Wayne E. Ferson
University of Washington

Robert A. Korajczyk
Northwestern University

Do Arbitrage Pricing Models Explain the Predictability of Stock Returns?*

This article studies predictability in U.S. stock returns for multiple investment horizons. We measure to what extent predictability is driven by premiums for economywide risk factors, comparing two standard methods for factor selection. We study single-beta models and multiple-beta models. We show how to estimate the fraction of the predictability in returns captured by the model, simultaneously with the other parameters. Our analysis indicates that the models capture a large fraction of the predictability for all of the investment horizons. The performance of the principal components and the prespecified-factor approaches are broadly similar.

I. Introduction

Recent studies find evidence that the rates of return to holding common stocks and bonds are, to some extent, predictable over time. While measurement errors can induce spurious predictability, there seems to be an emerging consensus that predictable variation is an important stylized fact about asset returns. There are two competing views of this predictability. Some interpret

---

* This article has benefited from the helpful comments of a number of colleagues, including Torben Andersen, John Cochrane, Eugene Fama, Kenneth French, Allaudeen Hameed, Lars Hansen, Campbell Harvey, Ravi Jagannathan, Narasimhan Jegadeesh, Andrew Lo, John Long, Jianping Mei, Kenneth Singleton, Charles Trzcinka, and two anonymous referees. This article was presented at the 1992 Berkeley Program in Finance; the 1992 Conference on Financial Economics and Accounting; the 1992 Pacific Northwest Finance Conference; the 1993 American Finance Association; the 1993 Western Finance Association; the 1994 Conference on Multivariate Times Series and Financial Econometrics sponsored by the University of California, San Diego; and at the following universities: University of California, Berkeley; University of Maryland; McGill University; University of Michigan; Northwestern University; Ohio State University; University of Rochester; University of Southern California; University of South Carolina; Tulane University; and University of Wisconsin—Madison. We acknowledge the financial support of the New York Stock Exchange. Ferson acknowledges financial support from the Pigott-Paccar professorship at the University of Washington, and Korajczyk acknowledges financial support from Merrill Lynch Capital Markets.

(*Journal of Business,* 1995, vol. 68, no. 3)

0021-9398/95/6803-0002$01.50

predictability as evidence of market inefficiencies, while others look to time-varying equilibrium expected returns, driven by economic fundamentals. Rational expectations implies that security returns should be predictable only to the extent that *expected* returns are related to the predictor variables. To the extent that standard models of the expected returns can capture predictability, the time-varying fundamentals view is encouraged.

Traditional asset pricing models imply that the expected returns of securities are determined by their "beta coefficients" or factor loadings and by the associated marketwide risk premiums. Therefore, in a rational model any predictability of returns should be driven by changes in the betas and changes in the expected risk premiums.

Ferson and Harvey (1991) provide evidence that much of the predictability in a sample of monthly portfolio returns can be related to standard risk factors in a multiple-beta model. They use prespecified economic factors similar to Chen, Roll, and Ross (1986). However, the ability of these models to capture the predictability in long-horizon returns has not been examined. This article uses the beta pricing framework to address the predictability in long-horizon returns.

Long-horizon returns are interesting for several reasons. First, estimates of the fraction of return variance that is predictable for multimonth returns are larger than for 1-month returns. Such evidence remains controversial, and much of the controversy over predictability in the literature has centered on long-horizon returns.[1] Second, returns measured for longer holding periods are less susceptible to the effects of thin and nonsynchronous trading, bid-ask spreads, and microstructural issues. The dynamic properties of returns that beta pricing models are designed to address may therefore represent a larger fraction of the variance in longer horizon return data.

Early studies of the capital asset pricing model (CAPM) were based on unconditional returns. Some of these studies used annual data (see, e.g., Douglas 1969; Miller and Scholes 1972), but most of the empirical work has used returns measured over weekly or monthly holding periods. Levhari and Levi (1977) note that empirical estimates of the CAPM parameters depend on the investment horizon and claim that some of the anomalies in the early studies are consistent with the measured return horizon being too short. More recently, Handa, Kothari, and Wasley (1989) find that evidence on the firm-size effect is sensitive to the return horizon, and Kothari, Shanken, and Sloan

---

[1] See, e.g., Fama and French (1988*a*), Goetzmann and Jorion (1993), and Nelson and Kim (1993) for evidence focusing on dividend yields. Fama and French (1988*b*), Lo and MacKinlay (1988), Richardson and Stock (1989), and Richardson (1993) focus on autocorrelations. Foster and Smith (1992) and Kirby (1993) provide multivariate analyses of predictability.

(1993) find that the book-to-market effect documented by Fama and French (1992) is sensitive to the return horizon. However, there is very little evidence on the performance of multiple-beta pricing models over long investment horizons and virtually no evidence using conditional expected returns. This article provides tests of beta pricing models for conditional expected returns, using investment horizons from 1 month to 2 years. Our article therefore provides an important link between the static asset pricing literature and the literature on predictable returns.

We compare two sets of risk factors in our models. The first are prespecified economic factors similar to Chen et al. (1986) and Ferson and Harvey (1991). The second are asymptotic principal components, following Connor and Korajczyk (1986). Ours is the first study to compare the two approaches to factor selection on conditional asset returns. As there is uncertainty among academics and practitioners about which approach is better, a comparison of the two is of practical importance.

We use the generalized method of moments (GMM) to estimate the fraction of the predictability in returns that is captured by a model, simultaneously with the other model parameters. This avoids the multistep procedure used by Ferson and Harvey (1991). The approach represents a useful methodological contribution which can be used to address other asset pricing questions.

We find that models with one to five factors and constant betas do not explain 100% of return predictability, but they do capture a large fraction of the predictability. For example, single-factor models explain about 60% of the predictability in our sample of industry-grouped stock portfolios, while five-factor models capture about 80%, on average. These results are not highly sensitive to the return horizon. The main empirical result of our article is that the models capture a large fraction of the predictability, irrespective of the return horizon.

We find that the performance of a five-principal-components model and a five-economic-factor model are broadly similar for capturing predictability in returns for all of the horizons. This is an important result, since the prespecified-factor approach is often easier to use in practice. However, this conclusion must be tempered by the fact that data limitations do not allow us to conduct tests which fully account for cross-sectional dependencies in the returns.

The article is organized as follows. Section II describes the models, and Section III describes our methodology. Section IV describes the data, while Section V presents the main empirical results. Section VI provides some evidence on the importance of time variation in betas for our results and examines sensitivity to the empirical methods. The final section is the conclusion.

II. The Models

A. *Asset Pricing Equations*

The asset pricing hypothesis is a multiple-beta or exact arbitrage pricing (APT) model of the following form:[2]

$$E(R_{i,t+1}|Z_t) = \lambda_0(Z_t) + \sum_{j=1}^{K} b_{ijt}\lambda_j(Z_t), \qquad (1)$$
$$i = 0, \ldots, N, t = 0, \ldots, T-1,$$

where $R_{i,t+1}$ is the rate of return on asset i between times t and $t+1$, and Z_t is a vector of instruments for the information available when prices are set at time t. The $b_{i1t}, \ldots, b_{iKt}$ are the time t conditional betas or factor loadings which measure the systematic risk of asset i relative to the K risk factors. The $\lambda_j(Z_t); j = 1, \ldots, K$ are the market prices of systematic risk, or expected risk premiums. The conditional betas with respect to the risk factors $F_{j,t+1}, j = 1, \ldots, K$ are defined by a factor model regression:

$$R_{i,t+1} = a_{it} + \sum_{j=1}^{K} b_{ijt}F_{j,t+1} + u_{i,t+1}, \qquad (2)$$
$$i = 0, \ldots, N, t = 0, \ldots, T-1,$$

where $E(u_{i,t+1}|Z_t) = E(u_{i,t+1}F_{j,t+1}|Z_t) = 0$ for all i, j, and t.

For any set of asset returns there is some (conditional) mean variance efficient portfolio such that equation (1) describes expected returns using this portfolio as the single factor (Roll 1977). The content of any beta pricing model is therefore the discipline imposed in modeling the factors. The goal of this study is to use standard factors, used in previous studies of unconditional mean returns, and to assess their usefulness in explaining asset return predictability over different investment horizons. Given the factors, we first examine models which assume that the betas are fixed parameters over time ($b_{ijt} = b_{ij}$ for all t). Stambaugh (1983) and Connor and Korajczyk (1989) derive versions of the APT with conditioning information, in which the risk premiums are time-varying and the betas are fixed parameters. We also test the constant beta hypothesis, and in Section VI we allow for time variation in the betas.

---

[2] The multiple-beta equilibrium model was developed by Merton (1973), Long (1974), and Breeden (1979). The approximate APT was developed by Ross (1976), and an equilibrium APT was developed by Connor (1984).

B. Choosing the Risk Factors

The first set of risk factors are prespecified economic factors similar to Chen et al. (1986) and Ferson and Harvey (1991), who studied monthly data after World War II. We extend the sample back to 1927 to study longer-horizon rates of return. Mimicking portfolios are constructed from individual common stocks for the factors which are not asset returns. A mimicking portfolio is defined as a portfolio that may be used in the factor model regression (2) to determine the betas and whose expected excess return is the risk premium in equation (1) (see Huberman, Kandel, and Stambaugh 1987). Mimicking portfolios are used because if the risk factors are not traded assets, the model implies no specific relation between the expected values of the factors and the expected risk premiums. The construction of the mimicking portfolios is described in the appendix.

The second approach to selecting the factors is motivated by previous studies of the APT. The APT assumes that the factors capture the important comovements in asset returns. We use the asymptotic principal components methods of Connor and Korajczyk (1986, 1988*a*, 1988*b*) to estimate the common factors and select the first five principal components ordered by their eigenvalues.[3] This is representative of previous studies which examined APT pricing for unconditional expected returns and allows comparisons with the five economic variables.

Equation (1) should describe expected returns using any set of variables for which the error terms in (2) are uncorrelated with marginal utility. While there is no claim that a unique set of variables captures the relevant economic risks, the performance of the two approaches remains an interesting empirical question.

III. Methodology

A. Regression Tests

Our first set of tests uses time-series regressions of returns on the factor-mimicking portfolios and a vector of predetermined instru-

---

[3] We extract principal components from the $T \times T$ *cross-sectional* moment matrix $(rr')/N$, where there are N assets and T time-series observations, and r is a $T \times N$ matrix of returns in excess of a Treasury bill. This allows the factor premia to be time-varying but assumes that the factor loadings of the individual securities are fixed parameters. This does not imply that portfolios formed with time-varying weights must have fixed betas. We use returns in excess of the Treasury bill. If the bill return is a function only of systematic risk factors, then asymptotic principal components deliver factors which approximate mimicking portfolio returns in excess of a zero beta rate (see Korajczyk and Viallet 1989).

ments:

$$r_{it} = \alpha_{i0} + \sum_{p=1}^{L} \alpha_{ip} Z_{p,t-1} + \sum_{j=1}^{K} \beta_{ij} F_{jt} + u_{it}, \quad t = 1, \ldots, T. \tag{3}$$

We use excess returns $r_{it} = R_{it} - R_{ft}$, where R_{ft} is the return of a 1-month Treasury bill. For multiple-month returns, we use the return to a strategy of rolling over the 1-month bill each month as R_{ft}. The subscript i indicates the regression for a given asset. The symbol $Z_{p,t-1}$ denotes the value of predetermined variable p at period $t - 1$, and F_{jt} denotes the excess return of the mimicking portfolio for factor j at period t. We assume that the conditional betas β_{ij} are fixed parameters over time, where $\beta_{ij} = b_{ijt} - b_{fjt}$, and b_{fjt} is the conditional beta of the Treasury bill. The model implies that $r_{it} - \Sigma_{j=1}^{K} \beta_{ij} F_{jt}$ should have a conditional mean equal to zero given the information Z_{t-1}. In particular, α_{i0} should be zero, and the predetermined variables should not enter the regression. The joint hypothesis may be examined by testing the restrictions that the α_{i0} and the α_{ip}'s are equal to zero in (3).[4]

If the predetermined variables were not in the regression (3), the α_{i0} should be zero if the model explains the assets' *unconditional* expected returns. Such tests are conducted in previous studies, which found that $\alpha_{i0} = 0$ can be rejected in models with a small number of factors.[5] However, our main interest in this article is the ability of the models to capture the predictable variation in returns over time. We therefore focus on the restrictions that the $\alpha_{ip} = 0, p = 1, \ldots, L$. In essence, these tests examine whether the violations of the model vary through time and are correlated with the lagged variables.

If we reject the hypothesis that the lagged variables may be excluded from the regression (3), we reject a joint hypothesis which includes constant betas. If the betas for a given set of factors actually depend on Z_{t-1}, this could cause a rejection of the exclusion restrictions. Thus, a rejection of the model (3) may be interpreted as indicating either time-varying betas or the need for additional factors. In general, there is an isomorphic relation between a fixed-beta model with a larger number of factors and a time-varying beta model with a smaller number. Additional factors may often be constructed to "control" the time variation in beta. See Cochrane (1992) for further discussion of this point.

---

[4] Since the mimicking portfolios F_{jt} are estimated using a finite sample of assets, the small sample estimation error implies an errors-in-variables problem for the regressions. However, the simulation evidence in Connor and Korajczyk (1988*a*) and McCulloch and Rossi (1989) suggests that estimation error should be small in samples with as many assets as we use. See Wheatley (1989) for an analysis of estimation errors in mimicking portfolios.

[5] See Huberman and Kandel (1987), Connor and Korajczyk (1988*b*), Lehmann and Modest (1988), Harvey and Zhou (1990), McCulloch and Rossi (1990), and Wei, Lee, and Chen (1991).

B. Decomposing Return Predictability

The regression tests ask the model to explain 100% of the variance of the returns that is predictable from regressions on the lagged variables Z_{t-1}. This is an extreme test, since no model is perfect. It is interesting to measure the fraction of the predictability that is captured by the model. We therefore estimate a ratio, VR1, which is the predictable variation in a return that is explained by the product of its betas and the expected premia, divided by an "unrestricted" estimate of the predictable variation. We estimate a system of moment conditions:

$$u1_{it} = (r_{it} - Z'_{t-1}\delta_i), \tag{4a}$$

$$u2_{it} = (F'_t - Z'_{t-1}\gamma_i)' \tag{4b}$$

$$u3_{it} = [(u2_{it}u2'_{it})\beta_i - (F_t u1_{it})]Z'_{t-1}, \tag{4c}$$

$$u4_{it} = (Z'_{t-1}\delta_i - \theta_i), \tag{4d}$$

$$u5_{it} = (Z'_{t-1}\gamma_i\beta_i + \alpha_i - \theta_i), \tag{4e}$$

$$u6_{it} = (u4^2_{it})\mathrm{VR1}_i - u5^2_{it}. \tag{4f}$$

The symbol F_t denotes a $K \times 1$ vector of factor-mimicking portfolios, β_i denotes a $K \times 1$ vector of the betas for asset i, and Z_{t-1} denotes an $(L + 1)$ vector of the lagged instruments, including the constant. The parameters are $\{\theta_i, \alpha_i, \mathrm{VR1}_i, \delta_i, \beta_i, \gamma_i\}$ for each asset i. The first three parameters are scalars, δ_i is $(L + 1) \times 1$, β_i is $K \times 1$, and γ_i is $(L + 1) \times K$. Under the assumption that the conditional betas are fixed parameters, the β_i in system (4) are the same as the β_i in the regression model (3). The model (4) implies the orthogonality conditions $E(u1_{it}Z_{t-1}, u2_{it}Z'_{t-1}, u3_{it}, u4_{it}, u5_{it}, u6_{it}) = 0$, for each i.

The intuition for the system of moment conditions is as follows. The first equation, (4a), describes a regression of the excess asset return r_{it} on the lagged instruments. The "unrestricted" conditional expected return then is $E(r_{it}|Z_{t-1}) = Z'_{t-1}\delta_i$, and the predictable variance of the return is defined as $\mathrm{var}[E(r_{it}|Z_{t-1})] = \mathrm{var}[Z'_{t-1}\delta_i]$. The second equation, (4b), is a system of regressions for the factor-mimicking portfolios on the lagged instruments. The fitted values are used to model the expected risk premiums. Equation (4c) defines the conditional betas, which are assumed to be fixed parameters. The predictable variance of the asset return that is captured by the model is the part attributed to betas and risk premiums: $\mathrm{var}[E(\Sigma_j \beta_{ij}F_{jt}|Z_{t-1})] = \mathrm{var}[Z'_{t-1}\gamma_i\beta_i]$. The last three equations ([4d]–[4f]) define the variance ratio:

$$\begin{aligned}\mathrm{VR1}_i &= \mathrm{var}\left[E\left(\sum_j \beta_{ij}F_{jt}|Z_{t-1}\right)\right] \Big/ \mathrm{var}[E(r_{it}|Z_{t-1})], \\ &= \mathrm{var}[Z'_{t-1}\gamma_i\beta_i]/\mathrm{var}[Z'_{t-1}\delta_i].\end{aligned} \tag{5}$$

The ratio VR1 is the predictable variance in the return that is attributed to the model, relative to the total predictable variance in the return, given Z_{t-1}.

Ferson and Harvey (1991) estimate variance ratios similar to VR1, using a multiple-step regression procedure. They are forced to use simulations to estimate standard errors for the variance ratios. In our formulation the GMM provides consistent estimates of the model parameters, including the variance ratios, their asymptotic standard errors, and a goodness-of-fit test.[6] The approach can be further extended to examine other functional forms for the conditional betas and other asset-pricing hypotheses. Ferson and Harvey (1993) and Section VI below provide examples.

The advantages of the system (4) come with some costs. Since the number of moment conditions in the system is large, we are unable to pool the equations across the assets. The empirical model therefore does not impose the restriction that the conditional risk premiums are common across assets ($\gamma_i = \gamma$). More powerful tests could in principle be constructed by using the cross-sectional restrictions. We address this issue by checking the sensitivity of our inferences using cross-sectional methods, which do estimate a common risk premium. Although asset-by-asset estimation of the model means that we cannot impose all of the restrictions implied by the model, asset-pricing models do impose restrictions for individual assets. In addition, an asset-by-asset examination allows us to see which of the asset returns present more of a challenge to the model.

IV. The Data

A. Common Stock Returns

For the purposes of extracting asymptotic principal components and forming mimicking portfolios, we use monthly stock return data for all individual firms listed on the New York Stock Exchange (NYSE) and

---

[6] The number of orthogonality conditions in the system (4) is $(L + 1)(2K + 1) + 3$, and the number of parameters is $(L + 1)(K + 1) + K + 3$ for a one-asset system. The model is estimated separately for each asset to keep the size of the system tractable. The model is overidentified, with degrees of freedom equal to the difference between the number of orthogonality conditions and the number of parameters, which is $K \times L$. The overidentifying conditions are used to construct a χ^2 test of the model specification, as in Hansen (1982). Given the simulation evidence of Ferson and Foerster (1994), we use an iterated GMM procedure for improved finite sample properties, and we adjust the standard errors for finite sample bias using the correction factor suggested by Ferson and Foerster (1994). The Ferson-Foerster correction is to multiply the asymptotic variance matrix by the term $[(N + L)T/(N + L)T - Q]$, where N is the number of assets ($N = 1$ in our tables), L is the number of instruments, T is the number of time-series observations, and Q is the sum of the number of model parameters plus the number of unique elements of the GMM weighting matrix.

the American Stock Exchange (AMEX). These data are provided by the Center for Research in Security Prices (CRSP). We use the return of a firm in each month for which it is available, over the 1926–89 period. Missing observations are handled for the principal components as in Connor and Korajczyk (1988*a*), by averaging over the available observations for each month.

We test the models using portfolios grouped according to market capitalization and industry affiliation. Ten value-weighted portfolios are formed by CRSP according to size deciles of the combined NYSE + AMEX sample, based on the market value of equity outstanding at the beginning of each year. We also form 12 portfolios according to the 2-digit Standard Industrial Classification code, using the same groupings as Breeden, Gibbons, and Litzenberger (1989) and Ferson and Harvey (1991). These portfolios are value-weighted each month.

It is common in the empirical literature on multiple-horizon returns to use continuously compounded returns. This is convenient because a long-horizon return is then the sum of the short-horizon returns.[7] However, in our context continuously compounded returns are not appropriate for a number of reasons. The discrete time first-order condition which motivates (1) states that expected returns are related to the conditional covariances of returns with the marginal utility for wealth. Wealth depends directly on the simple, arithmetic rate of return to an optimal portfolio. Moreover, continuously compounded portfolio returns are not the portfolio-weighted averages of the securities' continuously compounded returns. Continuous compounding implies that the mimicking portfolios and the estimators in cross-sectional models lose their interpretation as portfolio returns. Therefore, in this article we use simple, arithmetic returns.[8]

Using a model like equation (1) to "explain" the predictability of

---

[7] For example, when returns are continuously compounded, there are at least two alternatives to estimating a regression of long-horizon (sum of) returns on a predetermined variable, z. One is to regress the single-period return on a sum of the lagged zs (Hodrick 1992). Another is to regress the single-period return on the single period z, posit an autoregressive model, and infer the long-horizon coefficient (Kandel and Stambaugh 1990). These approaches can provide the same information because the long-horizon return is the sum of the short-horizon returns.

[8] While it is clear that continuously compounded returns are not appropriate in our setting, it may not be clear how important is the distinction for our results. We therefore conduct an experiment, in which we regress $R_t - \ln(1 + R_t)$ on Z_{t-1}, where R_t represents the simple, arithmetic return, and the Z_{t-1} are our predetermined instruments. The regressions show significant predictability in the difference between the simple and continuously compounded returns. Therefore, evidence on the predictability of one type of return (simple) cannot be used to make direct inferences about the predictability of the other type (continuously compounded). Indeed, finding predictability in the difference should not be surprising in view of the fact that the expected difference is, to a first approximation, proportional to the conditional variance of the return. There is much evidence in the literature for time variation in conditional variances (for a review, see Bollerslev, Chou, and Kroner 1992).

asset returns with data for different investment horizons raises other interesting issues. The risk premiums and betas are not in general invariant to the return measurement interval (Levhari and Levy 1977; Handa et al. 1989; Longstaff 1989). Therefore, our analyses for different investment horizons should not be interpreted as the same model applied to various horizons but as alternative specifications of the asset pricing model. When we study monthly (quarterly, annual, 2-year) returns, the model assumes that the representative investor makes consumption and investment decisions at monthly (quarterly, annual, 2-year) frequencies.

B. *The Predetermined Information Variables*

We study predictability based on a collection of variables suggested by previous studies. They are: (1) the level of the 1-month Treasury-bill rate (TB), (2) the dividend yield of the CRSP value-weighted NYSE stock index (VWYLD), (3) a detrended stock index price level (PLEV), (4) a measure of the slope of the term structure (TERM), (5) a quality-related yield spread in the corporate bond market (QUAL), and (6) a dummy variable for the month of January (used in monthly and quarterly data).[9]

The 1-month bill rate used to predict the future returns is from the CRSP RISKFREE files.[10] The lagged dividend yield is the stock index level at the end of the previous month, divided into the previous year's dividend payments for the index. The lagged price level PLEV is formed by dividing the average level of the CRSP equally weighted stock index with dividends over the past 12 months by the level at the end of the previous month. TERM is the difference between the lagged yields-to-maturity of bonds in the composite bond index with more than 15 years to maturity, less the yield of bonds with 5–15 years to maturity. QUAL is the lagged yield to maturity of Baa-rated bonds, less the yield of Aaa-rated bonds. The data for the TERM and QUAL variables are from the Ibbotson Corporate Bond Module.[11]

C. *The Economic Risk Factors*

Our list of economic risk factors is similar to Chen et al. (1986). First, we use the return of the Standard and Poor's 500 (SP500) Stock Index

---

[9] Studies which document predictability using these variables include Fama and Schwert (1977) and Ferson (1989), for short-term interest rates; Campbell and Shiller (1988), Fama and French (1988*a*, 1989), and Poterba and Summers (1988) for dividend yields; Keim and Stambaugh (1986) for detrended stock price levels; Fama and French (1989) and Fama (1990) for the corporate term spread; Fama (1990) for the quality-related yield spread; and Keim (1983) for the January dummy variable.

[10] We express the excess returns relative to 1-month bills from Ibbotson Associates. These are the monthly returns for the shortest maturity bill with at least 1 month to maturity. We use different bill series to minimize the risk of spurious correlation between the excess returns and the lagged instruments which could arise from common measurement errors in the prices of Treasury bills.

[11] We are grateful to Roger Ibbotson for making these data available.

as a stock market factor. Second is a real interest rate factor, measured as the nominal 1-month Treasury-bill rate less the rate of change in the consumer price index. Third is an unexpected inflation factor, measured as the difference between the ex post monthly percentage change in the consumer price index and a simple, time-series model for the expected rate of inflation.[12] Fourth is a corporate default risk factor, measured as the difference between the monthly returns of low-grade corporate bonds and high-grade corporate bonds. The high-grade corporate bond return is the Ibbotson and Sinquefield series from CRSP, and the low-grade bond return is from Blume, Keim, and Patel (1991).[13] Fifth is a term structure risk factor, measured as the difference between the return of a long-term U.S. government bond and a 1-month Treasury bill. We form mimicking portfolios for the inflation and real interest rate factors using monthly individual common stock returns, as described in the appendix. For longer horizons we compound the monthly rates of return and express them net of the return to rolling over a 1-month Treasury bill.

V. Empirical Results

A. *The Predictability of Stock Returns*

All of our results are based on nonoverlapping returns data. The alternative strategy of using overlapping data can improve estimation efficiency by increasing the number of observations (Hansen and Hodrick 1980). However, the efficiency gains are likely to be small when the regressors are highly autocorrelated, as is the case in our data (Boudoukh and Richardson 1994). Overlapping data also cause problems. They induce autocorrelation into the error terms, which can bias regression R^2s. In the presence of overlapping data, the consistent estimators of the covariance matrix require the estimation of a number of autocovariances, which causes the standard errors to be understated in finite samples (Richardson and Smith 1991; Goetzmann and Jorion 1993; Nelson and Kim 1993). Compared with the simple regression settings that have been studied in the previous literature, our GMM setting is more complex and is nonlinear. The finite sample problems caused by overlapping data are likely to be more severe in such settings. We therefore use nonoverlapping data for all of the return horizons.

Table 1 summarizes regressions of the asset returns on the predetermined variables. For the horizons longer than 1 month, the return is

---

[12] Expected inflation is measured as the 1-month, nominal, U.S. Treasury-bill rate less the fitted expected real return on the nominal bill. Our time-series model for the real return of the bill is based on regressing the real return on 12 of its past monthly lagged values (an AR[12]).

[13] We are grateful to Don Keim for making these data available to us.

TABLE 1 Summary Statistics of Regressing Portfolio Excess Returns and Factor Excess Returns on the Predetermined Instrumental Variables

	Return Horizon											
	Monthly			Quarterly			Annual			2-Year		
Portfolio	F-Test p-Value	Wald Test p-Value	$\hat{R}^2$	F-Test p-Value	Wald Test p-Value	$\hat{R}^2$	F-Test p-Value	Wald Test p-Value	$\hat{R}^2$	F-Test p-Value	Wald Test p-Value	$\hat{R}^2$
A. Sample period 1927–88:												
Petroleum	.044	.048	.02	.047	.075	.05	.028	.005	.20	.030	.019	.382
Finance	.016	.005	.02	.034	.244	.06	.028	.001	.20	.135	.000	.279
Consumer durables	.006	.008	.02	.006	.162	.07	.072	.010	.16	.120	.000	.288
Basic industries	.059	.041	.02	.104	.411	.04	.031	.043	.20	.279	.003	.217
Food/tobacco	.021	.035	.02	.039	.452	.05	.199	.099	.12	.267	.002	.220
Construction	.003	.013	.03	.013	.212	.07	.011	.015	.23	.034	.000	.375
Capital goods	.042	.030	.02	.094	.501	.04	.182	.160	.12	.546	.009	.142
Transportation	.001	.000	.03	.005	.220	.07	.076	.010	.16	.154	.002	.268
Utilities	.071	.010	.02	.371	.666	.03	.190	.042	.12	.491	.175	.156
Textiles/trade	.016	.069	.02	.059	.468	.05	.268	.116	.11	.550	.236	.141
Services	.017	.003	.02	.360	.334	.03	.359	.173	.09	.211	.017	.242
Leisure	.011	.010	.02	.048	.271	.05	.334	.132	.10	.464	.167	.162
Size 1	.000	.000	.12	.001	.002	.10	.065	.086	.17	.154	.003	.268
Size 2	.000	.000	.08	.001	.002	.09	.111	.075	.15	.225	.001	.236
Size 3	.000	.000	.07	.003	.007	.08	.100	.030	.15	.172	.000	.259
Size 4	.000	.000	.05	.010	.037	.07	.134	.059	.14	.288	.000	.213
Size 5	.000	.000	.05	.004	.072	.08	.074	.035	.16	.168	.000	.261
Size 6	.000	.000	.04	.014	.135	.07	.091	.033	.15	.186	.000	.252
Size 7	.001	.000	.03	.025	.200	.06	.045	.022	.18	.110	.000	.295
Size 8	.002	.001	.03	.017	.228	.06	.064	.012	.17	.132	.000	.281
Size 9	.013	.005	.02	.029	.238	.06	.059	.005	.17	.122	.000	.287
Size 10	.065	.039	.02	.084	.321	.05	.076	.011	.16	.201	.000	.246
Factor 5	.010	.002	.02	.007	.096	.07	.044	.024	.18	.001	.000	.548
Factor 4	.000	.000	.04	.044	.065	.05	.461	.152	.08	.202	.116	.245
Factor 3	.000	.000	.04	.000	.000	.12	.053	.000	.18	.016	.000	.418
Factor 2	.548	.717	.01	.420	.563	.03	.334	.173	.10	.059	.000	.341
Factor 1	.000	.000	.05	.000	.089	.10	.009	.011	.24	.028	.000	.387
B. Sample period 1961–88:												
Petroleum	.005	.017	.06	.001	.000	.19	.090	.023	.34	. . .	. . .	. . .

Finance	.000	.001	.09	.000	.000	.21	.021	.000	.45	...	...	...
Consumer durables	.000	.000	.12	.000	.000	.25	.000	.000	.64	...	...	...
Basic industries	.000	.001	.08	.001	.000	.20	.048	.000	.39	...	...	...
Food/tobacco	.000	.000	.09	.002	.000	.19	.034	.000	.41	...	...	...
Construction	.000	.000	.11	.000	.001	.24	.000	.000	.65	...	...	...
Capital goods	.000	.000	.09	.000	.000	.22	.074	.000	.36	...	...	...
Transportation	.000	.000	.11	.000	.000	.22	.005	.000	.53	...	...	...
Utilities	.000	.000	.11	.000	.000	.23	.003	.000	.55	...	...	...
Textiles/trade	.000	.001	.08	.001	.001	.19	.039	.000	.41	...	...	...
Services	.000	.000	.12	.000	.000	.23	.055	.000	.38	...	...	...
Leisure	.000	.001	.09	.000	.004	.23	.013	.000	.47	...	...	...
Size 1	.000	.000	.28	.000	.000	.34	.184	.000	.28	...	...	...
Size 2	.000	.000	.23	.000	.000	.32	.065	.000	.37	...	...	...
Size 3	.000	.000	.20	.000	.000	.30	.021	.000	.45	...	...	...
Size 4	.000	.000	.18	.000	.000	.30	.010	.000	.49	...	...	...
Size 5	.000	.000	.16	.000	.000	.28	.011	.000	.48	...	...	...
Size 6	.000	.000	.16	.000	.000	.28	.003	.000	.54	...	...	...
Size 7	.000	.000	.14	.000	.000	.25	.003	.000	.55	...	...	...
Size 8	.000	.000	.13	.000	.000	.26	.001	.000	.61	...	...	...
Size 9	.000	.000	.12	.000	.000	.24	.001	.000	.61	...	...	...
Size 10	.000	.001	.09	.000	.000	.23	.003	.000	.55	...	...	...
Factor 5	.000	.000	.12	.111	.152	.10	.548	.070	.16	...	...	...
Factor 4	.000	.000	.18	.000	.000	.22	.113	.000	.32	...	...	...
Factor 3	.000	.000	.13	.001	.001	.20	.028	.000	.43	...	...	...
Factor 2	.000	.000	.19	.000	.000	.31	.052	.000	.39	...	...	...
Factor 1	.000	.000	.15	.000	.000	.27	.002	.000	.57	...	...	...

NOTE.—All returns are simple, arithmetic rates of return, net of the return to a strategy of rolling over 1-month Treasury bills. Size 1 is the portfolio of common stocks of the 10% smallest size group, where size is measured by the market value of equity, and size 10 is the portfolio of the largest firms. The F-test p-value is the right tail area for an F-test of the hypothesis that excess returns are not predictable by the instruments. The test has an exact F-distribution under the assumption that the prediction errors are normal, independent, and identically distributed. The Wald test p-value is the tail area for a Wald test of the hypothesis that excess returns are not predictable by the instruments. The test has an asymptotic χ^2 distribution and allows for conditional heteroscedasticity in the prediction errors, following White (1980) and Hansen (1982). $\hat{R}^2$ is the coefficient of determination of the regression with a heteroscedasticity-consistent adjustment for bias derived in Kirby (1993). The regressions use nonoverlapping data for each horizon. The predetermined instrumental variables are a constant, the level of the 1-month Treasury-bill rate, the dividend yield of the Center for Research in Security Prices value-weighted stock index, a stock price level variable, the term structure yield spread, the quality yield spread, and (for monthly and quarterly horizons) a dummy variable for the month of January. For the full sample the monthly data are for January 1927–January 1988 (733 observations); the quarterly data are for first quarter 1927–first quarter 1988 (245 observations); the annual data are for 1927–88 (62 observations); and the 2-year data are for 1927–88 (31 observations). For the 1961–88 subsample the monthly data are for January 1961–January 1988 (325 observations); the quarterly data are for first quarter 1961–first quarter 1988 (109 observations); and the annual data are for 1961–88 (28 observations). For each group of monthly, quarterly, annual, and 2-year statistics, the p-values of joint Wald tests taken across industries = .000, the p-values across the size portfolios = .000, and the joint p-values across the factors = .000.

the compounded multimonth return, and the lagged instrument is the level observed in the last month of the prior period. The regressors are therefore the same for each horizon, but they are observed less frequently for the longer-horizon regressions. We conduct much of our analysis over the 1927–88 period. We also form subperiods by dividing our sample roughly in half. Panel B of table 1 reports the regressions for the second subperiod, which is 1961–88. The column denoted "*F*-Test *p*-Value" reports exact small-sample tests of predictability under the hypothesis of normality and independent, identically distributed prediction errors. The column denoted "Wald Test *p*-Value" reports tests of predictability that are robust to conditional heteroscedasticity (see White 1980; Hansen 1982).

For the monthly regressions, the *F*-tests in table 1 reject the null hypothesis that there is no predictability at a 5% level, for 10 of 12 (1927–88) and all 12 (1961–88) industries. The tests reject the null for nine of 10 (1927–88) and all 10 (1961–88) size portfolios. The Wald tests tell essentially the same story. For quarterly horizons, the *F*-tests reject the null hypothesis for seven of 12 (1927–88) and all 12 (1961–88) industry portfolios and for nine of 10 (1927–88) and all 10 (1961–88) of the size portfolios. The Wald test has a tendency to reject the null hypothesis more often than the *F*-test. For the longer horizons the adjusted R^2s generally increase, but the two tests more frequently disagree. The tests strongly reject the null hypothesis of no predictability in the second subperiod, but the evidence for the full sample period is weaker.

The regressions for different portfolios are correlated, and inferences using a number of portfolios which do not account for this correlation can be flawed. Table 1 reports joint tests for the size and industry portfolio groups, using heteroscedasticity-consistent Wald statistics which account for the correlation across the equations. These tests strongly reject the hypothesis that there is no predictability.

The predictability in the returns that is summarized by these regressions is the raw material for our study. We want to know to what extent the predictability is captured by the beta pricing models. It is therefore important to establish to what extent predictability actually exists in these data. The question of predictability raises a number of issues and has been the focus of recent controversy. Skeptics have raised a number of questions about the evidence of predictability. We can group the issues into those involving (1) small-sample problems in the statistics used to evaluate the predictability regressions, and (2) "data mining."

The first set of issues refers to the fact that the asymptotic distribution theory for the predictive regressions may be a poor approximation in the sample sizes encountered in practice. There may be finite-sample biases in both the regression coefficients and their standard errors, as well as in the regression R^2s. Kirby (1993) presents a finite sample correction for the regression R^2s. Table 1 uses his adjustment in the column labeled $\hat{R}^2$. In our data, the adjustment is typically smaller than the adjustment embodied in the usual adjusted R^2s.

Recent simulation evidence of Goetzmann and Jorion (1993) and Nelson and Kim (1993) addresses the finite sample issues, for regressions of returns on the large stocks of the Standard and Poor's composite index against a single lagged predictor, the dividend yield. Nelson and Kim (1993) find that the regression R^2s for long horizons using 1872–1987 data are significant at the 5%–10% level, but the evidence of predictability is weaker than it appears using the asymptotic distribution. Goetzmann and Jorion (1993) find similar results using 1927–90 data.

In private communication, Allaudeen Hameed reports simulation results which address the issue of finite sample bias in our data set. He generates 1,000 samples of independent first-order autoregressive variables representing the returns and the vector of the predictor variables that we use, with autocorrelations matching our sample statistics. He replicates the regressions in table 1 for the same sample sizes as ours, generating empirical 5% confidence levels for the R^2s of the regressions. He finds that there is a bias in the R^2s when there is no relation between the returns and the predictor variables. However, his estimates of the 5% critical values are still much lower than the R^2s we report for monthly and quarterly data, and also for the annual data in the second subperiod, which supports the significance of these regressions. In annual and 2-year data for the full sample period, his 5% critical values for the R^2s are comparable to those found in the actual data.[14]

The second criticism of the predictability regressions is data mining. Data mining refers to the fact that many researchers in finance use the same data, and a chance correlation of future returns with a predictor variable is likely to be discovered as an "interesting" phenomenon. (See Foster and Smith [1992], Lo and MacKinlay [1990], and Black [1993] for analyses of data-mining biases in financial studies.) We choose our predictor variables precisely because they have been discovered by previous authors. For our purposes, predictability that is the result of a data-mining bias should be hard to "explain" using the beta pricing models. That is because the risk factors are selected following studies which largely predated the predictability studies in the literature. The factor selection should therefore be independent of any data-mining bias in the predetermined instruments. Since our main

---

[14] We are grateful to Professor Hameed for providing us with the results of his analysis. One caveat is that Hameed's simulations do not take account of heteroscedasticity. Kirby (1993) presents results which suggest that heteroscedasticity can affect the distributional properties of regression R^2s.

result is that the models do a fairly good job of capturing predictability based on the predetermined instruments, any bias due to previous authors mining the data for instruments should be a conservative bias.

B. Regression Tests

Table 2 presents the tests based on the time-series regression equation (3). The right-tail probability values for two statistics are presented. The first is the F-test for the incremental explanatory power of the lagged instruments, when the factors are in the regressions. The second test is a heteroscedasticity-consistent Wald test of the same hypothesis. The F-test is based on the sums of squared residuals in ordinary least squares (OLS) regressions and is motivated by assuming normality. The Wald test is based on only the regression which includes the lagged instruments, and an estimate of the covariance matrix of the parameters, following White (1980) and Hansen (1982), which is consistent under conditional heteroscedasticity. The Wald statistic is appropriate if heteroscedasticity is an important feature, but it also requires the estimation of many more sample moments.

We conduct simulation experiments to assess the two test statistics. The simulations are described in the appendix. Empirical p-values are reported in the second row for each portfolio and horizon in table 2. The empirical p-value is the fraction of 1,000 trials in which the test statistic in the simulation for a given asset is larger than the sample value of the statistic. Each case in table 2 therefore has four p-values: two for the F-test and two for the Wald test. By comparing the theoretical p-values from the F distribution and χ^2 distribution with the empirical p-values, we obtain information about the finite sample biases of the tests.

The tests for the individual portfolios are correlated, and joint tests would be appropriate. However, given the large number of parameters relative to the number of observations, joint estimation is problematic. A partial solution is to examine "Bonferroni p-values" for the tests. Consider the event that any of N statistics for a test of size α rejects the model. Given dependent events, the joint probability is less than or equal to the sum of the individual probabilities, αN. The simple Bonferroni p-value places a conservative upper bound on the joint p-value, equal to the smallest of the N p-values multiplied by N.

Panel A of table 2 presents results for the 1927–88 sample period (733 monthly observations). The first five principal components are the risk factors in regression (3). The tests for the monthly models provide strong evidence that the factors do not capture all of the predictability, when used in a constant-beta model. The F-test and Wald test produce similar results, with no evidence of finite sample problems. The Bonferroni p-values for $N = 25$ are zero to three decimal places. To assess the importance of the January effect for these results, we ran the tests excluding the January dummy variable. Using the F-statistic, a 5% test rejects the model for six of the 12 industries and the smallest size portfolio, a total of seven out of 22 cases.

Panel B of table 2 reports tests for the second subperiod, 1961–88 (325 monthly observations). With monthly data the models are rejected using a 10% significance level, for all except three of the industries and from two to four of the size portfolios, depending on the test. There is good agreement between the F-test and Wald test, and the empirical p-values are close to the asymptotic ones. The Bonferroni p-values are again zero to three decimal places. Excluding the January dummy variable and using the F-statistic, only two industries and the smallest size portfolio produce rejections of the model.

Using different methodologies, Ferson (1990), Shanken (1990), and Ferson, Foerster, and Keim (1993) reject conditional beta pricing models that have from one to three risk factors. Shanken (1990) and Ferson, Foerster, and Keim (1993) use monthly data, and Ferson (1990) uses quarterly data. The bad news for the beta pricing models in table 2 is that even with five principal components as the factors, the constant-beta models are rejected for monthly data.

In quarterly data over the full sample period (245 observations), the tests reject the models for from four to seven of the twelve industry portfolios and from seven to 10 of the ten size portfolios, at the 5% level. The Wald test and the F-test imply consistent inferences for most of the cases. The Bonferroni p-values are less than 0.001. In the second subperiod (109 quarterly observations), the overall rejection frequencies are similar and the Bonferroni p-values are less than 0.001, but the F-tests and the Wald tests begin to disagree more frequently. The simulations imply that the Wald test rejects too often using the asymptotic p-value, while the F-test is reasonably well specified.

Moving to annual data (61 years), the F-test rejects the models for only from two to four of the 12 industries and one of the size portfolios using a 5% significance level. The Wald test using the empirical p-values rejects for only five of the industries. With 2-year data the F-test produces only one asymptotic p-value less than 10% and four empirical p-values that are that small. The Wald test begins to show evidence of severe bias in these small samples.

It is difficult to compare the regression tests across investment horizons for a number of reasons. One problem, emphasized by Richardson (1993), is that the tests for different horizons are likely to be correlated. A second issue arises in view of the Jefferies-Lindley paradox, which observes that a test with a fixed size has power approaching unity as the number of observations grows. Stated another way, the tests may have lower power given the smaller number of observations for the longer return horizons. A Bayesian view of the tests may be less susceptible to this problem. Schwartz (1978), Klein and Brown

TABLE 2 Tests of Exclusion Restrictions for the Lagged Information Variables When Five Principal Components Are the Factors

	Return Horizon							
	Monthly		Quarterly		Annual		2-Year	
Portfolio	*F* $a_j = 0$	Wald $a_j = 0$	*F* $a_j = 0$	Wald $a_j = 0$	*F* $a_j = 0$	Wald $a_j = 0$	*F* $a_j = 0$	Wald $a_j = 0$
A. Sample period 1927–88:								
Petroleum	.001	.001	.019	.064	.349	.339	.321	.271
	.000	.000	.011	.128	.350	.562	.336	.653
Finance	.000	.003	.082	.116	.078	.007	.164	.069
	.000	.000	.074	.202	.060	.084	.099	.471
Consumer durables	.000	.000	.089	.101	.473	.260	.394	.003
	.000	.000	.075	.160	.503	.532	.429	.195
Basic industries	.000	.000	.001	.003	.022	.014	.559	.253
	.000	.000	.000	.013	.010	.117	.653	.709
Food/tobacco	.000	.000	.034	.035	.057	.000	.108	.000
	.000	.000	.027	.069	.031	.027	.056	.030
Construction	.004	.016	.531	.662	.064	.013	.255	.010
	.000	.000	.542	.729	.042	.134	.280	.246
Capital goods	.000	.000	.112	.136	.513	.368	.298	.076
	.000	.000	.099	.210	.517	.572	.327	.520
Transportation	.001	.006	.013	.028	.091	.001	.144	.000
	.000	.000	.014	.079	.055	.023	.108	.107
Utilities	.000	.002	.039	.029	.058	.000	.416	.004
	.000	.000	.026	.051	.037	.005	.437	.222
Textiles/trade	.001	.012	.364	.476	.348	.031	.153	.015
	.000	.000	.371	.574	.361	.178	.099	.238
Services	.002	.018	.003	.005	.116	.183	.217	.093
	.000	.000	.002	.010	.099	.415	.187	.449
Leisure	.000	.000	.019	.006	.030	.000	.073	.000
	.000	.000	.013	.012	.011	.006	.021	.010
Industry averages	.001	.005	.109	.138	.183	.101	.259	.066
	.000	.000	.105	.186	.173	.221	.253	.320

Size 1	.000	.000	.000	.000	.913	.887	.969	.952
	.000	.000	.013	.012	.920	.952	.974	.986
Size 2	.000	.000	.000	.000	.155	.027	.661	.312
	.000	.000	.013	.012	.163	.151	.709	.663
Size 3	.000	.000	.000	.000	.048	.003	.183	.001
	.000	.000	.013	.012	.032	.072	.150	.136
Size 4	.000	.050	.000	.023	.097	.065	.195	.017
	.000	.000	.013	.044	.079	.259	.152	.363
Size 5	.085	.653	.117	.119	.974	.932	.615	.135
	.000	.000	.102	.194	.962	.957	.684	.614
Size 6	.000	.023	.001	.013	.577	.547	.287	.068
	.000	.000	.102	.032	.592	.732	.319	.450
Size 7	.000	.000	.000	.000	.709	.787	.272	.107
	.000	.000	.102	.032	.735	.897	.286	.544
Size 8	.000	.000	.027	.010	.734	.719	.289	.049
	.000	.000	.021	.021	.769	.858	.279	.404
Size 9	.000	.000	.025	.006	.165	.127	.205	.054
	.000	.000	.024	.019	.160	.363	.172	.464
Size 10	.000	.000	.004	.007	.112	.039	.539	.363
	.000	.000	.000	.011	.097	.174	.599	.794
Size averages	.009	.073	.017	.018	.448	.413	.422	.206
	.000	.000	.040	.034	.451	.452	.432	.542
B. Sample period 1961–88:								
Petroleum	.116	.129	.048	.011	.179	.082	. . .	. . .
	.110	.194	.035	.061	.156	.453	. . .	. . .
Finance	.008	.003	.079	.021	.187	.065	. . .	. . .
	.001	.002	.070	.089	.148	.395	. . .	. . .
Consumer durables	.002	.005	.000	.000	.147	.022	. . .	. . .
	.001	.002	.070	.002	.086	.282	. . .	. . .
Basic industries	.207	.222	.622	.423	.353	.012	. . .	. . .
	.145	.238	.631	.564	.380	.360	. . .	. . .
Food/tobacco	.024	.012	.059	.004	.072	.000	. . .	. . .
	.003	.042	.044	.024	.022	.015	. . .	. . .

TABLE 2 (*Continued*)

	Return Horizon							
	Monthly		Quarterly		Annual		2-Year	
Portfolio	F $a_j = 0$	Wald $a_j = 0$	F $a_j = 0$	Wald $a_j = 0$	F $a_j = 0$	Wald $a_j = 0$	F $a_j = 0$	Wald $a_j = 0$
Construction	.012	.004	.175	.007	.155	.000	...	...
	.028	.001	.154	.047	.107	.094	...	...
Capital goods	.072	.093	.175	.089	.405	.000	...	...
	.045	.114	.174	.229	.439	.184	...	...
Transportation	.208	.256	.418	.132	.470	.029	...	...
	.223	.325	.435	.302	.520	.416	...	...
Utilities	.001	.000	.111	.056	.092	.000	...	...
	.000	.000	.098	.157	.047	.103	...	...
Textiles/trade	.092	.049	.307	.154	.602	.199	...	...
	.074	.047	.287	.288	.653	.620	...	...
Services	.000	.001	.006	.000	.267	.003	...	...
	.000	.001	.000	.010	.258	.208	...	...
Leisure	.002	.000	.347	.051	.344	.000	...	...
	.000	.000	.333	.138	.352	.117	...	...
Industry averages	.062	.064	.196	.079	.273	.035	...	...
	.053	.081	.194	.159	.264	.271	...	...
Size 1	.000	.000	.000	.000	.056	.000	...	...
	.000	.000	.333	.138	.007	.078	...	...
Size 2	.000	.000	.004	.001	.211	.067	...	...
	.000	.000	.001	.024	.177	.432	...	...

Size 3	.093	.329	.603	.740	.931	.821	. . .	. . .
	.118	.434	.618	.838	.957	.956	. . .	. . .
Size 4	.693	.688	.583	.628	.338	.030	. . .	. . .
	.684	.705	.595	.755	.372	.361	. . .	. . .
Size 5	.057	.143	.247	.248	.382	.137	. . .	. . .
	.027	.112	.254	.413	.386	.532	. . .	. . .
Size 6	.001	.009	.017	.000	.205	.000	. . .	. . .
	.027	.000	.016	.014	.176	.112	. . .	. . .
Size 7	.002	.006	.003	.000	.454	.128	. . .	. . .
	.027	.015	.001	.004	.505	.597	. . .	. . .
Size 8	.000	.000	.000	.000	.315	.001	. . .	. . .
	.027	.015	.001	.004	.325	.178	. . .	. . .
Size 9	.000	.000	.005	.000	.598	.175	. . .	. . .
	.027	.015	.002	.003	.673	.603	. . .	. . .
Size 10	.045	.131	.021	.011	.183	.000	. . .	. . .
	.056	.258	.011	.058	.137	.132	. . .	. . .
Size averages	.089	.130	.148	.163	.367	.136	. . .	. . .
	.099	.155	.183	.225	.372	.398	. . .	. . .

NOTE.—The model is

$$r_{it} = \alpha_{i0} + \sum_{p=1}^{L} \alpha_{ip} Z_{p,t-1} + \sum_{j=1}^{K} \beta_{ij} F_{jt} + u_{it}, \quad t = 1, \ldots, T,$$

where the F_{jt} are the excess returns on mimicking portfolios for the factors, and the $Z_{p,t-1}$ are the lagged instruments. The r_{it} are discrete arithmetic rates of return, net of the return to a strategy of rolling over 1-month Treasury bills each month. The tests examine the hypothesis that the $a_j = 0, j = 1, \ldots, L$. F stands for the step-down F-statistic, which compares the sum of squares of regressions that include the Zs with regressions that exclude the Zs. Wald is the heteroscedasticity-consistent Wald test that $a_j = 0, j = 1, \ldots, L$. Two right-tail probability values of the test statistics are reported for each case. The first is based on the asymptotic F and χ^2 distributions. The second are empirical p-values from a simulation experiment.

(1984), and Christie (1990) examine an approximate odds ratio which is a simple transformation of the F-test. The odds ratio for the model which excludes $Z_{p,t-1}$, $p = 1, \ldots, K$ in (3), relative to the model which includes the lagged instruments is given by

$$T^{L/2}\left[1 + \frac{LF}{T - (K + L + 1)}\right]^{(-T/2)}, \tag{6}$$

where T is the sample size, L is the number of lagged instruments excluding the constant, K is the number of factors, and F is the F-statistic. To save space, we do not report the odds ratios in table 2 (they are available on request), but we summarize the results. The odds ratios favor including the lagged Zs only in monthly and quarterly samples, and mainly for the smaller size-based portfolios. In annual data, 2-year data, and in all of the other cases, the odds ratios favor the models which exclude the lagged instruments.

We also conduct exclusion tests using mimicking portfolios for the economic variables as the five factors. The results are remarkably similar in all respects to those in table 2, and we do not report them here. The performance of the models, at least as indicated by these tests, is not sensitive to the choice of the factors.

The results of the regression tests suggest the following conclusions about the constant-beta pricing models with five factors. Tests reject the models for monthly and quarterly data, but the evidence for the longer-horizon returns is more favorable to the models. On the one hand, the tests may be of low power for the longer-horizon models. The odds ratio results favor the restricted models, which indicates that the failure to reject is not solely an issue of power. On the other hand, the tests examine the strict hypothesis that the models capture 100% of the predictability in the returns. This motivates our measures of the fraction of the predictability which can be explained by a model.

C. An Analysis of Predictable Variance

Tables 3 and 4 present the results of the GMM system (4). Table 3 summarizes results using the principal components as the factors. In Table 4, the economic factors are used. In order to complement the information provided by the estimates of VR1, we modify the regression system (4) to estimate the variance ratio

$$\begin{aligned}\mathrm{VR2}_i &= \mathrm{var}\left[E\left(r_{it} - \sum_j \beta_{ij}F_{jt}|Z_{t-1}\right)\right]\Big/\mathrm{var}[E(r_{it}|Z_{t-1})],\\ &= \mathrm{var}[Z'_{t-1}\delta_i - Z'_{t-1}\gamma_i\beta_i]/\mathrm{var}[Z'_{t-1}\delta_i].\end{aligned}$$

To estimate VR2 we replace the model-fitted part of the predictable return, $Z'_{t-1}\gamma_i\beta_i$, in (4e) with the predictable part of the model residual, $Z'_{t-1}\delta_i - Z'_{t-1}\gamma_i\beta_i$. The ratio VR2 measures the predictable variation

TABLE 3 **Estimates of the Fraction of the Predicted Variance of Common Stock Excess Returns That Is Explained by Constant-Beta Models with Principal Components as the Factors**

	Return Horizon							
	Monthly		Quarterly		Annual		2-Year	
Portfolio	VR1	VR2	VR1	VR2	VR1	VR2	VR1	VR2
A. One-factor models:								
Sample period 1927–88:								
Petroleum	1.49	1.55	1.21	.284	.557	.295	.543	.553
	(1.03)	(1.24)	(.339)	(.184)	(.646)	(.435)	(.503)	(.247)
Finance	3.11	1.07	1.10	.0770	1.82	.938	.464	.743
	(1.56)	(.888)	(.338)	(.0258)	(1.82)	(1.10)	(.375)	(.242)
Consumer durables	3.22	1.37	1.13	.0890	1.18	.193	.470	.456
	(2.12)	(1.48)	(.301)	(.0874)	(.845)	(.265)	(.302)	(.188)
Basic industries	4.08	1.96	.862	.0240	1.02	.416	.866	.523
	(3.06)	(2.13)	(.152)	(.0157)	(.865)	(.522)	(.415)	(.254)
Food/tobacco	2.07	1.12	.472	.153	1.21	1.39	.574	.716
	(1.35)	(.984)	(.098)	(.0482)	(1.77)	(1.68)	(.346)	(.240)
Construction	3.35	1.39	.480	.183	.360	.273	.265	.600
	(1.96)	(1.28)	(.114)	(.0347)	(.153)	(.110)	(.315)	(.465)
Capital goods	6.76	3.35	.715	.0540	2.86	.528	1.08	1.14
	(6.17)	(4.18)	(.123)	(.0258)	(3.76)	(1.69)	(.826)	(.670)
Transportation	1.07	.356	.645	.0910	1.97	.856	.955	.751
	(.503)	(.255)	(.241)	(.110)	(2.05)	(1.33)	(.823)	(.519)
Utilities	2.06	1.07	.343	.873	.591	.822	.0650	.993
	(1.70)	(.85)	(.050)	(1.06)	(.733)	(.988)	(.0710)	(.221)
Textiles/trade	2.22	1.15	.615	.0610	2.99	1.59	.628	.540
	(1.52)	(1.17)	(.125)	(.0370)	(4.08)	(2.79)	(.427)	(.261)
Services	1.48	.250	1.38	.229	.811	2.80	.638	2.18
	(.545)	(.218)	(.466)	(.224)	(1.85)	(3.23)	(.535)	(1.00)
Leisure	2.61	1.18	.640	.0680	3.96	3.31	.793	.983
	(1.96)	(1.04)	(.108)	(.0291)	(5.54)	(4.83)	(.522)	(.648)
Industry averages	2.79	1.32	.799	.182	1.61	1.12	.612	.847
	(1.96)	(1.31)	(.205)	(.157)	(2.01)	(1.58)	(.455)	(.413)

TABLE 3 (*Continued*)

	Return Horizon							
	Monthly		Quarterly		Annual		2-Year	
Portfolio	VR1	VR2	VR1	VR2	VR1	VR2	VR1	VR2
Size 1	.279	.312	1.22	.319	1.28	.0590	1.62	.408
	(.0932)	(.0798)	(.243)	(.106)	(.807)	(.199)	(.597)	(.284)
Size 2	.579	.149	.843	.118	1.32	.0570	1.83	.359
	(.164)	(.0539)	(.134)	(.0358)	(.799)	(.164)	(.620)	(.334)
Size 3	.744	.0600	1.13	.114	1.30	.0740	1.66	.330
	(.0818)	(.0207)	(.203)	(.0549)	(.675)	(.141)	(.761)	(.398)
Size 4	.959	.0290	1.32	.098	1.36	.0650	1.60	.338
	(.0766)	(.0145)	(.221)	(.0549)	(.700)	(.145)	(.508)	(.152)
Size 5	.943	.0190	1.09	.024	1.26	.0340	.984	.111
	(.148)	(.0186)	(.184)	(.0224)	(.607)	(.0909)	(.475)	(.0740)
Size 6	1.19	.0150	1.28	.049	1.54	.0950	.861	.0420
	(.127)	(.0145)	(.232)	(.0414)	(.843)	(.221)	(.309)	(.0210)
Size 7	1.24	.0520	.957	.009	1.35	.0760	.998	.211
	(.169)	(.0331)	(.114)	(.0056)	(.586)	(.166)	(.444)	(.174)
Size 8	2.57	.563	1.02	.081	1.65	.179	1.37	.416
	(.803)	(.344)	(.220)	(.0437)	(.731)	(.277)	(.839)	(.371)
Size 9	3.23	1.06	1.44	.144	1.74	.263	1.07	.450
	(1.60)	(.880)	(.358)	(.112)	(.930)	(.381)	(.663)	(.309)
Size 10	3.04	2.11	.500	.128	1.69	.414	.720	.600
	(1.99)	(.183)	(.0695)	(.036)	(1.08)	(.528)	(.420)	(.183)
Size averages	1.48	.436	1.08	.108	1.45	.132	1.27	.327
	(.525)	(.329)	(.198)	(.0513)	(.775)	(.231)	(.564)	(.230)
Overall averages	2.20	.917	.927	.149	1.54	.669	.912	.611
	(1.31)	(.864)	(.202)	(.109)	(1.45)	(.968)	(.504)	(.330)
Sample period 1961–88:								
Petroleum	1.13	.424	.837	.636	.158	.500	. . .	. . .
	(.535)	(.376)	(.586)	(.494)	(.273)	(.311)		
Finance	2.08	.527	.907	.400	.434	.232	. . .	. . .
	(.707)	(.348)	(.471)	(.243)	(.166)	(.124)		

Consumer durables	1.22	.211	.938	.194	.339	.246	. . .	. . .
	(.284)	(.118)	(.252)	(.110)	(.0970)	(.105)		
Basic industries	2.79	.531	1.40	.0740	.525	.0820	. . .	. . .
	(.914)	(.396)	(.381)	(.0787)	(.187)	(.0660)		
Food/tobacco	1.91	.601	1.29	.537	.571	.767	. . .	. . .
	(.675)	(.394)	(.659)	(.391)	(.304)	(.322)		
Construction	1.48	.225	1.31	.0780	.294	.269	. . .	. . .
	(.374)	(.148)	(.380)	(.0895)	(.0960)	(.0900)		
Capital goods	2.18	.456	1.02	.172	1.06	.204	. . .	. . .
	(.726)	(.322)	(.382)	(.138)	(.602)	(.256)		
Transportation	1.25	.0760	1.09	.0940	.255	.350	. . .	. . .
	(.274)	(.0619)	(.365)	(.0719)	(.137)	(.136)		
Utilities	1.00	.487	.560	.536	.120	.706	. . .	. . .
	(.469)	(.331)	(.304)	(.263)	(.120)	(.201)		
Textiles/trade	1.45	.516	1.07	.295	.585	.435	. . .	. . .
	(.603)	(.375)	(.506)	(.277)	(.240)	(.299)		
Services	1.19	.264	1.14	.243	.0100	.833	. . .	. . .
	(.379)	(.160)	(.425)	(.176)	(.0180)	(.146)		
Leisure	1.23	.364	1.56	.301	1.07	1.48	. . .	. . .
	(.416)	(.217)	(.667)	(.278)	(.807)	(2.28)		
Industry averages	1.58	.390	1.09	.297	.452	.509	. . .	. . .
	(.530)	(.271)	(.448)	(.217)	(.254)	(.361)		
Size 1	.271	.351	.596	.228	.988	.0310	. . .	. . .
	(.0739)	(.0815)	(.194)	(.125)	(.285)	(.0940)		
Size 2	.424	.192	.710	.126	.562	.0930	. . .	. . .
	(.0891)	(.0587)	(.182)	(.0854)	(.179)	(.0650)		
Size 3	.566	.0950	.818	.0500	.592	.0760	. . .	. . .
	(.0967)	(.0391)	(.179)	(.0434)	(.174)	(.0540)		
Size 4	.686	.0440	.810	.0320	.538	.104	. . .	. . .
	(.0826)	(.0196)	(.152)	(.0326)	(.146)	(.0580)		
Size 5	.828	.0150	.923	.0130	.606	.0730	. . .	. . .
	(.0869)	(.0120)	(.142)	(.0176)	(.142)	(.0460)		
Size 6	.894	.0180	.938	.0250	.509	.0960	. . .	. . .
	(.0826)	(.0109)	(.151)	(.0231)	(.110)	(.0480)		
Size 7	.994	.0290	1.06	.0350	.761	.0340	. . .	. . .
	(.106)	(.0152)	(.142)	(.0285)	(.151)	(.0200)		

TABLE 3 (*Continued*)

	Return Horizon							
	Monthly		Quarterly		Annual		2-Year	
Portfolio	VR1	VR2	VR1	VR2	VR1	VR2	VR1	VR2
Size 8	1.13	.0960	1.12	.0960	.728	.0430	. . .	. . .
	(.156)	(.0413)	(.193)	(.0624)	.0860	.0230		
Size 9	1.40	.217	1.15	.139	.572	.0680	. . .	. . .
	(.317)	(.115)	(.231)	(.0949)	(.0840)	(.0310)		
Size 10	1.60	.336	.960	.227	.436	.186	. . .	. . .
	(.474)	(.202)	(.369)	(.156)	(.192)	(.0900)		
Size averages	.879	.139	.908	.0971	.629	.0804	. . .	. . .
	(.157)	(.056)	(.193)	(.0669)	(.155)	(.0529)		
Overall averages	1.26	.276	1.01	.206	.532	.314	. . .	. . .
	(.360)	(.175)	(.332)	(.149)	(.209)	(.221)		
B. Five-factor models:								
Sample period 1927–88:								
Petroleum	1.08	.910	1.02	.453	.565	.387	. . .	. . .
	(2.63)	(1.00)	(.238)	(.138)	(.222)	(.143)		
Finance	1.00	.112	.704	.297	.567	.532	. . .	. . .
	(.209)	(.0665)	(.296)	(.190)	(.202)	(.316)		
Consumer durables	1.31	.425	1.97	2.06	1.08	.0440	. . .	. . .
	(.966)	(.548)	(1.91)	(1.96)	(.217)	(.0490)		
Basic industries	2.03	1.35	.791	.147	.706	.163	. . .	. . .
	(3.38)	(2.51)	(.151)	(.0610)	(.125)	(.0690)		
Food/tobacco	.401	.751	.344	.355	1.56	4.59	. . .	. . .
	(.418)	(.757)	(.121)	(.119)	(1.01)	(3.34)		
Construction	1.78	.383	1.08	.0950	.543	.238	. . .	. . .
	(1.17)	(.527)	(.199)	(.0570)	(.121)	(.0870)		
Capital goods	1.97	.497	2.29	.435	.875	2.34	. . .	. . .
	(.987)	(.709)	(.858)	(.381)	(.786)	(1.35)		
Transportation	1.15	.196	.875	.507	2.13	.831	. . .	. . .
	(.665)	(.272)	(.311)	(.202)	(.895)	(.463)		

Utilities	.282	.702	.609	.306	1.53	.833	. . .	. . .
	(.268)	(.660)	(.128)	(.0870)	(.420)	(.459)		
Textiles/trade	1.06	1.45	.917	.289	3.72	1.40	. . .	. . .
	(2.95)	(3.55)	(.400)	(.191)	(.725)	(.565)		
Services	1.37	.561	.730	.484	.527	.603	. . .	. . .
	(1.07)	(.804)	(.310)	(.269)	(.210)	(.254)		
Leisure	2.34	1.06	1.18	.310	1.52	.581	. . .	. . .
	(2.46)	(1.62)	(.263)	(.185)	(.472)	(.226)		
Industry averages	1.31	.699	1.04	.478	1.28	1.04	. . .	. . .
	(1.43)	(1.09)	(.432)	(.320)	(.451)	(.610)		
Size 1	.599	.240	.951	.203	1.81	.298	. . .	. . .
	(.243)	(.110)	(.148)	(.0580)	(.433)	(.238)		
Size 2	.745	.0810	.809	.119	1.36	.0540	. . .	. . .
	(.135)	(.0342)	(.0820)	(.0350)	(.152)	(.0290)		
Size 3	.815	.0330	.895	.0740	.718	.0350	. . .	. . .
	(.127)	(.0209)	(.0810)	(.0250)	(.0730)	(.00900)		
Size 4	.940	.0300	1.08	.122	.965	.00300	. . .	. . .
	(.148)	(.0342)	(.148)	(.0540)	(.0950)	(.00400)		
Size 5	1.02	.0010	1.32	.0680	.928	.00500	. . .	. . .
	(.105)	(.0038)	(.285)	(.0410)	(.107)	(.00600)		
Size 6	1.29	.0470	1.25	.123	.983	.0250	. . .	. . .
	(.245)	(.0513)	(.154)	(.0930)	(.115)	(.0140)		
Size 7	1.32	.0700	1.45	.304	.639	.0610	. . .	. . .
	(.335)	(.0646)	(.393)	(.160)	(.0740)	(.0250)		
Size 8	.879	.0240	1.30	.135	.872	.00700	. . .	. . .
	(.0837)	(.0076)	(.337)	(.0970)	(.149)	(.0130)		
Size 9	1.27	.122	1.13	.194	.641	.136	. . .	. . .
	(.519)	(.131)	(.178)	(.0750)	(.107)	(.0370)		
Size 10	.783	.569	.985	.652	1.29	.555	. . .	. . .
	(.797)	(.702)	(.686)	(.485)	(.805)	(.608)		
Size averages	.967	.122	1.12	.199	1.02	.118	. . .	. . .
	(.274)	(.116)	(.249)	(.112)	(.211)	(.0983)		
Overall averages	1.154	.437	1.076	.351	1.162	.653	. . .	. . .
	(.904)	(.647)	(.349)	(.225)	(.342)	(.377)		

NOTE.—VR1 = $\mathrm{var}(E[\Sigma_j \beta_j F_{jt} | Z_{t-1}])/\mathrm{var}(E[r|Z_{t-1}])$, and VR2 = $\mathrm{var}(E[r - \Sigma_j \beta_j F_{jt} | Z_{t-1}])/\mathrm{var}(E[r|Z_{t-1}])$, where the F_{jt} are the excess returns on the factors, r is the asset excess return, and Z_{t-1} are the lagged instruments. The one-factor model uses the first principal component as the factor. The variance ratios are estimated by the generalized method of moments. The asymptotic standard errors, adjusted for finite sampling bias by using the Ferson-Foerster correction factor, are in parentheses. Results for a four-factor model are reported in the annual data for the bottom panel.

TABLE 4 Estimates of the Fraction of the Predicted Variance of Common Stock Excess Returns That Is Explained by Constant-Beta Models, Using Mimicking Portfolios for Economic Factors

	Return Horizon							
	Monthly		Quarterly		Annual		2-Year	
Portfolio	VR1	VR2	VR1	VR2	VR1	VR2	VR1	VR2
A. One-factor models:								
Sample period 1927–88:								
Petroleum	.824	.393	1.27	.392	.541	.178	.875	.209
	(.461)	(.381)	(.640)	(.267)	(.493)	(.335)	(.448)	(.187)
Finance	1.20	.323	.794	.413	.746	.596	.644	.291
	(.404)	(.223)	(.414)	(.296)	(1.06)	(.588)	(.260)	(.165)
Consumer durables	1.07	.141	1.53	.242	.856	.169	.666	.301
	(.330)	(.117)	(.576)	(.319)	(.662)	(.207)	(.233)	(.183)
Basic industries	1.73	.393	1.33	.331	1.48	.470	1.39	.370
	(1.12)	(.397)	(.618)	(.274)	(1.08)	(.781)	(.759)	(.263)
Food/tobacco	.523	.658	.612	.577	.662	1.45	.271	.927
	(.309)	(.367)	(.318)	(.258)	(1.12)	(1.79)	(.160)	(.315)
Construction	.952	.166	.797	.272	.695	.634	.294	.330
	(.451)	(.162)	(.296)	(.169)	(.660)	(.673)	(.070)	(.058)
Capital goods	1.20	.177	1.63	.343	3.19	1.22	2.01	.408
	(.539)	(.181)	(.717)	(.336)	(5.45)	(2.91)	(1.31)	(.621)
Transportation	.358	.408	.570	.473	1.03	.259	.692	.373
	(.158)	(.186)	(.375)	(.361)	(.787)	(.435)	(.346)	(.212)
Utilities	.452	.727	.732	.433	.460	.757	.461	.610
	(.283)	(.448)	(.310)	(.423)	(.437)	(.783)	(.157)	(.227)
Textiles/trade	.554	.510	1.33	.912	1.82	1.51	.955	.643
	(.370)	(.537)	(.873)	(.642)	(3.14)	(3.06)	(.494)	(.386)
Services	.546	.373	1.29	.505	.888	.158	.349	1.25
	(.332)	(.270)	(.544)	(.285)	(.329)	(.137)	(.266)	(.623)
Leisure	.421	.312	.827	.581	1.29	.735	.765	.679
	(.161)	(.171)	(.575)	(.467)	(1.56)	(1.29)	(.427)	(.536)
Industry averages	.819	.382	1.06	.456	1.14	.677	.781	.533
	(.410)	(.287)	(.521)	(.341)	(1.40)	(1.08)	(.411)	(.315)

Size 1	.0610	.819	.233	.796	.788	.087	.267	.379
	(.0435)	(.113)	(.170)	(.270)	(.810)	(.193)	(.207)	(.142)
Size 2	.144	.669	.316	.669	.728	.083	1.02	.756
	(.0911)	(.129)	(.174)	(.245)	(.619)	(.178)	(.703)	(.687)
Size 3	.173	.593	.308	.662	.669	.083	.351	.370
	(.0839)	(.141)	(.194)	(.287)	(.511)	(.143)	(.174)	(.208)
Size 4	.283	.515	.671	.492	.787	.056	.278	.335
	(.169)	(.161)	(.272)	(.258)	(.542)	(.126)	(.127)	(.160)
Size 5	.274	.474	.512	.436	.866	.144	.817	.259
	(.115)	(.165)	(.239)	(.234)	(.814)	(.219)	(.392)	(.235)
Size 6	.328	.412	.690	.416	.918	.051	.888	.223
	(.136)	(.161)	(.311)	(.267)	(.638)	(.124)	(.341)	(.183)
Size 7	.585	.283	1.01	.232	1.10	.130	.806	.091
	(.279)	(.155)	(.309)	(1.95)	(.913)	(.248)	(.344)	(.122)
Size 8	.742	.131	.872	.181	.943	.159	.720	.079
	(.194)	(.0849)	(.288)	(.141)	(.689)	(.219)	(.322)	(.077)
Size 9	.585	.143	.766	.100	.924	.039	.548	.089
	(.191)	(.0963)	(.277)	(.0829)	(.441)	(.075)	.113	.046
Size 10	1.16	.0220	1.03	.0320	1.12	.028	1.18	.017
	(.115)	(.0135)	(.138)	(.0280)	(.265)	(.0387)	(.179)	(.014)
Size averages	.433	.406	.640	.402	.884	.0860	.688	.260
	(.142)	(.122)	(.237)	(.201)	(.624)	(.156)	(.290)	(1.87)
Overall averages	.644	.393	.869	.431	1.02	.409	.738	.409
	(.288)	(.212)	(.392)	(.277)	(1.05)	(.661)	(.356)	(.257)
Sample period 1961–88:								
Petroleum	1.10	.214	.696	.196	.434	.671	. . .	. . .
	(.481)	(.234)	(.339)	(.180)	(.203)	(.246)		
Finance	1.21	.192	1.10	.0890	1.04	.100	. . .	. . .
	(.313)	(.126)	(.283)	(.0814)	(.302)	(.0610)		
Consumer durables	.671	.0710	.684	.0510	.413	.155	. . .	. . .
	(.147)	(.0391)	(.229)	(.0420)	(.0860)	(.0730)		
Basic industries	1.71	.155	1.07	.107	1.62	.142	. . .	. . .
	(.447)	(.133)	(.245)	(.0814)	(.368)	(.0970)		
Food/tobacco	1.37	.353	1.49	.537	.588	.577	. . .	. . .
	(.343)	(.212)	(.841)	(.469)	(.353)	(.316)		
Construction	.769	.114	.722	.158	.315	.262	. . .	. . .
	(.197)	(.0609)	(.260)	(.0936)	(.112)	(.112)		

TABLE 4 *(Continued)*

	Return Horizon							
	Monthly		Quarterly		Annual		2-Year	
Portfolio	VR1	VR2	VR1	VR2	VR1	VR2	VR1	VR2
Capital goods	.961	.0430	.990	.0280	1.39	.289	. . .	. . .
	(.284)	(.0478)	(.262)	(.0366)	(.594)	(.245)		
Transportation	.670	.0800	.755	.0760	.303	.232	. . .	. . .
	(.161)	(.0532)	(.271)	(.0868)	(.155)	(.112)		
Utilities	.727	.557	.630	.350	.282	.540	. . .	. . .
	(.296)	(.281)	(.279)	(.218)	(.113)	(.183)		
Textiles/trade	1.06	.189	.983	.237	.903	.255	. . .	. . .
	(.353)	(.139)	(.439)	(.202)	(.545)	(.259)		
Services	.614	.143	.766	.165	1.60	.254	. . .	. . .
	(.185)	(.0685)	(.380)	(.110)	(.902)	(.303)		
Leisure	.969	.173	1.48	.249	1.60	2.29	. . .	. . .
	(.347)	(.118)	(.697)	(.286)	(1.53)	(1.76)		
Industry averages	.985	.190	.947	.187	.875	.480	. . .	. . .
	(.296)	(.126)	(.377)	(.157)	(.439)	(.314)		
Size 1	.0920	.726	.333	.609	.344	.293	. . .	. . .
	(.0402)	(.123)	(.175)	(.256)	(.247)	(.149)		
Size 2	.162	.581	.369	.482	.230	.372	. . .	. . .
	(.0630)	(.128)	(.183)	(.210)	(.137)	(.137)		
Size 3	.234	.501	.418	.353	.220	.356	. . .	. . .
	(.0902)	(.136)	(.201)	(.163)	(.133)	(.143)		
Size 4	.289	.405	.446	.314	.225	.381	. . .	. . .
	(.0967)	(.121)	(.205)	(.153)	(.122)	(.139)		
Size 5	.388	.311	.494	.236	.218	.349	. . .	. . .
	(.118)	(.111)	(.212)	(.125)	(.122)	(.142)		
Size 6	.478	.229	.560	.180	.232	.326	. . .	. . .
	(.129)	(.0869)	(.225)	(.104)	(.113)	(.128)		
Size 7	.571	.165	.712	.134	.404	.213	. . .	. . .
	(.135)	(.0717)	(.262)	(.0963)	(.172)	(.105)		

Size 8	.685	.0930	.773	.0670	1.36	.360	. . .	. . .
	(.133)	(.0500)	(.254)	(.0637)	(.477)	(.160)		
Size 9	.826	.0570	.942	.0420	.899	.119	. . .	. . .
	(.127)	(.0380)	(.228)	(.0420)	(.282)	(.069)		
Size 10	1.01	.0070	.992	.0060	1.09	.013	. . .	. . .
	(.0728)	(.00543)	(.0936)	(.0041)	(.127)	(.007)		
Size averages	.474	.308	.604	.242	.523	.278	. . .	. . .
	(.101)	(.0870)	(.204)	(.122)	(.193)	(.118)		
Overall averages	.752	.244	.791	.212	.715	.388	. . .	. . .
	(.207)	(.108)	(.298)	(.141)	(.327)	(.225)		
B. Five-factor models:*								
Sample period 1927–88:								
Petroleum	.939	.544	.509	.364	1.06	.266	. . .	. . .
	(.810)	(.728)	(.115)	(.188)	(.195)	(.0570)		
Finance	.993	.573	1.20	.360	.810	.497	. . .	. . .
	(.901)	(.607)	(.403)	(.205)	(.254)	(.168)		
Consumer durables	.843	.108	.813	.073	1.59	.237	. . .	. . .
	(.534)	(.184)	(.123)	(.044)	(.292)	(.0930)		
Basic industries	1.23	.238	1.84	.355	.628	.235	. . .	. . .
	(1.12)	(.454)	(.309)	(.185)	(.134)	(.0790)		
Food/tobacco	.627	.587	.608	.955	.507	1.40	. . .	. . .
	(.492)	(.506)	(.243)	(.393)	(.256)	(.447)		
Construction	.615	.211	.683	.174	.562	.507	. . .	. . .
	(.449)	(.257)	(.182)	(.079)	(.136)	(.147)		
Capital goods	1.09	.090	1.45	.615	1.04	.275	. . .	. . .
	(1.07)	(.122)	(.858)	(.506)	(.263)	(.0900)		
Transportation	.655	.230	1.20	.575	1.55	.400	. . .	. . .
	(.354)	(.219)	(.336)	(.506)	(.301)	(.136)		
Utilities	.581	.516	1.84	1.25	.635	1.40	. . .	. . .
	(.395)	(.376)	(1.19)	(.904)	(.310)	(.539)		
Textiles/trade	1.03	.266	1.17	.307	1.46	.933	. . .	. . .
	(1.02)	(.428)	(.460)	(.221)	(.766)	(.507)		
Services	.225	.537	1.13	1.02	1.76	.689	. . .	. . .
	(.213)	(.500)	(.923)	(.499)	(.671)	(.415)		
Leisure	.716	.229	1.19	.089	.675	.467	. . .	. . .
	(.553)	(.259)	(.177)	(.036)	(.172)	(.0990)		

TABLE 4 (*Continued*)

	Return Horizon							
	Monthly		Quarterly		Annual		2-Year	
Portfolio	VR1	VR2	VR1	VR2	VR1	VR2	VR1	VR2
Industry averages	.795	.344	1.14	.511	1.02	.609	. . .	. . .
	(.660)	(.387)	(.443)	(.314)	(.313)	(.231)		
Size 1	.0980	.688	.372	.580	1.59	.328	. . .	. . .
	(.0761)	(.188)	(.131)	(.143)	(.443)	(.193)		
Size 2	.141	.574	.493	.564	.733	.137	. . .	. . .
	(.112)	(.224)	(.150)	(.154)	(.130)	(.0540)		
Size 3	.165	.513	.695	.132	1.03	.137	. . .	. . .
	(.120)	(.234)	(.106)	(.0380)	(.175)	(.0520)		
Size 4	.258	.425	.798	.161	.801	.130	. . .	. . .
	(.181)	(.228)	(.155)	(.0640)	(.155)	(.0480)		
Size 5	.276	.393	.839	.170	1.11	.224	. . .	. . .
	(.203)	(.264)	(.162)	(.0670)	(.212)	(.0930)		
Size 6	.384	.291	1.00	.0690	1.04	.126	. . .	. . .
	(.249)	(.226)	(.123)	(.0260)	(.189)	(.0710)		
Size 7	.683	.121	1.04	.0600	.775	.166	. . .	. . .
	(.272)	(.127)	(.149)	(.0410)	(.157)	(.0730)		
Size 8	.896	.092	1.20	.0670	.395	.360	. . .	. . .
	(.376)	(.124)	(.225)	(.0600)	(.0750)	(.0640)		
Size 9	.749	.080	1.03	.0910	.588	.209	. . .	. . .
	(.350)	(.127)	(.166)	(.0600)	(.108)	(.0650)		
Size 10†	. . .	. . .	1.06	.00400	1.34	.0090	. . .	. . .
	. . .	. . .	(.114)	(.00600)	(.190)	(.0040)		
Size averages‡	.406	.353	.829	.210	.895	.202	. . .	. . .
	(.215)	(.194)	(.152)	(.0726)	(.183)	(.0792)		
Overall averages‡	.628	.348	1.00	.382	.968	.435	. . .	. . .
	(.469)	(.304)	(.318)	(.210)	(.257)	(.166)		

NOTE.—VR1 = $\text{var}(E[\Sigma_j \beta_j F_{jt} | Z_{t-1}])/\text{var}(E[r | Z_{t-1}])$ and VR2 = $\text{var}(E[r - \Sigma_j \beta_j F_{jt} | Z_{t-1}])/\text{var}(E[r | Z_{t-1}])$, where the F_{jt} are the excess returns on the factors, r is the asset excess return, and Z_{t-1} are the lagged instruments. The one-factor models use the excess return of the Standard and Poor's 500 index as the factor. The variance ratios are estimated by the Generalized Method of Moments (GMM). The asymptotic standard errors, adjusted for finite sample bias by using the Ferson-Foerster correction factor, are in parentheses.

* A four-factor model is shown for annual data, where the real interest rate factor is not included.

† The GMM system was singular for the monthly data.

‡ Size and overall averages do not include the largest decile of firms for the monthly results.

that is not captured by the model, relative to the total predictable variance.[15] The A panels of tables 3 and 4 report the variance ratios for one-factor models, and the B panels summarize results for the five-factor models. To conserve space, the five-factor models are shown only for the full sample period. The results for the second subperiod are similar.

Given recently renewed interest in the empirical performance of the CAPM, the one-factor results are interesting. The one-factor models use the first principal component as the factor in table 3, and the SP500 excess return is the factor in table 4. Although the standard errors indicate that the precision with which the variance ratios are estimated is often low, several patterns emerge. First, using a single principal component in table 3, the ratios VR1 are in most cases larger than the ratios VR2. This indicates that the models capture more of the predictable variance of the returns than they leave in the residuals. The standard errors confirm this impression. For example, in the full sample period, 84 of the 176 estimates of VR1 are more than 2 standard errors from zero. Only 44 of the VR2s are more than 2 standard errors from zero. These differences are more pronounced in the second subperiod, where 106 of the VR1s are larger than 2 standard errors and only 36 of the VR2s are more than 2 standard errors in magnitude.

A second impression from the one-factor models in tables 3 and 4 is that the explanatory power of the models is broadly similar for all of the return horizons. This is an important finding, since beta pricing models may be used in practical applications that involve long investment horizons. Our results provide some of the first direct empirical evidence on the performance of the models for long-horizon conditional expected returns.

There are some interesting differences in the performances of the one-factor, principal-component model, and the single-factor model using the SP500. The "size effect," as represented by the returns to portfolios of small firms, has long presented a challenge to the CAPM. The variance ratios reveal a size effect in the SP500 model for the conditional returns. In monthly and quarterly data, the ratios VR1 are often less than the VR2s for the smaller size portfolios, and most of the statistically significant VR2s for size portfolios occur in the smaller

---

[15] While we could simply define VR2 = 1 − VR1 and avoid the separate estimation, the separate estimates provide additional information because the two variance ratios are correlated. The ratios are correlated for the same reason that the fitted values from regressing Y on Z and from regressing X on Z are correlated, even if Y and X are uncorrelated (see Ferson and Harvey [1991] for more discussion). When the sum of the variance ratios, VR1 + VR2, is less than 1.0, the correlation is positive; if it is greater than 1.0, the correlation is negative. It would be possible, in principle, to append additional moment conditions to the system to estimate VR1 and VR2 simultaneously. Given the small sample sizes implied by our focus on long-horizon returns, we do not pursue this or other potential refinements of the approach in this article.

size deciles. In contrast, the one-principal-component model performs well for the size portfolios at all return horizons. However, the explanatory power of the one-principal-component model for the industry portfolios at the longer return horizons is not as good as the SP500 model. The preferred method of factor selection in a one-factor model is therefore ambiguous, as the performance depends on the assets under study.

We use the GMM goodness-of-fit statistic to examine the fit of the models with the constant beta assumption. The single-factor models can be rejected. However, rejections are found only for the monthly and quarterly data and only for the longer sample period. In these cases, the tests produce right-tail p-values smaller than 0.05 for 11 (nine) of the industries and eight (five) size portfolios using monthly (quarterly) data. The tests do not reject the single-factor models with constant betas for the second subperiod or for annual or 2-year returns.

The B panels of tables 3 and 4 summarize five-factor models.[16] Similar to the single-factor case, we find that the models with the constant-beta specification can be rejected over the full sample period. The individual-asset p-values are less than 0.05 in about two-thirds of the cases. However, there are relatively few instances where the model is rejected in the second subperiod. Compared with the single-factor models, the estimates of the variance ratios appear more precise, and the ability of the model to explain the predictable variance is improved. This is especially so for the industry portfolios. The multiple-factor results using the economic variables are similar to the results using the principal components. Taken together with table 2, these results show that the multiple-beta models perform similarly with either choice of risk factors. This is an interesting result, given that previous research provides little information about which method of factor selection is better for longer-horizon returns. Our results suggest that the performance of the two methods is similar for capturing time variation in the expected returns of equity portfolios.

VI. Robustness of the Results

A. Alternative Econometric Methods

Ferson and Harvey (1991) estimate variance ratios using a multistep generalization of the approach of Fama and MacBeth (1973). The portfolio betas are first obtained from rolling time-series regressions of returns on the factors, and mimicking portfolios are formed as the

---

[16] In annual data, four-factor models are used because the system is too large to estimate with five factors. The first four principal components are used, and the real interest rate factor is dropped from the economic-factor models.

coefficients in cross-sectional regressions of the returns on the lagged betas. The model-fitted part of the return is the product of the betas and the mimicking portfolio excess returns. The difference between the return and its model-fitted component is the model residual. Variance ratios are calculated by regressing the returns and their two components separately over time on the lagged instruments and taking ratios of the sample variances of the fitted values. This approach has the advantage of allowing (ad hoc) time variation in betas, and it imposes the same risk premium for each asset return in the cross-section. However, the statistical properties of the approach are not well understood.

Using the methods of Ferson and Harvey (1991) on our data confirms our finding that five-factor models capture much of the predictable variation in the monthly and quarterly asset returns (details of these results are available on request). Other main features of the results confirm our evidence based on the GMM system (4). For example, the performance of the principal-component models for the industry portfolios is worse than it is for the size portfolios.

We also use a multistep, time-series regression approach to estimate variance ratios in models with constant betas. This provides a further check on the sensitivity of the results to the econometric technique, while using the constant beta assumption as in the GMM system. In the first step we decompose each asset return into model-fitted and model-residual parts by regressing each return on the factors. The fitted values of the regression are the model-fitted component of the return, and the residual from the time series regression is the model-residual component. Each component is regressed on the lagged instruments in a second step. The variance ratios are calculated in a third step, from the sample variances of the regressions on the lagged instruments. These experiments produce a similar impression of the model's performance as when using the GMM with fixed betas.

In a third experiment we construct alternative mimicking portfolios for two of the economic variables using the time-series methods of Breeden, Gibbons, and Litzenberger (1989). We repeat the monthly, multistep, time-series regression approach and find that the overall results are similar.

Although data limitations do not allow us to pool information across assets for estimating the GMM variance ratios, we conduct some informal cross-asset analysis of the ratios. We first compute time-series correlations between the model-fitted expected returns, which are the betas multiplied by the expected factor premia, and the "unrestricted" expected returns of the assets, which are the fitted values of the regressions in table 1. These correlations should be high if the model does a good job of tracking the predictable component of the return over time. They range from less than 0.4 to larger than 0.99, with a mean of .806

in the one-factor models and .819 in the five-factor models. We should expect that the correlations are higher in cases where the variance ratio VR1 is relatively high, if the variance ratios provide good cross-sectional measures of the model's ability to fit conditional returns. The cross-sectional correlations between the time-series correlations and the ratios VR1/(VR1 + VR2) in monthly (quarterly, annual, 2-year) data are 0.83 (0.85, 0.65, 0.80), so the variance ratios and the correlations present a similar picture. This increases our confidence that the variance ratios are informative measures of the model performance.

B. The Effects of Changing Betas

If the betas relative to a given set of factors vary over time and are correlated with Z_{t-1}, then the model residuals may be correlated with Z_{t-1} when constant betas are assumed. Our GMM tests reject the models with the constant-beta specification for monthly and quarterly data over the full sample period. We therefore allow for time variation in the betas. We modify the system (4), approximating the conditional betas by linear functions of the lagged Zs and substitute this linear function into the moment condition (4c). The new system has equal numbers of parameters and orthogonality conditions, so it is exactly identified. The overall impression is that the performance of the model is not improved by allowing the betas to vary over time. However, the standard errors are large, which makes it difficult to draw precise inferences from these models.

We also compute direct measures of the contribution of time-varying betas, using the cross-sectional methods of Ferson and Harvey (1991). We decompose the variance of the model fitted part of the expected return for each asset as

$$\begin{aligned}\operatorname{var}\{E(\beta'\gamma|Z)\} = {} & E(\beta)' \operatorname{var}\{E(\gamma|Z)\}E(\beta) \\ & + E(\gamma)' \operatorname{var}\{E(\beta|Z)\}E(\gamma) \\ & + \text{interaction effects,}\end{aligned} \tag{7}$$

where $E(\cdot)$ and var$\{\cdot\}$ are estimated by the sample means and variances, and $E(\cdot|Z)$ is estimated by linear regression on the lagged Zs. (Tables of these results are available on request.) In the monthly returns for our second subperiod, the decomposition attributes less than 1% of the predicted variation to movements in the betas. A small beta effect is similar to the findings of Ferson and Harvey (1991), who examined monthly data for 1964–86. A small beta effect is also consistent with the results of our GMM tests of the constant-beta models, which do not often reject the models in the second subperiod. The interaction effects between time-varying betas and risk premiums are less than one-third of the variance for every asset but usually are very close to zero. The interaction effects arise because the betas and the risk premiums move together through time.

Over the full sample period, the direct contribution of the time-varying betas is small, but 40% or more of the expected return variation is attributed to interaction effects. The covariation between time-varying betas and expected risk premiums seems to be an important component of monthly time-varying expected returns over the early parts of the sample. This is consistent with the results of our GMM tests, which reject the models with constant betas for the full sample in monthly data.

The large interaction effects over the longer sample period are interesting, as they reflect an important distinction between conditional and unconditional models. If there are no interaction effects, then a conditional beta pricing model implies an unconditional model using the expected betas and the unconditional expected risk premiums for the same factors. Covariances between conditional betas and risk premiums, which cause interaction effects, imply additional terms in the cross-sectional relation between unconditional mean returns and average betas (unless the covariances are perfectly correlated with the unconditional betas). Jagannathan and Wang (1994) note that covariation between monthly betas and expected risk premiums can be empirically important for explaining the cross-section of unconditional expected returns using a version of the Capital Asset Pricing Model.[17]

VII. Conclusions

We present an approach for estimating how much of the predictability in security returns is explained by an asset-pricing model and use the approach to evaluate conditional models for multiple return horizons. We compare the performance of economic factors, similar to Chen et al. (1986), and principal components, similar to Connor and Korajczyk (1986, 1988*a*, 1988*b*). Ours is the first study of these models using long holding period returns, focusing on conditional returns and risk. We find that a large fraction of the predictability in returns can be explained by the models, for all of the investment horizons. Our statistical tests reject a constant-beta assumption for the shorter-horizon returns over long sample periods but cannot reject a constant-beta model for long-horizon returns. When we allow for time variation in the betas,

---

[17] We repeat the varying-beta analysis using quarterly rates of return. Sixty past quarters are used to obtain the betas for each cross-sectional regression. Once again,the decomposition implies that time variation in the expected risk premiums is much more important than variation in the betas. On average, the risk premiums account for more than 90% of the predictability, and the direct beta effects account for less than 10%. Although there are interaction effects for some of the industries, the magnitudes of the quarterly interaction effects are generally small.

we confirm that independent movements in betas are less important than risk premiums for time-varying expected portfolio returns. Single- and multiple-beta models are compared, and the multiple-beta models perform better. Models using five economic variables, and models using five principal components as the risk factors, have similar overall ability to explain the predictability in portfolio returns.

Appendix

Mimicking Portfolios and Simulations to Evaluate the Regression Tests

Mimicking Portfolios

When an economic risk factor is measured in the form of an excess return, such as the SP500, the term structure, and default risk variables, we use the excess return directly as a mimicking portfolio. Shanken (1992) argues that such an approach delivers the most efficient estimate of the risk premiums. When the economic factor is not an excess return, such as the real interest rate and inflation risk variables, we construct mimicking portfolios as follows. We run time-series regressions of individual common stock returns on the economic variables and the lagged instruments. The slope coefficients on the economic variables are the estimates, B, of the $K \times n$ matrix of conditional betas, assuming that the betas are constant parameters. The regression residuals are estimates of the conditional nonfactor-related, or idiosyncratic, returns. We use the residuals to form estimates of the $n \times n$ conditional idiosyncratic variance matrix, V. We assume for this purpose that the matrix is diagonal, as suggested by Lehmann and Modest (1988). The regressions use all available data for each firm. As firms enter and leave the sample, the dimensions of B and V change. For each of the months in the sample we form the factor-mimicking portfolios by finding a weight vector w_j that satisfies

$$\min w_j' V w_j \quad \text{subject to } w_j' B_{[j]}' = \underline{0} \text{ and } w_j' \underline{1} = 1, \tag{A1}$$

where $B_{[j]}$ excludes the jth row from B, $\underline{0}$ is a $K - 1$ vector of zeros, $\underline{1}$ is an n vector of ones, and w_j is an n vector. The jth mimicking portfolio is formed from the individual stocks, using the portfolio weight w_j. The conditional beta of the jth mimicking portfolio on the jth economic factor may change as B and V change over time. We adjust the mimicking portfolios to have constant factor betas by combining them with the Treasury bill so that the combined portfolio has a beta equal to the time-series average of the betas that are produced by the constrained optimizations.

Simulations to Evaluate the Regression Tests

We construct simulated asset returns and factor-mimicking portfolios which satisfy the null hypothesis that the lagged instruments have coefficients equal to zero in the regression (3). We use a bootstrapping procedure similar to Efron (1982). Let $X = [r, F]$ denote the $T \times (N + K)$ data matrix of the excess returns on the N assets and K factor-mimicking portfolios. For each trial τ ($\tau = 1, \ldots, 1{,}000$) of a simulation we construct X_τ by sampling the rows of X, randomly with replacement. For each trial, we sample a number of rows equal to the sample size of the regression that we wish to evaluate. This procedure preserves the cross-sectional relation between the test assets and the factors. For each regression in the simulations, we use the actual sample of the instruments, Z, as the regressors. This guarantees that in the simulations, the error terms of the regression (3) with the instruments excluded will have conditional mean zero given Z. This also implies that the artificial data will not display conditional heteroscedasticity as a function of Z. We do not think that this is a major factor in our data, since ordinary least squares and White's (1980) heteroscedasticity-consistent standard errors for the coefficients of the actual regressions are similar in magnitude. Reusing the instruments retains their time-series properties but ignores the random aspect of the instruments across simulation trials. For each trial of the simulation, we estimate regression (3) using the artificial data, and we calculate the F-test and χ^2 statistics.

References

Black, Fischer. 1993. Estimating expected return. *Financial Analysts Journal* (September–October): 36–38.

Blume, M.; Keim, D. B.; and Patel, S. A. 1991. Returns and volatility of low-grade bonds, 1977–1989. *Journal of Finance* 46:49–74.

Bollerslev, Tim; Chou, Ray Y.; and Kroner, Kenneth. 1992. ARCH modeling in finance: A review of the theory and empirical evidence. *Annals of the Journal of Econometrics* 52:5–60.

Boudoukh, Jacob, and Richardson, Matthew. 1994. The statistics of long-horizon regressions revisited. *Mathematical Finance* 4:103–19.

→ Breeden, Douglas T. 1979. An intertemporal asset pricing model with stochastic consumption and investment opportunities. *Journal of Financial Economics* 7:265–96.

Breeden, Douglas T.; Gibbons, Michael R.; and Litzenberger, Robert H. 1989. Empirical tests of the consumption-oriented CAPM. *Journal of Finance* 44:231–62.

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

Chen, Nai-fu.; Roll, Richard; and Ross, Stephen A. 1986. Economic forces and the stock market. *Journal of Business* 59:383–403.

Christie, Andrew A. 1990. Aggregation of test statistics. *Journal of Accounting and Economics* 12:15–36.

Cochrane, John H. 1992. A cross-sectional test of a production-based asset pricing model. Working paper. Chicago: University of Chicago.

Connor, Gregory. 1984. A unified beta pricing theory. *Journal of Economic Theory* 34:13–31.

Connor, Gregory, and Korajczyk, Robert A. 1986. Performance measurement with the arbitrage pricing theory: A new framework for analysis. *Journal of Financial Economics* 15:373–94.

Connor, Gregory, and Korajczyk, Robert A. 1988*a*. Estimating pervasive economic factors with missing observations. Working Paper no. 34. Evanston, Ill.: Northwestern University, Department of Finance.

Connor, Gregory, and Korajczyk, Robert A. 1988*b*. Risk and return is an equilibrium APT: Application of a new test methodology. *Journal of Financial Economics* 21: 255–89.

Connor, Gregory, and Korajczyk, Robert A. 1989. An intertemporal equilibrium beta pricing model. *Review of Financial Studies* 2, no. 3:255–89.

Douglas, G. W. 1969. Risk in the equity markets: An empirical appraisal of market efficiency. *Yale Economic Essays* 9 (Spring): 3–45.

Efron, Bradford. 1982. *The Jackknife, the Bootstrap, and Other Resampling Plans.* Monograph, no. 38. Philadelphia: Society for Industrial and Applied Mathematics.

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

Fama, Eugene F., and French, Kenneth R. 1988*a*. Dividend yields and expected stock returns. *Journal of Financial Economics* 22:3–25.

Fama, Eugene F., and French, Kenneth R. 1988*b*. Permanent and temporary components of stock prices. *Journal of Political Economy* 96:246–73.

Fama, Eugene F., and French, Kenneth R. 1989. Business conditions and expected stock returns. *Journal of Financial Economics* 25:23–50.

Fama, Eugene F., and French, Kenneth R. 1992. The cross-section of expected stock returns. *Journal of Finance* 47, no. 2:427–66.

Fama, Eugene F., and MacBeth, James D. 1973. Risk, return and equilibrium: Empirical tests. *Journal of Political Economy* 81:607–36.

Fama, Eugene F., and Schwert, G. William. 1977. Asset returns and inflation. *Journal of Financial Economics* 5:115–46.

Ferson, Wayne E. 1989. Changes in expected security returns, risk and the level of interest rates. *Journal of Finance* 44:1191–1218.

Ferson, Wayne E. 1990. Are the latent variables in time-varying expected return compensation for consumption risk? *Journal of Finance* 45:397–430.

Ferson, Wayne E., and Foerster, Stephen R. 1994. Finite sample properties of the generalized method of moments in tests of conditional asset pricing models. *Journal of Financial Economics* 36:29–56.

Ferson, Wayne E.; Foerster, Stephen R.; and Keim, Donald B. 1993. General tests of latent variable models and conditional mean variance spanning. *Journal of Finance* 48:131–56.

Ferson, W., and Harvey, Campbell R. 1991. The variation of economic risk premiums. *Journal of Political Economy* 99:385–415.

Ferson, W. E., and Harvey, Campbell R. 1993. The risk and predictability of international equity returns. *Review of Financial Studies* 6:527–66.

Foster, Douglas, and Smith, Thomas. 1992. Assessing the goodness-of-fit of asset pricing models: The distribution of the maximal R^2. Working paper. Durham, N.C.: Duke University.

Goetzmann, William, and Jorion, Philippe. 1993. Testing the predictive power of dividend yields. *Journal of Finance* 48:663–80.

Handa, Puneet; Kothari, S. P.; and Wasley, Charles. 1989. The relation between the return interval and beta: Implications for the size effect. *Journal of Financial Economics* 23:79–100.

Hansen, Lars P. 1982. Large sample properties of generalized method of moments estimators. *Econometrica* 50:1029–54.

Hansen, Lars P., and Hodrick, Robert J. 1980. Forward exchange rates as predictors of future spot rates: An econometric analysis. *Journal of Political Economy* 88:829–53.

Harvey, Campbell R., and Zhou, Guofu. 1990. Bayesian inference in asset pricing tests. *Journal of Financial Economics* 26:221–54.

Hodrick, Robert J. 1992. Dividend yields and expected stock returns: Alternative procedures for inference and measurement. *Review of Financial Studies* 5, no. 3:357–86.

Huberman, Gur, and Kandel, Shmuel A. 1987. Mean-variance spanning. *Journal of Finance* 42:383–88.

Huberman, Gur; Kandel, Shmuel A; Stambaugh, Robert F. 1987. Mimicking portfolios and exact arbitrage pricing. *Journal of Finance* 42:1–10.

Jagannathan, Ravi, and Wang, Zhenyu. 1994. The CAPM is alive and well. Working paper. Minneapolis: University of Minnesota.

Kandel, Shmuel A., and Stambaugh, Robert F. 1990. Expectations and volatility of consumption and asset returns. *Review of Financial Studies* 3, no. 2:207–32.

Keim, Donald B. 1983. Size-related anomalies and stock return seasonality: Further empirical evidence. *Journal of Financial Economics* 12:13–32.

Keim, Donald B., and Stambaugh, Robert F. 1986. Predicting returns in the bond and stock markets. *Journal of Financial Economics* 17:357–90.

Kirby, Christopher M. 1993. Measuring the predictable variation in stock and bond returns. Working paper. Durham, N.C.: Duke University.

Klein, R. W., and Brown, S. J. 1984. Model selection when there is "minimal" prior information. *Econometrica* 52:1291–1312.

Korajczyk, R., and Viallet, Claude J. 1989. An empirical investigation of international asset pricing. *Review of Financial Studies* 2:553–85.

Kothari, S. P.; Shanken, Jay; and Sloan, Richard G. 1993. Another look at the cross-section of expected stock returns. Working paper. Rochester: University of Rochester.

Lehmann, Bruce N., and Modest, David M. 1988. The empirical foundations of the arbitrage pricing theory. *Journal of Financial Economics* 21:213–54.

Levhari, David, and Levy, H. 1977. The capital asset pricing model and the investment horizon. *Review of Economics and Statistics* 59:92–104.

Lo, Andrew W., and MacKinlay, A. Craig. 1988. Stock market prices do not follow random walks: Evidence from a simple specification test. *Review of Financial Studies* 1:41–66.

→ Lo, Andrew W., and MacKinlay, A. Craig. 1990. Data snooping biases in tests of financial asset pricing models. *Review of Financial Studies* 3, no. 3:431–68.

Long, John B. 1974. Stock prices, inflation, and the term structure of interest rates. *Journal of Financial Economics* 1:131–70.

Longstaff, Francis. 1989. Temporal aggregation and the continuous-time CAPM. *Journal of Finance* 44:871–88.

McCulloch, Robert, and Rossi, Peter E. 1989. Posterior, predictive and utility-based approaches to testing the arbitrage pricing theory. Working paper no. 269. Chicago: University of Chicago Graduate School of Business.

McCulloch, Robert, and Rossi, Peter E. 1990. Posterior, predictive and utility-based approaches to testing the arbitrage pricing theory. *Journal of Financial Economics* 28:7–38.

Merton, Robert C. 1973. An intertemporal capital asset pricing model. *Econometrica* 41:867–87.

Miller, Merton H., and Scholes, M. 1972. Rates of return in relation to risk: A re-examination of some recent findings. In M. C. Jensen (ed.), *Studies in the Theory of Capital Markets*. New York: Praeger.

Nelson, Charles R., and Kim, Myung J. 1993. Predictable stock returns: The role of small sample bias. *Journal of Finance* 48:641–61.

→ Poterba, J., and Summers, L. 1988. Mean reversion in stock prices: Evidence and implications. *Journal of Financial Economics* 22:27–60.

Richardson, Matthew. 1993. Temporary components of stock prices: A skeptic's view. *Journal of Business and Economic Statistics* 11:199–207.

Richardson, Matthew, and Smith, Thomas. 1991. Tests of financial models in the presence of overlapping observations. *Review of Financial Studies* 4, no. 2:227–54.

Richardson, Matthew, and Stock, James H. 1989. Drawing inferences from statistics based on multiyear asset returns. *Journal of Financial Economics* 25, no. 2:323–48.

Roll, Richard. 1977. A critique of the asset pricing theories' tests, part 1: On past and potential testability of the theory. *Journal of Financial Economics* 4:129–76.

→ Ross, S. 1976. The arbitrage theory of capital market asset pricing. *Journal of Economic Theory* 13:341–60.

→ Schwartz, G. 1978. Estimating the dimension of a model. *Annals of Statistics* 6:461–64.

Shanken, Jay. 1990. Intertemporal asset pricing: An empirical investigation. *Journal of Econometrics* 45:99–120.

Shanken, Jay. 1992. On the estimation of beta pricing models. *Review of Financial Studies* 5:1–34.

Sharpe, William F. 1964. Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance* 19:425–42.

Stambaugh, Robert F. 1983. Arbitrage pricing with information. *Journal of Financial Economics* 12, no. 3:357–70.

Wei, K. C. John; Lee, Cheng-few; and Chen, Andrew H. 1991. Multivariate regression tests of the arbitrage pricing theory: The instrumental variables approach. *Review of Quantitative Finance and Accounting* 1:191–208.

Wheatley, Simon. 1989. Testing asset pricing models with infrequently measured factors. Working paper. Seattle: University of Washington.

White, H. 1980. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica* 48:817–38.