



ELSEVIER

Journal of Financial Economics 43 (1997) 153–193

**JOURNAL OF  
Financial  
ECONOMICS**

## Industry costs of equity

Eugene F. Fama<sup>a</sup>, Kenneth R. French<sup>\*.b</sup>

<sup>a</sup>Graduate School of Business, University of Chicago, Chicago, IL 60637, USA

<sup>b</sup>School of Management, Yale University, New Haven, CT 06520, USA

(Received March 1994; final version received August 1996)

---

### Abstract

Estimates of the cost of equity for industries are imprecise. Standard errors of more than 3.0% per year are typical for both the CAPM and the three-factor model of Fama and French (1993). These large standard errors are the result of (i) uncertainty about true factor risk premiums and (ii) imprecise estimates of the loadings of industries on the risk factors. Estimates of the cost of equity for firms and projects are surely even less precise.

*Key words:* Cost of equity; Asset pricing models; Risk loadings

*JEL classification:* G12; G31

---

### 1. Introduction

Textbooks in corporate finance advise managers to evaluate an investment project by comparing the required outlay to the present value of the expected future cash flows. Most textbooks emphasize the uncertainty in projections of cash flows. Our main point is that the cost of capital estimates used to discount cash flows are also unavoidably imprecise.

There are at least three cost of capital problems. First, it is not clear which asset pricing model should be used. The capital asset pricing model (CAPM) of Sharpe (1964) and Lintner (1965) is the common choice. Recent evidence suggests, however, that the CAPM is not a good description of expected returns. As an alternative, Fama and French (1993, 1995) propose a three-factor pricing model. But some argue that this model is empirically inspired and lacks strong

---

\*Corresponding author.

We acknowledge the helpful comments of Kent Daniel, Ed George, Mark Mitchell, Stewart Myers, Stephen Ross, G. William Schwert, Ross Stevens, and Jay Shanken, the referee.

theoretical foundations. Other multifactor models have been used to estimate the cost of capital (Bower, Bower, and Logue, 1984; Goldenberg and Robin, 1991; Bower and Schink, 1994; Elton, Gruber, and Mei, 1994), but there is no consensus about which is best. And the choice of model is important. In the tests below, differences of 2% per year between estimates of the cost of equity from the CAPM and our three-factor model are common.

We do not take a stance on which is the right asset pricing model. Instead we use both the CAPM and our three-factor model to estimate industry costs of equity (*CE*'s). Our goal is to illustrate in detail two problems that plague *CE* estimates from any asset pricing model.

The first problem is imprecise estimates of risk loadings. Estimates of CAPM and three-factor risk loadings for industries would be precise if the loadings were constant. We find, however, that there is strong variation through time in the CAPM and three-factor risk loadings of industries. As a result, if we are trying to measure an industry's current risk loadings and cost of equity, estimates from full sample (1963–1994) regressions are no more accurate than the imprecise estimates from regressions that use only the latest three years of data. And industries give an understated picture of the problems that will arise in estimating risk loadings for individual firms and investment projects.

The second problem is imprecise estimates of factor risk premiums. For example, the price of risk in the CAPM is the expected return on the market portfolio minus the risk-free interest rate,  $E(R_M) - R_f$ . The annualized average excess return on the Center for Research in Security Prices (CRSP) value-weight market portfolio of NYSE, AMEX, and NASDAQ stocks for our 1963–1994 sample period is 5.16%; its standard error is 2.71%. Thus, if we use the historical market premium to estimate the expected premium, the traditional plus-and-minus-two-standard-error interval ranges from less than zero to more than 10.0%.

Our message is that uncertainty of this magnitude about risk premiums, coupled with the uncertainty about risk loadings, implies woefully imprecise estimates of the cost of equity.

We start with a brief discussion of the CAPM and the three-factor model (Section 2). Section 3 explores variation through time in the CAPM and three-factor risk loadings of industries. Sections 4 and 5 compare different ways to estimate the loadings and the cost of equity. Section 6 examines uncertainty about factor risk premiums. In Section 7 we present standard errors for *CE* estimates that allow for uncertainty about both risk loadings and risk premiums. Section 8 concludes.

## 2. The CAPM and the three-factor model

In the CAPM, the expected return on stock *i* or, equivalently, the cost of equity for firm *i* is

$$E(R_i) = R_f + \beta_i[E(R_M) - R_f], \quad (1)$$

where  $R_f$  is the risk-free interest rate,  $E(R_M)$  is the expected return on the value-weight market portfolio, and  $\beta_i$ , the CAPM risk of stock  $i$ , is the slope in the regression of its excess return on the market's excess return,

$$R_i - R_f = \alpha_i + \beta_i [R_M - R_f] + e_i. \quad (2)$$

Recent empirical work questions the adequacy of the CAPM as a model for expected returns. Specifically, many papers argue that market beta does not

Table 1  
Factor risk premiums for the CAPM and the three-factor model: 7/63–12/94

$$R_i - R_f = a_i + b_i [R_M - R_f] + e_i, \quad R_i - R_f = a_i + b_i [R_M - R_f] + s_i SMB + h_i HML + e_i$$

The returns here and in all following tables are in percents.  $R_f$  is the one-month Treasury bill rate observed at the beginning of the month. The explanatory returns  $R_M$ ,  $SMB$ , and  $HML$  are formed as follows. At the end of June of each year  $t$  (1963–1994), NYSE, AMEX, and NASDAQ stocks are allocated to two groups (small or big, S or B) based on whether their June market equity ( $ME$ , stock price times shares outstanding) is below or above the median  $ME$  for NYSE stocks. NYSE, AMEX, and NASDAQ stocks are allocated in an independent sort to three book-to-market-equity ( $BE/ME$ ) groups (low, medium, or high; L, M, or H) based on the breakpoints for the bottom 30%, middle 40%, and top 30% of the values of  $BE/ME$  for NYSE stocks.  $BE$  is the Compustat book value of stockholders' equity, plus balance sheet deferred taxes and investment tax credit (if available), plus post-retirement benefit liability (if available), minus the book value of preferred stock. Depending on availability, we use the redemption, liquidation, or par value (in that order) to estimate the book value of preferred stock. The  $BE/ME$  ratio used to form portfolios in June of year  $t$  is then 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 size- $BE/ME$  portfolios (S/L, S/M, S/H, B/L, B/M, B/H) are defined as the intersections of the two  $ME$  and the three  $BE/ME$  groups. Value-weight monthly returns on the portfolios are calculated from July of year  $t$  to the following June.  $SMB$  is the difference, each month, between the average of the returns on the three small-stock portfolios (S/L, S/M, and S/H) and the average of the returns on the three big-stock portfolios (B/L, B/M, and B/H).  $HML$  is the difference between the average of the returns on the two high- $BE/ME$  portfolios (S/H and B/H) and the average of the returns on the two low- $BE/ME$  portfolios (S/L and B/L). We do not use negative  $BE$  firms, which are rare prior to 1980, when calculating the breakpoints for  $BE/ME$  or when forming the six size- $BE/ME$  portfolios. Also, only ordinary common equity (as classified by CRSP) is included in the tests. This means that ADR's, REIT's, and units of beneficial interest are excluded. The market return  $R_M$  is the value-weight average of the returns on all stocks in the six size- $BE/ME$  portfolios, plus the negative  $BE$  stocks excluded from the portfolios. The sample size is 378 months.

	$R_M - R_f$	$SMB$	$HML$
<i>Monthly</i>			
Average premium	0.43	0.27	0.45
Standard deviation (SD)	4.39	2.86	2.56
Standard error (SD/378 <sup>1/2</sup> )	0.23	0.15	0.13
<i>Annualized (12 times monthly)</i>			
Average premium	5.16	3.24	5.40
Standard error	2.71	1.77	1.58

suffice to explain expected stock returns. (See Fama and French, 1992, and the references therein.) Multifactor variants of Merton's (1973) intertemporal asset pricing model (ICAPM) or Ross' (1976) arbitrage pricing theory (APT) seem to give better descriptions of expected stock returns (e.g., Chen, Roll, and Ross, 1986; Fama and French, 1993, 1996).

Fama and French (1993) propose a three-factor model in which a security's expected return depends on the sensitivity of its return to the market return and the returns on two portfolios meant to mimic additional risk factors. The mimicking portfolios are *SMB* (small minus big), which is the difference between the returns on a portfolio of small stocks and a portfolio of big stocks, and *HML* (high minus low), the difference between the returns on a portfolio of high-book-to-market-equity (*BE/ME*) stocks and a portfolio of low-*BE/ME* stocks. (See Table 1.) The expected-return equation of the three-factor model is

$$E(R_i) - R_f = b_i[E(R_M) - R_f] + s_i E(SMB) + h_i E(HML), \quad (3)$$

where  $b_i$ ,  $s_i$ , and  $h_i$  are the slopes in the regression

$$R_i - R_f = a_i + b_i[R_M - R_f] + s_i SMB + h_i HML + e_i. \quad (4)$$

Using *SMB* to explain returns is in line with the evidence of Huberman and Kandel (1987) that there is covariation in the returns on small stocks that is not captured by the market return and is compensated in average return. Using *HML* to explain returns is in line with the evidence of Chan and Chen (1991) that there is return covariation related to relative distress (proxied here by *BE/ME*, the ratio of the book value of a firm's common stock to its market value) that is missed by the market return and is compensated in average return. Fama and French (1993, 1996) show that the three-factor model captures much of the spread in the cross-section of average returns on portfolios formed on size, *BE/ME*, and other variables (earnings/price, cash flow/price, and long-term past return) known to cause problems for the CAPM.

### 3. Time-varying risk loadings

A manager using an asset pricing model to measure the discount rate for a project must estimate the project's sensitivities to the model's risk factors. Table 2 shows estimates of CAPM and three-factor risk loadings for 48 value-weight industries. The industries are defined with the goal of having a manageable number of distinct industries that cover all NYSE, AMEX, and NASDAQ stocks. (See Appendix A.) Because the sample of firms on Compustat is rather limited in earlier years, the sample period is July 1963 to December 1994.

The full-period risk loadings in Table 2 seem to be estimated precisely. The average standard error for the CAPM market slopes is only 0.04. The average standard errors for the market, *SMB*, and *HML* slopes in the three-factor model

Table 2  
CAPM and three-factor industry regressions: 7/63–12/94

$$R_i - R_f = a_i + b_i[R_M - R_f] + e_i, \quad R_i - R_f = a_i + b_i[R_M - R_f] + s_i SMB + h_i HML + e_i$$

The industries are defined in Appendix A. The monthly explanatory returns,  $R_M - R_f$ ,  $SMB$ , and  $HML$ , are described in Table 1.  $t(a)$  is the  $t$ -statistic for the regression intercept. The regression  $R^2$  are adjusted for degrees of freedom. Mean is the average across the 48 industries. The average standard error of the  $R_M - R_f$  slopes in the CAPM regressions is 0.04. The average standard errors of the  $R_M - R_f$ ,  $SMB$ , and  $HML$  slopes in the three-factor regressions are 0.05, 0.07, and 0.07.

Industry	CAPM				Three-factor					
	$a$	$t(a)$	$b$	$R^2$	$a$	$t(a)$	$b$	$s$	$h$	$R^2$
Drugs	0.23	1.29	0.92	0.59	0.61	3.88	0.84	-0.25	-0.63	0.68
MedEq	0.11	0.57	1.17	0.67	0.39	2.24	0.99	0.26	-0.60	0.73
Hlth	0.28	0.91	1.56	0.56	0.43	1.54	1.24	0.93	-0.59	0.66
Comps	-0.11	-0.55	1.04	0.59	0.13	0.66	0.90	0.17	-0.49	0.63
Chips	0.07	0.32	1.38	0.69	0.15	0.83	1.15	0.69	-0.39	0.77
BusSv	0.12	0.76	1.34	0.80	0.14	1.26	1.13	0.72	-0.29	0.89
LabEq	-0.15	-0.91	1.29	0.77	-0.08	-0.56	1.13	0.49	-0.29	0.82
Hshld	-0.00	-0.02	0.97	0.72	0.14	1.04	0.91	0.00	-0.27	0.73
Meals	0.25	1.18	1.32	0.66	0.25	1.30	1.12	0.74	-0.24	0.74
Beer	0.37	2.12	0.92	0.59	0.51	2.90	0.90	-0.13	-0.22	0.60
PerSv	-0.08	-0.35	1.25	0.59	-0.16	-0.79	1.00	1.00	-0.20	0.74
Cnstr	-0.28	-1.50	1.28	0.70	-0.27	-1.43	1.21	0.21	-0.09	0.71
Rtail	0.07	0.48	1.11	0.73	0.06	0.37	1.04	0.27	-0.06	0.75
Fun	0.21	0.91	1.35	0.64	0.08	0.40	1.17	0.83	-0.04	0.73
Food	0.32	2.36	0.87	0.68	0.35	2.51	0.88	-0.07	-0.03	0.68
Agric	-0.07	-0.27	1.00	0.44	-0.18	-0.77	0.85	0.71	-0.02	0.53
Mach	-0.11	-0.86	1.16	0.82	-0.15	-1.22	1.11	0.25	-0.00	0.83
Books	0.12	0.73	1.17	0.71	0.04	0.26	1.08	0.45	0.00	0.75
Aero	0.03	0.14	1.26	0.68	-0.07	-0.34	1.15	0.51	0.00	0.72
Coal	0.04	0.12	0.96	0.36	-0.05	-0.18	0.86	0.46	0.01	0.39
Guns	0.17	0.80	1.04	0.55	0.09	0.42	0.95	0.41	0.01	0.59
Whlsl	-0.10	-0.81	1.15	0.81	-0.24	-2.89	1.01	0.71	0.01	0.92
Fin	0.19	1.14	1.16	0.72	0.12	0.75	1.11	0.30	0.02	0.74
ElcEq	0.06	0.42	1.15	0.75	0.05	0.34	1.15	-0.00	0.02	0.74
Boxes	0.13	0.78	1.03	0.65	0.09	0.51	0.99	0.17	0.02	0.66
BldMt	-0.01	-0.09	1.13	0.83	-0.06	-0.55	1.11	0.15	0.05	0.84
Insur	0.08	0.39	1.01	0.58	0.03	0.14	1.00	0.09	0.06	0.58
Gold	0.33	0.78	0.78	0.15	0.21	0.50	0.71	0.40	0.08	0.16
Misc	-0.28	-1.00	1.26	0.50	-0.54	-2.31	1.03	1.19	0.08	0.67
Trans	-0.07	-0.43	1.21	0.75	-0.71	-1.09	1.16	0.30	0.09	0.77
Rubbr	0.05	0.37	1.21	0.78	-0.08	-0.61	1.12	0.49	0.09	0.83
FabPr	-0.13	-0.55	1.31	0.63	-0.37	-2.16	1.11	1.10	0.09	0.80
Clths	0.08	0.39	1.24	0.66	-0.13	-0.78	1.09	0.83	0.11	0.78
Chem	-0.02	-0.17	1.09	0.81	-0.10	-0.85	1.13	-0.03	0.17	0.81
Toys	-0.01	-0.04	1.34	0.54	-0.28	-1.11	1.17	0.97	0.17	0.65
Ships	0.17	0.61	1.19	0.50	-0.05	-0.18	1.09	0.66	0.17	0.56
Soda	0.30	1.32	1.24	0.60	0.13	0.55	1.19	0.44	0.18	0.63

Table 2 (continued)

Industry	CAPM				Three-factor					
	<i>a</i>	<i>t(a)</i>	<i>b</i>	<i>R</i> <sup>2</sup>	<i>a</i>	<i>t(a)</i>	<i>b</i>	<i>s</i>	<i>h</i>	<i>R</i> <sup>2</sup>
Enrgy	0.13	0.71	0.85	0.50	0.08	0.45	0.96	−0.35	0.21	0.54
Mines	0.30	1.24	0.98	0.45	0.08	0.34	0.91	0.53	0.23	0.50
Smoke	0.40	1.82	0.80	0.40	0.28	1.24	0.86	−0.04	0.24	0.41
Paper	−0.05	−0.32	1.11	0.75	−0.22	−1.54	1.14	0.16	0.27	0.77
Txtls	0.05	0.27	1.12	0.65	−0.24	−1.50	1.03	0.71	0.30	0.76
Banks	−0.04	−0.26	1.09	0.76	−0.25	−1.84	1.13	0.13	0.35	0.79
Telecm	0.13	0.92	0.66	0.52	−0.02	−0.11	0.79	−0.23	0.35	0.59
Util	−0.00	−0.02	0.66	0.55	−0.17	−1.33	0.79	−0.20	0.38	0.62
RIEst	−0.58	−2.32	1.17	0.53	−1.01	−5.45	1.01	1.18	0.40	0.75
Steel	−0.22	−1.06	1.16	0.61	−0.53	−2.64	1.17	0.40	0.43	0.67
Autos	−0.04	−0.21	1.01	0.56	−0.40	−2.09	1.10	0.17	0.60	0.63
Mean	0.05	0.25	1.11	0.63	−0.03	−0.21	1.04	0.39	0.02	0.68

are 0.05, 0.07, and 0.07. These small standard errors are misleading, however, because they assume the true slopes are constant. Industry risk loadings wander through time, and estimates of period-by-period loadings are much less precise.

### 3.1. The implied volatility of true risk loadings

One way to document the temporal variation in risk loadings is with rolling CAPM and three-factor regressions (estimated monthly using five years of past returns). The idea is that, if the true CAPM and three-factor slopes for industries vary through time, the time-series variation of the rolling-regression slopes should exceed that implied by estimation error. Specifically, under the standard assumption that the sampling error of a slope is uncorrelated with the true value of the slope, the time-series variance of a rolling-regression slope is just the sum of the variance of the true slope and the variance of the estimation error,

$$\sigma^2(\text{Time Series}) = \sigma^2(\text{True}) + \sigma^2(\text{Estimation Error}). \quad (5)$$

Table 3 reports estimates of  $\sigma(\text{True})$  for the market, *SMB*, and *HML* slopes in five-year rolling CAPM and three-factor regressions. The estimates document substantial temporal variation in the CAPM betas of industries. All but five of the implied standard deviations of the true CAPM market slopes are greater than zero, 28 are greater than 0.10, and nine are greater than 0.20. The average is 0.12. Thus, if the typical industry has a long-term average beta of 1.0, the traditional two-standard-deviation rule of thumb suggests that its current true beta might be anywhere between 0.76 and 1.24. If we use the average market premium during our sample period, 5.16% per year, as the expected premium,

Table 3  
Implied standard deviations of true market, *SMB*, and *HML* slopes in five-year rolling CAPM and three-factor regressions

$$R_i - R_f = a_i + b_i[R_M - R_f] + e_i, \quad R_i - R_f = a_i + b_i[R_M - R_f] + s_iSMB + h_iHML + e_i$$

The industries are defined in Appendix A.  $R_M - R_f$ , *SMB*, and *HML* are defined in Table 1. CAPM and three-factor regressions are estimated each month of the 6/68 to 12/94 period, using a rolling window of 60 prior monthly returns. The implied standard deviation of an industry's true market, *SMB*, or *HML* slope,  $\hat{\sigma}(True)$ , is the square root of the difference between the time-series variance of the industry's five-year slope estimates and the average of the estimation-error variances (squared standard errors) of its five-year slope estimates,

$$\hat{\sigma}(True) = [\hat{\sigma}^2(Time\ Series) - \hat{\sigma}^2(Estimation\ Error)]^{1/2}.$$

If the average estimation-error variance exceeds the time-series variance,  $\hat{\sigma}(True)$  is set to zero. Mean is the average of  $\hat{\sigma}(True)$  across the 48 industries.

Industry	CAPM	Three-factor		
	<i>b</i>	<i>b</i>	<i>s</i>	<i>h</i>
Drugs	0.088	0.101	0.141	0.262
MedEq	0.079	0.116	0.000	0.000
Hlth	0.251	0.148	0.000	0.181
Comps	0.000	0.000	0.038	0.277
Chips	0.025	0.051	0.030	0.314
BusSv	0.114	0.067	0.155	0.255
LabEq	0.043	0.063	0.000	0.169
Hshld	0.092	0.082	0.084	0.209
Meals	0.227	0.142	0.176	0.331
Beer	0.218	0.226	0.142	0.201
PerSv	0.120	0.000	0.382	0.104
Cnstr	0.098	0.031	0.332	0.293
Rtail	0.116	0.083	0.113	0.158
Fun	0.163	0.000	0.000	0.181
Food	0.104	0.042	0.226	0.245
Agric	0.248	0.133	0.038	0.000
Mach	0.000	0.000	0.201	0.155
Books	0.149	0.136	0.187	0.184
Aero	0.168	0.096	0.161	0.229
Coal	0.172	0.140	0.373	0.000
Guns	0.131	0.069	0.255	0.325
Whlsl	0.113	0.011	0.173	0.053
Fin	0.106	0.169	0.040	0.052
ElcEq	0.071	0.108	0.000	0.213
Boxes	0.084	0.069	0.077	0.190
BldMt	0.000	0.000	0.086	0.071
Insur	0.074	0.000	0.086	0.233
Gold	0.415	0.425	0.090	0.000
Misc	0.250	0.120	0.084	0.152
Trans	0.078	0.023	0.148	0.118
Rubbr	0.080	0.042	0.144	0.102

Table 3 (continued)

Industry	CAPM	Three-factor		
	<i>b</i>	<i>b</i>	<i>s</i>	<i>h</i>
FabPr	0.253	0.124	0.179	0.211
Clths	0.144	0.131	0.135	0.290
Chem	0.041	0.000	0.000	0.154
Toys	0.082	0.079	0.321	0.144
Ships	0.114	0.000	0.000	0.313
Soda	0.215	0.186	0.272	0.000
Enrgy	0.180	0.156	0.171	0.365
Mines	0.170	0.129	0.211	0.000
Smoke	0.118	0.056	0.179	0.365
Paper	0.000	0.063	0.000	0.148
Txtls	0.000	0.076	0.088	0.137
Banks	0.121	0.093	0.126	0.126
Telcm	0.138	0.162	0.197	0.000
Util	0.037	0.000	0.107	0.000
RIEst	0.274	0.091	0.136	0.206
Steel	0.021	0.000	0.201	0.148
Autos	0.106	0.135	0.196	0.314
Mean	0.123	0.087	0.135	0.170

the industry's current cost of equity (in excess of the risk-free rate) might be anywhere between 3.92% and 6.40% per year.

The industries' true sensitivities to the market, size, and distress risks of the three-factor model are also volatile. The average of the implied standard deviations of the true *SMB* slopes is 0.14. Forty of 48 are positive, 29 are greater than 0.10, and ten are greater than 0.20. Thus, many industries' value-weight returns behave like small-stock returns during some periods and like big-stock returns in others. Similarly, 40 of the standard deviations of the true *HML* slope are positive, 37 are greater than 0.10, and 20 are greater than 0.20. The average is 0.17. In comparison, the standard deviation of the cross-section of the 48 full-period *HML* slopes in Table 2 is 0.27. Thus, the variation through time in the true *HML* slopes of many industries is almost as large as the cross-sectional standard deviation of the long-term average *HML* slopes of the 48 industries.

Table 3 suggests that the true market betas of the CAPM are more variable than the true market slopes of the three-factor model. For 33 of 48 industries, the implied standard deviation of the true CAPM market beta is higher than that of the true three-factor market slope. The average  $\hat{\sigma}(True)$  for beta falls from 0.12 in the CAPM regressions to 0.09 in the three-factor regressions. This is consistent with the evidence that the *SMB* and *HML* slopes in three-factor regressions typically reduce the cross-sectional variation in market slopes. For example, the

standard deviation of the 48 market slopes in Table 1 is 0.19 in the CAPM regressions versus 0.13 in the three-factor regressions.

A note of caution. The population time-series variances of the rolling-regression slopes are well-defined if the true slopes are stationary (mean-reverting). Permanent changes in the supply and demand conditions facing an industry may, however, produce permanent changes in its risk loadings. If the true slopes are not stationary, the implied standard deviations in Table 3 are still descriptive evidence that the true slopes change through time. But comparisons of the estimates across industries, and comparisons of the estimates for different risk factors, may not be meaningful.

### 3.2. Conditional regressions

An alternative to the rolling regressions, both for documenting temporal variation in risk loadings and for estimating an industry's cost of equity at a specific time, is to use instruments to track the wandering risks. Though size is surely not a perfect proxy for sensitivity to *SMB*, we expect that an industry's *SMB* loading will increase if firms in the industry become smaller. We also expect that if an industry becomes distressed, its book-to-market ratio and its *HML* loading will increase. Thus, we try to track time-varying sensitivities to *SMB* and *HML* with conditional regressions in which an industry's *SMB* and *HML* slopes vary with the average size and book-to-market-equity of firms in the industry,<sup>1</sup>

$$R_i - R_f = a_i + b_i[R_M - R_f] + [s_{i1} + s_{i2} \ln(ME)_i]SMB + [h_{i1} + h_{i2} \ln(BE/ME)_i]HML + e_i. \quad (6)$$

The estimates of (6) in Table 4 confirm that the loadings of industries on *SMB* and *HML* wander through time. As predicted, *SMB* loadings fall when the average size of firms in an industry increases; the  $\ln(ME)_i$  *SMB* slope,  $s_{i2}$ , is negative for all but seven of the 48 industries, and 20 of the  $s_{i2}$  are more than two standard errors below zero. As predicted, *HML* loadings are positively related to the measure of relative distress,  $\ln(BE/ME)_i$ ; all but one of the 48  $\ln(BE/ME)_i$  *HML* slopes are positive, and 31 are more than two standard errors above zero.

Table 4 also reports the time-series standard deviations of the monthly conditional loadings on *SMB* and *HML*. These standard deviations, estimated as the absolute value of the conditional slope ( $s_{i2}$  or  $h_{i2}$ ) multiplied by the time-series standard deviation of the conditioning variable [ $\ln(ME)_i$  or  $\ln(BE/ME)_i$ ], are similar to the implied standard deviations of the true *SMB* and

<sup>1</sup>We thank Jay Shanken for suggesting this approach, which is like that in Shanken (1990).

Table 4

Conditional industry regressions for 7/65-12/94

$$R_{i,t} - R_{f,t} = a_i + b_i[R_M - R_{f,t}] + [s_{1i} + s_{12} \ln(ME)_t]SMB + [h_{1i} + h_{12} \ln(BE/ME)_t]HML + e_{i,t}$$

The industries are defined in Appendix A.  $R_{f,t}$ ,  $R_{M,t}$ ,  $SMB$ , and  $HML$  are defined in Table 1.  $\ln(ME)_t$  is the value-weight average of the natural log of market equity for all firms in an industry, measured at the end of the month preceding the dependent and explanatory returns.  $\ln(BE/ME)_t$  is the value-weight average of the natural log of  $BE/ME$  for industry firms that have Compustat data for  $BE$ .  $BE/ME$  is measured once each calendar year (as described in Table 1), and it is used to explain 12 monthly returns starting in July of the following year. To control for marketwide variation,  $\ln(ME)$  and  $\ln(BE/ME)$  are measured net of their average values (across industries) each month. In some industries we do not have accounting data for any firms before 1964. Thus the start date for the regressions is 7/65. The regression  $R^2$  are adjusted for degrees of freedom. Std. dev. is the standard deviation of the monthly conditional loading on  $SMB$  or  $HML$ , calculated as the absolute value of the conditional slope ( $s_{12}$  or  $h_{12}$ ) times the time-series standard deviation of the conditioning variables  $[\ln(ME)_t$  or  $\ln(BE/ME)_t]$ .

Industry	Coefficients				t-statistics							Std. dev.		
	a	b	s <sub>1</sub>	s <sub>2</sub>	h <sub>1</sub>	h <sub>2</sub>	t(a)	t(s <sub>1</sub> )	t(s <sub>2</sub> )	t(h <sub>1</sub> )	t(h <sub>2</sub> )	R <sup>2</sup>	SMB	HML
Drugs	0.58	0.86	0.40	-0.39	0.40	1.09	3.64	1.72	-2.72	2.04	5.38	0.72	0.12	0.26
MedEq	0.40	1.01	0.20	-0.26	-0.20	0.67	2.24	2.61	-1.49	-1.06	2.28	0.74	0.09	0.15
Hlth	0.47	1.20	0.41	-0.39	-0.59	0.21	1.68	1.70	-2.22	-4.44	1.00	0.69	0.24	0.10
Comps	0.13	0.90	0.67	-0.17	-0.39	0.18	0.65	2.86	-2.20	-2.94	1.01	0.65	0.16	0.08
Chips	0.20	1.15	0.67	-0.22	-0.02	0.95	1.03	8.93	-0.71	-0.15	4.17	0.78	0.05	0.28
BusSv	0.13	1.11	0.30	-0.37	-0.08	0.59	1.19	2.62	-3.62	-0.97	3.17	0.91	0.15	0.14
LabEq	-0.09	1.12	0.48	-0.24	-0.22	0.23	-0.59	8.58	-1.15	-1.95	0.70	0.83	0.05	0.04
Hshld	0.13	0.93	0.02	-0.00	0.00	0.51	0.91	0.11	-0.06	0.06	4.60	0.75	0.00	0.23
Meals	0.26	1.12	0.60	-0.30	0.10	0.84	1.33	7.14	-2.39	0.83	3.67	0.76	0.14	0.24
Beer	0.46	0.92	0.10	-0.15	0.27	0.73	2.55	0.51	-1.19	2.00	4.34	0.62	0.08	0.25
PerSv	-0.21	0.99	0.18	-0.44	-0.21	0.09	-1.04	0.77	-3.70	-1.99	0.50	0.75	0.24	0.04
Cnstr	-0.17	1.23	0.20	-0.50	0.02	0.65	-0.90	2.88	-4.77	0.20	3.75	0.74	0.33	0.23
Rtail	0.08	1.05	0.24	0.05	0.21	1.18	0.50	1.96	0.36	2.50	4.83	0.76	0.02	0.25
Fun	0.13	1.17	0.79	-0.14	-0.02	0.24	0.60	9.78	-1.10	-0.26	0.67	0.73	0.07	0.05
Food	0.37	0.87	0.21	-0.38	0.13	1.30	2.70	2.57	-3.79	2.19	6.24	0.72	0.17	0.30
Agric	-0.22	0.84	0.21	-0.43	-0.05	0.11	-0.90	0.76	-1.83	-0.44	0.70	0.56	0.13	0.07
Mach	-0.17	1.12	0.29	-0.41	-0.08	0.91	-1.37	6.27	-2.44	-1.40	3.47	0.84	0.11	0.14

Aero	-0.03	1.15	0.65	-0.51	-0.31	0.93	-0.17	7.51	-3.46	-2.86	3.81	0.74	0.23
Books	0.04	1.06	0.30	-0.58	0.40	1.31	0.22	4.69	-4.12	3.07	3.89	0.78	0.24
Coal	-0.05	0.87	0.26	-0.19	-0.10	0.35	-0.15	1.05	-1.02	-0.50	0.94	0.39	0.12
Guns	0.20	0.93	0.45	-0.61	-0.56	1.41	0.95	5.69	-3.18	-4.13	5.10	0.63	0.36
Whisl	-0.16	0.99	-0.29	-0.80	-0.04	0.37	-2.03	-1.64	-5.74	-1.08	2.14	0.94	0.06
EleEq	0.05	1.16	0.00	-0.05	0.05	0.07	0.33	0.25	-0.30	0.48	0.22	0.75	0.04
Fin	0.20	1.10	0.29	-0.10	0.00	0.12	1.19	4.85	-0.82	0.02	0.35	0.75	0.02
Boxes	0.12	1.00	0.19	-0.15	-0.04	0.27	0.69	2.73	-0.80	-0.42	0.92	0.66	0.05
BldMt	-0.03	1.12	0.28	-0.19	0.15	0.43	-0.23	2.70	-1.40	2.30	2.14	0.84	0.05
Insur	0.12	0.99	0.21	-0.69	-0.09	0.57	0.57	2.16	-1.97	-0.98	3.02	0.59	0.13
Gold	0.22	0.68	0.63	0.64	0.26	0.36	0.49	2.74	1.23	0.82	0.76	0.16	0.17
Misc	-0.60	1.03	1.48	0.14	0.04	0.30	-2.42	4.46	0.94	0.37	1.62	0.68	0.16
Rubbr	-0.18	1.15	-0.04	-0.67	0.06	1.20	-1.38	-0.35	-4.50	1.06	3.25	0.85	0.08
Trans	-0.16	1.16	0.32	-0.38	-0.31	0.88	-0.96	5.56	-1.46	-1.74	2.31	0.78	0.24
FabPr	-0.40	1.09	0.67	-0.22	-0.02	1.15	-2.22	2.22	-1.38	-0.29	2.85	0.81	0.18
Clths	-0.11	1.07	0.27	-0.37	-0.21	1.17	-0.69	1.26	-2.36	-2.65	6.97	0.81	0.11
Chem	-0.11	1.13	-0.14	0.11	0.15	0.15	-0.87	-0.70	0.58	2.71	0.69	0.81	0.02
Toys	-0.22	1.16	0.73	-0.25	0.25	0.80	-0.86	4.56	-1.26	2.29	3.05	0.67	0.11
Ships	-0.03	1.03	0.47	-0.26	-0.43	1.30	-0.11	3.03	-1.56	-2.58	4.56	0.58	0.14
Soda	0.08	1.18	0.22	-0.32	0.12	0.46	0.33	1.68	-2.32	1.31	2.16	0.64	0.20
Enrgy	0.03	0.94	-1.57	0.65	-0.35	1.90	0.16	2.21	1.88	-3.30	7.28	0.60	0.12
Mines	0.00	0.91	0.51	-0.21	0.19	0.36	0.03	5.06	-1.05	1.71	0.88	0.51	0.08
Smoke	0.28	0.83	0.19	-0.51	-0.00	0.86	1.19	1.53	-2.35	-0.01	3.68	0.43	0.18
Paper	-0.19	1.14	0.13	0.09	0.32	-0.18	-1.23	2.16	0.29	2.55	-0.48	0.77	0.01
Txtls	-0.28	1.02	0.14	-0.49	-0.16	0.74	-1.71	0.73	-3.12	-0.86	2.57	0.78	0.15
Banks	-0.18	1.15	0.17	-0.21	0.09	0.86	-1.27	3.10	-1.28	0.87	3.12	0.80	0.06
Telcm	-0.06	0.79	-0.66	0.16	0.13	0.52	-0.42	-2.91	1.96	1.04	2.12	0.60	0.10
Util	-0.09	0.78	0.01	-0.82	0.10	0.53	-0.64	0.13	-2.75	0.59	1.65	0.63	0.12
RIEst	-1.02	1.00	0.36	-0.38	0.38	0.21	-5.22	0.84	-1.93	4.80	0.83	0.76	0.13
Steel	-0.55	1.16	0.34	-0.40	-0.49	1.42	-2.64	3.92	-1.56	-1.73	3.43	0.68	0.12
Autos	-0.40	1.12	0.64	-0.23	0.20	0.88	-2.06	2.77	-2.13	1.46	3.65	0.65	0.13

*HML* slopes in Table 3. The correlation between the two estimates of the volatility of an industry's sensitivity is 0.51 for the *SMB* slopes and 0.63 for the *HML* slopes. Moreover, the averages of the standard deviations of the conditional loadings in Table 4 (0.13 for *SMB* and 0.17 for *HML*) are almost identical to the averages of the implied standard deviations in Table 3 (0.14 and 0.17). Thus, the conditional regressions track meaningful variation in industry loadings on *SMB* and *HML*, and the magnitude of this variation is similar to that inferred from the rolling regressions.

#### 4. Estimating CAPM and three-factor risk loadings for industries

Tables 3 and 4 say that, for many industries, true sensitivities to CAPM and three-factor risks are quite volatile. Moreover, the variation in true risk loadings, and the implied variation in the cost of equity, are surely larger for individual firms. Appendix B explores the effects of time-varying risk loadings on tests of asset pricing models. Here we ask how a manager who is trying to measure the cost of equity should estimate wandering loadings.

For near-term cash flows, the manager wants a current *CE* and thus current true risk loadings. Conditional or rolling regressions, designed to track wandering risk loadings, are likely candidates. The answer for more distant cash flows depends on the behavior of the true risk loadings. If the true loadings are mean-reverting, full-period constant-slope regressions like those in Table 2 are probably best. At the other extreme, if the true risk loadings follow a random walk, current loadings are the best forecasts of all future loadings, and conditional or rolling regressions may be better than full-period constant-slope regressions.

We compare three approaches to estimating risk loadings for near-term and long-term *CE*'s: (i) full-period estimates of the CAPM and three-factor regressions (2) and (4); (ii) three-, four-, and five-year rolling estimates of (2) and (4); and (iii) full-period estimates of the conditional regression (6). We evaluate the precision of the competing risk loadings, for the purpose of estimating near-term and long-term *CE*'s, by examining their ability to explain industry returns next month and at more distant horizons. The idea is that more precise loadings will produce less disperse forecast errors (Gonedes, 1973).

##### 4.1. One-month forecasts

Part A of Table 5 compares the in-sample fits of the full-period, constant-slope, and conditional regressions for July 1968 to December 1994. The results confirm that the conditioning variables improve the fit of the three-factor regressions. Adding  $\ln(ME)SMB$  and  $\ln(BE/ME)HML$  to the full-period constant-slope regressions raises the average  $R^2$  (across industries) from 0.70 to 0.71. The average mean absolute unexplained return, *MA*, falls from 2.72% in the

Table 5  
Comparisons of techniques for estimating risk loadings

Part A.  $R_M - R_f$ ,  $SMB$ , and  $HML$  are defined in Table 1;  $\ln(ME)$  and  $\ln(BE/ME)$  are defined in Table 4.  $MA$  is the mean absolute value of the sum of the intercept and the residuals from an industry regression;  $\hat{\sigma}(e)$  is the standard error of the residuals.  $R^2$ ,  $MA$ , and  $\hat{\sigma}(e)$  are adjusted for degrees of freedom. The  $t$ -statistics are ratios of coefficients to their standard errors, except for the market slope  $b$  where the test is that the true slope is one. The statistics shown are averages across industries.

Part B. The one-factor and three-factor regressions are estimated for each industry for each month beginning in 6/68, using a rolling window of three, four, or five years of past monthly returns. The in-sample regression coefficients, and the values of the explanatory variables for the month following the in-sample period, are used to make conditional out-of-sample forecasts of industry returns. The table shows averages, across industries, of the mean ( $M$ ), mean absolute ( $MA$ ), and the standard deviation ( $\hat{\sigma}(e)$ ) of the out-of-sample forecast errors for different methods of estimating the in-sample regressions. In the first three columns (OLS, intercept), the forecasts use simple OLS regression coefficients. In the second three columns (OLS, no intercept), the OLS intercept is dropped in the forecasts. In the next three columns (shrunk, no intercept), the forecasts use regression slopes that are shrunk to correct for sampling error. The Bayes shrinkage method is described in Appendix C.

Part A: Full-period constant-slope and conditional regressions: Coefficients,  $t$ -statistics (in parentheses), and summary statistics, averaged across 48 industries, 7/68–12/94

$$R_i - R_f = a_i + b_i[R_M - R_f] + [s_{i1} + s_{i2}\ln(ME)_i] SMB + [h_{i1} + h_{i2}\ln(BE/ME)_i] HML + e_i$$

$b$	$s_1$	$s_2$	$h_1$	$h_2$	$R^2$	$a$	$MA$	$\hat{\sigma}(e)$
1.11 (2.35)					0.64	-0.02 (-0.07)	2.96	3.83
1.04 (0.92)	0.38 (5.35)		0.01 (0.15)		0.70	-0.04 (-0.24)	2.72	3.50
1.03 (0.81)	0.21 (2.25)	-0.25 (-1.59)	-0.03 (-0.08)	0.68 (2.56)	0.71	-0.04 (-0.21)	2.68	3.43

Part B: One-month-ahead forecast errors from rolling regressions, averaged across 48 industries, 7/68–12/94

	OLS						Shrunk, no intercept		
	Intercept			No intercept			$M$	$MA$	$\hat{\sigma}(e)$
	$M$	$MA$	$\hat{\sigma}(e)$	$M$	$MA$	$\hat{\sigma}(e)$			
$R_i - R_f = a_i + b_i[R_M - R_f] + \varepsilon_i$									
3-year rolling	-0.02	3.00	3.94	0.02	2.97	3.87	0.02	2.95	3.84
4-year rolling	-0.05	2.99	3.92	0.00	2.97	3.87	0.00	2.95	3.85
5-year rolling	-0.05	2.98	3.90	0.00	2.96	3.86	0.00	2.96	3.85
$R_i - R_f = a_i + b_i[R_M - R_f] + s_i SMB + h_i HML + \varepsilon_i$									
3-year rolling	-0.00	2.81	3.67	-0.04	2.78	3.62	-0.04	2.72	3.52
4-year rolling	-0.00	2.79	3.63	-0.03	2.77	3.59	-0.04	2.72	3.52
5-year rolling	0.00	2.76	3.60	-0.03	2.75	3.57	-0.03	2.72	3.52

constant-slope regressions to 2.68% in the conditional regressions. The average residual standard deviation,  $\hat{\sigma}(e)$ , drops from 3.50% to 3.43%. These improvements are small, but they are consistent. The conditional  $MA$  is lower than the constant-slope  $MA$  for 35 of 48 industries; the conditional  $\hat{\sigma}(e)$  is lower than the constant-slope  $\hat{\sigma}(e)$  for 39 industries.

In-sample regression fits (not shown) suggest that rolling-regression estimates of risk loadings are a bit better than full-period or conditional estimates. For the 1968–1994 period, the average of the adjusted  $R^2$  for rolling CAPM regressions estimated monthly with five years of returns is 0.65 versus 0.64 for full-period estimates. Similarly, the average  $R^2$  for rolling five-year three-factor regressions, 0.72, is slightly higher than the averages for the full-period, constant-slope, or conditional regressions, 0.70 and 0.71. Because the true CAPM and three-factor slopes vary through time, however, in-sample fits probably exaggerate the precision of the rolling regressions for our problem – estimating risk loadings for future  $CE$ 's. For this purpose, out-of-sample forecast errors give better perspective on the precision of rolling-regression slopes.

We start with one-month forecasts. To construct them, we estimate CAPM and three-factor regressions each month for each industry using rolling windows of three, four, and five years of past returns. We then use the in-sample regression coefficients and next month's (out-of-sample) explanatory returns to generate out-of-sample forecasts. Because the slope estimates from the rolling regressions are so imprecise, we add a Bayesian wrinkle. More extreme estimates are likely to have more error, so in principle we can improve the rolling-regression slopes by shrinking them toward a grand mean. We use the Bayes shrinkage method of Blattberg and George (1991). The details are in Appendix C.

We consider several versions of the rolling regressions to judge (i) whether out-of-sample forecasts improve when estimation-period intercepts are dropped, (ii) whether Bayes shrinkage of the slopes helps, (iii) whether three-factor regressions forecast better than one-factor regressions, and (iv) whether longer estimation periods improve the out-of-sample forecasts. On the first three questions, the results in part B of Table 5 are clear, but not always overwhelming. (i) In every case, and for both the one-factor and three-factor regressions, suppressing estimation-period intercepts (and so imposing CAPM or three-factor asset pricing) produces forecast errors with less dispersion. The improvements, however, are small (1%–2%). (ii) If we suppress the intercepts, the least disperse forecast errors are obtained when the regression slopes are shrunk to correct for estimation error. But again, the improvements are small. (iii) In every comparison, three-factor regressions produce less disperse forecast errors than one-factor regressions. Here the improvements are larger; the averages of the mean absolute and standard deviations of the one-factor forecast errors are about 10% greater than those of the three-factor regressions.

The one-month forecasts are clean evidence that the three-factor rolling regressions capture return variation missed by the one-factor regressions. The

asset pricing evidence from the rolling regressions is, however, weak. In particular, the fact that out-of-sample forecast errors are less disperse when the regression intercepts (one-factor or three-factor) are dropped is not much evidence that the true intercepts are zero. The average standard errors of the intercepts in the in-sample regressions are large (for example, 0.46 for the five-year three-factor regressions). If the true intercepts are small relative to their estimation errors, suppressing the intercepts is likely to improve the out-of-sample forecasts.

The most important result from the rolling regressions is that, for both the CAPM and the three-factor model, forecast quality is insensitive to the length of the regression estimation period. Focusing on the shrunk-no-intercept regressions, which produce the best forecasts, the averages of the mean absolute and standard deviations of the forecast errors are remarkably similar for three-, four-, and five-year estimation periods. Six- to ten-year estimation periods (not shown) produce forecast errors like those of three- to five-year estimation periods. In fact, Table 5 suggests that the forecast power of the regressions does not change if we lengthen the estimation period to the full sample. The averages of  $MA$  and  $\hat{\sigma}(e)$  for the rolling three-factor regressions, 2.72 and 3.52, are almost identical to those for the full-period constant-slope regressions, 2.72 and 3.50. Similarly, the average  $MA$  for the full-period CAPM regressions (2.96) basically matches the rolling-regression estimates (2.95 and 2.96), while the average  $\hat{\sigma}(e)$  for the full-period regressions (3.83) is slightly lower than the rolling-regression estimates (3.84 and 3.85).

The bottom line for the CAPM is that, on average, full-period estimates of current industry betas are no better or worse than estimates from three-, four-, and five-year rolling regressions. The insensitivity of forecast quality to the regression estimation period says that noise in the forecasts, caused by increased smoothing of variation in the true betas, just about offsets the increase in precision obtained by extending the estimation period. These results for industries are like those in Gonedes (1973) for individual firms.

The implications of Table 5 for estimates of current three-factor risk loadings are similar. For the typical industry, estimates of three-factor loadings from full-period constant-slope regressions produce forecasts of returns one month ahead that are as accurate as those from rolling regressions, and only slightly less accurate than the forecasts from conditional regressions. The next section asks which approach is best for estimating risk loadings for longer horizons.

#### 4.2. *Forecasts for distant horizons*

Table 6 examines forecasts of monthly returns up to five years ahead. For the CAPM, the table shows that rolling-regression market slopes are about as good as full-period slopes for forecasts one month ahead, but the full-period slopes dominate at longer forecast horizons. This suggests that the typical industry's

CAPM beta is mean-reverting. The rolling regressions track risk loadings that wander through time. If an industry's true beta is mean-reverting, deviations from the long-term mean are temporary, and estimates from the full-period constant-slope regressions provide better estimates of distant betas. Since, on average, the full-period estimates are as good as the rolling-regression estimates at short horizons, the prescription for the CAPM is simple. Full-period market slopes are typically good choices for estimating CAPM betas and *CE*'s for all future periods.

Table 6

Summary statistics for forecast errors from conditional and rolling regressions and for residuals from full-period constant-slope regressions

$R_M - R_f$ , *SMB*, and *HML* are defined in Table 1;  $\ln(ME)$  and  $\ln(BE/ME)$  are defined in Table 4. The conditional regression forecasts for month  $t + i$  ( $i = 1, 12, 24, 36, 48$ , and  $60$ ) combine slopes ( $b$ ,  $s_1$ ,  $s_2$ ,  $h_1$ , and  $h_2$ ) estimated from 7/68 to 12/94 with values of  $\ln(ME)$  and  $\ln(BE/ME)$  for month  $t$ , and explanatory returns ( $R_M - R_f$ , *SMB*, and *HML*) for month  $t + i$ . The rolling-regression forecasts for month  $t + i$  use the slopes for months  $t - n$  ( $n = 36, 48$ , and  $60$ ) to  $t$  and market, *SMB*, and *HML* returns for month  $t + i$ . The slopes in the rolling regressions are shrunk using the Bayes shrinkage method in Appendix C. The one-month-ahead results for the conditional and rolling regressions summarize monthly forecast errors for 7/68 to 12/94. The one-year-ahead results summarize monthly forecast errors for 6/69 to 12/94, and the five-year-ahead results are for 6/73 to 12/94. The full-period constant-slope regression results summarize residuals from regressions estimated over the same periods as the conditional and rolling-regression forecast errors. For example, the summary statistics in the one-year column describe the residuals from constant-slope regressions estimated using monthly observations for 6/69 to 12/94. Since they summarize regression residuals, we adjust for degrees of freedom when calculating the mean absolute and standard deviation measures for the constant-slope regressions and for the one-month-ahead conditional regressions. The results are averages across the 48 industries.

One-factor regressions:  $R_i - R_f = a_i + b_i [R_M - R_f] + e_i$

	Forecast horizon					
	1 month	1 year	2 years	3 years	4 years	5 years
<i>Average mean absolute forecast error (or residual plus intercept)</i>						
3-year rolling	2.95	2.99	2.99	2.99	3.02	3.02
4-year rolling	2.95	2.98	2.98	2.99	3.01	3.02
5-year rolling	2.96	2.98	2.99	2.99	3.01	3.02
Full period	2.96	2.97	2.96	2.95	2.97	2.97
<i>Average standard deviation of forecast errors (or residuals)</i>						
3-year rolling	3.84	3.89	3.88	3.89	3.92	3.93
4-year rolling	3.85	3.87	3.88	3.89	3.92	3.93
5-year rolling	3.85	3.88	3.89	3.88	3.91	3.93
Full period	3.83	3.84	3.83	3.82	3.84	3.84

Table 6 (continued)

## Three-factor regressions

Rolling and full-period constant-slope:  $R_i - R_f = a_i + b_i[R_M - R_f] + s_i SMB + h_i HML + e_i$ Conditional:  $R_i - R_f = a_i + b_i[R_M - R_f] + [s_{i1} + s_{i2} \ln(ME)_i] SMB + [h_{i1} + h_{i2} \ln(BE/ME)_i] HML + e_i$ 

## Forecast horizon

	1 month	1 year	2 years	3 years	4 years	5 years
<i>Average mean absolute forecast error (or residual plus intercept)</i>						
Conditional	2.68	2.67	2.71	2.72	2.74	2.77
3-year rolling	2.72	2.78	2.80	2.84	2.86	2.89
4-year rolling	2.72	2.77	2.80	2.83	2.85	2.89
5-year rolling	2.72	2.77	2.80	2.82	2.85	2.88
Full period	2.72	2.73	2.74	2.74	2.75	2.77
<i>Average standard deviation of forecast error (or residual)</i>						
Conditional	3.43	3.45	3.49	3.50	3.53	3.58
3-year rolling	3.52	3.60	3.64	3.68	3.72	3.75
4-year rolling	3.52	3.59	3.64	3.67	3.70	3.74
5-year rolling	3.52	3.59	3.63	3.66	3.69	3.73
Full period	3.50	3.50	3.50	3.51	3.52	3.54

The three-factor regression results in Table 6 also suggest some mean reversion in industry loadings on  $R_M - R_f$ ,  $SMB$ , and  $HML$ . The performance of the conditional and rolling three-factor regressions deteriorates for forecasts further into the future. Since the quality of the full-period constant-slope forecasts also falls a bit, part of the deterioration of the conditional- and rolling-regression forecasts occurs simply because the sample months differ across forecast horizons. But the full-period constant-slope forecasts do not deteriorate as much as the conditional- and rolling-regression forecasts, which suggests some mean reversion in risk loadings. Consistent with this conclusion, beyond two years the forecasts from the full-period constant-slope regressions are about as good as (but no better than) those from the conditional regressions (which are always a bit better than the forecasts from the rolling regressions).

Table 6 also suggests, however, that reversion to constant means is not a universal property of industry three-factor risk loadings. With true mean reversion, the constant-slope regressions should provide better forecasts than the conditional regressions at distant horizons. At least for horizons out to five years, they do not. A likely explanation is that the conditional three-factor regressions capture some permanent changes in risk loadings that are missed by the full-period constant-slope regressions.

### 4.3. Complications

The message from Tables 5 and 6 about the choice of betas to be used in estimates of CAPM costs of equity (*CE*'s) is relatively simple. For the typical industry, the full-period market slope is a good choice for estimating both near-term and distant betas. (We leave open, of course, the possibility that for some industries shifting demand and supply conditions produce nonstationarity in true betas that would favor the rolling regressions for estimates of both near-term and distant *CE*'s.)

The implications of Tables 5 and 6 for estimating three-factor *CE*'s are more complicated. The tables suggest that for the typical industry the time-varying slopes from the conditional three-factor regression (6) should be used to estimate three-factor *CE*'s for near horizons (up to two years). At longer horizons, estimates from the constant-slope regression (4) are as precise as those from the conditional regression. Statistical or economic evidence of mean reversion in an industry's three-factor slopes would push the choice for longer horizons toward constant-slope estimates. Evidence of permanent shifts in demand or supply conditions that produce nonstationarity in true three-factor slopes would favor the conditional (or rolling) regressions for estimates of both near-term and distant *CE*'s.

Even for near horizons other considerations might make three-factor *CE*'s from full-period constant-slope regressions better than conditional estimates. Table 6 says that, for forecasts one month ahead, the full-period constant-slope regressions are only a bit worse than the conditional regressions. This suggests that using full-period constant-slope *CE*'s to value near-term cash flows will not produce substantially less precise estimates of value than conditional-regression *CE*'s. Moreover, spurious variability in conditional *CE*'s can create startup and shutdown costs that are avoided with estimates from full-period constant-slope regressions. Such costs favor full-period constant-slope *CE*'s over conditional *CE*'s.

The conditional regressions have another weakness if used at the firm level. Managers have at least partial control of a firm's size (*ME*) and book-to-market-equity (*BE/ME*). If a firm uses a conditional regression to track its wandering risk loadings, management could change the estimated cost of equity by changing the firm's *ME* or *BE/ME*. Such gaming of the conditional regression would create more noise in *CE* estimates and bias the projects that are accepted. To avoid such problems, firms might be better off using full-period constant-slope *CE*'s for capital budgeting.

## 5. Industry costs of equity

Tables 5 and 6 do not identify clear winners among alternative slope estimates for CAPM and three-factor *CE*'s. For long horizons, CAPM betas from

full-period regressions produce return forecasts that are a bit more accurate than those from rolling regressions, but for all horizons the advantages of one approach over the other are small. Similarly, for near horizons, three-factor conditional regressions produce slightly more accurate return forecasts than full-period constant-slope regressions, but the differences between the two approaches, and their advantages over rolling regressions, are always small. Although the statistical evidence from the forecasts does not clearly identify the best approach to estimating risk loadings, we show next that the choice is of some consequence. Competing approaches often produce much different *CE*'s, especially for the three-factor model.

### 5.1. Comparisons of alternative CAPM and three-factor *CE*'s

Table 7 shows two estimates of the risk premiums (expected returns in excess of the risk-free rate) in CAPM *CE*'s and three estimates of the risk premiums in three-factor *CE*'s for each of the 48 industries. Two estimates use the slopes from the full-period constant-slope CAPM and three-factor regressions, (2) and (4). Two *CE*'s use shrunk CAPM and three-factor slopes estimated on the five years of monthly returns ending in December 1994. The fifth estimate combines slopes from the conditional regression (6) with values of the conditioning variables,  $\ln(ME)$  and  $\ln(BE/ME)$ , for December 1994.

For the CAPM, most of the differences between the *CE*'s from the full-period constant-slope regressions and the end-of-period rolling regressions are modest. The two CAPM *CE*'s differ by more than 1% per year for only 11 of 48 industries; the two *CE*'s never differ by more than 2%. Differences among the alternative three-factor *CE*'s tend to be larger. The full-period estimate from the constant-slope regression (4) and the end-of-period estimate from the conditional regression (6) differ by more than 1% per year for 25 of 48 industries and by more than 2% for eight industries. Thus, for many industries, the essentially arbitrary choice between constant-slope and conditional-regression three-factor *CE*'s can lead to substantially different valuations of investment projects.

The differences in three-factor *CE*'s are driven by differences in the estimates of *SMB* and *HML* slopes. Health Care, Personal Services, and Gold are examples. Their end-of-sample values of  $\ln(ME)$  are below their full-period averages. As a result, their conditional *SMB* slopes for December 1994 are above the estimates from full-period constant-slope regressions, and their conditional *CE*'s are above their constant-slope *CE*'s. Computers, Machinery, and Coal illustrate the opposite case. Firms in these industries are relatively large in December 1994, their conditional *SMB* slopes are below their full-period slopes, and their conditional three-factor *CE*'s are below their full-period *CE*'s.

For many industries, the conditional *SMB* and *HML* slopes for December 1994 both differ a lot from their full-period counterparts. Business Services and

Table 7

Regression slopes and risk premiums in CAPM and three-factor costs of equity

The industries are defined in Appendix A. The risk premiums ( $CE$ 's) in the cost of equity are obtained by substituting the regression slopes in the table and the average monthly  $R_M - R_t$ ,  $SMB$ , and  $HML$  returns for 7/63 to 12/94 (Table 1) into (2) and (4) and then multiplying by 12. The full-period  $CE$ 's use the slopes from (2) and (4) for 7/63 to 12/94 in Table 2. The five-year  $CE$ 's use shrunk CAPM and three-factor slopes (see Appendix C) estimated on the five years of monthly returns ending in 12/94. The conditional  $CE$ 's use the slopes from (6) for 7/65 to 12/94. For these estimates,  $s = s_1 + s_2 \ln(ME)$  and  $h = h_1 + h_2 \ln(BE/ME)$ , where  $s_1, s_2, h_1,$  and  $h_2$  are slopes from Table 4, and  $\ln(ME)$  and  $\ln(BE/ME)$  are for 12/94.

Industry	CAPM						Three-factor						Conditional						Five-year	
	Full-period		Five-year		Full-period		Full-period		Five-year		Full-period		Five-year		Full-period		Five-year			
	$CE$	$b$	$CE$	$b$	$CE$	$b$	$CE$	$b$	$CE$	$b$	$CE$	$b$	$CE$	$b$	$CE$	$b$	$s$	$h$	$CE$	$CE$
Drugs	4.71	0.92	5.35	1.05	0.09	0.84	0.09	0.84	-0.25	-0.63	2.60	0.86	2.60	0.86	2.60	0.86	-0.26	-0.18	-0.46	-0.46
MedEq	5.99	1.17	6.09	1.19	2.64	0.99	2.64	0.99	0.26	-0.60	4.73	1.01	4.73	1.01	4.73	1.01	0.41	-0.32	1.25	1.25
Hlth	7.95	1.56	6.77	1.33	6.14	1.24	6.14	1.24	0.93	-0.59	4.97	1.20	4.97	1.20	4.97	1.20	0.62	-0.59	4.29	4.29
Comps	5.29	1.04	5.47	1.07	2.49	0.90	2.49	0.90	0.17	-0.49	4.34	0.90	4.34	0.90	4.34	0.90	0.54	-0.37	5.66	5.66
Chips	7.04	1.38	6.14	1.20	6.01	1.15	6.01	1.15	0.69	-0.39	7.07	1.15	7.07	1.15	7.07	1.15	0.56	-0.11	9.23	9.23
BusSv	6.83	1.34	6.27	1.23	6.51	1.13	6.51	1.13	0.72	-0.29	4.21	1.11	4.21	1.11	4.21	1.11	0.26	-0.43	4.99	4.99
LabEq	6.59	1.29	5.92	1.16	5.80	1.13	5.80	1.13	0.49	-0.29	6.31	1.12	6.31	1.12	6.31	1.12	0.55	-0.22	6.56	6.56
Hshld	4.96	0.97	5.24	1.03	3.19	0.91	3.19	0.91	0.00	-0.27	3.97	0.93	3.97	0.93	3.97	0.93	0.01	-0.15	5.12	5.12
Meals	6.75	1.32	6.73	1.32	6.81	1.12	6.81	1.12	0.74	-0.24	7.23	1.12	7.23	1.12	7.23	1.12	0.56	-0.05	8.24	8.24
Beer	4.69	0.92	5.16	1.01	2.99	0.90	2.99	0.90	-0.13	-0.22	1.78	0.92	1.78	0.92	1.78	0.92	-0.30	-0.36	1.98	1.98
PerSv	6.40	1.25	5.20	1.02	7.26	1.00	7.26	1.00	1.00	-0.20	5.92	0.99	5.92	0.99	5.92	0.99	0.66	-0.23	5.10	5.10
Cnstr	6.52	1.28	6.42	1.26	6.42	1.21	6.42	1.21	0.21	-0.09	9.69	1.23	9.69	1.23	9.69	1.23	0.77	0.17	8.68	8.68
Rtail	5.68	1.11	5.96	1.17	5.88	1.04	5.88	1.04	0.27	-0.06	5.65	1.05	5.65	1.05	5.65	1.05	0.28	-0.12	4.51	4.51
Fun	6.91	1.35	6.04	1.18	8.43	1.17	8.43	1.17	0.83	-0.04	7.68	1.17	7.68	1.17	7.68	1.17	0.73	-0.12	9.32	9.32
Food	4.44	0.87	4.97	0.97	4.09	0.88	4.09	0.88	-0.07	-0.03	2.48	0.87	2.48	0.87	2.48	0.87	-0.30	-0.18	0.90	0.90
Agric	5.11	1.00	4.98	0.97	6.51	0.85	6.51	0.85	0.71	-0.02	6.03	0.84	6.03	0.84	6.03	0.84	0.60	-0.04	5.65	5.65
Mach	5.93	1.16	5.49	1.08	6.46	1.11	6.46	1.11	0.25	-0.00	7.54	1.12	7.54	1.12	7.54	1.12	0.45	0.07	9.09	9.09
Books	5.98	1.17	5.59	1.10	6.96	1.08	6.96	1.08	0.45	0.00	7.08	1.06	7.08	1.06	7.08	1.06	0.23	0.17	7.23	7.23
Aero	6.43	1.26	4.98	0.97	7.54	1.15	7.54	1.15	0.51	0.00	7.57	1.15	7.57	1.15	7.57	1.15	0.20	0.20	5.25	5.25

Coal	4.90	0.96	4.41	0.86	5.97	0.86	0.46	0.01	7.42	0.87	0.71	0.13	5.39
Guns	5.29	1.04	4.01	0.79	6.25	0.95	0.41	0.01	6.43	0.93	0.21	0.19	4.24
Whlsl	5.90	1.15	5.33	1.04	7.52	1.01	0.71	0.01	7.03	0.99	0.69	-0.04	5.42
Fin	5.95	1.16	6.99	1.37	6.72	1.11	0.30	0.02	6.95	1.10	0.35	0.03	7.58
ElcEq	5.86	1.15	5.91	1.16	5.98	1.15	-0.00	0.02	6.05	1.16	-0.04	0.05	6.43
Boxes	5.24	1.03	3.94	0.77	5.77	0.99	0.17	0.02	5.70	1.00	0.17	0.01	4.14
BldMt	5.76	1.13	5.19	1.02	6.40	1.11	0.15	0.05	6.74	1.12	0.16	0.10	5.93
Insur	5.14	1.01	5.97	1.17	5.72	1.00	0.09	0.06	6.23	0.99	0.08	0.17	6.44
Gold	3.98	0.78	3.33	0.65	5.35	0.71	0.40	0.08	4.29	0.68	0.25	0.00	6.13
Misc	6.43	1.26	5.81	1.14	9.56	1.03	1.19	0.08	8.13	1.03	1.26	-0.22	8.92
Trans	6.17	1.21	5.92	1.16	7.39	1.16	0.30	0.09	7.30	1.16	0.38	0.03	7.78
Rubbr	6.16	1.21	6.37	1.25	7.78	1.12	0.49	0.09	6.73	1.15	0.37	-0.06	7.11
FabPr	6.71	1.31	5.81	1.14	9.69	1.11	1.10	0.09	8.82	1.09	0.92	0.05	5.24
Clths	6.33	1.24	6.80	1.33	8.85	1.09	0.83	0.11	6.37	1.07	0.68	-0.24	6.28
Chem	5.57	1.09	5.44	1.06	6.58	1.13	-0.03	0.17	6.60	1.13	-0.05	0.18	6.13
Toys	6.83	1.34	5.11	1.00	10.01	1.17	0.97	0.17	7.75	1.16	0.77	-0.12	5.43
Ships	6.07	1.19	4.86	0.95	8.63	1.09	0.66	0.17	5.85	1.03	0.69	-0.31	8.99
Soda	6.35	1.24	6.04	1.18	8.46	1.19	0.44	0.18	7.15	1.18	0.34	0.00	7.04
Enrgy	4.32	0.85	3.22	0.63	4.93	0.96	-0.35	0.21	6.22	0.94	-0.49	0.55	4.80
Mines	4.99	0.98	3.38	0.66	7.65	0.91	0.53	0.23	6.96	0.91	0.66	0.03	7.85
Smoke	4.08	0.80	5.01	0.98	5.56	0.86	-0.04	0.24	4.72	0.83	-0.27	0.25	2.02
Paper	5.68	1.11	5.22	1.02	7.78	1.14	0.16	0.27	7.42	1.14	0.12	0.22	6.67
Txils	5.71	1.12	6.00	1.17	9.18	1.03	0.71	0.30	7.60	1.02	0.79	-0.03	9.12
Banks	5.55	1.09	6.53	1.28	8.08	1.13	0.13	0.35	9.19	1.15	0.12	0.54	9.88
Telec	3.39	0.66	4.49	0.88	5.17	0.79	-0.23	0.35	3.74	0.79	-0.38	0.17	5.34
Util	3.39	0.66	3.25	0.64	5.41	0.79	-0.20	0.38	6.74	0.78	-0.06	0.54	3.79
REst	5.99	1.17	5.19	1.02	11.16	1.01	1.18	0.40	11.55	1.00	1.27	0.43	9.00
Steel	5.94	1.16	5.43	1.06	9.61	1.17	0.40	0.43	9.12	1.16	0.57	0.25	9.13
Autos	5.13	1.01	5.24	1.03	9.39	1.10	0.17	0.60	5.12	1.12	0.37	-0.33	11.74

Construction are extreme examples. The December 1994 slopes on *SMB* and *HML* from the conditional regression for Business Services (0.26 and  $-0.43$ ) are much lower than the full-period slopes (0.72 and  $-0.29$ ), so the conditional three-factor *CE* (4.21%) is much lower than the constant-slope *CE* (6.51%). Similarly, the December 1994 conditional *SMB* and *HML* slopes for Construction (0.77 and 0.17) are much higher than the full-period slopes (0.21 and  $-0.09$ ), so the December 1994 conditional *CE* (9.69%) is much higher than the full-period *CE* (6.42%).

Finally, there are also large differences between the three-factor *CE*'s for December 1994 from the conditional and five-year rolling regressions. The estimates differ by more than 1% per year for 27 industries and by more than 2% for 13 industries. For Drugs, Medical Equipment, Fabricated Products, Shipping, and Autos, the conditional and rolling-regression *CE*'s for December 1994 differ by more than 3%. The forecast tests in Tables 5 and 6 suggest that the conditional regressions provide slightly more precise estimates of near-term *CE*'s than the rolling regressions, but there is much uncertainty about which is better. Unfortunately, Table 7 says the choice can have large consequences.

### 5.2. CAPM and three-factor *CE*'s

As one might expect, there are also large differences between CAPM and three-factor *CE*'s. The full-period CAPM and three-factor *CE*'s differ by more than 2% for 17 industries and by more than 3% for eight industries. The five-year CAPM and three-factor *CE*'s differ by more than 2% for 19 industries and by more than 3% for 15 industries. For many industries, the choice of a CAPM or three-factor cost of equity will have a large impact on the valuation of investments.

The differences between the three-factor and CAPM *CE*'s are largely determined by the *SMB* and *HML* slopes in the three-factor regressions. Some industries have *SMB* and *HML* slopes close to zero, so their CAPM and three-factor *CE*'s are similar. Focusing on the full-period constant-slope regressions, this group includes Food, Machinery, Electrical Equipment, Boxes, Building Materials, and Insurance. Other industries have similar full-period CAPM and three-factor *CE*'s because their *SMB* and *HML* slopes offset. This group includes Business Services, Meals (Restaurant Services), Construction, and Retailers.

More numerous and interesting are the industries for which three-factor and CAPM *CE*'s differ a lot. For example, the health industries (Health Services, Medical Equipment, and Drugs) and the high-tech industries (Computers, Chips, and Laboratory Equipment) have lower full-period three-factor *CE*'s, largely due to strong negative loadings on *HML*. The three-factor model identifies these as industries with strong growth prospects during the sample

period and rewards them with three-factor costs of equity that are lower than their CAPM  $CE$ 's. On the other hand, many industries have full-period three-factor  $CE$ 's that are at least 2% higher than their CAPM  $CE$ 's. Mining, Textiles, Banking, Real Estate, Steel, and Autos are examples. The three-factor model assigns high costs of equity to these industries because their returns covary with the returns on small stocks (they have large positive slopes on  $SMB$ ) and because they behave like distressed stocks (they have large positive slopes on  $HML$ ).

Finally, there is more cross-industry variation in three-factor  $CE$ 's than in CAPM  $CE$ 's. For example, the range of the full-period three-factor estimates, 0.09% to 11.16%, dwarfs that of the full-period CAPM estimates, 3.39% to 7.95%. Part of the dispersion of the three-factor  $CE$ 's is caused by estimation error in  $SMB$  and  $HML$  slopes. For the purpose of estimating an industry's average  $CE$  over the sample period, however, the full-period  $SMB$  and  $HML$  slopes are rather precise; their average standard errors are 0.07. Thus, much of the higher dispersion of the three-factor  $CE$ 's reflects true differences in the risk loadings of the two asset pricing models.

## 6. Estimating factor risk premiums

Imprecise risk loadings imply economically important uncertainty about  $CE$ 's. But the risk loadings are in fact a small part of the  $CE$  estimation problem. Uncertainty about the market,  $SMB$ , and  $HML$  premiums in the CAPM and the three-factor model is more important.

The  $CE$ 's in Table 7 estimate the annual factor risk premiums by annualizing the average monthly premiums for July 1963 to December 1994 in Table 1. The annualized standard errors (12 times the monthly standard errors) of the market,  $SMB$ , and  $HML$  premiums are 2.71%, 1.77%, and 1.58%. These large standard errors are, of course, examples of the general imprecision of estimates of expected stock returns (Merton, 1980). They imply that, even if we knew true risk loadings, estimates of CAPM and three-factor  $CE$ 's would be unavoidably imprecise.

For example, the annualized average market premium for 1963–1994 is 5.16% per year. The standard error of the premium, 2.71%, implies that the one-standard-error bounds on the CAPM  $CE$  of a project known to have a true beta equal to 1.0 are 2.45% to 7.87% per year. The two-standard-error bounds are  $-0.26\%$  to 10.58%. These estimates clearly imply extreme imprecision in the values assigned to investment projects.

At the risk of beating a horse already dead, there is another problem in estimates of factor risk premiums. There is evidence that the expected market premium,  $E(R_M) - R_f$ , varies through time. (See Fama and French, 1989, and the references therein.) Predictable variation in  $E(R_M) - R_f$  and in the expected

*SMB* and *HML* returns<sup>2</sup> should in principle be incorporated in *CE* estimates (Brennan, 1995). There is, however, much controversy about the extent of the variation in the expected market premium, due (once again) to the imprecision of the relevant parameter estimates. As a result, the evidence is consistent with a wide range of estimates of the time variation in the expected premium. Again, different estimates can produce big differences in estimated *CE*'s and in the values assigned to investment projects.

## 7. Standard errors for *CE* estimates

Though the preceding discussion gives strong clues, it is interesting to examine more formally how uncertainty about risk loadings combines with uncertainty about factor risk premiums to determine the overall imprecision of *CE* estimates.

The error in the estimate of an industry's CAPM *CE* is

$$\begin{aligned} e(\text{CAPM } CE) &= \hat{b} \overline{R_M - R_f} - b E(R_M - R_f) \\ &= (b + u_b) [E(R_M - R_f) + u_M] - b E(R_M - R_f) \\ &= u_b E(R_M - R_f) + u_M b + u_b u_M, \end{aligned}$$

where  $\overline{R_M - R_f}$  is the estimated market premium,  $u_M$  is the error in this estimate,  $\hat{b}$  is the estimated market slope, and  $u_b$  is the error in the slope. If the joint distribution of security returns is multivariate normal,  $u_M$  and  $u_b$  are uncorrelated, and the standard error of the *CE* is

$$\begin{aligned} se(\text{CAPM } CE) &= [E(R_M - R_f)^2 \text{var}(u_b) + b^2 \text{var}(u_M) \\ &\quad + \text{var}(u_b) \text{var}(u_M)]^{1/2}. \end{aligned} \quad (7)$$

Similarly, the standard error of a three-factor *CE* depends on the covariance matrix of the estimation errors of the factor loadings and the covariance matrix of the factor risk premiums. (See Table 8.)

Table 8 reports averages (across industries) of the standard errors of CAPM and three-factor *CE*'s under a variety of assumptions about the precision of estimates of factor risk loadings and risk premiums. At one extreme, the table shows that if there is no uncertainty about the market premium, and if the CAPM betas of industries are constant, then estimates of CAPM *CE*'s are precise. In particular, using the full-period CAPM betas and ignoring uncertainty about the market premium, the average standard error of CAPM industry

<sup>2</sup>Without showing the details, we can report that *SMB* and especially *HML* seem less predictable than the market premium, at least with dividend yields and the common interest rate variables that seem to forecast the market premium.

Table 8

Averages across industries of standard errors of CAPM and three-factor CE's

The standard error for an industry's CAPM CE is

$$se(CAPM CE) = [(\overline{R_M - R_f})^2 \text{var}(u_b) + \hat{b}^2 \text{var}(u_M) + \text{var}(u_b) \text{var}(u_M)]^{1/2},$$

where  $\overline{R_M - R_f}$  is the annualized average market premium for 7/63 to 12/94,  $\text{var}(u_M)$  is the annualized variance of the average market premium,  $\hat{b}$  is the market slope from the industry's full-period CAPM regression, and  $\text{var}(u_b)$  is the variance of the error of the slope. The standard error for a three-factor CE is

$$se(3\text{-factor CE}) = [Prem' \text{var}(u_\gamma) Prem + \gamma' \text{var}(u_{Prem}) \gamma + I' [\text{var}(u_{Prem}) \square \text{var}(u_\gamma)] I]^{1/2},$$

where  $Prem = [\overline{R_M - R_f} \overline{SMB} \overline{HML}]'$  is a vector of the annualized average market, *SMB*, and *HML* premiums for 7/63 to 12/94,  $\text{var}(u_{Prem})$  is the annualized covariance matrix of the premiums,  $\gamma$  is the vector of the market, *SMB*, and *HML* slopes from the industry's full-period regression,  $\text{var}(u_\gamma)$  is the covariance matrix of the sampling errors in the slopes,  $I$  is a vector of 1's, and  $a \square b$  means each element of  $a$  is multiplied by the corresponding element of  $b$ .

We consider two measures of  $\text{var}(u_M)$  or  $\text{var}(u_{Prem})$ . The first (rows 1 and 2) assumes the expected premiums are measured perfectly, so  $\text{var}(u_M)$  and  $\text{var}(u_{Prem})$  are zero. The second (rows 3–5) uses the variance or the covariance matrix of the historical average premiums. We also consider three estimators of  $\text{var}(u_b)$  or  $\text{var}(u_\gamma)$  for each industry. The first (row 3) assumes the slopes are measured perfectly, so  $\text{var}(u_b)$  and  $\text{var}(u_\gamma)$  are zero. The second (rows 1 and 4) uses the variance of beta, or the covariance matrix of the slopes, from full-period regressions. The third (rows 2 and 5) uses the average variance of beta, or the average covariance matrix of the slopes, from each industry's three-year rolling CAPM or three-factor regressions.

	CAPM	Three-factor
<i>No error in risk premiums</i>		
Covariance matrix of slopes from:		
(1) Full-period regressions	0.23	0.54
(2) Rolling 3-year regressions	0.78	1.99
<i>Risk premiums estimated with error</i>		
(3) No error in slopes	3.01	3.17
Covariance matrix of slopes from:		
(4) Full-period regressions	3.03	3.23
(5) Rolling 3-year regressions	3.15	3.85

CE's is only 0.23% per year. Still ignoring uncertainty about the market premium, but using three-year rolling regressions to allow for time-varying true betas, the average standard error of CAPM CE's rises from 0.23% to 0.78% per year. Even if the market premium is known, wandering betas in themselves produce substantial uncertainty about CE's, and thus about estimates of project values.

Uncertainty about CAPM  $CE$ 's due to imprecise beta estimates is, however, small relative to the problem caused by the imprecision of the market risk premium. Table 8 shows that if we treat industry betas as known constants, the estimation error of the market premium in itself produces standard errors of CAPM  $CE$ 's that average a whopping 3.01% per year. [The average industry beta in (7) is greater than one.] Moreover, given the uncertainty about the market premium, the marginal effect of beta uncertainty is small. Using the three-year rolling regressions to allow for uncertainty about beta due to estimation error and time-varying true betas only raises the average standard error of the CAPM  $CE$ 's from 3.01% to 3.15%.

Since the imprecision of CAPM  $CE$ 's is largely due to uncertainty about the market premium, all the industry  $CE$ 's have large standard errors. When the full-period regressions are used to measure beta uncertainty, the standard errors of the CAPM  $CE$ 's for 29 of 48 industries (not shown) are above 3.0%. When the rolling three-year regressions are used, 31 industries have CAPM standard errors above 3.0% per year.

Table 8 shows that uncertainty about risk loadings is somewhat more important in the imprecision of three-factor  $CE$ 's. Again, however, uncertainty about factor risk premiums in itself creates massive uncertainty about three-factor  $CE$ 's, and thus about project values. When we only allow for the imprecision of the premiums, the average standard error of the three-factor industry  $CE$ 's is huge, 3.17% per year. When we use the three-year rolling-regression slopes to also allow for uncertainty about risk loadings, the average standard error of the three-factor  $CE$ 's rises to 3.85% per year. Thus, given the uncertainty about factor risk premiums, the marginal effect of uncertainty about risk loadings is again relatively small.

## 8. Conclusions

Estimates of the cost of equity are distressingly imprecise. Standard errors of more than 3.0% per year are typical when we use the CAPM or the three-factor model to estimate industry  $CE$ 's. These large standard errors are driven primarily by uncertainty about true factor risk premiums, with some help from imprecise estimates of period-by-period risk loadings. Since the risk loadings for individual firms or projects are less precise than those of industries, the standard errors of  $CE$ 's for firms or projects are even larger.

Uncertainty about the true asset pricing model adds further to the uncertainty about project values. For example, though they share the same estimate of the market risk premium, the CAPM and three-factor  $CE$ 's of many industries differ by more than 2.0% per year. And denominator uncertainty is, of course, only half of the project valuation problem. Uncertainty about the cashflow estimates in the numerator also creates first-order imprecision in estimates of project values.

Project valuation is central to the success of any firm. Our message is that the task is beset with massive uncertainty. The question then is whether there is an approach that values projects with less error than its competitors. Is the net-present-value approach, advocated with zeal by textbooks, typically more accurate than a less complicated approach, like payback? And how would one tell? Our guess is that whatever the formal approach two of the ubiquitous tools in capital budgeting are a wing and a prayer, and serendipity is an important force in outcomes.

## Appendix A

We use four-digit SIC codes to assign firms to 48 industries. The industries (short name, long name, and SIC codes) are:

Agric	Agriculture	0100–0799, 2048–2048
Food	Food Products	2000–2046, 2050–2063, 2070–2079, 2090–2095, 2098–2099
Soda	Candy and Soda	2064–2068, 2086–2087, 2096–2097
Beer	Alcoholic Beverages	2080–2085
Smoke	Tobacco Products	2100–2199
Toys	Recreational Products	0900–0999, 3650–3652, 3732–3732, 3930–3949
Fun	Entertainment	7800–7841, 7900–7999
Books	Printing and Publishing	2700–2749, 2770–2799
Hshld	Consumer Goods	2047–2047, 2391–2392, 2510–2519, 2590–2599, 2840–2844, 3160–3199, 3229–3231, 3260–3260, 3262–3263, 3269–3269, 3630–3639, 3750–3751, 3800–3800, 3860–3879, 3910–3919, 3960–3961, 3991–3991, 3995–3995
Clths	Apparel	2300–2390, 3020–3021, 3100–3111, 3130–3159, 3965–3965
Hlth	Healthcare	8000–8099
MedEq	Medical Equipment	3693–3693, 3840–3851
Drugs	Pharmaceutical Products	2830–2836
Chems	Chemicals	2800–2829, 2850–2899
Rubbr	Rubber and Plastic Products	3000–3000, 3050–3099
Txtls	Textiles	2200–2295, 2297–2299, 2393–2395, 2397–2399
BldMt	Construction Materials	0800–0899, 2400–2439, 2450–2459, 2490–2499, 2950–2952, 3200–3219, 3240–3259, 3261–3261, 3264–3264,

		3270–3299, 3420–3442, 3446–3452, 3490–3499, 3996–3996
Cnstr	Construction	1500–1549, 1600–1699, 1700–1799
Steel	Steel Works, Etc.	3300–3369, 3390–3399
FabPr	Fabricated Products	3400–3400, 3443–3444, 3460–3479
Mach	Machinery	3510–3536, 3540–3569, 3580–3599
ElcEq	Electrical Equipment	3600–3621, 3623–3629, 3640–3646, 3648–3649, 3660–3660, 3691–3692, 3699–3699
Misc	Miscellaneous	3900–3900, 3990–3990, 3999–3999, 9900–9999
Autos	Automobiles and Trucks	2296–2296, 2396–2396, 3010–3011, 3537–3537, 3647–3647, 3694–3694, 3700–3716, 3790–3792, 3799–3799
Aero	Aircraft	3720–3729
Ships	Shipbuilding, Railroad Eq	3730–3731, 3740–3743
Guns	Defense	3480–3489, 3760–3769, 3795–3795
Gold	Precious Metals	1040–1049
Mines	Nonmetallic Mining	1000–1039, 1060–1099, 1400–1499
Coal	Coal	1200–1299
Enrgy	Petroleum and Natural Gas	1310–1389, 2900–2911, 2990–2999
Util	Utilities	4900–4999
Telcm	Telecommunications	4800–4899
PerSv	Personal Services	7020–7021, 7030–7039, 7200–7212, 7215–7299, 7395–7395, 7500–7500, 7520–7549, 7600–7699, 8100–8199, 8200–8299, 8300–8399, 8400–8499, 8600–8699, 8800–8899
BusSv	Business Services	2750–2759, 3993–3993, 7300–7372, 7374–7394, 7397–7397, 7399–7399, 7510–7519, 8700–8748, 8900–8999
Comps	Computers	3570–3579, 3680–3689, 3695–3695, 7373–7373
Chips	Electronic Equipment	3622–3622, 3661–3679, 3810–3810, 3812–3812
LabEq	Measuring and Control Equip	3811–3811, 3820–3830
Paper	Business Supplies	2520–2549, 2600–2639, 2670–2699, 2760–2761, 3950–3955
Boxes	Shipping Containers	2440–2449, 2640–2659, 3210–3221, 3410–3412
Trans	Transportation	4000–4099, 4100–4199, 4200–4299, 4400–4499, 4500–4599, 4600–4699, 4700–4799

Whlsl	Wholesale	5000–5099, 5100–5199
Rtail	Retail	5200–5299, 5300–5399, 5400–5499, 5500–5599, 5600–5699, 5700–5736, 5900–5999
Meals	Restaurants, Hotel, Motel	5800–5813, 5890–5890, 7000–7019, 7040–7049, 7213–7213
Banks	Banking	6000–6099, 6100–6199
Insur	Insurance	6300–6399, 6400–6411
RIEst	Real Estate	6500–6553
Fin	Trading	6200–6299, 6700–6799

## Appendix B

Temporal variation in true risk loadings complicates tests of asset pricing models. This appendix explores this problem in the context of tests of the three-factor model on industries.

### B.1. The negative correlation between intercepts and slopes

The full-period regressions in Table 2 reject the three-factor model. The  $F$ -test of Gibbons, Ross, and Shanken (1989) rejects the hypothesis that the intercepts in (4) are zero for all industries at the 0.0003 level.

The rejection of the three-factor model seems to be driven by the strong negative correlation between the three-factor intercepts,  $a_i$ , and the  $HML$  slopes,  $h_i$ . The observed correlation,  $-0.65$ , is much stronger than the correlation,  $-0.22$ , implied by the sample covariance matrix of the explanatory variables under the hypothesis that the regression coefficients are constant through time and the same for all industries. To highlight the correlation between  $h_i$  and  $a_i$ , Table 2 sorts industries on their  $HML$  slopes. All positive three-factor intercepts more than two standard errors from zero are associated with negative  $HML$  slopes; all negative intercepts more than two standard errors from zero are associated with positive  $HML$  slopes. Real Estate produces the most extreme negative intercept,  $-1.01\%$  per month ( $t = -5.45$ ), and the third largest  $HML$  slope,  $0.40$ . The drug industry produces the largest positive intercept,  $0.61\%$  per month ( $t = 3.88$ ), and the lowest  $HML$  slope,  $-0.63$ .

Negative correlation between abnormal returns and slopes is not special to the  $HML$  slopes in Table 2. The correlation between the intercepts and market slopes in the CAPM regressions,  $-0.26$ , is more negative than the estimated correlation of their sampling errors,  $-0.10$ . Similarly, the correlations between the intercepts and the  $R_M - R_f$  and  $SMB$  slopes in the three-factor regressions,

– 0.27 and – 0.44, are also more negative than the estimated correlations of their sampling errors, – 0.14 and – 0.06.

There are at least two interpretations of these negative correlations. First, perhaps the risk premiums in the CAPM and the three-factor model are overstated. For example, if the three-factor model overstates the risk premium associated with distress, it will overpredict the returns on industries with high *HML* loadings and underpredict the returns on industries with low *HML* loadings.

The alternative interpretation is that the negative correlations between intercepts and slopes are driven by time-varying risk loadings. Specifically, we argue that the negative relation between abnormal returns and *HML* loadings is a natural consequence of the dynamics of growth and distress. Consider an industry that becomes distressed. In the three-factor model, one result of distress is an increase in  $h_i$ , the industry's loading on *HML*. If the industry's bad times are unexpected, the increase in  $h_i$  is probably accompanied by negative abnormal returns in (4). Conversely, the surprise onset of good times likely implies a decline in an industry's *HML* loading and positive abnormal returns. Extending the argument, industries that on balance have more surprise bad times than good times during our sample period are more likely to have positive *HML* slopes and negative intercepts in estimates of (4). Industries whose cumulative shocks are positive probably have negative *HML* slopes and positive intercepts.

This argument is not limited to the *HML* slopes. The negative correlations between the intercepts and the market and *SMB* slopes are consistent with the hypothesis that bad news about future cash flows also tends to raise an industry's risk loadings on these factors. In addition, increases in risk loadings raise discount rates and lower current prices. Both effects create negative correlation between abnormal returns and factor risks. Chan (1988) and Ball and Kothari (1989) make a similar point about CAPM market betas.

In sum, we hypothesize that in three-factor regressions like those in Table 2, industries with large positive *HML* slopes are more likely to have experienced surprise distress and negative abnormal average returns during the regression estimation period. Conversely, industries with large negative *HML* slopes are more likely to have experienced surprise good times and positive abnormal average returns. The result is a negative correlation between intercepts and *HML* slopes in the three-factor model. The alternative to this dynamics-of-distress story is the bad-model hypothesis that the three-factor model exaggerates the premium for distress. It overestimates expected returns on industries with high *HML* slopes and underestimates expected returns on industries with low *HML* slopes. The next two sections try to distinguish between these two hypotheses by looking at portfolios with roughly constant risk loadings.

## B.2. Tests using deciles formed on past industry *HML* slopes

Suppose we allocate industries to deciles based on past *HML* slopes. Both the bad-model hypothesis and the dynamics-of-distress hypothesis predict negative correlation between intercepts and *HML* slopes during the portfolio-formation period. But they make different predictions about the post-formation returns on the deciles. The bad-model hypothesis says that the negative relation between intercepts and *HML* slopes is a model specification problem (an exaggerated premium for distress) that will persist in post-formation returns. In contrast, the dynamics-of-distress story predicts that the three-factor model will look better in post-formation returns. Specifically, if we re-form the deciles frequently based on past *HML* slopes, there should be little unexpected drift in their post-formation *HML* slopes. Thus, if the expected-return Eq. (3) holds, the intercepts in three-factor regressions on post-formation returns should be close to zero and largely uncorrelated with post-formation *HML* slopes.

To test these predictions, we sort industries into two sets of deciles, using three- and five-year past *HML* slopes. We re-form the portfolios monthly and weight industries equally in the deciles. Table A.1 shows average values of the formation-period three-factor regression coefficients for the deciles. The predicted negative relation between intercepts and *HML* slopes is clear. Growth portfolios (strong negative formation-period *HML* slopes) have strong positive formation-period abnormal average returns. Distress portfolios (strong positive *HML* slopes) have strong negative formation-period abnormal returns.

The formation-period regressions are important because they show that allocating industries to portfolios based on their *HML* slopes emphasizes (rather than diversifies away) the negative formation-period relation between three-factor regression intercepts and *HML* slopes. Thus, if the negative formation-period relation between intercepts and *HML* slopes is a bad-model problem (an exaggerated premium for distress), it should reappear in the post-formation returns on the portfolios. Table A.2 shows that this does not happen. Instead, as predicted by the dynamics-of-distress story, the three-factor intercepts for post-formation returns are close to zero and largely unrelated to the post-formation *HML* slopes. The Gibbons, Ross, and Shanken (GRS) tests in Table A.2 confirm that the post-formation intercepts are consistent with the three-factor model.

The success of the three-factor model in the tests on post-formation decile returns (Table A.2) is also consistent with the evidence in Tables 3 and 4 that industry slopes on *HML* wander through time. If the slopes were constant, the negative correlation between intercepts and *HML* slopes in Tables 2 and A.1 would be a bad-model problem. With constant slopes, industries would not move much across the deciles and the tests on post-formation returns (Table A.2) would tend to reproduce the strong negative correlation between

intercepts and *HML* slopes observed in formation-period returns (Table A.1). Again, this does not happen.

### B.3. Deciles formed on industry *BE/ME*

Tables A.1 and A.2 also summarize tests on deciles formed in June each year using industry book-to-market ratios (*BE/ME*) for the fiscal year ending in the preceding calendar year. We hypothesize that, like high past *HML* slopes, high *BE/ME* is likely to identify industries that on balance experienced recent surprise distress (and low stock returns), while low *BE/ME* is likely to be associated with recent surprise growth. Table A.1 confirms this prediction. There is the usual strong negative relation between intercepts and *HML* slopes in five-year regressions estimated on returns preceding formation of the *BE/ME* deciles.

Table A.1

Summary statistics for portfolios formed on *BE/ME* or on three- and five-year past *HML* slopes from three-factor regressions

$$R_i - R_f = a_i + b_i[R_M - R_f] + s_iSMB + h_iHML + e_i$$

$R_M - R_f$ , *SMB*, and *HML* are defined in Table 1. In parts A and B, industries are allocated to deciles (five industries per decile, except for the fifth and sixth, which have four) every month beginning in 6/68 based on their *HML* slopes estimated from three or five years of past monthly returns. Returns on the deciles are calculated for the following month with equal weighting of the industries in a decile. In part C, industries are allocated to *BE/ME* deciles in June of each year  $t$  from 1968 to 1994. Monthly returns on the deciles are calculated for the following year (July to June) with equal weighting of the industries in a decile. *BE* is the sum of book equity (defined in Table 1) for the firms in an industry that have positive Compustat *BE* for the fiscal year ending in calendar year  $t - 1$ . *ME* is the sum of *ME* at the end of December of year  $t - 1$  for these firms. Thus, *BE/ME* uses only Compustat firms, but the industry returns include all NYSE, AMEX, and NASDAQ firms on CRSP. The table shows the average values of the ranking period (6/68–11/94, 318 months) three-factor regression coefficients, and means, standard deviations (Std. dev.), and  $t$ -statistics for the means [ $t(\text{Mean})$ ] of the decile returns for the post-ranking period (7/68–12/94).

Decile	1	2	3	4	5	6	7	8	9	10
Part A: Portfolios formed on three-year pre-ranking <i>HML</i> slopes										
<i>Ranking-period regression coefficients</i>										
Mean $a$	0.41	0.22	0.11	0.12	0.02	-0.08	-0.14	-0.19	-0.31	-0.44
Mean $b$	1.00	1.01	1.03	1.00	1.02	1.02	1.02	1.03	1.05	1.08
Mean $s$	0.26	0.29	0.30	0.39	0.41	0.44	0.38	0.42	0.39	0.44
Mean $h$	-0.82	-0.48	-0.29	-0.14	-0.03	0.06	0.17	0.29	0.46	0.73
<i>Post-ranking returns</i>										
Means	0.66	0.91	0.82	0.92	1.07	1.02	0.99	1.02	1.06	1.10
Std. dev.	6.10	5.81	5.80	5.56	5.56	5.61	5.28	5.31	5.23	5.27
$t(\text{Mean})$	1.92	2.78	2.53	2.94	3.43	3.23	3.32	3.41	3.60	3.72

Table A.1 (continued)

Decile	1	2	3	4	5	6	7	8	9	10
<b>Part B: Portfolios formed on five-year pre-ranking <i>HML</i> slopes</b>										
<i>Ranking-period regression coefficients</i>										
Mean <i>a</i>	0.41	0.25	0.10	0.08	-0.02	-0.06	-0.10	-0.22	-0.31	-0.44
Mean <i>b</i>	1.04	1.02	1.03	1.04	1.00	1.03	1.04	1.04	1.01	1.03
Mean <i>s</i>	0.30	0.31	0.27	0.40	0.46	0.47	0.43	0.37	0.36	0.44
Mean <i>h</i>	-0.71	-0.42	-0.25	-0.11	-0.00	0.07	0.15	0.26	0.39	0.62
<i>Post-ranking returns</i>										
Means	0.93	0.74	0.71	0.91	0.99	1.07	0.10	1.04	0.94	1.10
Std. dev.	6.28	5.71	5.81	5.73	5.67	5.72	5.22	5.22	5.15	5.03
<i>t</i> (Mean)	2.64	2.32	2.19	2.84	3.10	3.33	3.77	3.56	3.27	3.89
<b>Part C: Portfolios formed on <i>BE/ME</i></b>										
<i>Pre-ranking five-year regression coefficients</i>										
Mean <i>a</i>	0.53	0.31	0.03	0.06	0.04	-0.12	-0.21	-0.25	-0.25	-0.44
Mean <i>b</i>	0.93	1.03	1.05	1.10	1.03	1.03	1.03	1.04	1.04	1.00
Mean <i>s</i>	0.14	0.40	0.38	0.36	0.37	0.42	0.48	0.34	0.41	0.44
Mean <i>h</i>	-0.49	-0.28	-0.15	-0.07	-0.02	0.06	0.07	0.16	0.33	0.32
<i>Post-ranking returns</i>										
Means	0.76	0.91	0.72	0.94	0.91	1.03	0.85	1.01	1.19	1.21
Std. dev.	5.52	6.01	5.84	5.56	5.55	5.51	5.43	5.29	5.25	5.47
<i>t</i> (Mean)	2.45	2.71	2.21	3.00	2.92	3.34	3.78	3.38	4.02	3.93

Again, however, since we re-form the *BE/ME* portfolios annually, there should be little unexpected drift in their true post-formation *HML* slopes. If the three-factor model holds, three-factor regressions on post-formation returns should produce intercepts close to zero, and the negative correlation between *HML* slopes and intercepts should largely disappear. Table A.2 confirms these predictions.

The portfolios formed on industry *BE/ME* also provide evidence on whether survivor bias drives the book-to-market effect in average returns. Kothari, Shanken, and Sloan (1995) argue that average returns on high-*BE/ME* portfolios of Compustat stocks are overstated because Compustat is more likely to include distressed firms that survive and miss distressed firms that fail. Since our industries contain all NYSE, AMEX, and NASDAQ stocks, the strong spread in post-formation average returns (0.45% per month, or about 5.5% per year) for portfolios formed on industry *BE/ME* (Table A.1) cannot be attributed to survivor bias. Our results are thus consistent with the more detailed evidence of

Chan, Jegadeesh, and Lakonishok (1995) that Compustat survivor bias cannot explain the strong positive relation between  $BE/ME$  and average return.

#### B.4. The unresolved testing problem

The portfolios formed on past  $HML$  slopes and  $BE/ME$  in Tables A.1 and A.2 refute the bad-model interpretation of the strong negative correlation between

Table A.2

Three-factor and one-factor regressions for post-formation returns on portfolios formed on industry  $HML$  slopes or industry book-to-market equity ratios: 7/68–12/94

$$R_i - R_f = a_i + b_i[R_M - R_f] + s_iSMB + h_iHML + e_i, \quad R_i - R_f = a_i + b_i[R_M - R_f] + e_i$$

$R_M - R_f$ ,  $SMB$ , and  $HML$  are defined in Table 1. The formation of deciles on industry  $BE/ME$  or  $HML$  slopes is described in Table A.1. One- and three-factor regressions are estimated on the post-formation returns for 7/68–12/94 (318 months). The regression  $R^2$  are adjusted for degrees of freedom.  $t(a)$  is the  $t$ -statistic for an intercept. The standard errors of  $b_i$ ,  $s_i$ , and  $h_i$  are about 0.03, 0.04, and 0.05.  $F(a)$  is the  $F$ -statistic of Gibbons, Ross, and Shanken (1989) for tests of the hypothesis that the intercepts for all deciles are zero;  $p(F)$  is its  $p$ -value (the probability of a value of  $F$  as large or larger than the sample value if the true intercepts are all zero).  $Mn(a)$ ,  $Mn(|a|)$ , and  $Mn(a^2)$  are the average, average absolute, and average squared values of the intercepts for a set of deciles.

Decile	Three-factor						One-factor			
	$a$	$b$	$s$	$h$	$R^2$	$t(a)$	$a$	$b$	$R^2$	$t(a)$
Regressions for portfolios formed on three-year past $HML$ slopes										
1	-0.11	1.03	0.37	-0.45	0.89	-0.95	-0.35	1.21	0.83	-2.44
2	0.05	1.04	0.32	-0.27	0.88	0.45	-0.08	0.17	0.85	-0.64
3	-0.06	1.07	0.36	-0.23	0.92	-0.66	-0.17	1.19	0.88	-1.56
4	-0.08	1.09	0.31	0.00	0.92	-0.94	-0.07	1.15	0.90	-0.67
5	0.10	1.02	0.43	-0.02	0.88	0.85	0.10	1.11	0.84	0.81
6	-0.04	1.09	0.40	0.12	0.90	-0.45	0.04	1.14	0.87	0.34
7	-0.09	0.98	0.50	0.22	0.88	-0.89	0.05	1.03	0.80	0.35
8	-0.05	1.04	0.38	0.18	0.90	-0.56	0.06	1.07	0.86	0.55
9	-0.02	1.00	0.38	0.22	0.86	-0.15	0.12	1.03	0.81	0.93
10	-0.05	1.00	0.40	0.37	0.82	-0.37	0.18	1.00	0.75	1.19
10 - 1	0.07	-0.03	0.03	0.82	0.34	0.36	0.52	-0.21	0.06	2.49
Regressions for portfolios formed on five-year past $HML$ slopes										
1	-0.17	1.05	0.38	-0.49	0.88	1.39	-0.09	1.24	0.82	-0.58
2	-0.09	1.03	0.32	-0.31	0.91	-0.85	-0.24	0.16	0.87	-2.10
3	-0.25	1.08	0.40	-0.09	0.90	-2.34	-0.28	1.18	0.87	-2.34
4	-0.07	1.07	0.45	-0.05	0.92	-0.73	-0.08	1.17	0.88	-0.70
5	0.02	1.09	0.34	0.00	0.90	-0.15	0.00	1.16	0.88	0.01
6	-0.00	1.08	0.45	0.13	0.88	0.00	0.09	1.13	0.83	0.68
7	0.11	0.98	0.40	0.06	0.88	1.09	0.16	1.04	0.84	1.36
8	-0.09	1.04	0.38	0.32	0.89	-0.90	0.10	1.04	0.83	0.84
9	-0.14	1.02	0.37	0.24	0.90	-1.47	0.00	1.03	0.85	0.05
10	-0.02	0.95	0.37	0.36	0.81	-0.17	0.20	0.94	0.73	1.33
10 - 1	-0.19	-0.10	-0.01	0.86	0.38	-1.05	0.28	-0.30	0.11	1.31

Table A.2 (continued)

Decile	Three-factor					One-factor				
	<i>a</i>	<i>b</i>	<i>s</i>	<i>h</i>	$R^2$	<i>t(a)</i>	<i>a</i>	<i>b</i>	$R^2$	<i>t(a)</i>
Regressions for portfolios formed on industry <i>BE/ME</i>										
1	0.09	0.95	0.07	-0.53	0.87	0.81	-0.20	1.09	0.82	-1.50
2	0.10	1.02	0.46	-0.38	0.91	1.00	-0.09	1.20	0.84	-0.68
3	-0.23	1.07	0.45	-0.11	0.92	-2.37	-0.27	1.19	0.87	-2.31
4	-0.02	1.04	0.40	-0.06	0.90	-0.21	-0.04	1.13	0.87	-0.34
5	-0.11	1.08	0.33	0.05	0.90	-1.04	-0.06	1.13	0.87	-0.58
6	0.04	1.00	0.48	0.02	0.87	0.38	0.07	1.09	0.82	0.57
7	-0.21	1.03	0.46	0.14	0.91	-2.18	-0.11	1.09	0.85	-0.96
8	-0.08	1.04	0.35	0.21	0.89	-0.74	0.06	1.06	0.85	0.47
9	0.08	1.04	0.40	0.27	0.90	0.78	0.24	1.05	0.84	2.04
10	-0.06	1.09	0.46	0.55	0.89	-0.61	0.26	1.05	0.77	1.78
10 - 1	-0.16	0.13	0.39	1.08	0.54	-1.02	0.46	-0.04	-0.00	2.09
Summary of regression intercepts (excluding 10 - 1)										
Explanatory variables				$F(a)$	$p(F)$	$Mn(a)$	$Mn( a )$	$Mn(a^2)$		
Regressions for portfolios formed on three-year past <i>HML</i> slopes										
$R_M - R_f$				1.350	0.203	-0.012	0.122	0.0227		
$R_M - R_f$ <i>SMB</i> <i>HML</i>				0.637	0.782	-0.037	0.067	0.0052		
Regressions for portfolios formed on five-year past <i>HML</i> slopes										
$R_M - R_f$				2.074	0.026	-0.014	0.124	0.0234		
$R_M - R_f$ <i>SMB</i> <i>HML</i>				1.762	0.067	-0.038	0.095	0.0144		
Regressions for portfolios formed on <i>BE/ME</i>										
$R_M - R_f$				2.384	0.010	-0.015	0.141	0.0275		
$R_M - R_f$ <i>SMB</i> <i>HML</i>				1.513	0.134	-0.039	0.102	0.0145		

intercepts and *HML* slopes in Table 2. Although there is a large spread in the portfolios' post-formation *HML* slopes, the three-factor intercepts are close to zero, and there is little relation between the intercepts and *HML* slopes. Moreover, the portfolios in Tables A.1 and A.2 have roughly constant *HML* slopes. In contrast, the implied volatilities in Table 3 and the conditional regressions in Table 4 say that there is substantial variation in the true *HML* slopes for industries. Thus, it appears that the large negative correlation between the intercepts and *HML* slopes in Table 2 – and the strong rejection of the three-factor model – are caused by the dynamics of growth and distress.

Because the conditional regressions in Table 4 seem to capture the wandering *SMB* and *HML* slopes, one might hope that they also absorb the intercepts produced by the dynamics of growth and distress. In fact, the intercepts in these regressions are close to those of the constant-slope regressions in Table 2. A naive application of the GRS test (the explanatory variables in the conditional regressions are not the same for all industries) rejects the three-factor model as strongly in the conditional regressions as in the constant-slope regressions.

Why don't the conditional regressions absorb the regression intercepts? Even if we knew the true model for expected returns and the true risk loadings at the beginning of each month, the dynamics-of-distress story predicts a negative correlation between intercepts and unexpected drift in risk loadings. Bad news that increases risk loadings is likely to produce a negative abnormal return; good news is likely to produce lower risk loadings and a positive abnormal return. Thus, in tests of the true asset pricing model that use the true time-varying risk loadings, we can predict the abnormal average return for a test period based on the unexpected drift in the true risk loadings during the period. We do not predict that an industry's average abnormal return is zero unless, by chance, its risk loadings at the end of the test period are the same as at the beginning.

Negative correlation between abnormal returns and unexpected drift in risk loadings does not by itself imply, however, that tests of asset pricing models are biased toward rejection. If the intercepts generated by the dynamics of growth and distress are just average values of normally distributed return shocks, they satisfy the assumptions of the GRS test, and they do not bias the tests in Tables 2 and 4 toward rejection of the three-factor model. In fact, if we maintain the other assumptions of the GRS test, such as multivariate normality and a constant covariance matrix for the residuals, wandering risk loadings increase the residual variances in constant-slope regressions like those in Table 2, and so make false rejection of an asset pricing model less likely.

Then why do we reject the three-factor model in Tables 2 and 4? Perhaps the three-factor model misses important industry factors in expected returns that happen to be negatively related to *HML* slopes. For example, Real Estate produces the most extreme intercepts in Tables 2 and 4,  $-1.01\%$  ( $t = -5.45$ ) and  $-1.02\%$  ( $t = -5.22$ ) per month. Applying Bonferroni's inequality to the  $t$ -statistics of the Real Estate intercepts [multiplying their univariate  $p$ -values by 48(!)] in itself produces a comfortable rejection of the hypothesis that the three-factor regression intercepts are zero for all industries. It is then tempting to argue that Real Estate stocks are a hedge against the relative price of housing services (a potential state variable in the ICAPM), so the equilibrium expected return on Real Estate is less than predicted by the three-factor model. Unfortunately, this hedging argument also seems to predict negative intercepts for the health industries (Drugs, Medical Equipment, and Health Services), but their intercepts in Tables 2 and 4 are positive, and large. Indeed, the intercepts for

Drugs, 0.61 ( $t = 3.88$ ) and 0.58 ( $t = 3.64$ ), suffice to reject the hypothesis that the three-factor regression intercepts are zero for all industries.

Another possibility is that the rejection of the three-factor model in Tables 2 and 4 is due to violations of the assumptions of the GRS test. For example, the residuals from the three-factor model are not multivariate normal. In Table 2, the average kurtosis of the residuals for the 48 industries is 4.22. In simulations of the GRS test, Affleck-Graves and McDonald (1989) find that departures from normality of this magnitude can substantially increase the probability of rejecting the true asset pricing model.

Until we better understand the testing problem created by negative correlation between abnormal returns and stochastic drift in the risk loadings of industries and firms, perhaps the best solution to the problem is to test asset pricing models on portfolios formed to have constant risk loadings. This is, of course, the approach in Tables A.1 and A.2, and in most of the existing literature.

## Appendix C

We can represent the one-and three-factor regressions (2) and (4) as

$$R_i - R_f = X\beta_i + \varepsilon_i,$$

where  $\beta_i$  ( $2 \times 1$  or  $4 \times 1$ ) is the vector of true regression coefficients for industry  $i$  in (2) or (4),  $X$  is the ( $N \times 2$  or  $N \times 4$ ) matrix of explanatory returns for  $N$  months, and  $R_i - R_f$  and  $\varepsilon_i$  are ( $N \times 1$ ) vectors of excess returns and disturbances. If the joint distribution of security returns is multivariate normal (MVN) and if  $\beta_i$  is constant, then the ordinary least squares (OLS) estimate of  $\beta_i$ ,  $B_i$ , is MVN with mean vector  $\beta_i$  and covariance matrix  $\sigma^2(\varepsilon_i)(X'X)^{-1}$ ,

$$B_i \sim \text{MVN}[\beta_i, \sigma^2(\varepsilon_i)(X'X)^{-1}]. \quad (\text{A.1})$$

In a Bayesian framework,  $\beta_i$  is a random vector. We assume a prior distribution for  $\beta_i$  that is MVN with mean vector  $\beta$  and covariance matrix  $\Sigma(\beta)$  ( $2 \times 2$  or  $4 \times 4$ ),

$$\beta_i \sim \text{MVN}[\beta, \Sigma(\beta)]. \quad (\text{A.2})$$

With these assumptions, Blattberg and George (1991) show that the mean of the posterior distribution of  $\beta_i$  is

$$E(\beta_i) = D(i)^{-1}[X'X/(\sigma^2(\varepsilon_i)B_i) + \Sigma(\beta)^{-1}\beta], \quad (\text{A.3})$$

where

$$D(i) = X'X/\sigma^2(\varepsilon_i) + \Sigma(\beta)^{-1}. \quad (\text{A.4})$$

The posterior mean is thus a weighted average of the prior mean  $\beta$  and the sample estimates  $B_i$ , where the weights are relative precisions. Equivalently, rearranging (A.3), the posterior mean shrinks the OLS  $B_i$  toward the prior mean  $\beta$  to correct the sample estimates for sampling error,

$$E(\beta_i) = B_i - W(i)(B_i - \beta), \quad (\text{A.5})$$

where the shrinkage matrix  $W(i)$  is

$$W(i) = D(i)^{-1} \Sigma(\beta)^{-1}. \quad (\text{A.6})$$

To implement (A.5), we estimate  $\beta$  as the mean, across industries, of the  $B_i$  for an estimation period, and we estimate  $\sigma^2(e_i)$  as the variance of the OLS residuals for industry  $i$ ,  $\hat{\sigma}^2(e_i)$ . The estimates of  $\beta$  and  $X'X$  are specific to an estimation period, and  $\hat{\sigma}^2(e_i)$  is specific to industries and estimation periods.

Invertibility problems force us to use an estimate of  $\Sigma(\beta)$  for the overall sample period. Specifically, we can estimate the covariance matrix of the  $B_i$  with the sample covariance matrix ( $2 \times 2$  or  $4 \times 4$ ) of the  $B_i$  for an estimation period, denoted by  $C(B)$ . We can then estimate  $\Sigma(\beta)$  as

$$C(\beta) = C(B) - \hat{\sigma}^2(e)(X'X)^{-1}, \quad (\text{A.7})$$

where  $\hat{\sigma}^2(e)$  is the mean of  $\hat{\sigma}^2(e_i)$  across industries. When we use  $C(\beta)$  to estimate  $\Sigma(\beta)$  for each regression estimation period, there are occasional periods when  $C(\beta)$  is not invertible. Intuitively, there are periods when the dispersion of the  $B_i$  estimates across industries [ $C(B)$ ] is too low, relative to the average estimated covariance matrix of the sampling errors of the slopes [ $C(B - \beta) = \hat{\sigma}^2(e)(X'X)^{-1}$ ], to produce a meaningful covariance matrix for the true slopes. To bypass this problem, we produce one overall estimate of  $\Sigma(\beta)$ , using average values across estimation periods of the three components of  $C(\beta)$  in (A.7).

There is one final problem in estimating  $\Sigma(\beta)$ . The estimate for the overall period is typically not invertible if the intercepts in (4) are included. Intuitively, the average cross-sectional variance of the estimated intercepts is not large enough, relative to the sampling-error variances of the estimates, to produce a meaningful estimate of the cross-sectional variance of the true intercepts. We interpret this as evidence that a dogmatic prior for the intercepts (zero) must be used to get Bayes estimates of the regression slopes. In other words, to get Bayes estimates of the slopes in (4) we must impose the three-factor expected-return Eq. (3). Thus, the intercepts are dropped in all the inputs in (A.5).

For perspective on the Bayes estimates, Table A.3 shows averages, over industries and estimation periods, of the key shrinkage inputs for five-year rolling estimates of the slopes in (2) and (4). Note that estimates of the standard deviation of the true market slopes in the three-factor regressions are small (around 0.16) relative to the those for the *SMB* and *HML* slopes (around 0.40 and 0.33). As a result, although the market slopes are estimated more reliably (they have smaller standard errors), the weighting matrix that produces shrunk

Table A.3

Shrinkage inputs for five-year one-factor and three-factor regression slopes

$$R_i - R_f = a_i + b_i[R_M - R_f] + e_i, \quad R_i - R_f = a_i + b_i[R_M - R_f] + s_i \text{SMB} + h_i \text{HML} + e_i$$

$R_M - R_f$ ,  $\text{SMB}$ , and  $\text{HML}$  are defined in Table 1. One-factor and three-factor regressions are estimated for each industry for each month of the 6/68 to 11/94 period, using a rolling window of five years of past monthly returns. In the three-factor regressions,  $C(B)$  is an overall estimate of the covariance matrix of the estimated regression slopes.  $C(B)$  is the average, across regression estimation periods, of the  $(3 \times 3)$  matrix of the covariances (across industries) of the regression slopes for an estimation period.  $C(B - \beta)$  is an overall estimate of the covariance matrix of the estimation errors of the estimated slopes.  $C(B - \beta)$  is the average, first across industries and then across estimation periods, of the residual variances in the three-factor regressions, times the average, over the estimation periods, of  $(X'X)^{-1}$ , the inverse of the covariance matrix of the explanatory variables.  $C(\beta) = C(B) - C(B - \beta)$  is an overall estimate of the covariance matrix of the true regression slopes.  $s(B)$ ,  $s(B - \beta)$ , and  $s(\beta)$  are the square roots of the diagonal elements of  $C(B)$ ,  $C(B - \beta)$ , and  $C(\beta)$ . The three-factor shrinkage matrix,  $W$ , uses overall averages (over industries and then regression estimation periods) of each of the terms in (A.4) and (A.6). This  $W$  matrix is illustrative, but it is not used in Tables 5 to 7. We do use the overall  $C(\beta)$  shown here in Tables 5 to 7, but the other elements of (A.4) and (A.6) used to calculate shrinkage weights are specific to industries and estimation periods.  $s^2(b)$ ,  $s^2(b - \beta)$ , and  $s^2(\beta)$  are the one-factor analogues of  $C(B)$ ,  $C(B - \beta)$ , and  $C(\beta)$ .

	Three-factor			One-factor		
	$b$	$s$	$h$			
		$C(B)$		$s(B)$	$s^2(b)$	$s(b)$
$b$	0.042	0.012	-0.000	0.205	0.067	0.258
$s$	0.012	0.194	0.018	0.441		
$h$	-0.000	0.018	0.155	0.394		
		$C(B - \beta)$		$s(B - \beta)$	$s^2(b - \beta)$	$s(b - \beta)$
$b$	0.015	-0.006	0.010	0.124	0.013	0.115
$s$	-0.006	0.036	0.002	0.189		
$h$	0.010	0.002	0.044	0.208		
		$C(\beta)$		$s(\beta)$	$s^2(\beta)$	$s(\beta)$
$b$	0.027	0.018	-0.010	0.163	0.053	0.231
$s$	0.018	0.159	0.015	0.399		
$h$	-0.010	0.015	0.111	0.334		
		$W$			$W$	
$b$	0.658	0.059	-0.062		0.816	
$s$	0.206	0.839	0.028			
$h$	-0.213	0.038	0.764			

slopes gives more weight to the sample estimates of the  $\text{SMB}$  and  $\text{HML}$  slopes than to the sample estimates of the market slopes. Roughly speaking (i.e., ignoring the off-diagonal terms in the weighting matrix) about 66% of the difference between an industry's OLS market slope and the average market slope is allocated to its shrunk market slope, whereas about 84% and

76% of the OLS estimates of the *SMB* and *HML* slopes make it into the shrunk slopes.

## References

- Affleck-Graves, John and Bill McDonald, 1989, Nonnormalities and tests of asset pricing theories, *Journal of Finance* 44, 889–908.
- Ball, Ray and S. P. Kothari, 1989, Non-stationary expected returns: Implications for tests of market efficiency and serial correlation of returns, *Journal of Financial Economics* 25, 51–74.
- Blattberg, Robert C. and Edward I. George, 1991, Shrinkage estimation of price and promotional elasticities: Seemingly unrelated equation, *Journal of the American Statistical Association* 86, 305–315.
- Bower, Dorothy A., Richard S. Bower, and Dennis E. Logue, 1984, Arbitrage pricing and utility stock returns, *Journal of Finance* 39, 1041–1054.
- Bower, Richard S. and George R. Schink, 1994, Application of the Fama–French model to utility stocks, *Financial Markets, Institutions and Instruments* 3, 74–96.
- Brennan, Michael, 1995, The term structure of discount rates, Working paper no. 10–93 (Anderson Graduate School of Management, University of California, Los Angeles, CA).
- Chan, K.C., 1988, On the contrarian investment strategy, *Journal of Business* 61, 147–163.
- Chan, K.C. and Nai-fu Chen, 1991, Structural and return characteristics of small and large firms, *Journal of Finance* 46, 1467–1484.
- Chan, Louis K.C., Narasimhan Jegadeesh, and Josef Lakonishok, 1995, Evaluating the performance of value versus glamour stocks: The impact of selection bias, *Journal of Financial Economics* 38, 269–296.
- Chen, Nai-fu, Richard Roll, and Stephen A. Ross, 1986, Economic forces and the stock market, *Journal of Business* 59, 383–403.
- Elton, Edwin J., Martin J. Gruber, and Jianping Mei, 1994, Cost of capital using arbitrage pricing theory: A case study of nine New York utilities, *Financial Markets, Institutions and Instruments* 3, 46–73.
- Fama, Eugene F. and Kenneth R. French, 1989, Business conditions and expected returns on stocks and bonds, *Journal of Financial Economics* 25, 23–49.
- 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.
- Gibbons, Michael R., Stephen A. Ross, and Jay Shanken, 1989, A test of the efficiency of a given portfolio, *Econometrica* 57, 1121–1152.
- Goldenberg, David H. and Ashok J. Robin, 1991, The arbitrage pricing theory and cost-of-capital estimation: The case of electric utilities, *Journal of Financial Research* 14, 181–196.
- Gonedes, Nicholas, 1973, Evidence on the information content of accounting numbers: Accounting-based and market-based estimates of systematic risk, *Journal of Financial and Quantitative Analysis* 8, 407–444.
- Huberman, Gur and Shmuel Kandel, 1987, Mean-variance spanning, *Journal of Finance* 42, 873–888.
- Kothari, S.P., Jay Shanken, and Richard G. Sloan, 1995, Another look at the cross-section of expected stock returns, *Journal of Finance* 50, 185–224.

- 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.
- Merton, Robert C., 1973, An intertemporal capital asset pricing model, *Econometrica* 41, 867–887.
- Merton, Robert C., 1980, On estimating the expected return on the market: An exploratory investigation, *Journal of Financial Economics* 8, 323–361.
- Ross, Stephen A., 1976, The arbitrage theory of capital asset pricing, *Journal of Economic Theory* 13, 341–360.
- Shanken, Jay, 1990, Intertemporal asset pricing: An empirical investigation, *Journal of Econometrics* 45, 99–120.
- Sharpe, William F., 1964, Capital asset prices: A theory of market equilibrium under conditions of risk, *Journal of Finance* 19, 425–42.