TESTS OF ASSET PRICING MODELS WITH CHANGING EXPECTATIONS

by

Wayne E. Ferson
Stephen R. Foerster
Donald B. Keim

1-91

RODNEY L. WHITE CENTER FOR FINANCIAL RESEARCH
The Wharton School
University of Pennsylvania
Philadelphia, PA 19104-6367

Tests of Asset Pricing Models with Changing Expectations

Wayne E. Ferson
Graduate School of Business, University of Chicago

Stephen R. Foerster
School of Business Administration, University of Western Ontario

Donald B. Keim
Wharton School, University of Pennsylvania

November 1986
Revised: December 1990
Comments are welcome.

We thank Phillip Braun, Robert Korajczyk, Andrew Lo, Craig MacKinlay, an anonymous referee, and seminar participants at Northwestern University and the Western Finance Association meetings for helpful comments. The research assistance of Raymond Chan and Edward Nelling is gratefully acknowledged. Financial support (Keim) was provided by the Geewax-Terker Research Program in Financial Instruments at the Wharton School.

Tests of Asset Pricing Models with Changing Expectations

ABSTRACT

This paper studies predictable variation through time in the returns of common stocks and bonds, using linear models with constant "beta" coefficients. Both individual stocks and portfolios, formed on the basis of firm size and industry, are examined. The methods of Gibbons and Ferson (1985) are nested in a sequence of more general latent variables models. This generalization allows the examination of the separate elements of their joint hypothesis. The results indicate that more than a single premium is needed to model expected returns. The number of latent variables in the time-varying expected returns is similar for daily and monthly returns, and is small. Two or three latent variables are indicated in all cases. There is strong evidence that conditional expected risk premiums are nonzero in months other than January.

In this paper we provide further evidence on the predictable behavior of stock and bond portfolio returns using a framework consistent with multiple state variables in a "beta" pricing model. If several state variables explain expected returns, then the identification of the state variables is not unique in the usual formulations of financial valuation models. Yet the models can still impose structure on the expected returns. Expected returns differ across assets, according to a linear relation in the assets' conditional beta coefficients. The betas measure the sensitivity to the relevant state variables. Thus, it is useful to ask if the behavior of expected returns is consistent with simple linear models, even if the exact nature of the state variables is unknown. If the number of "latent variables" required to capture the predictability of the returns is small, the results are encouraging for the potential usefulness of conditional linear asset pricing theories.

This paper examines time series regressions for returns, using predetermined information variables, and conducts tests of the restrictions that are implied by linear asset pricing models. The methodology follows Gibbons and Ferson (GF, 1985). They do not reject a "single-beta" model of the daily returns for the Dow Jones 30 common stocks. Subsequent studies using stock portfolios and bond returns find evidence that more complicated models are required.[1] We apply the GF methodology to obtain further insights about the ability of simple models to capture the predictable patterns in returns.

We first extend the GF design to include more data. We find that a single-beta model can be rejected even for the Dow Jones 30 common stocks.

---

[1] Examples include Campbell (1987), Chan (1988), Stambaugh (1988) and Ferson (1989, 1990).

By varying the econometric specification, we produce evidence about the sensitivity of the tests to different elements of the joint hypothesis examined by GF. Specifically, we examine the assumptions that expected returns are linear and that the conditional covariance matrix is fixed over time. We find evidence against the homoskedasticity assumption used by GF; however, the failure of the single beta model to explain the Dow Jones 30 is not attributed to the failure of their maintained assumptions.

Tests using common stock and bond portfolios provide further evidence on the number of common factors in expected returns. Similar evidence is found in an industry portfolio design and in a design based on size-ranked portfolios. Two or three latent variables are indicated in each case. The results for monthly and daily data are also similar, when we allow for heteroskedasticity and nonlinearities in the returns generating process. A by-product of the analysis is strong evidence that expected risk premiums are non-zero and time-varying, and that the time-variation is not confined to January.

In Section I we review the methodology. Section II examines the sensitivity of inferences, in the context of the GF empirical example, to the choice of instrumental variables and sample period. Section III provides the analysis of monthly returns. Section IV summarizes and concludes.

I. Methodology

The essence of the approach is to test restrictions that are imposed by an asset pricing model with the familiar form[2]

$$E(R_{it}|Z_{t-1}) = [1- \Sigma_{i=1}^{K} b_{ih}]\ E(R_{0t}|Z_{t-1}) + \Sigma_{i=1}^{K} b_{ih}\ E(\lambda_{ht}|Z_{t-1}) \ ; \qquad (1)$$

where λ_{ht} ≡ one of k (possibly unobservable) risk factor-mimicking portfolios;

b_{ih} ≡ risk measure ("beta") of security i relative to risk factor h, conditional on the information Z_{t-1} (The beta is assumed to be a constant parameter); and

R_{0t} ≡ the return on a "zero-beta" security.

For any set of asset returns, there will almost always exist a mean-variance efficient portfolio [Roll (1977)]. This means that equation (1) will hold, with K=1, when λ_{ht} is the return of the efficient portfolio. But the betas (b_{ih}) of assets relative to the efficient portfolio will in general be time-varying when (1) describes conditional expected returns. Following much of the empirical work on the Capital Asset Pricing Model (CAPM) and Arbitrage Pricing Theory (APT), equation (1) assumes the b_{ih} are constant and specifies an asset pricing hypothesis of a given dimension, K. We call the hypothesis that equation (1) holds with constant betas, for a given K ≥ 1, the "**K Latent Variable Model.**"[3] If a K latent variable model is accepted, the results may be interpreted as indicating the number of

[2] Examples of asset pricing models like equation (1) include those of Sharpe (1964), Black (1972), Merton (1973), Long (1974), Ross (1976) and Breeden (1979).

[3] It is sufficient to assume that ratios of the conditional betas are constant parameters. See Connor and Korajczyck (1989) for an equilibrium model in which conditional betas are constant over time. See Lehmann (1990) for conditions under which approximately constant factor betas are implied in a conditional APT setting.

"factors," or time-varying risk premiums in the expected returns. An alternative interpretation is to view the tests as indicating the behavior of conditional covariances of returns with a benchmark pricing variable.[4]

Given rational expectations in a K latent variable model, the GF approach derives its power from the assumption that expected returns are changing over time and are correlated with observable instruments Z_{t-1}. Assume that conditional expected returns given Z_{t-1} are linear with fixed coefficients, so that returns obey the regression model:

$$R_{it} = \delta_i' Z_{t-1} + u_{it} \; ; \qquad E(u_{it} \mid Z_{t-1}) = 0, \tag{2}$$

where Z_{t-1} is an L-vector of predetermined variables, contained in the market's information set at time t - 1, and which includes a constant term. δ_i is the regression coefficient vector for asset i. Equation (2) will be called the **"Linear Expectations Assumption."** Given the linear expectations assumption, then the following expression

$$E(R_{it} \mid Z_{t-1}) = \delta_i' Z_{t-1} \tag{3}$$

may be substituted into Equation (1). Gibbons and Ferson (1985) show that the following parameter restrictions on system (2) are implied:

---

[4] See Gibbons and Ferson (1985), Campbell (1987), Ferson (1989) and Wheatley (1989) for discussions.

$$\delta_i = \Sigma_{j=0}^{K} c_{ij}\,\delta_j\,, \tag{4}$$
$$1 = \Sigma_{j=0}^{K} c_{ij};\quad i = K + 1, \ldots, N.$$

The δ_j ; j = 0, ..., K are the regression coefficients for K + 1 assets chosen as "**reference assets.**"[5] The restriction (4) states that the coefficients of all N + 1 assets may be replicated from only K + 1 assets if the K-latent variable model characterizes expected returns. The c_{ij} may be interpreted as ratios of the betas for assets i and j in equation (1).[6] The restriction that they must sum to 1.0 for each asset follows from the fact that $[1 - \Sigma_{i=1}^{K} b_{ih}]$ is the coefficient on the "zero-beta" factor in equation (1). The information variables, $\mathbf{Z}_{t-1}$, should be correlated with changes in investor expectations and must be known when the market sets prices at t - 1. The number of information variables L must equal at least the number of latent variables, K.

GF implement their tests as a restricted multivariate regression model for a system of regression equations like (2). They assume that the residual covariance matrix is fixed over time and examine the likelihood ratio test statistic (LRT). Thus, their tests examine a joint hypothesis which we characterize as consisting of three parts. The first is the specification of K, the dimension of the latent variable model (the K latent variable

---

[5] The reference assets must be chosen so that the matrix of their betas and a unit vector is nonsingular; that is, they must span all the relevant risk factors and cannot have identical betas on any combination of risk factors. Given these conditions, the tests are invariant to the choice of reference assets [see Ferson (1990) or Ferson and Foerster (1990)].

[6] This is strictly true if the unobserved risk factors are mutually uncorrelated and the reference assets are mutually uncorrelated; otherwise, the c_{ij} are related to the assets' betas by a linear transformation.

hypothesis). The second is the linear expectations assumption. The third is the fixed residual covariance matrix assumption. The GF tests can be interpreted as examining:

$$H_0: \{ \text{K latent variables, linear expectations, fixed covariance} \}. \quad (5)$$

We conduct experiments to assess the sensitivity to these three components of the joint hypothesis. Accordingly, the following additional hypotheses are examined:

$$H_1: \{ \text{K latent variables, linear expectations} \}, \quad (6)$$

and

$$H_2: \{ \text{K latent variables} \}. \quad (7)$$

To examine H_1, we relax the fixed covariance matrix assumption by estimating the restricted regression system (2), imposing the restrictions (4), by the generalized method of moments [GMM, see Hansen (1982)]. We examine the minimized value of the GMM objective function as a goodness-of-fit statistic. This approach allows the residual covariance matrix to be conditionally heteroskedastic, and thus vary over time as a function of Z_{t-1}. The statistic is asymptotically distributed as a χ^2 variable. We call this the **"GMM1"** test statistic.[7]

To examine H_2 we use the GMM, but relax the assumption that expected

---

[7] The GMM criterion function is the quadratic form Tg'Wg, where g=vec(u'Z/T), u is the matrix of the error terms from system (2) with the restrictions (3) imposed, Z is the matrix of the instruments, T is the sample size, and W is the inverse of the covariance matrix of the orthogonality conditions, g.

returns are formed by a linear regression on Z_{t-1}. This is accomplished by reformulating the restricted model as:

$$
\begin{aligned}
&R_{it} - \Sigma_{j=0}^{K} c_{ij} R_{jt} = \varepsilon_{it}, \\
&1 = \Sigma_{j=0}^{K} c_{ij}; \quad i = K + 1, \ldots, N. \qquad (8) \\
&E(\varepsilon_{it} \mid Z_{t-1}) = 0.
\end{aligned}
$$

Equation (8) is similar to, but less restrictive than tests of mean-variance spanning [Huberman and Kandel (1987)] because the c_{ij}'s are not restricted to be the regression coefficients of the N-K test asset returns on the (K+1) reference assets, as they are in tests of mean-variance spanning. The regression coefficients δ_i do not appear in system (8), and there is no assumption made about the functional relation of expected returns to the lagged instruments. When our tests are based on system (8), we call the test statistic the "GMM2" test statistic.[8]

The GMM tests for H_0 through H_2 are tests against vague alternative hypotheses. It should be possible to obtain more powerful tests using more specific alternatives. Therefore we construct additional tests in which H_0 is the null and H_1 is the alternative hypothesis, for a given number of latent variables. In these tests, we use a statistic described in Eichenbaum, Hansen and Singleton (1988, appendix C), which is similar to a

---

[8] Hansen (1982) shows that sufficient conditions to apply the GMM include the assumption that the data are strictly stationary and ergodic. We assume that the data satisfy the conditions needed to apply the GMM in all of our tests. The matrix C is also assumed to be fixed over time. This is implied by the asset pricing model (1) if the conditional betas are fixed parameters, since the matrix of the c_{ij}'s is $C=\beta_1^{-1}\beta_2$, where β_1 is the K+1 square matrix of the reference asset betas and β_2 is the (K+1)x(N-K) matrix of the remaining asset betas.

likelihood ratio test statistic but is more general. The GMM criterion function is first minimized under H_0, imposing homoskedasticity through a set of auxiliary moment conditions. The moment conditions are of the form: $e = u^{\cdot 2} - \underline{\sigma}^2$, $E(e|Z)=0$, where $\underline{\sigma}^2$ is a vector of the conditional variances of the individual asset returns, which are hypothesized to be constant, and (.2) denotes the element by element squares. Then, the system is estimated under the alternative, H_1. Under H_1 the subset of the moment conditions implied by homoskedasticity are not imposed. The difference of the two quadratic forms is asymptotically distributed as a chi-squared variable, with degrees of freedom equal to the number of additional restrictions imposed under H_0, compared with H_1, less the number of additional parameters. We call this test statistic the ΔJ statistic.[9]

It is more difficult to form similar tests for H_1 against H_2, because there is no readily identifiable subset of moment conditions which hold under H_1 in equation (2) and are relaxed under H_2 in equation (8). Given the empirical results reported below for the tests against vague alternatives however, the motivation for constructing such tests seems limited, and we do not do so.

Of course, it can be hazardous to draw strong inferences from a comparison of p-values for different test statistics. The power of the tests may differ and there could be finite sample problems which distort the results. Our results do suggest that the GMM2 test statistic, which requires

[9] In the ΔJ statistic the weighting matrix under H_1 is fixed at the inverse of the corresponding partition of the covariance matrix of the orthogonality conditions under H_0 [see Eichenbaum, Hansen and Singleton (1988), appendix C]. To keep the size of the system manageable, we introduce additional orthogonality conditions only for the individual assets' conditional variances under H_0, excluding the moment conditions for the off-diagonal terms of the covariance matrix.

fewer asset return equations, may be more powerful than the GMM1 test statistic. Finite sample problems should not be important in our large samples of daily data, but could be important in monthly sample sizes.

Foerster (1987) provides simulation evidence on the finite sample properties of the LRT and of the Lagrange multiplier (LMT) tests for H_0 which we use to help interpret our results. Ferson and Foerster (1990) provide simulation evidence on the finite sample properties of the GMM1 tests. Their evidence suggests that the size of the GMM1 tests should be well-specified, using the asymptotic distribution, for monthly samples only half the size of ours. They found that a two-stage GMM approach, as described in Hansen and Singleton (1982), tends to reject a correct null hypothesis too often while an iterated GMM approach provides more accurate test statistics and has higher power in finite samples. Following Ferson and Foerster (1990), we use an iterated GMM approach in our tests.[10]

II. The Gibbons-Ferson Empirical Application

This section explores the sensitivity of GF's empirical results for the daily returns of the Dow Jones 30 common stocks. GF chose the Dow Jones 30 stocks to avoid predictable patterns in the daily returns which arise spuriously because of infrequent trading. They used a lagged stock index return and a dummy indicator for Mondays as instruments. We extend this design to include additional time series observations and a dummy variable

---

[10] Specifically, we construct the weighting matrix W using the parameter estimates from the n-th stage minimization, use this matrix to find parameters for stage n+1 which minimize the criterion function, and then use the new parameters to update the weighting matrix. The iterations continue until either a minimum value is obtained or the objective function converges.

for the month of January. In addition to the LRT examined by GF, we conduct tests using the LMT, GMM1, ΔJ, and GMM2 test statistics.

A. The Daily Dow Jones 30 Data

The data consist of returns for the thirty stocks in the Dow Jones Industrial Index (DJ30) for the period January 2, 1963 to December 31, 1985.[11] These are obtained from the Center for Research in Security Prices (CRSP) daily files. To highlight the patterns evident in the DJ30 returns, panel A of table I presents the mean daily rates of return and standard errors of the mean for an equal-weighted portfolio of the DJ30 stocks for four subperiods. Averages over all months, January only, February-December, Mondays, and Fridays are reported.

Panel A of table 1 illustrates the day-of-the-week patterns in returns, including the "Monday effect" exploited by GF in their tests. In each subperiod, average Monday returns are negative. Also, Friday returns are greater than average daily returns. Mean daily returns are larger in January than in non-January months, excepting the second subperiod (1969 to 1973), which exhibits a negative average return in January.

---

[11]Securities are included if they are in the Dow Jones Index at the beginning of the period. There is one exception. General Foods was acquired by Phillip Morris on November 2, 1985. Phillip Morris is used instead of General Foods during our fourth subperiod.

B. Test Results

GF specialize equation (2) as

$$R_{it} = \delta_{i0} + \delta_{i1}D_t^m + \delta_{i2}RVM_{t-1} + u_{it}\,, \qquad (9)$$
$$i=1, \ldots, 30, \; t = 1, \ldots, T;$$

where D_t^m is equal to one if day t is Monday and zero otherwise, and RVM_{t-1} is the lagged return on the CRSP value-weighted index. A multivariate regression system is formed by combining the 30 equations. The number of restrictions on the regression system that are implied by equation (4) is (N-K-1)(L-K) where N is the number of assets (N = 30), K is the number of latent variables and L is the number of predetermined variables including the intercept (L = 3). The LRT and LMT statistics used to examine H_0 have an asymptotic chi-square distribution with degrees of freedom equal to the number of restrictions. The GMM goodness-of-fit statistics are asymptotically chi-square with degrees of freedom equal to the number of orthogonality conditions less the number of parameters. For the GMM1 (GMM2) statistic with 30 assets and L instruments, there are 30×L ([30-K-1]×L) orthogonality conditions. There are [(K+1)×L + (N-K-1)] parameters for the GMM1 case, based on equations (2) and (4), and K×(N-K-1) parameters in the GMM2 case, based on equation (8). The ΔJ statistic, used to test H_0 against H_1, has degrees of freedom equal to the number of additional orthogonality conditions under H_0, less the number of additional parameters under H_0. With L instruments and 30 assets, the degrees of freedom in this case is 30×(L-1).

An examination of the unrestricted regressions confirms GF's observation that equation (9) appears to be a reasonably well-specified model for the Dow Jones 30. We examine the individual regressions, but we summarize the regressions in panel B of Table I using an equally-weighted portfolio of the Dow Jones 30 stocks.[12] Like GF, we find that the regressions can detect predictable variation in the returns in the first three subperiods, covering January 1963 to December 1979. (This corresponds roughly to the overall period of GF: August 1962 to December 1980.) However, in the fourth subperiod, the R-squares are smaller and the standard F-test for the regression produces right-tail p-values smaller than .05 in only twelve of the thirty cases. For every regression in each subperiod, the absolute value of the first-order autocorrelation of the residuals is less than .12, although 22% of these were greater than two standard errors from zero. Chow tests indicate that the hypothesis of constant regression parameters within a subperiod is rejected at the .01 level (.05) in only five (fifteen) of the 120 regressions and these rejections occur uniformly across all subperiods. Finally, White's (1980) test shows little evidence of heteroskedasticity in the residuals of (9).[13]

Table II reports the test results for the latent variables models. In subperiods one through three, the LRT does not reject the joint hypothesis H_0, with a single latent variable (K=1). The p-values range from .13 to .65. Aggregating the statistics over 1963 to 1979 by summing across the first

---

[12]Foerster (1987) presents diagnostics for regression models of the Dow Jones 30 daily expected returns, and presents a more detailed analysis of their expected return behavior.

[13]Of course, these diagnostics may be of low power. For example, White's (1980) test only detects heteroskedasticity that is related to the instruments, and the significance of the first order autocorrelation coefficient is not the most powerful test for detecting persistence in regression residuals.

three subperiods yields a p-value of .211. These are very similar to the results obtained by GF. We also compute the Lagrange multiplier test (LMT) statistic. The LRT and the LMT produce virtually identical inferences about H_0, which is consistent with the assumption that the daily sample sizes are large enough for inferences based on the asymptotic distributions.[14] The GMM1 and GMM2 statistics confirm these results, producing no evidence against the single latent variable model in the first three subperiods.

Given that a single latent variable model is not rejected under H_1 in the first three subperiods, the ΔJ statistic is used to test the hypothesis of homoskedasticity in the individual error terms for the DJ30, against the alternative hypothesis of a heteroskedastic single latent variable model ($K=1$, with linear expectations). The tests provide some evidence against homoskedasticity in the second and third subperiods, and the aggregate statistic for the first three subperiods implies a p-value of less than 0.002. Thus, there is evidence against the homoskedasticity assumption used by GF, but the LRT and LMT do not detect the heteroskedasticity.

Very different results are observed in the fourth subperiod (1980 to 1985). In this subperiod the LRT statistic testing H_0 is 82.3 (p-value = .013). As a result of the large statistics in the fourth subperiod, the aggregate (1963 to 1985) LRT and LMT strongly reject the single latent variable under H_0. Additionally, both the GMM1 and the GMM2 test statistics reject (at the 0.05 level) a single latent variable in the fourth subperiod. Thus, the rejections of a single latent variable model in the fourth subperiod are not attributable to the assumption that the residual

[14] Asymptotically the LMT and LRT statistics are equivalent, but the LMT statistic is smaller (produces a larger p-value) in any finite sample.

covariance matrix is fixed or the assumption that the expected returns, conditional on the instruments, are linear regressions with constant coefficients.

The results of the tests in the first panel of Table II which examine the joint hypothesis H_0, with a two latent variable model (K=2), conform very closely to the results of GF. The p-values of the LRT, LMT, GMM1 and the GMM2 test statistics are in excess of .62 in every subperiod and in the combined samples. There is no evidence that more than two-latent variables are necessary to describe the expected returns. The ΔJ statistic produces evidence against homoskedasticity, given a two-latent variable model, which is similar to the evidence that it produces when a single latent variable is assumed. Thus, there is evidence of conditional heteroskedasticity which is not controlled simply by moving from a single latent variable to a two latent variable model of the conditional means.

We extend the GF example, adding to equation (9) a dummy variable to capture the January seasonal in returns:

$$R_{it} = \delta_{i0} + \delta_{i1} D_t^m + \delta_{i2} RVM_{t-1} + \delta_{i3} DJAN_t + u_{it}, \qquad (10)$$
$$i=1, \ldots, 30, \; t = 1, \ldots, T;$$

where $DJAN_t$ equals 1 if date t is in January and zero otherwise. We include the January dummy to check the sensitivity of GF's results to the choice of instruments for the daily DJ30 returns. The regressions for an equally-weighted portfolio of the DJ30 stocks are summarized in panel C of Table I. The regressions (10) show a small increase in the adjusted R-squares relative to the regressions (9). The average adjusted R-square in equation

(9) across the four subperiods is 4.96%. For equation (10), containing the January dummy, it increases to 5.01%. White's specification test and the other diagnostics suggest a well-specified regression model.

The lower panel of Table II reports tests of the K=1 and K=2 models using equation (10). The LRT and LMT tests now reject the K=1 model in each subperiod except the second, and in the overall sample. The GMM1 and GMM2 tests reject the K=1 model in the fourth subperiod, using a .10 significance level. It is interesting that when a January dummy is included as an instrument the LRT and LMT statistics strongly reject the single latent variable model under H_0, which imposes homoskedasticity, while the GMM1 and GMM2 statistics do not provide such strong evidence against a single latent variable under H_1 and H_2. This suggests that the predictable, seasonal variation in the returns that is uncovered by the January dummy may be related to seasonal changes in the conditional covariance matrix.[15] The ΔJ statistic provides further evidence for this interpretation. The test rejects the homoskedasticity hypothesis in nearly every case that we examine.

Given that other studies have rejected single-beta models using different predictive variables and portfolio returns, it is interesting to find that GF's failure to reject such a model for the individual common

---

[15] Only in the second subperiod is the increment to the R-square due to the January dummy not statistically significant. Only in this subperiod is the single latent variable model not rejected under H_0. Observations at the turn of the year do not seem to be particularly influential for the rejections indicated in the fourth subperiod, but they are influential in the first and third subperiods. Omitting the observations for the last trading day of the year and the first five trading days of the subsequent year from the tests and using equation (9) we obtain a p-value, using the LRT for the single latent variable model, of .001 in the fourth subperiod. P-values in excess of .185 are observed in the other three subperiods.

stocks of the DJ30 is reversed by expanding the sample. Thus, the evidence against a single latent variable model is fortified. It is also interesting that the evidence for more than two latent variables is weak. This suggests that the structure of time-varying, conditional expected returns may be captured using relatively simple, linear models.

III. Tests with Portfolio Returns

One might question the generality of tests of asset pricing models using 30 "blue chip" securities. To draw valid inferences about the number of latent variables in the expected returns, the test-assets' risk sensitivities must span all of the relevant risks. GF (p. 231) suggest that "expected returns on the Dow Jones 30 stocks may be better explained by a single-factor model than would the returns on a broader sample of assets." In this section, therefore, we conduct tests using portfolios constructed from a broader sample of assets. Our common stock portfolios are formed on the basis of two common grouping methods: firm size and industry affiliation.

Several studies document biases in the daily portfolio returns of common stocks, especially for portfolios of small stocks. For example, Blume and Stambaugh (1983) document a bid-ask related bias in the returns of equal-weighted portfolios. Keim (1989) and Porter (1990) find that seasonal patterns in returns (e.g. turn of the year, day of the week) are related to systematic concentrations of closing trades at bid and ask prices. Also, spurious cross-correlation of daily portfolio returns at various lags [Reinganum (1982), Lo and MacKinlay (1990a)] may influence regression models, like equations (9) and (10), that contain lagged market returns. In each of these cases, the incidence of the bias is especially important for low-price, thinly-traded stocks that are concentrated in portfolios of small stocks.[16]

We examine both daily and monthly returns for size-based and industry-grouped common stock portfolios. Results for each are summarized below, but we focus on the results for monthly returns. Monthly returns are less susceptible to the biases from bid-ask effects and thin trading. Furthermore, monthly returns should be influenced less by short term, predictable patterns that result from specialist behavior and other "microstructure" effects, than are daily returns.

Market microstructure models typically assume that any predictable patterns due to expected risk premiums are trivial enough to be safely ignored at the short time intervals involved. Studies do this, for example, by assuming that traders are risk neutral. A common setting posits exogenous (e.g. uninformed "liquidity") demand shocks in order to focus on specialist behavior in the presence of asymmetric information, strategic trading and other such issues. The microstructure literature is still in an early stage

---

[16] There is of course, a chance that the DJ30 stocks are susceptible to such biases that might contaminate the inferences drawn from the tests in section II. To explore a possible connection between the seasonal patterns in the DJ30, documented in Table I, and the trading pattern bias discussed in Keim (1989), we examine the DJ30 stocks in the ten days surrounding the turn of the year. The average daily return for the DJ30 stocks over this period was 0.24 percent in 1988-89 and the average daily proportional bid-ask spread (i.e., (ask-bid)/bid prices) for these stocks at the end of 1988 was 0.48 percent. In constrast, for the smallest decile of NYSE stocks, the mean proportional bid-ask spread reported by Keim (1989) is 6.6%. Furthermore, we find no systematic movement of DJ30 closing prices within the spread (i.e. from the bid toward the ask) at the turn of the year. The average "bias," measured as the difference between returns computed with transaction prices and returns computed with bid prices, is a small negative number (-0.006 percent) during this period, and it displays none of the systematic patterns found by Keim (1989) for smaller stocks.

of development. The "exogenous" demands must ultimately come from some endogenous source. Possibly, they are the result of risk averse investors' optimization, and therefore driven by the same underlying factors that determine demands in equilibrium models of expected returns. We find it interesting, therefore, that our tests for the number of latent variables produce broadly similar results for both the daily and the monthly returns.

A. The Portfolio Return Data

The sample of monthly returns consists of common stocks of NYSE firms, beginning in 1928:1 and ending in 1987:12, a total of 720 monthly observations. We conduct the monthly analysis over 240-month subperiods. Ten common stock portfolios are formed according to size deciles, based on the market value of equity outstanding at the beginning of each year. The ten "size" portfolios are value-weighted averages of the firms. (Value-weighing approximates a "buy-and-hold" investment strategy.) The daily size portfolio sample is similar, but the daily data are only available beginning in 1963. When we form portfolios using daily data, we rank the firms by size each year and we weight the individual returns using the previous day's gross relative returns.[17]

We also examine 12 portfolios of NYSE firms grouped by 2-digit SIC

---

[17] When we use daily data, the portfolio returns are computed as the weighted average of the individual gross daily returns, where the weights are the gross relative returns for the previous day. Blume and Stambaugh (1983) and Roll (1983) show that equal weighted portfolio returns are subject to a statistical bias related to bid-ask spreads. The use of a buy-and-hold portfolio reduces the bias; using the lagged gross relative return is an approximation to a buy-and-hold strategy. Foerster and Porter (1990) study the effectiveness of such an approach in reducing the bias in measured portfolio returns and conclude that the approximation to buy-and-hold is accurate enough for our purposes.

industry code. Unlike the size portfolios, the number of firms in each industry portfolio is not (approximately) the same.[18] We include a firm in the portfolio for its industry, in every month for which a return, a price per common share, and the number of shares outstanding is recorded by CRSP. The monthly industry portfolios are value-weighted each month. The daily industry portfolios are from Foerster (1987), and are weighted within each industry by the lagged gross return relatives.

In the monthly size and industry portfolio samples, we also include a long-term government bond and a long-term, low-grade corporate bond (i.e. "junk" bond) portfolio. The junk bond portfolio returns are provided by Ibbotson Associates for 1928-1976 and by Blume, Keim and Patel (1989) for 1977 to 1987. The government bond returns are from CRSP.

B. Selection of Predetermined Variables

With monthly data, an expanded set of predetermined instrumental variables is available. The instruments used below consist of a constant, the lagged return of the CRSP equally-weighted stock index (denoted EW), the detrended average price level of the Standard and Poors 500 stock index (PLEV), the level of the nominal one-month treasury bill rate (TB), the lagged spread between three-month and one-month treasury bills (HB3), and a dummy variable for the month of January (DJAN). We include the January dummy variable to capture seasonal patterns in the returns [Keim (1983)] and for

---

[18] The industry classification follows Sharpe (1982), Breeden, Gibbons and Litzenberger (1989), and others. The number of firms in a portfolio varies from a low of 8 (Services industry before September, 1960) to a high of 300 (Finance/Real Estate in October, 1986). The mean number of firms over the 1959-1987 sample period varies across the industries from 33.6 (Services) to 213.7 (Basic Industries).

further analysis of possible seasonal changes in risk, as suggested by our results for the DJ30 stocks. The motivation for including the other variables and a brief description of each follows.

EW is the one month lagged return of the equal-weighted NYSE index from CRSP. Such a variable may capture a common factor in the autocorrelations of returns, related to mean-reverting behavior in the stock market. Results of Fama and French (1988a) suggest that a common factor explains much of the autocorrelation of stock portfolio returns.

HB3 is the one-month return of a three-month Treasury bill less the one-month return of a one-month bill. Campbell (1987) finds that such measures of the short-maturity term structure can predict monthly returns in both the bond and the stock markets.

PLEV is the inverse of a detrended price level of common stocks, a variable studied by Keim and Stambaugh (1986). Such a variable is highly correlated with the aggregate dividend yield, a variable studied by Fama and French (1988b), Poterba and Summers (1988) and others. Fama and French (1989) argue that dividend yields and related variables may capture cyclical patterns in expected returns related to business conditions.

TB is the nominal, one-month Treasury bill rate. The ability of short-term bills to predict monthly returns of bonds and stocks is documented by Fama and Schwert (1977), Ferson (1989) and others.

The predetermined variables used in the monthly regressions follow previous empirical work on predicting portfolio returns.[19] There is a

---

[19] When we examine daily portfolio returns, the instruments are a constant, dummy variables for Friday, Monday, and January, the return of the CRSP equally-weighted stock index lagged once and thrice, and the return of the equally-weighted index of the Dow Jones 30 stocks lagged twice. This predictive model is examined by Ferson and Keim (1984) over a shorter sample

natural concern about predictability uncovered through collective "data snooping" by a series of researchers.[20] However, some evidence to support the view that the predictability is not spurious is available from studies using international data.[21] Furthermore, Ferson and Harvey (1990) found that beta pricing models using a set of five specific economic risk factors could "explain" much of the predictability in monthly U.S. data. Some theoretical support for the predictability is also available. For example, Grossman (1981) argued that the parameters of the CAPM should be conditional on the prices of assets. Bossaerts and Green (1989) developed a model in which conditional expected returns are inversely-related to the price of an asset. Kandel and Stambaugh (1990) developed a model economy in which a default-related yield spread and a measure of the term structure slope, track time-varying expected risk premiums.[22]

---

period. Our results for more recent data therefore provide some out-of-sample evidence on the specification.

[20] Such concerns are raised by Merton (1985) and Lakonishok and Smidt (1988), and analyzed by Lo and MacKinlay (1990b).

[21] Cutler, Poterba and Summers (1988) found that dividend yields have predictive power for future stock returns in many countries. Campbell and Hamao (1989) found that predictable components of bond and stock returns were highly correlated between the U.S. and Japan. Harvey (1990) found that a related set of lagged instruments for the U.S. had predictive power for stock returns in many countries.

[22] We have replicated many of our tests using alternative choices for the monthly instruments. We replaced the price level variable PLEV with a dividend yield variable, following Fama and French (1988b) and others. We also examined a default-related yield spread similar to Keim and Stambaugh (1986). None of the broad features of the monthly results are affected by these alternative choices of instruments.

C. Regression Estimates and Diagnostics

Our monthly regression model takes the following form:

$$R_{pt} = \alpha_p + \delta_{p1}EW_{t-1} + \delta_{p2}HB3_{t-1} + \delta_{p3}PLEV_{t-1} + \delta_{p4}TB_{t-1} + \delta_{p5}DJAN_t + u_{it}, \quad (11)$$

where $DJAN_t$ equals one if month t is January and zero otherwise.

Table III reports OLS estimates for regression (11), with heteroskedasticity-consistent t-statistics, for selected size-based and each industry-grouped stock portfolio, the government bond and the junk bond portfolio for the 1968-1987 subperiod. (To conserve space, we report estimates only for the most recent subperiod.) The adjusted R-squares of the regressions range from less than 1% to over 14% across the portfolios. The coefficients of the regressions are similar to the findings of other studies over similar periods. For example, the inverse of the price level, PLEV, enters with a positive coefficient and TB enters with a negative coefficient in all of the regressions for 1968-1987. The January dummy is prominent in the regressions for the smaller firms and the Junk bond, but is less important for most of the industry portfolios. The other variables have more complex patterns across the portfolios.

Recall that the restrictions of the latent variable models under H_0 and H_1 [equation (4)] imply that the regression coefficients, including the intercept, for all of the test assets are linear combinations of the coefficients for the reference assets. The rich patterns in the coefficient estimates and the R-squares in Table III suggest that the sample design should provide some power. We conduct tests of various linear hypotheses on these coefficients, which confirm this impression. The tests reject, for most of the instruments and subperiods, the hypotheses that the coefficients are zero or are equal across the portfolios.[23]

Tinic and West (1984), following earlier observations of Rozeff and Kinney (1976), suggest that the expected market risk premium is not different from zero in months other than January. Their tests are based on sample average returns (i.e., estimates of **unconditional** expected returns) and a specific market proxy. If expected risk premiums are zero in months other than January, then the difference between the expected returns on two portfolios should be zero and the regression coefficients for each portfolio on the predetermined instruments should be the same, given that the January dummy variable $DJAN_t = 0$. Such a test is not dependent on a specific market proxy, and it should be more powerful to detect nonzero risk premiums if conditional expected returns move over time.

We conduct tests of the hypothesis of zero expected risk premiums outside of January. The tests are constructed by forming the differences between the returns of each portfolio and a reference portfolio, which is the first asset. The return differences are regressed on a constant and the predetermined variables TB, HB3, EW and PLEV, and these instruments multiplied by a dummy variable for January. The hypothesis is that the coefficients on the variables without the dummy are jointly equal to zero. The hypothesis is examined using a standard, heteroskedasticity-consistent Wald test. The tests strongly reject the hypothesis for each sample of assets in every subperiod, both in the daily and the monthly data. The

---

[23] Similar results were found in the daily data. Examining the coefficients of the predictive regressions in the daily data revealed a more complex pattern of the coefficients across the industry portfolios than in the size portfolios, where many of the coefficients vary smoothly across the size spectrum.

evidence shows that the Dow Jones 30 common stocks, the size and the industry portfolios all display cross-sectional dispersion in expected returns and nonzero expected risk premiums, both in January and in non-January months.

B. Tests of Asset Pricing Models

Table IV summarizes the tests of the latent variable models using monthly data for the size and industry portfolios. In contrast to the case of the DJ30 stocks, the single latent variable model (K=1) is strongly rejected under H_0, in each of the 240-month subperiods, using both the LRT and the LMT. Rejections of the K=2 latent variables models are also indicated, under H_0, and the test results are similar for the size and the industry portfolio samples.[24] There is also evidence against H_0, even when K=3. The LRT and the LMT reject the K=3 model at standard significance levels, although the industry portfolio design provides the stronger evidence. Using daily returns of the size and industry portfolios, we found similar results.[25]

Foerster's (1987) Monte Carlo experiments indicate some tendency for both the LRT and the LMT to reject H_0 too often using samples as large as 480 monthly observations. Adjusting for small sample bias on the order of that suggested by Foerster, the rejections of K=1 latent variable under H_0

---

[24] To assess how influential the junk bond return is for these rejections, we replaced the junk bond portfolio with the high grade corporate bond return series from Ibbotson Associates, as provided by CRSP, and we obtained similar results.

[25] In the daily data there was little evidence against H_0, K=3 using the size portfolio design, but K=3 was rejected at standard significance levels using the industry portfolios.

reported in Table III are not reversed.[26]

The right-hand columns of Table IV summarize the tests for H_1 and H_2, using the GMM1 and GMM2 test statistics respectively. Recall that these tests examine latent variable models, with K=1,2 and 3 factors, while relaxing the assumptions that the conditional covariance matrix of the returns is fixed (under both H_1 and H_2) and that the conditional expected returns are given by linear regressions with fixed coefficients (under H_2). The tests reject a single latent variable (K=1) under both H_1 and H_2 at standard significance levels. The rejections are not as dramatic as the rejections under H_0, and they occur only in certain subperiods. There is no strong evidence that more than two or three latent variables are required, under H_1.[27]

The GMM2 test statistic, under H_2, produces smaller p-values in every case than does the GMM1 test statistic, despite the weaker assumptions of H_2 compared with H_1. This suggests that the GMM2 test statistic has higher

---

[26] Foerster (1987) studies a system with K=1, L=7, and N=10 size-based portfolios. With 500 simulations of this system under the null hypothesis H_0, he finds that the largest differences between the asymptotic and the simulated p-values occur near the 0.250 tail area of the asymptotic distribution, where the LMT (LRT) overrejects with tail area of 0.314 (0.322). Using a critical value from the asymptotic distribution for a size of 0.050, the LMT (LRT) produce simulated rejection rates of 0.068 (0.084). When the size of the test is nominally set at 0.010, the simulations reveal no bias in the rejection frequencies. Adjusting the test statistics in Table IV according to Foerster's simulation results does not change our inferences. In our daily data samples, the large numbers of observations imply that the distributions of the test statistics should be well-approximated by the asymptotic distributions.

[27] Given the strong rejections of H_0 against a vague alternative and the evidence against the alternative H_1, we do not report the ΔJ statistics in Table IV.

power.[28] The rejections of K=2 and K=3 implied by the GMM2 test statistics are driven by especially large values of the statistic in the second subperiod (the smallest p-value in the first and third subperiods is 0.09).

The evidence confirms the rejection of a single latent variable model, and shows that the evidence against the model is robust to both conditional heteroskedasticity and to any assumption about the functional form of the conditional expectations. Furthermore, the evidence suggests that when the statistical methods allow for conditional heteroskedasticity in the returns, then a small number of latent variables may be able to capture the predictable variation in the asset returns. This is encouraging for the potential usefulness of simple, linear asset pricing models. However, there is some evidence of more complex patterns in the expected returns, especially in the industry portfolios, and the rejections of the two- and three-latent variable models under H_0, which imposes homoskedasticity, begs further explanation.

C. Extensions of the Tests[29]

We examine several modifications of the latent variable models' restrictions in an attempt to learn more about what is driving the

---

[28] Recall that the GMM1 test examines restrictions on the regressions of the assets' returns on a set of predetermined variables. Since the explanatory power of these regressions is low (see Table III), the error variances are similar in magnitude to the variances of the asset returns. Higher power is expected under the GMM2 test because the test asset returns are "regressed" in this test on the contemporanous values of the reference asset returns, and the error variance is therefore substantially smaller than the error variance of the GMM1 test. A smaller number of equations are also involved for a given asset sample and number of latent variables in the GMM2 test.

[29] The results discussed in this section are available upon request to the authors.

rejections of the single latent variable models and the higher order models under H_0. We first allow the regression intercepts to be unrestricted in each equation. If the model can be rejected with unrestricted intercepts, we conclude that differences across assets in the predictable component of returns is important. If a rejection is not observed when the intercepts are unrestricted, it suggests that the cross-section of unconditional mean returns is important. The results using both the size and industry portfolios with unrestricted intercepts are similar to the restricted-intercept results. This is the case in both the daily and in the monthly data. Differences between the structure of unconditional and conditional return variation do not appear to be a driving factor in the rejections.[30]

A second extension of the tests allows for the possibility that the size portfolios' conditional betas have seasonal shifts. Rogalski and Tinic (1986) and Keim and Stambaugh (1986) suggest that betas may shift in January. Such a model can be examined by a simple modification of the cross-equation restrictions. Assume that the "true" beta coefficients follow a simple switching model: Betas in January (b^J_{ij}), are possibly different from betas during the rest of the year (b_{ij}). We modify the predictive regressions in the same way that we did to test for nonzero expected risk premiums in months other than January. Specifically, the regressions now include the original instruments and additional interactive terms that result from mulitplying each of the original instruments by the January dummy. The restrictions on the coefficients of these regressions

[30] Campbell (1987, table 4) also finds little difference in tests with the intercept restricted or unrestricted, using a different sample of monthly returns and instruments.

[equation (4)] are modified. We estimate, possibly different, relative beta coefficients (c^{J}_{ij}) in January and for the other months (c^{NJ}_{ij}). In tests using either daily or monthly data, we find that the models with more than one latent variable (K=2 and K=3) are no longer rejected under H_0 when betas are allowed to be different in January.[31] Thus, the rejections of models with more than a single latent variable for the size portfolios, under H_0, can be explained by allowing for seasonal shifts in the conditional betas.

IV. Concluding Remarks

This paper examines the behavior of conditional expected returns over time on common stocks and bonds, extending the methods of Gibbons and Ferson (1985). We include an expanded sample of daily returns, monthly returns and instrumental variables for the market's predetermined conditioning information. We generalize the test methodology to examine the separate elements of the joint hypothesis, in order to determine the source of rejections. Gibbons and Ferson (1985) did not reject a single latent variable model, but we find that a single latent variable model can be rejected for the same sample of assets studied by Gibbons and Ferson (the Dow Jones 30 common stocks), when the time period is extended. The rejections are robust to conditional heteroskedasticity and to any assumption about the functional form of conditional expected returns.

The evidence suggests that two time-varying premiums are needed to "explain" the expected returns of the Dow Jones 30 common stocks, in a constant-beta model. In tests using size and industry portfolios, with

---

[31] The only exception to this is the 1963-68 subperiod for the DJ30, where a two-latent variable model is rejected at standard significance levels.

either monthly or daily data, we find similar results. The tests indicate that more than a single premium is needed to model expected returns, but there is no evidence of more than two or three latent variables in the time-varying expected returns. If the conditional betas of the size portfolios are allowed to display a January seasonal, we reject a single-premium model but do not reject a two premium model. Finally, there is evidence that expected risk premiums are nonzero and differ across portfolios, even in months other than January.

We interpret our results as optimistic for the potential usefulness of simple, linear asset pricing theories to capture the predictable variation of security returns over time. A small number of sources of predictable variation seems to be indicated, and interestingly, the results for monthly and daily returns are broadly similar.

REFERENCES

Black, F., 1972, Capital market equilibrium with restricted borrowing, Journal of Business 45, 444-454.

Blume, M., D.B. Keim and S.A. Patel, 1990, Returns and Volatility of Low-grade Bonds, 1977-1989, forthcoming, Journal of Finance.

Blume, M. and R. Stambaugh, 1983, Biases in computed returns: An application to the size effect, Journal of Financial Economics 12, 387-404.

Bossaerts, P. and R.C. Green, 1989, A General Equilibrium Model of Changing Risk Premia: Theory and Tests, Review of Financial Studies 2, 467-494.

Breeden, D., 1979, An intertemporal asset pricing model with stochastic consumption and investment opportunities, Journal of Financial Economics 7, 265-296.

Breeden, D., M. Gibbons, and R. Litzenberger, 1989, Empirical tests of the consumption-oriented CAPM, Journal of Finance 44, 231-262.

Campbell, J., 1987, Stock returns and the term structure, Journal of Financial Economics 18, 373-399.

Campbell, J. and Y. Hamao, 1989, Predictable bond and stock returns in the United States and Japan: A study of long-term capital market integration, Unpublished working paper (Princeton University).

Chan, K. C., 1988, Autocorrelation of stock returns and market efficiency, unpublished paper, Ohio State University.

Connor, G. and R. A. Korajczyck, 1989, An intertemporal equilibrium beta pricing model, Review of Financial Studies 2, 373-392.

Cutler, D. M., J. M. Poterba and L. H. Summers, 1988, International Evidence on the Predictability of Stock Returns," Mimeographed. Boston: Massachusetts Institute of Technology.

Eichenbaum, M. S., L. P. Hansen, and K. J. Singleton, 1988, A Time Series Analysis of Representative Agent Models of Consumption and Leisure Choices under Uncertainty, Quarterly Journal of Economics 103, 51-78.

Fama, E. and K. French, 1988a, Permanent and temporary components of stock prices, Journal of Political Economy 96, 246-273.

Fama, Eugene F. and French, Kenneth R., 1988b, Dividend Yields and Expected Stock Returns, Journal of Financial Economics 22, 3-25.

Fama, Eugene F. and French, Kenneth R., 1989, Business Conditions and Expected Stock Returns," Journal of Financial Economics 25, 23-50.

Fama, E. and G. Schwert, 1977, Asset returns and inflation, Journal of

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

Ferson, W. E., 1990, Are the latent variable in time-varying expected returns compensation for consumption risk?, Journal of Finance, 45, 397-430.

Ferson, W. E., 1989, Changes in expected security returns, risk, and the level of interest rates, Journal of Finance 44, 1191-1218.

Ferson, W. E. and Foerster, S. R., 1990, Finite sample properties of the Generalized Method of Moments is Latent variable tests of asset pricing models, working paper, University of Chicago and University of Western Ontario.

Ferson, W. E. and C. Harvey, 1990, Seasonality and Consumption-based Asset Pricing, working paper, Duke University and the University of Chicago.

Ferson, W. E. and D. Keim, 1984, Empirical regularities in stock returns involving day, size and season, Proceedings of the Seminar on the Analysis of Security Prices, Center for Research in Security Prices, University of Chicago, May.

Foerster, S. R., 1987, Asset pricing models with changing expectations: An empirical study, unpublished dissertation (The Wharton School, Phila., PA).

Foerster, S. R. and D. Porter, 1990, Calendar- and Size-based anomalies in Canadian Stock Returns, Unpublished manuscript (University of Western Ontario).

Gibbons, M. and W. Ferson, 1985, Testing asset pricing models with changing expectations and an unobservable market portfolio, Journal of Financial Economics 14, 217-236.

Grossman, S., 1981, Further results on the informational efficiency of stock markets, working paper, Stanford University.

Hansen, L. P., 1982, Large Sample Properties of Generalized Method of Moments Estimators, Econometrica 50, 1029-1054.

Hansen, L.P., and K. Singleton, 1982, Generalized instrumental variables estimation of nonlinear rational expectations models, Econometrica 50, 1269-1285.

Harvey, C. R., 1990, The world price of covariance risk, working paper, Duke University.

Huberman, G. and S. Kandel, 1987, Mean variance spanning, Journal of Finance 42, 383-88.

Kandel, S. and R. Stambaugh, 1990, Modelling Expected Stock Returns for Long and Short Horizons, Review of Financial Studies (forthcoming).

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

Keim, D., 1989, Trading patterns, bid-ask spreads, and estimated security returns: the case of common stocks at calendar turning points, Journal of Finance 45, 75-97.

Keim, D. and R. Stambaugh, 1986, Predicting returns in the stock and bond markets, Journal of Financial Economics 17, 357-390.

Lakonishok, J. and S. Smidt, 1988, Are seasonal anomalies real? A ninety year perspective, Review of Financial Studies 1, 403-425.

Lehmann, Bruce N., 1990, Notes on Dynamic Factor Pricing Models, working paper, Columbia University.

Lo, A. W. and A. C. MacKinlay, 1990a, When are contrarian profits due to stock market overreaction? Review of Financial Studies 3, 175-206.

Lo, A. W. and A. C. MacKinlay, 1990b, Data Snooping and Biases in Tests of Financial Models, Review of Financial Studies 3, 431-468.

Long, J., 1974, Stock prices, inflation, and the term structure of interest rates, Journal of Financial Economics 1, 131-170.

Merton, R., 1973, An intertemporal capital asset pricing model, Econometrica 41, 867-887.

Merton, R. C., 1985, On the current state of the stock market rationality hypothesis, in R. Dornbusch and S. Fischer (eds.) Macroeconomics and Finance, MIT press, Cambridge, Ma.

Porter, D., 1990, The probability of a trade at the ask: an examination of interday and intraday behavior, working paper, Marquette University.

Poterba, J. and L. Summers, 1988, Mean reversion in stock prices: evidence and implications, Journal of Financial Economics 22, 27-60.

Reinganum, M., 1982, A direct test of Roll's conjecture on the firm size effect, Journal of Finance 37, 27-36.

Rogalski, R. and S. Tinic, 1986, The January size effect: Anomaly or risk mismeasurement?, Financial Analysts Journal, November-December 1986.

Roll, R., 1977, A critique of the asset pricing theory's tests; Part I: On past and potential testability of the theory, Journal of Financial Economics 4, 129-176.

Roll, R., 1983, On computing mean returns and the small firm premium, Journal of Financial Economics 12, 371-386.

Ross, S., 1976, The arbitrage theory of capital market asset pricing, Journal of Economic Theory 13, 341-360.

Rozeff, M. and W. Kinney, 1976, Capital market seasonality: The case of stock returns, Journal of Financial Economics 3, 379-402.

Sharpe, W., 1964, Capital asset prices: A theory of market equilibrium under conditions of risk, Journal of Finance, 19, 425-442.

Sharpe, William F., 1982, Some Factors in New York Stock Exchange security returns, 1931-1979." Journal of Portfolio Management 8, 5-19.

Stambaugh, R., 1988, The information in forward rates: Implications for models of the term structure, Journal of Financial Economics 21, 41-70.

Tinic, S., and R. West, 1984, Risk and return: January vs. the rest of the year, Journal of Financial Economics 13, 561-574.

Wheatley, S., 1989, "A Critique of latent variable tests of asset pricing models," Journal of Financial Economics 23, 325-338.

White, H., 1980, A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity, Econometrica 48, 817-838.

TABLE I

Summary statistics and predictive regression results for daily rates of return for an equally-weighted portfolio of the Dow Jones 30 Common Stocks, including overall, January, February-December, Monday and Friday averages for each of four subperiods.[a]

Panel A: Mean Daily Return × 10^4 (standard error × 10^4)

Time Period	1963 to 1968		1969 to 1973		1974 to 1979		1980 to 1985	
All days	4.75	(1.49)	0.99	(2.46)	3.48	(2.47)	6.58	(2.42)
January	13.46	(4.52)	-3.16	(6.46)	22.47	(9.58)	13.54	(9.49)
Feb-Dec	3.93	(1.58)	1.37	(2.62)	1.71	(2.55)	5.95	(2.50)
Monday	-10.02	(3.54)	-23.09	(6.04)	-0.56	(6.05)	-2.70	(6.27)
Friday	9.35	(2.78)	13.10	(5.21)	9.23	(5.23)	11.27	(4.82)

Panel B: Regression Results for the regression model:[b]

$$R_{it} = \delta_{i0} + \delta_{i1}D_t^m + \delta_{i2}RVM_{t-1} + u_{it}, \quad t = 1, \ldots, T. \tag{9}$$

Time Period	1963 to 1968	1969 to 1973	1974 to 1979	1980 to 1985
Adjusted R-squared	.0438	.1068	.0414	.0064
ρ_1	.009	.050	-.001	-.009

Panel C: Regression Results for the regression model:[c]

$$R_{it} = \delta_{i0} + \delta_{i1}D_t^m + \delta_{i2}RVM_{t-1} + \delta_{i3}DJAN_t + u_{it}, \quad t = 1, \ldots, T. \tag{10}$$

Time Period	1963 to 1968	1969 to 1973	1974 to 1979	1980 to 1985
Adjusted R-squared	.0446	.1062	.0434	.0060
ρ_1	.010	.050	-.001	-.009

[a] The subperiods have 1484, 1258, 1514 and 1504 observations, respectively.
[b] D_t^m is equal to one if day t is Monday and zero otherwise, and RVM_{t-1} is the lagged return on the CRSP value-weighted index. ρ_1 is the first order autocorelation of the regression residual.
[c] $DJAN_t$ is a dummy variable equal to one if day t is in January and zero otherwise.

TABLE II

Tests of Asset Pricing Models with K=1 and K=2 Latent Variables.

Daily Data for the Dow Jones 30 Common Stocks are used. The model is:

$$r_1 = Z\delta_1 + \epsilon_1$$

$$r_2 = Z\delta_1 C + \epsilon_2,$$

$$\ell' C = \ell,$$

where $r=(r_1, r_2)$ is a vector of monthly returns, Z is a vector of predetermined instrumental variables, and ℓ is a vector of ones. δ_1 and C are parameters. The number of observations for the subperiods are:

1963-68	1487
1969-73	1262
1974-79	1518
1980-85	1519

The LRT and LMT statistics assume that the covariance matrix of the error terms is fixed. The GMM1 statistic is the minimized value of the Generalized method of moments criterion function for the system, based on the implication of the model that the error term $u=(\epsilon_1, \epsilon_2)$ has conditional mean zero given the instruments Z. The orthogonality condition tested is $E(u'Z)=0$. The ΔJ statistic examines the null hypothesis of the LRT and LMT tests, against the alternative hypothesis, H_1, by appending additional moment conditions to the system. The moment conditions are of the form: $e = u^2 - \sigma^2$, $E(e|Z)=0$, where σ^2 is a vector of the conditional variances of the individual asset returns, which are hypothesized to be constant, and (.2) denotes the element by element squares. The GMM2 statistic is the value of GMM criterion function for the reformulated model:

$$\eta = r_2 - r_1 C,$$

$$\ell' C = \ell.$$

In the first panel, the instruments are a constant, the lagged return on a value weighted stock index and a dummy variable for Monday. In the second panel, a dummy variable for the month of January is included as an additional instrument. The right-tail probability values for the test statistics are reported in the table.

No. Latent Vars.	Subperiod	LRT	LMT	ΔJ	GMM1	GMM2
1	1963-68	0.125	0.143	0.586	0.362	0.291
1	1969-73	0.645	0.666	0.069	0.675	0.268
1	1974-79	0.254	0.273	0.000	0.565	0.350
1	1980-85	0.013	0.017	na	0.036	0.030
1	1963-85[a]	0.000	0.000	0.001[c]	0.221	0.048
2	1963-68	0.977	0.978	0.455	0.987	0.987
2	1969-73	0.863	0.869	0.088	0.940	0.907
2	1974-79	0.621	0.631	0.000	0.802	0.724
2	1980-85	0.903	0.907	0.020	0.980	0.972
2	1963-85[a]	0.992	0.993	0.000	0.999	0.999

Tests including a January Dummy Variable in the Instrument Set:

No. Latent Vars.	Subperiod	LRT	LMT	ΔJ	GMM1	GMM2
1	1963-68	0.000	0.000	0.018	0.497	0.436
1	1969-73	0.846	0.861	0.023	0.703	0.344
1	1974-79	0.002	0.003	0.001	0.329	0.077
1	1980-85	0.022	0.030	na	0.096	0.005
1	1963-85[a]	0.000	0.000	0.000[c]	0.281	0.011
2	1963-68	0.001	0.002	0.009	0.976	0.614
2	1969-73	0.926	0.931	0.000	0.890	0.839
2	1974-79	0.268	0.288	0.000	0.681	0.257
2	1980-85	0.944	0.948	0.000	0.991	0.862
2	1963-85[a]	0.999	0.999	0.000	0.999	0.814

[a] The aggregate test statistic sums the chi-square values across the subperiods.
[b] The ΔJ statistic is not applicable when the alternative hypothesis is rejected by the GMM1 test statistic.
[c] The aggregate test statistic is based on the first three subperiods only.

TABLE III

Predictive regression results for monthly rates of return for size- and industry-grouped common stock portfolios. The data are for 1968-1987 (240 observations). Ordinary Least squares regression coefficients are shown, with heteroskedasticity-consistent t-statistics in parentheses.

Regression Results for the regression model:[a]

$$R_{pt} = \alpha_p + \delta_{p1}TB_{t-1} + \delta_{p2}HB3_{t-1} + \delta_{p3}EW_{t-1} + \delta_{p4}PLEV_{t-1} + \delta_{p5}DJAN_t + u_{it}, \quad (11)$$

Portfolio[b]	TB	HB3	EW	PLEV	DJAN	adj. R^2	ρ_1
Decile 1	-0.50 (-2.18)	-0.17 (-0.20)	0.16 (1.67)	0.04 (2.14)	0.10 (3.01)	14.0%	-0.04
Decile 2	-0.51 (-2.36)	0.46 (0.59)	0.13 (1.84)	0.05 (2.65)	0.07 (2.79)	12.4	-0.00
Decile 5	-0.48 (-2.67)	0.57 (0.79)	0.11 (1.89)	0.04 (2.58)	0.04 (2.04)	7.5	0.01
Decile 10	-0.36 (-2.33)	0.95 (1.42)	0.04 (0.72)	0.02 (1.83)	0.01 (0.05)	2.7	-0.07
Petroleum	-0.43 (-1.8)	0.58 (0.58)	0.02 (0.27)	0.02 (1.66)	-0.01 (-0.49)	0.7	-0.03
Fin/RE	-0.33 (-1.90)	0.49 (0.75)	0.01 (0.22)	0.03 (1.79)	0.01 (0.27)	0.7	0.05
Cons. Dur.	-0.54 (-3.07)	1.51 (1.82)	0.11 (1.97)	0.04 (2.53)	0.01 (0.60)	6.9	-0.01
Basic Ind.	-0.36 (-2.17)	1.01 (1.51)	0.06 (0.96)	0.02 (1.39)	0.01 (0.39)	2.0	-0.08
Food/Tob.	-0.23 (-1.50)	1.35 (2.37)	0.03 (0.42)	0.02 (1.50)	0.00 (0.34)	2.3	0.04
Constr.	-0.61 (-3.28)	0.91 (1.15)	0.13 (2.28)	0.04 (2.32)	0.02 (0.85)	5.9	-0.00
Cap. Goods	-0.50 (-3.10)	1.20 (1.55)	0.12 (1.72)	0.03 (1.96)	0.01 (0.45)	5.6	-0.03
Trans.	-0.51 (-2.55)	1.52 (1.80)	0.12 (1.67)	0.03 (2.05)	0.02 (1.04)	5.4	-0.01
Utilities	-0.19 (-1.52)	0.19 (0.43)	-0.05 (-0.95)	0.02 (2.23)	0.02 (1.69)	2.9	0.05
Textiles /Trade	-0.29 (-1.66)	1.19 (1.58)	0.08 (1.17)	0.02 (1.54)	0.02 (0.77)	2.0	0.11
Services	-0.50 (-2.17)	1.89 (2.12)	0.11 (1.47)	0.04 (2.19)	0.02 (0.75)	5.3	0.05
Leisure	-0.54 (-2.57)	1.76 (2.14)	0.17 (2.21)	0.04 (2.19)	0.01 (0.26)	6.3	0.07
Govt. Bond	-0.04 (-0.26)	0.05 (0.98)	-0.07 (-1.82)	0.01 (1.67)	-0.01 (-1.18)	1.1	0.03
Junk Bond	-0.12 (-1.14)	-0.40 (-1.12)	0.01 (0.39)	0.02 (3.24)	0.02 (3.47)	5.0	0.13

[a] EW is the one month lagged return of the equal-weighted NYSE index from CRSP. HB3 is the one-month return of a three-month Treasury bill less the one-month return of a one-month bill. PLEV is the inverse of a detrended price level of the S&P 500 stock index. TB is the nominal, one-month Treasury bill rate. DJAN is a dummy variable equal to one if day t is in January and zero otherwise. adj. R^2 is the adjusted R-square and ρ_1 is the first order autocorrelation of the regression residual.

[b] Decile 1 is the smallest common stock portfolio and Decile 10 is the largest stock portfolio; a subset of the ten decile portfolios are shown.

TABLE IV

Tests of Asset Pricing Models with K=1, 2 and 3 Latent Variables.

The model is:

$$r_1 = Z\delta_1 + \varepsilon_1$$

$$r_2 = Z\delta_1 C + \varepsilon_2,$$

$$\ell' C = \ell,$$

where $r=(r_1, r_2)$ is a vector of monthly returns, Z is a vector of predetermined instrumental variables, and ℓ is a vector of ones. δ_1 and C are parameters. Each twenty-year subperiod has 240 monthly observations. The LRT and LMT statistics assume that the covariance matrix of the error terms is fixed. The GMM1 statistic is the minimized value of the Generalized method of moments criterion function for the system, based on the implication of the model that the error term $u=(\varepsilon_1, \varepsilon_2)$ has conditional mean zero given the instruments Z. The orthogonality condition tested is $E(u'Z)=0$. The GMM2 statistic is the value of GMM criterion function for the reformulated model:

$$\eta = r_2 - r_1 C,$$

$$\ell' C = \ell.$$

The instruments are a constant, the detrended price level of the Standard and Poors stock index, the level of the one-month treasury bill, the lagged spread of a three-month over a one-month bill, the lagged return of the CRSP equally-weighted stock index, and a dummy variable for the month of January. The right-tail probability values for the test statistics are reported in the table.

No. Latent Varis.	Subperiod	Assets	LRT	LMT	GMM1	GMM2
1	1928-47	10 Size, GB,	0.000	0.000	0.058	0.049
1	1948-67	Junkret	0.000	0.000	0.053	0.032
1	1968-87		0.001	0.003	0.252	0.198
1	1928-87[a]		0.000	0.000	0.015	0.007
1	1928-47	12 Industry,	0.000	0.000	0.301	0.252
1	1948-67	GB, Junkret	0.000	0.000	0.063	0.004
1	1968-87		0.009	0.024	0.369	0.227
1	1928-87[a]		0.000	0.000	0.091	0.009
2	1928-87[a]	10 Size, GB, Junkret	0.000	0.000	0.208	0.033
2	1928-87[a]	12 Industry, GB, Junkret	0.000	0.000	0.667	0.013
3	1928-87[a]	10 Size, GB, Junkret	0.019	0.031	0.567	0.069
3	1928-87[a]	12 Industry, GB, Junkret	0.001	0.001	0.241	0.017

[a] The aggregate test statistic sums the chi-square values across the subperiods. Only the aggregate statistic is shown for the K=2 and K=3 models.