderings on the corresponding risk loadings. The post-formation SMB loadings in Table II decrease with increasing size, and the spread in SMB slopes from small to big stock portfolios is about 1.3. Post-formation HML loadings increase with BE/ME, and the spread between the HML slopes for high and low BE/ME portfolios is approximately 0.8. The strong correlation between the BE/ME characteristic and the HML risk loading limits our ability to produce variation in the risk loading independent of BE/ME. And this lowers the power of the tests to distinguish the characteristics model from the risk model. Still, we show later that the tests have power.

Tests of the three-factor model (1) center on the intercepts in regression (4), which, if the model holds, should be indistinguishable from zero. The F-test of Gibbons, Ross, and Shanken (1989) rejects the zero-intercepts hypothesis for the full 68-year sample period (p-value = 0.005) and for both 34-year subperiods (p-values of 0.060 and 0.035). The three-factor model's difficulties in the second half of the sample are similar to those observed in

Table II

Three-Factor Regressions for Portfolios Formed from Independent Sorts on Size and BE/ME

$$R_i - R_f = a_i + b_i(R_M - R_f) + s_i\mathrm{SMB} + h_i\mathrm{HML} + \epsilon_i$$

We form the portfolios here and in the following tables at the end of June of each year t (1929 to 1996) using all NYSE, AMEX, and Nasdaq stocks with nonnegative BE for year $t-1$, and at least 36 months of returns data in the five years ending in December of $t-1$. Here we allocate the stocks to three size groups (small, medium, or big; S, M, or B) each year based on their June market capitalization, ME. We allocate stocks in an independent sort to three book-to-market equity (BE/ME) groups (low, medium, or high; L, M, or H) based on BE/ME for December of year $t-1$. The breakpoints are the 33rd and 67th ME and BE/ME percentiles for the NYSE firms in the sample. We form nine portfolios (S/L, S/M, S/H, M/L, M/M, M/H, B/L, B/M, and B/H) as the intersections of the three size and the three BE/ME groups. The returns explained by the regressions, R_i, are the value-weight returns on the portfolios from July of year t to June of $t+1$. The t-statistics, $t(\,)$, for the regression coefficients (here and in all following tables) use the heteroskedasticity consistent standard errors of White (1980). Here and in all following tables: (i) BE/ME is the aggregate of BE for the firms in a portfolio divided by the aggregate of ME; (ii) Size is the value-weight average of the NYSE size percentiles for the firms in a portfolio; (iii) BE/ME and Size are averages of the annual values for the time periods shown; (iv) Ex Ret is the average monthly post-formation return in excess of R_f; (v) the regressions R^2 are adjusted for degrees of freedom.

	BE/ME	Size	Ex Ret	a	b	s	h	$t(a)$	$t(b)$	$t(s)$	$t(h)$	R^2
7/29–6/97												
S/L	0.55	22.39	0.61	−0.42	1.06	1.39	0.09	−4.34	30.78	19.23	1.73	0.91
S/M	1.11	22.15	1.05	−0.01	0.97	1.16	0.37	−0.18	53.55	19.49	9.96	0.96
S/H	2.83	19.05	1.24	−0.03	1.03	1.12	0.77	−0.73	67.32	39.21	26.97	0.98
M/L	0.53	55.85	0.70	−0.06	1.04	0.59	−0.12	−1.29	55.83	18.01	−4.30	0.96
M/M	1.07	55.06	0.95	−0.01	1.05	0.47	0.34	−0.15	32.98	17.50	9.50	0.96
M/H	2.18	53.21	1.13	−0.04	1.08	0.53	0.73	−0.90	47.85	8.99	11.12	0.97
B/L	0.43	94.65	0.58	0.02	1.02	−0.10	−0.23	0.88	148.09	−6.88	−13.52	0.98
B/M	1.04	92.06	0.72	−0.09	1.01	−0.14	0.34	−1.76	61.61	−4.96	13.66	0.95
B/H	1.87	89.53	1.00	−0.09	1.06	−0.07	0.84	−1.40	52.12	−0.86	21.02	0.93
7/29–6/63												
S/L	0.68	23.83	0.69	−0.53	1.01	1.47	0.23	−3.04	18.66	15.72	2.82	0.90
S/M	1.35	23.63	1.21	−0.01	0.96	1.24	0.38	−0.07	34.72	15.60	6.21	0.95
S/H	3.96	20.23	1.44	−0.03	1.02	1.17	0.83	−0.40	44.71	28.80	17.76	0.98
M/L	0.64	55.20	0.84	−0.08	0.98	0.56	0.01	−1.14	37.44	12.26	0.39	0.96
M/M	1.28	54.20	1.13	0.00	1.07	0.47	0.33	0.07	26.38	11.77	7.73	0.97
M/H	2.83	51.59	1.30	−0.07	1.07	0.50	0.79	−0.92	52.49	5.44	7.74	0.97
B/L	0.48	94.92	0.72	−0.01	1.02	−0.08	−0.20	−0.20	131.66	−4.89	−8.09	0.99
B/M	1.21	91.97	0.89	−0.09	1.00	−0.12	0.37	−1.20	43.96	−2.90	10.08	0.96
B/H	2.33	88.91	1.30	0.00	1.02	−0.12	0.97	−0.01	34.28	−0.96	17.99	0.94
7/63–6/97												
S/L	0.42	20.94	0.54	−0.22	1.06	1.22	−0.14	−3.31	60.47	39.87	−4.51	0.96
S/M	0.87	20.68	0.89	0.03	0.97	1.02	0.31	0.71	74.53	52.41	13.82	0.98
S/H	1.71	17.88	1.04	0.04	0.99	1.03	0.62	1.27	75.12	64.49	25.86	0.98
M/L	0.42	56.51	0.56	−0.02	1.07	0.58	−0.24	−0.33	71.73	27.08	−9.73	0.96
M/M	0.87	55.93	0.77	0.02	1.00	0.48	0.30	0.31	64.36	22.60	11.22	0.95
M/H	1.54	54.83	0.96	0.03	1.05	0.55	0.63	0.53	69.16	28.08	24.23	0.96
B/L	0.38	94.38	0.45	0.10	0.99	−0.15	−0.32	2.89	91.73	−8.92	−16.53	0.98
B/M	0.86	92.14	0.54	−0.04	0.99	−0.19	0.25	−0.70	55.19	−6.91	8.53	0.91
B/H	1.41	90.16	0.70	−0.13	1.04	−0.01	0.69	−2.59	76.64	−0.36	28.53	0.94

Fama and French (1993, 1996). Three of the nine intercepts in the July 1963 to June 1997 regressions are more than two standard errors from zero. The S/L portfolio (small growth stocks) and the B/H portfolio (big value stocks) have average returns that are too low given their risk loadings, while the return on the B/L portfolio (big growth stocks) is too high. To some extent, the t-statistics for these intercepts are large not because the differences between average returns and the predictions of the three-factor model are large, but rather because the regressions absorb so much return variance. All the regression R^2 for July 1963 to June 1997 are at least 0.91, and two of the three intercepts that differ from zero on a statistical basis are only 0.10 and -0.13 percent per month.

Interestingly, only one portfolio produces an intercept much different from zero in the regressions for the earlier July 1929 to June 1963 period. But the aberrant portfolio is S/L, which is also the biggest problem for the model in the later period. In fact, the S/L intercept for the first period, -0.53 percent per month, is more extreme than the intercept for the second, -0.22 percent. Thus, the pricing of small growth stocks presents problems for the three-factor model throughout the 68-year sample period.

Like the tests in Fama and French (1993, 1996), Table II unmasks the three-factor model for what it is, a model, and so necessarily false. But the model does provide a reasonable approximation for the returns on portfolios formed on the size and BE/ME characteristics. It is thus a viable risk story against which the predictions of the characteristics model can be tested. Moreover, since all models are false, the three-factor model should only be discarded in favor of a better model. And we argue in Section III that the three-factor risk model provides a more accurate description of average returns than the characteristics model.

B. Is the Extended Sample Unusual?

The tests of the characteristics model against the risk model hinge on the enhanced power obtained from our long sample period. In addition to the number of years and the number of firms per year, the power of the tests depends on (i) the spreads in the size and (especially) the BE/ME characteristics, and (ii) the spreads in the SMB and (especially) the HML risk loadings. It is thus interesting to examine whether the samples for July 1929 to June 1963 (which until 1954 are restricted to NYSE industrials) and the samples for July 1963 to June 1997 (which include all COMPUSTAT firms and any NYSE industrials not on COMPUSTAT) have systematically different characteristics and risk loadings.

Table II shows the value-weight averages of the NYSE size percentiles of the firms in each of the nine size-BE/ME portfolios. (The size percentile breakpoints are for all NYSE firms, not just those in our annual samples.) The firms in the small-stock portfolios (S/L, S/M, and S/H) of the July 1929 to June 1963 period are a tad larger (about three percentiles) than the firms in the corresponding portfolios of the July 1963 to June 1997 period. The average size percentiles of the stocks in the big-stock portfolios (B/L, B/M, and B/H) are quite similar in the two periods. Thus, limiting the samples of the earlier years to NYSE industrials does not produce an important size tilt relative to the more complete samples of later years.

The differences between the SMB slopes of the small-stock and big-stock portfolios are also similar in the two halves of the 68-year sample period, as are the differences between the HML slopes of the high and low book-to-market portfolios. The portfolios of the July 1929 to June 1963 period have higher BE/ME, which is not surprising given that this period includes the great depression. More important, the BE/ME spreads between the high and low book-to-market portfolios are larger during the earlier period.

In short, the spreads in the critical characteristics and risk loadings for July 1929 to June 1963 are similar to those of the July 1963 to June 1997 period. The samples of the earlier period cover fewer firms, but the samples of all years are large. And Table I says that like the July 1963 to June 1997 period, the July 1929 to June 1963 period produces a strong value premium (average HML return). Thus, we have every reason to expect that our extended sample will produce more powerful tests of the characteristics model against the risk model. We see next that this is indeed the case.

III. BE/ME versus HML Risk Loading

The three-factor risk model, equation (1), says that the intercept, a_i, in regression (4) is zero for all assets. The characteristics model says that nonzero a_i are to be expected when stocks have HML risk loadings that do not line up with their BE/ME. Thus, the trick to distinguishing between the two models is to isolate variation in the HML risk loading that is independent of BE/ME.

To this end, Daniel and Titman (1997) form nine portfolios as the intersections of independent sorts of firms into three size and three BE/ME groups. Each of these nine portfolios is then subdivided equally into five value-weight portfolios based on pre-formation values of the HML risk loading h_i in regression (4). The result is 45 portfolios constructed to produce variation in the HML risk loading independent of the size and BE/ME characteristics of the portfolios. Their tests of the characteristics model against the risk model then focus on the return on an arbitrage portfolio, Hh-Lh. This portfolio is the difference between the sum of the returns on the two high h_i portfolios of a size-BE/ME group minus the sum of the returns on the two low h_i portfolios of the same size-BE/ME group, averaged across the nine size-BE/ME groups.

Daniel and Titman (1997) report that prior to 1973, some of their 45 portfolios have only one stock. Because of this problem, we use a slightly different approach. Like them, we place stocks into nine groups based on independent size and BE/ME sorts. Unlike them, we subdivide each of the nine groups into three portfolios based on pre-formation HML slopes. The advantage of fewer third-pass sorts on h_i is that the resulting 27 portfolios always contain some stocks, so the tests need not be limited to the post-1973 period. Forming three (rather than five) h_i portfolios for each size-BE/ME group should be innocuous since Daniel and Titman (1997) calculate the critical Hh-Lh return as the difference between the sum of the returns on the two high h_i portfolios and the sum of the returns on the two low h_i portfolios in each size-BE/ME group, averaged over the nine groups. Our version of Hh-Lh simply takes the difference between the returns on the high h_i and the low h_i portfolio of each size-BE/ME group, and then averages the differences over the nine groups.

Even during the early years of the July 1929 to June 1997 period, almost all of our 27 portfolios have more than 10 stocks. After 1934, all portfolios have at least four stocks, and after 1965, all have at least 10 stocks. In 1930 and 1931, a few portfolios have only one stock. The Hh-Lh portfolio that is the focus of our inferences is, however, diversified across the nine size-BE/ME groups. Moreover, the standard errors of our regression coefficients are adjusted for heteroskedasticity, so they take account of the loss of power that results when portfolios are not diversified.

Finally, Daniel and Titman (1997) use special SMB and HML factors to estimate pre-formation risk loadings. When portfolios are formed on pre-formation risk loadings in June each year, the weights of securities in the pre-formation factors are fixed at their June values; security weights do not evolve with market values. The advantage of this approach is that it is likely to produce a wider spread in post-formation risk loadings (and thus more precise asset pricing tests) if the covariance matrix of security returns is relatively constant. To be consistent with Daniel and Titman (1997), we use fixed-weight factors to estimate pre-formation risk loadings, and standard Fama–French (1993) variable-weight versions of SMB and HML to estimate the three-factor model on post-formation returns. We can report, however, that using the variable-weight factors to estimate pre-formation risk loadings has little effect on the results.

Table III summarizes estimates of the three-factor regression (4) for our 27 triple-sorted portfolios for the full 68-year sample period. The table also shows the average size and BE/ME characteristics of the portfolios. For each of the nine size-BE/ME groups, the third-pass sort on HML risk loading, h_i, produces a large spread in post-formation HML risk loading but little variation in the size characteristic. The third-pass sorts on h_i do produce weak sorts on BE/ME. But the variation in BE/ME within the size-BE/ME groups is trivial relative to the variation in BE/ME across BE/ME groups. Like Daniel and Titman (1997), we conclude that the three-pass sorts succeed in producing substantial variation in post-formation HML risk loadings that is independent of the size and BE/ME characteristics.

Among the more striking results in favor of the characteristics model in Daniel and Titman (1997) is the evidence in their Table IV that, within the nine size-BE/ME groups, the third-pass sort on HML risk loading produces little variation in average return. Our results (not shown) for their 20.5-year (July 1973 to December 1993) period are similar to theirs. But the evidence

Table III

Regressions for Portfolios Formed from Sorts on Size, BE/ME, and HML slopes: July 1929 to June 1997

$$R_i - R_f = a_i + b_i(R_M - R_f) + s_i\text{SMB} + h_i\text{HML} + \epsilon_i$$

At the end of June of each year t (1929 to 1996), we allocate the NYSE, AMEX, and Nasdaq stocks in our sample to three size groups (small, medium, or big; S, M, or B) based on their June market capitalization, ME. We allocate stocks in an independent sort to three book-to-market equity (BE/ME) groups (low, medium, or high; L, M, or H) based on BE/ME for December of the preceding year. The breakpoints are the 33rd and 67th ME and BE/ME percentiles for the NYSE firms in the sample. We form nine portfolios (S/L, S/M, S/H, M/L, M/M, M/H, B/L, B/M, and B/H) as the intersections of the three size and the three BE/ME groups. The nine portfolios are each subdivided into three portfolios (Lh, Mh, or Hh) using pre-formation HML slopes. These slopes are estimated with five years (three years minimum) of monthly returns ending in December of year $t - 1$. The regressions explain R_i, the value-weight returns on the portfolios from July of year t to June of $t + 1$.

	BE/ME	Size	Ex Ret	a	b	s	h	$t(a)$	$t(b)$	$t(s)$	$t(h)$	R^2
Low BE/ME												
S/L/Lh	0.51	22.29	0.49	−0.56	1.21	1.25	−0.02	−2.88	16.53	6.93	−0.22	0.70
S/L/Mh	0.57	22.79	0.69	−0.34	1.07	1.21	0.15	−2.15	16.26	12.47	1.27	0.77
S/L/Hh	0.56	21.05	0.76	−0.38	1.03	1.64	0.26	−1.91	10.96	7.21	2.03	0.73
M/L/Lh	0.49	55.14	0.62	−0.04	1.09	0.60	−0.42	−0.53	49.96	14.44	−10.88	0.91
M/L/Mh	0.54	56.08	0.66	−0.07	0.97	0.59	−0.10	−1.02	50.53	14.55	−3.04	0.91
M/L/Hh	0.56	56.12	0.81	−0.08	1.06	0.57	0.15	−1.10	31.80	8.98	3.05	0.90
B/L/Lh	0.36	95.28	0.56	0.01	1.11	−0.10	−0.38	0.12	60.37	−2.66	−9.52	0.93
B/L/Mh	0.44	94.64	0.61	0.11	0.95	−0.11	−0.24	2.25	69.71	−6.36	−11.55	0.94
B/L/Hh	0.53	92.21	0.59	−0.05	0.98	−0.07	−0.01	−0.89	49.70	−1.94	−0.41	0.92
Medium BE/ME												
S/M/Lh	1.09	22.01	1.03	0.06	1.01	1.11	0.14	0.56	26.47	18.39	2.90	0.88
S/M/Mh	1.12	22.57	1.06	0.05	0.86	1.22	0.39	0.54	23.54	8.99	5.68	0.87
S/M/Hh	1.13	21.47	1.10	−0.13	1.04	1.20	0.62	−1.43	41.36	24.02	11.07	0.91
M/M/Lh	1.04	55.04	0.89	0.02	1.05	0.51	0.13	0.29	50.84	19.02	4.26	0.93
M/M/Mh	1.07	55.33	0.97	0.06	1.01	0.39	0.34	0.95	20.68	5.63	5.98	0.92
M/M/Hh	1.11	54.66	0.99	−0.10	1.09	0.51	0.55	−1.44	28.30	12.08	11.84	0.93
B/M/Lh	0.99	90.04	0.69	−0.08	1.02	−0.01	0.19	−0.96	29.75	−0.10	3.48	0.86
B/M/Mh	1.04	92.14	0.60	−0.17	1.01	−0.24	0.30	−2.16	45.14	−8.05	9.28	0.89
B/M/Hh	1.08	90.34	0.88	−0.07	1.04	−0.05	0.56	−0.81	35.50	−0.88	8.90	0.88
High BE/ME												
S/H/Lh	2.41	18.08	1.26	0.07	0.97	1.19	0.64	0.97	30.53	14.84	8.70	0.94
S/H/Mh	2.71	19.67	1.19	−0.04	1.04	1.04	0.70	−0.65	45.61	14.65	18.45	0.94
S/H/Hh	3.47	18.99	1.30	−0.13	1.09	1.18	0.99	−1.59	43.08	31.05	21.11	0.95
M/H/Lh	2.04	52.62	1.08	0.02	1.04	0.56	0.53	0.32	29.13	8.10	5.94	0.91
M/H/Mh	2.11	53.43	1.09	−0.06	1.02	0.64	0.72	−0.83	38.12	8.98	18.55	0.93
M/H/Hh	2.39	53.24	1.24	−0.07	1.19	0.40	0.93	−0.81	33.97	4.27	10.17	0.91
B/H/Lh	1.76	86.90	1.02	−0.06	1.15	−0.03	0.69	−0.47	16.50	−0.25	6.00	0.77
B/H/Mh	1.84	88.72	1.01	−0.04	0.99	0.00	0.85	−0.38	22.05	0.00	11.51	0.82
B/H/Hh	1.99	88.75	1.04	−0.07	1.04	−0.04	0.91	−0.63	21.12	−0.24	13.07	0.81

Table IV

Regressions for Hh-Lh Portfolios Formed from Sorts on Size, BE/ME, and HML Slopes

$$\text{Hh-Lh} = a + b(R_M - R_f) + s\text{SMB} + h\text{HML} + \epsilon$$

At the end of June of each year t (1929 to 1996), we allocate the NYSE, AMEX, and Nasdaq stocks in our sample to three size groups (small, medium, or big; S, M, or B) based on their June market capitalization, ME. We allocate stocks in an independent sort to three book-to-market equity (BE/ME) groups (low, medium, or high; L, M, or H) based on BE/ME for December of the preceding year. The breakpoints are the 33rd and 67th ME and BE/ME percentiles for the NYSE firms in the sample. We form nine portfolios (S/L, S/M, S/H, M/L, M/M, M/H, B/L, B/M, and B/H) as the intersections of the three size and the three BE/ME groups. The nine portfolios are each subdivided into three portfolios (Lh, Mh, or Hh) using pre-formation HML slopes. The slopes are estimated with five years (three years minimum) of monthly returns ending in December of year $t - 1$. Value-weight returns on the portfolios are calculated for July of year t to June of $t + 1$. Hh-Lh is ((S/L/Hh-S/L/Lh) + (M/L/Hh-M/L/Lh) + (B/L/Hh-B/L/Lh) + (S/M/Hh-S/M/Lh) + (M/M/Hh-M/M/Lh) + (B/M/Hh-B/M/Lh) + (S/H/Hh-S/H/Lh) + (M/H/Hh-M/H/Lh) + (B/H/Hh-B/H/Lh))/9. Ave is the average Hh-Lh return, and t(Ave) is its t-statistic. The 7/29–6/72 & 1/94–6/97 period includes all the months in the full 7/29–6/97 sample except those in the Daniel and Titman (1997) period, 7/73–12/93.

Period	Ave	t(Ave)	a	b	s	h	$t(a)$	$t(b)$	$t(s)$	$t(h)$	R^2
7/29–6/97	0.12	1.56	−0.06	−0.01	0.03	0.38	−0.83	−0.48	0.91	11.92	0.29
7/29–6/63	0.19	1.49	0.01	−0.01	0.06	0.35	0.11	−0.19	1.09	6.99	0.24
7/63–6/97	0.05	0.56	−0.14	0.01	−0.01	0.43	−2.07	0.48	−0.28	14.32	0.42
7/73–12/93	0.03	0.25	−0.22	0.02	0.03	0.46	−2.28	0.61	0.80	11.35	0.44
7/29–6/72 & 1/94–6/97	0.16	1.63	−0.00	−0.01	0.03	0.36	−0.01	−0.31	0.74	8.80	0.26

for our 68-year period (July 1929 to June 1997) in Table III is quite different, and in the direction predicted by the three-factor risk model. In particular, in each of the nine size-BE/ME groups, the average post-formation return on the high h_i portfolio is higher than the average return on the low h_i portfolio.

Like Daniel and Titman (1997), our formal inferences about the risk model are based on the intercepts in estimates of regression (4) for the Hh-Lh arbitrage portfolio (which, again, is the average of the differences between the returns on the high h_i and the low h_i portfolios of the nine size-BE/ME groups.) Table IV shows estimates of the Hh-Lh regression for several periods. The three-factor risk model (1) says that the regression intercepts should be indistinguishable from zero. The alternative hypothesis of the characteristics model is that since the positive HML loadings for Hh-Lh are largely unrelated to the BE/ME characteristic, they do not affect expected return. Thus, the characteristics model predicts that the intercepts in the Hh-Lh regressions are negative, to offset the positive return premiums implied by the product of the positive HML loadings and the positive expected HML return.

For the 20.5-year period studied by Daniel and Titman, our Hh-Lh regression produces an intercept, −0.22 (t-statistic = −2.28), similar to theirs, −0.18 (t-statistic = −2.30).[1] Thus, our Hh-Lh returns for July 1973 to December 1993 also seem to reject the three-factor model. When the Hh-Lh regression is estimated on our full 68-year period, however, the intercept, −0.06 (t-statistic = −0.83), is quite consistent with the risk model. Moreover, the standard error of the intercept for July 1929 to June 1997, 0.068, is 29 percent smaller than the standard error for July 1973 to December 1993, 0.096. Thus, the test for the overall period is more precise than the test for the period used by Daniel and Titman (1997).

Table IV shows, rather strikingly, that the risk model's problems are specific to the rather short 20.5-year period examined by Daniel and Titman (1997). The most relevant out-of-sample evidence is the Hh-Lh regression for the disjoint 47.5-year period (July 1929 to June 1973 and January 1994 to June 1997) that simply excludes the July 1973 to December 1993 period used by Daniel and Titman (1997) from the full sample. The intercept in this regression, −0.00 (−0.001 if the third decimal is added), is almost perfectly in line with the risk model. And the intercept for this full out-of-sample period is also more precise; its standard error is 0.087, versus 0.096 for the Daniel and Titman (1997) period.

The intercept, −0.06 (t-statistic = −0.83) in the Hh-Lh regression for July 1929 to June 1997 says that our 68-year sample period does not reject the three-factor risk model. But the negative intercept is in the direction predicted by the characteristics model. It is thus interesting to ask whether the data are also consistent with the characteristics model. Since the Hh-Lh portfolio is balanced with respect to the size and BE/ME characteristics (and it has market and SMB loadings close to zero), the characteristics model predicts that the expected return on the portfolio is zero. The alternative hypothesis provided by the risk model is that since Hh-Lh has strong positive loadings on HML, its expected return is positive.

Table IV shows the average Hh-Lh return and its t-statistic for various periods. Not surprisingly, the characteristics model does well in the 20.5-year period of Daniel and Titman (1997); the average Hh-Lh return for July 1973 to December 1993, 0.03 (t-statistic = 0.25), is close to the zero value predicted by the model. In contrast, the average Hh-Lh return for the 47.5-year period that excludes their 20.5-year period from our full sample is 0.16 (t-statistic = 1.63). In a one-sided test (relevant when the alternative is the positive expected return predicted by the three-factor risk model), this average Hh-Lh return rejects the characteristics model at about the 0.95 level.

[1] The intercept reported in Daniel and Titman (1997) is −0.354. They define the return on their arbitrage portfolio as the sum of the returns on two high h_i portfolios minus the sum of the returns on two low h_i portfolios. We divide their intercepts and slopes by two to make them comparable to the results for our arbitrage portfolios, which have only one dollar invested in the long and short portfolios. The t-statistics are not affected by the way the portfolios are standardized.

The most precise evidence and thus the best single test of the characteristics model is provided by our overall July 1929 to June 1997 period. The average Hh-Lh return for this period is 0.12 (t-statistic = 1.56), and the rejection level is again almost 0.95.

In short, rejection of the three-factor risk model in favor of the characteristics model is special to the 20.5-year sample period studied by Daniel and Titman (1997). In the tests for our 68-year period, the risk model outperforms the characteristics model. The risk model overpredicts the average Hh-Lh return by only 6 basis points per month (t-statistic = 0.83), whereas the characteristics model underpredicts the Hh-Lh return by 12 basis points per month (t-statistic = 1.56).

Finally, Tables III and IV closely replicate the portfolio formation approach in Daniel and Titman (1997). We can report, however, that small changes in the approach tilt the results even more toward the risk model. For example, the portfolios formed in June of each year t in Tables III and IV use pre-formation HML slopes estimated with five years of returns ending in December of $t - 1$. If the five-year period instead ends in June of t (the portfolio formation month), the intercepts in the Hh-Lh regressions for the periods shown in Table IV are closer to the zero value predicted by the risk model. For example, the intercept in the Hh-Lh regression for July 1929 to June 1997 drops from -0.06 (t-statistic = -0.83) in Table IV to -0.02 (t-statistic = -0.18). Thus, the results from our full 68-year period again say, but more loudly and clearly, that there is no reason to abandon the risk model in favor of the characteristics model.

IV. Sorts on Market Slopes

There is one issue on which our results agree entirely with Daniel and Titman (1997). Table V shows estimates of the three-factor regression (equation (4)) for three portfolios (Lb, Mb, and Hb) formed on five-year pre-formation three-factor market slopes, b_i. The spreads in the post-formation market slopes, from 0.25 to 0.50, are rather narrow. But the post-formation market slopes do reproduce the ordering of the pre-formation slopes, so pre-formation slopes are informative about post-formation slopes.

The spreads in average return from the high b_i portfolio (Hb) to the low b_i portfolio (Lb) in Table V are positive but tiny, 11 or 12 basis points per month. As a result, the intercepts in the estimates of regression (4) are positive for Lb and negative for Hb. The difference between the high b_i and low b_i returns, Hb-Lb, produces intercepts for the overall July 1929 to June 1997 period, -0.22, and for the earlier July 1929 to June 1963 subperiod, -0.29, that are -2.32 and -1.98 standard errors from zero. Although the Hb-Lb intercept for the later July 1963 to June 1997 subperiod is closer to zero, -0.11 (t-statistic = -0.99), overall Table V suggests that the average value of $R_M - R_f$ overstates the expected premium for differences in loadings on the market return.

Table V

Regressions for Portfolios Formed from a Simple Sort on Market Slope

$$R_i - R_f = a_i + b_i(R_M - R_f) + s_i\text{SMB} + h_i\text{HML} + \epsilon_i$$

At the end of June of each year t (1929 to 1996), we allocate equal numbers of stocks to three portfolios (Lb, Mb, or Hb) based on their three-factor $R_M - R_f$ slope, b_i, for the five years (three years minimum) ending in December of $t - 1$. The regressions explain R_i, the value-weight returns on the portfolios from July of year t to June of $t + 1$.

	BE/ME	Size	Ex Ret	a	b	s	h	$t(a)$	$t(b)$	$t(s)$	$t(h)$	R^2
7/29–6/97												
Lb	0.77	86.18	0.64	0.11	0.80	−0.06	−0.01	2.24	43.68	−2.34	−0.42	0.91
Mb	0.74	89.80	0.67	0.02	0.99	−0.08	−0.01	0.72	123.45	−5.92	−0.89	0.98
Hb	0.80	83.78	0.75	−0.11	1.21	0.13	0.04	−2.06	70.43	4.59	1.40	0.96
Hb − Lb			0.12	−0.22	0.41	0.19	0.06	−2.32	12.06	3.76	0.95	0.49
7/29–6/63												
Lb	0.84	86.46	0.79	0.15	0.77	−0.02	0.03	1.79	31.60	−0.63	0.75	0.91
Mb	0.81	92.21	0.81	0.02	0.97	−0.10	0.01	0.51	81.79	−5.43	0.58	0.98
Hb	0.86	86.30	0.90	−0.14	1.26	0.09	−0.01	−1.82	58.45	2.45	−0.28	0.97
Hb − Lb			0.11	−0.29	0.50	0.11	−0.05	−1.98	11.59	1.64	−0.54	0.57
7/63–6/97												
Lb	0.69	85.90	0.48	0.06	0.87	−0.14	−0.01	1.16	55.17	−5.96	−0.56	0.92
Mb	0.67	87.39	0.52	−0.01	1.04	−0.07	0.02	−0.40	121.70	−5.09	1.23	0.98
Hb	0.74	81.26	0.60	−0.05	1.13	0.22	0.04	−0.74	61.68	8.12	1.20	0.94
Hb − Lb			0.12	−0.11	0.26	0.36	0.06	−0.99	7.93	7.57	0.97	0.37
7/73–12/93												
Lb	0.78	85.37	0.47	0.10	0.87	−0.16	−0.04	1.30	44.02	−5.20	−1.08	0.93
Mb	0.79	85.79	0.54	0.01	1.04	−0.05	0.02	0.21	103.56	−3.29	1.01	0.98
Hb	0.87	79.55	0.59	−0.09	1.13	0.25	0.05	−1.04	48.82	6.65	1.21	0.95
Hb − Lb			0.12	−0.19	0.25	0.41	0.09	−1.22	6.19	6.26	1.21	0.38

These results are like those observed in tests of the Sharpe (1964)–Lintner (1965) CAPM, which typically find that the relation between average return and univariate market β is too flat (Black, Jensen, and Scholes (1972), Fama and MacBeth (1973), Fama and French (1992)). And the standard explanations can be invoked. (i) Perhaps the multifactor version of Merton's (1973) ICAPM that does not include a riskfree security (Fama (1996)), and is analogous to the Black (1972) version of the CAPM, is more relevant than the risk-free rate version. (ii) The problem may be one of implementation. For example, we use a market portfolio that includes only common stocks. (iii) The three-factor model is just a model, and this may be one of its shortcomings. Whatever the relevant story, the results in Table V are potentially important in applications of the three-factor model (for example, to evaluate portfolio performance or to estimate normal returns in event studies) since they imply that the expected returns predicted by equation (1) are distorted when loadings on the market factor differ a lot from 1.0.

V. Summary and Conclusions

The value premium in average stock returns is robust. Measured by HML (which is neutral with respect to size effects), the value premium for July 1929 to June 1963 is 0.50 percent per month (t-statistic = 2.80). This is close to the premium for July 1963 to June 1997, 0.43 percent per month (t-statistic = 3.38), observed in earlier work. The size effect in average returns is smaller. Measured by SMB (which is neutral with respect to value effects), the size premium for the full 68-year sample period is 0.20 percent per month (t-statistic = 1.78).

The three-factor risk model (1) explains the value premium better than a popular competitor, the characteristics model of Daniel and Titman (1997). Contradicting the characteristics model, the Hh-Lh returns for our 68-year sample period produce no evidence against the risk model's prediction that HML risk loading determines expected return, irrespective of the BE/ME characteristic. Specifically, the intercept in the Hh-Lh three-factor regression for July 1929 to June 1997, -6 basis points per month (t-statistic = -0.83), is economically and statistically close to zero. Moreover, if we omit the 20.5-year period examined by Daniel and Titman (1997), the intercept in the Hh-Lh regression for the rest of our 68-year sample period, -0.1 basis points per month, could hardly be closer to the zero value predicted by the risk model. Thus, the evidence of Daniel and Titman (1997) in favor of the characteristics model is special to their rather short sample period.

Finally, when portfolios are formed from independent sorts of stocks on size and BE/ME (Table II), the three-factor model (1) is rejected by the Gibbons et al. (1989) test. This result shows that the three-factor model is just a model and thus an incomplete description of expected returns. What the remaining tests say is that the model's shortcomings are just not those predicted by the characteristics model.

REFERENCES

Black, Fischer, 1972, Capital market equilibrium with restricted borrowing, *Journal of Business* 45, 444–454.

Black, Fischer, 1993, Beta and return, *Journal of Portfolio Management* 20, 8–18.

Black, Fischer, Michael C. Jensen, and Myron Scholes, 1972, The capital asset pricing model: Some empirical tests, in M. Jensen, ed.: *Studies in the Theory of Capital Markets* (Praeger, New York).

Capaul, Carlo, Ian Rowley, and William F. Sharpe, 1993, International value and growth stock returns, *Financial Analysts Journal* 49, 27–36.

Chan, Louis K.C., Yasushi Hamao, and Josef Lakonishok, 1991, Fundamentals and stock returns in Japan, *Journal of Finance* 46, 1739–1789.

Daniel, Kent, and Sheridan Titman, 1997, Evidence on the characteristics of cross sectional variation in stock returns, *Journal of Finance* 52, 1–33.

Davis, James, 1994, The cross-section of realized stock returns: The pre-COMPUSTAT evidence, *Journal of Finance* 49, 1579–1593.

DeBondt, Werner F. M., and Thaler, Richard H., 1987, Further evidence on investor overreaction and stock market seasonality, *Journal of Finance* 42, 557–581.

Fama, Eugene F., 1996, Multifactor portfolio efficiency and multifactor asset pricing, *Journal of Financial and Quantitative Analysis* 31, 441–465.

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

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

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

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

Fama, Eugene F., and Kenneth R. French, 1998, Value versus growth: The international evidence, *Journal of Finance* 53, 1975–1999.

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

Gibbons, Michael R., Stephen A. Ross, and Jay Shanken, 1989, A test of the efficiency of a given portfolio, *Econometrica* 57, 1121–1152.

Haugen, Robert, 1995, *The New Finance: The Case against Efficient Markets* (Prentice Hall, Englewood Cliffs, N.J.).

Lakonishok, Josef, Andrei Shleifer, and Robert W. Vishny, 1994, Contrarian investment, extrapolation, and risk, *Journal of Finance* 49, 1541–1578.

Lintner, John, 1965, The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets, *Review of Economics and Statistics* 47, 13–37.

Loughran, Tim, 1997, Book-to-market across firm size, exchange, and seasonality, *Journal of Financial and Quantitative Analysis* 32, 249–268.

MacKinlay, A. Craig, 1995, Multifactor models do not explain deviations from the CAPM, *Journal of Financial Economics* 38, 3–28.

Merton, Robert C., 1973, An intertemporal capital asset pricing model, *Econometrica* 41, 867–887.

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

Ross, Stephen A., 1976, The arbitrage theory of capital asset pricing, *Journal of Economic Theory* 13, 341–360.

Sharpe, William F., 1964, Capital asset prices: A theory of market equilibrium under conditions of risk, *Journal of Finance* 19, 425–442.

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